

The

Languages of  
Japan and Korea

Edited by  
Nicolas Tranter

ROUTLEDGE LANGUAGE FAMILY SERIES



---

---

THE

---

---

LANGUAGES OF

---

---

JAPAN AND KOREA

---

---

*The Languages of Japan and Korea* provides detailed descriptions of the major varieties of languages in the region, both modern and pre-modern, within a common format, producing a long-needed introductory reference source. Korean, Japanese, Ainu, and representative members of the main groupings of the Ryukyuan chain are discussed for the first time in great detail in a single work.

The volume is divided into language sketches, the majority of which are broken down into sections on phonology, orthography, morphology, syntax and lexicon. Specific emphasis is placed on aspects of syntactic interest, including speech levels, honorifics and classifiers. Each language variety is represented in Roman-based transcription, although its own script (where there is such orthography) and IPA transcriptions are used sparingly where appropriate. The dialects of both the modern and oldest forms of the languages are given extensive treatment, with a primary focus on the differences from the standard language. These synchronic snapshots are complemented by a discussion of both the genetic and areal relationships between languages in the region.

With contributions from a variety of scholars of the highest reputation, *The Languages of Japan and Korea* is a much needed and highly useful tool for professionals and students in linguistics, as well as area studies specialists.

**Nicolas Tranter** is Lecturer at the University of Sheffield, teaching courses on East Asian languages and linguistics.

## ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

The Austronesian Languages of  
Asia and Madagascar  
*Edited by Nikolaus Himmelmann  
& Sander Adelaar*

The Bantu Languages  
*Edited by Derek Nurse &  
Gérard Philippson*

The Celtic Languages 2<sup>nd</sup> Edition  
*Martin J. Ball & James Fife*

The Dravidian Languages  
*Edited by Sanford B. Steever*

The Germanic Languages  
*Edited by Ekkehard König &  
Johan van der Auwera*

The Indo-Aryan Languages  
*Edited by George Cardona &  
Dhanesh K. Jain*

The Indo-European Languages  
*Edited by Paolo Ramat &  
Anna Giacalone Ramat*

The Iranian Languages  
*Edited by Gernot Windfuhr*

The Khoesan languages (forthcoming)  
*Edited by Rainer Vossen*

The Languages of Japan and Korea  
*Edited by Nicolas Tranter*

The Mongolic Languages  
*Edited by Juha Janhunen*

The Munda Languages  
*Edited by Gregory D.S. Anderson*

The Oceanic Languages  
*Edited by John Lynch,  
Malcolm Ross  
& Terry Crowley*

The Romance Languages  
*Edited by Martin Harris &  
Nigel Vincent*

The Semitic Languages  
*Edited by Robert Hetzron*

The Sino-Tibetan Languages  
*Edited by Graham Thurgood &  
Randy J. Lapolla*

The Slavonic Languages  
*Bernard Comrie &  
Greville G. Corbett*

The Tai-Kadai Languages  
*Edited by Anthony Diller*

The Turkic Languages  
*Edited by Éva Csató &  
Lars Johanson*

The Uralic Languages  
*Edited by Daniel Abondolo*

THE  
LANGUAGES OF  
JAPAN AND KOREA

Edited by  
Nicolas Tranter

 **Routledge**  
Taylor & Francis Group  
LONDON AND NEW YORK

First published 2012  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada  
by Routledge  
711 Third Avenue, New York, NY 10017

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2012 Nicolas Tranter, selection and editorial matter; individual chapters, the contributors

The right of the editor to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*British Library Cataloguing in Publication Data*

A catalogue record for this book is available from the British Library

*Library of Congress Cataloging in Publication Data*

The languages of Japan and Korea / edited by Nicolas Tranter.

p. cm. — (Routledge language family series)

1. Ryukyu islands—Languages. 2. Japanese language—Grammar. 3. Korean language—Grammar. 4. Japan—Languages. 5. Korea—Languages. I. Tranter, Nicolas.

PL693.R9L36 2012  
495.6'7—dc23

2011037313

ISBN: 978-0-415-46287-7 (hbk)  
ISBN: 978-0-203-12474-1 (ebk)

Typeset in 9.5/11pt Times New Roman  
by Graphicraft Limited, Hong Kong

---

---

# CONTENTS

---

---

<i>List of illustrations</i>	vi
<i>List of contributors</i>	x
<i>Preface</i>	xi
<i>Acknowledgements</i>	xii
<i>List of abbreviations</i>	xiii
<i>Transcription systems</i>	xv
<b>PART I: PRELIMINARIES</b>	<b>1</b>
1 Introduction: typology and area in Japan and Korea <i>Nicolas Tranter</i>	3
2 The relationship between Japanese and Korean <i>John Whitman</i>	24
<b>PART II: KOREAN</b>	<b>39</b>
3 Old Korean <i>Nam Pung-hyun</i>	41
4 Middle Korean <i>Ho-min Sohn</i>	73
5 Modern Korean <i>Young-Key Kim-Renaud</i>	123
6 Korean dialects: a general survey <i>Jaehoon Yeon</i>	168
<b>PART III: JAPANESE</b>	<b>187</b>
7 Old Japanese <i>John R. Bentley</i>	189
8 Classical Japanese <i>Nicolas Tranter</i>	212
9 Late Middle Japanese <i>Mark Irwin and Heiko Narrog</i>	246
10 Modern Japanese <i>Nicolas Tranter and Mika Kizu</i>	268
11 Japanese dialects: focusing on Tsuruoka and Ei <i>Akiko Matsumori and Takuichiro Onishi</i>	313
<b>PART IV: RYUKYUAN</b>	<b>349</b>
12 Northern Ryukyuan <i>Michinori Shimoji</i>	351
13 Southern Ryukyuan <i>Wayne P. Lawrence</i>	381
14 Yonaguni <i>Atsuko Izuyama</i>	412
<b>PART V: AINU</b>	<b>459</b>
15 Southern Hokkaido Ainu <i>Anna Bugaeva</i>	461
<i>Index</i>	510

---

---

# ILLUSTRATIONS

---

---

## FIGURE

Figure 1	Dot-type kugyŏl	47
----------	-----------------	----

## TABLES

Table 0.1	Modern Japanese transcription systems where differences occur	xv
Table 0.2	Modern Korean transcription systems (South Korean han'gŭl alphabetical order)	xvi
Table 2.1	Proto-Japanese-Ryukyuan vowels	26
Table 2.2	Proto-Japanese-Ryukyuan *e	26
Table 2.3	Proto-Japanese-Ryukyuan *o	26
Table 2.4	Proto-Japanese-Ryukyuan consonants	27
Table 2.5	Old Korean (Silla) vowel system	28
Table 2.6	Proto-Korean consonants	28
Table 2.7	Obstruent lenition attested in Late Middle Korean	29
Table 2.8	Hypothesized lenitions for pre-Late Middle Korean	29
Table 2.9	Proto-Korean :: proto-Japanese-Ryukyuan vowel correspondences	30
Table 2.10	Proto-Japanese-Ryukyuan correspondences for proto-Korean *h	30
Table 2.11	Proto-Japanese-Ryukyuan correspondences for proto-Korean *c	31
Table 2.12	Syncopated proto-Korean *i, *ə	32
Table 2.13	Proto-Korean medial consonant lenition	32
Table 2.14	Proto-Japanese-Korean numerals	33
Table 2.15	Proto-Japanese-Korean pronouns	33
Table 2.16	Proto-Japanese-Korean verbal affixes and postverbal particles	34
Table 2.17	Proto-Japanese-Korean postnominal particles	34
Table 3.1	Dot-type kugyŏl systems of <i>Chubon Hwaŏm-gyŏng</i> and <i>Yugasaji-ron</i>	48
Table 3.2	Adaptation of Chinese characters to Old Korean	49
Table 3.3	Interpretive kugyŏl characters	52
Table 3.4	Mid Old Korean case and discourse particles	58
Table 3.5	Mid Old Korean verb enders	59
Table 3.6	Mid Old Korean auxiliary roots	60
Table 3.7	Mid Old Korean derivative suffixes	61
Table 3.8	Mid Old Korean auxiliary verbs	61
Table 3.9	Mid Old Korean quasi-grammatical forms	61
Table 3.10	Late Old Korean case and discourse particles	65
Table 3.11	Late Old Korean verb enders	66
Table 3.12	Late Old Korean auxiliary roots	67
Table 3.13	Late Old Korean derivative suffixes	68

Table 3.14	Late Old Korean auxiliary verbs	68
Table 4.1	Late Middle Korean consonants	78
Table 4.2	Late Middle Korean case particles	88
Table 4.3	Late Middle Korean demonstrative pronouns	98
Table 4.4	Late Middle Korean indefinite-interrogative pronouns	98
Table 4.5	Late Middle Korean speech-level sentence enders	111
Table 4.6	Late Middle Korean relativizer enders	113
Table 5.1	Consonants of Modern Korean	125
Table 5.2	Three-way contrast among Korean voiceless stops	125
Table 5.3	Korean consonantal groups according to their strength	125
Table 5.4	Vowel inventory of Modern Korean	126
Table 5.5	Korean diphthongs	126
Table 5.6	Meaning according to the intonation contour in Korean	127
Table 5.7	Allomorphy in Korean noun particles	136
Table 5.8	Allomorphy in Korean verb suffixes	139
Table 5.9	Korean sentential modifier endings	144
Table 5.10	Clause/sentence typing particles in Korean	151
Table 7.1	Old Japanese versus Eastern Old Japanese	190
Table 7.2	Old Japanese consonants	191
Table 7.3	Old Japanese vowels	191
Table 7.4	S-series in <i>Kojiki</i> and <i>Nihon shoki</i>	192
Table 7.5	Different strata of Chinese in Japanese records	194
Table 7.6	Old Japanese verb classes	197
Table 7.7	Old Japanese stative verbs	198
Table 7.8	Old Japanese numerals	199
Table 7.9	Noun versus classifier in Old Japanese	199
Table 7.10	Old Japanese pronouns	202
Table 7.11	Old Japanese imperfect suffixes	203
Table 7.12	Old Japanese perfect suffixes	204
Table 7.13	Old Japanese suffixes of possibility	205
Table 7.14	Old Japanese honorific verbs	208
Table 7.15	Paekche loans in Old Japanese	211
Table 8.1	Early Middle Japanese consonants	213
Table 8.2	Two-mora noun accent in Japanese-Ryukyuan	216
Table 8.3	Classical Japanese onbin	217
Table 8.4	Examples of <i>t</i> -sandhi in Classical Japanese quadrigrade verbs	218
Table 8.5	Variety of Middle Japanese orthographies	220
Table 8.6	Classical Japanese verb morphology	224
Table 9.1	Late Middle Japanese consonant phonemes	248
Table 9.2	Late Middle Japanese verb inflection	255
Table 9.3	Late Middle Japanese inflecting derivational suffixes	256
Table 10.1	Japanese consonants	269
Table 10.2	Japanese vowels	270
Table 10.3	Accent in compounds in Japanese	272
Table 10.4	Hiragana katakana syllabaries	276
Table 10.5	Japanese particles	280
Table 10.6	Japanese adjective morphology	281
Table 10.7	Polite forms of Japanese adjectives	281
Table 10.8	Japanese quinquagrade and monograde verbs and <i>-[a]nai</i>	282

Table 10.9	Japanese verb sandhi	283
Table 10.10	Japanese irregular verbs	283
Table 10.11	Older inflection of <i>dasu</i> ‘to take out’	284
Table 10.12	Japanese epistemic and deontic/evaluative modal expressions (for <i>yomu</i> ‘to read’)	293
Table 10.13	Japanese modality of utterance expressions	294
Table 10.14	Examples of other Japanese expression types using <i>toru</i> ‘to take’	296
Table 10.15	Japanese lexical causatives and their intransitive counterparts	300
Table 10.16	Japanese honorific and humble verbs (e.g. <i>yomi-masu</i> ‘read-POL’)	301
Table 10.17	Concordance between adverbial and predicate in Japanese	303
Table 11.1	Consonants in the Tsuruoka dialect	315
Table 11.2	Accent system of Standard Japanese	316
Table 11.3	Accent system in Tsuruoka dialect	317
Table 11.4	Accent patterns of 2-mora nouns in Standard Japanese and Tsuruoka	317
Table 11.5	Accent patterns of verbs in Tsuruoka dialect (based on Nitta 1994)	318
Table 11.6	Verb conjugation in the Tsuruoka dialect (I)	321
Table 11.7	Verb conjugation in the Tsuruoka dialect (II): irregular verbs	321
Table 11.8	Verb conjugation of irregular verbs in Standard Japanese	321
Table 11.9	Morphonological consonant changes in the Ei Dialect	329
Table 11.10	The four-way hiragana distinction in Ei dialect (contrasted with Tsuruoka dialect)	330
Table 11.11	Consonants in the Ei dialect	331
Table 11.12	Accent patterns of nouns in the Ei dialect (I)	332
Table 11.13	Distribution of noun accent patterns in ModJ and Ei	332
Table 11.14	Accent patterns of nouns in the Ei dialect (II)	333
Table 11.15	Accent patterns of verbs in the Ei dialect	333
Table 11.16	Accent patterns of adjectives in the Ei dialect	333
Table 11.17	Accent patterns of compound nouns in the Ei dialect	334
Table 11.18	Verb conjugation in the Ei dialect (I)	336
Table 11.19	Conjugation of vowel-stem verbs in ModJ	336
Table 11.20	Verb conjugation in the Ei dialect (II)	337
Table 11.21	Verb conjugation in the Ei dialect (III)	337
Table 12.1	Shuri consonants	352
Table 12.2	Word classes in Shuri	354
Table 12.3	Shuri pronoun system	356
Table 12.4	Shuri stem form and (pre-)inflection type	358
Table 12.5	Shuri finite inflection	359
Table 12.6	Shuri type A stem and inflection	359
Table 12.7	Shuri type A stem, pre-inflection (negative), and inflection	359
Table 12.8	Shuri type B stem, pre-inflection (politeness), and inflection	360
Table 12.9	Shuri type B stem, pre-inflection (imperfective), and inflection	360
Table 12.10	Shuri type C stem, pre-inflection (perfective), and inflection	361
Table 12.11	Shuri type C stem, pre-inflection (progressive), and inflection	361
Table 12.12	Shuri type C stem, pre-inflection (resultative), and inflection	361
Table 12.13	Shuri type B stem, pre-inflection (politeness + negative), and inflection	362
Table 12.14	Shuri non-finite inflection	362
Table 12.15	Shuri existential and copular verbs	364

---

Table 12.16	Shuri finite inflection for adjectives	367
Table 12.17	Shuri non-finite inflection for adjectives	367
Table 12.18	Sentence types in Shuri	373
Table 13.1	Consonant inventory of Hatoma	382
Table 13.2	Hatoma demonstrative pronominals	389
Table 13.3	Hatoma personal pronouns	391
Table 13.4	Hatoma pronouns and speech levels	403
Table 14.1	Yonaguni consonant inventory	414
Table 14.2	Yonaguni verb morphology	422
Table 14.3	Yonaguni copula	424
Table 14.4	Yonaguni adjective morphology	426
Table 14.5	Yonaguni personal (and reflexive), demonstrative and interrogative pronouns	431
Table 14.6	Yonaguni intransitive/transitive verbs: inflection	445
Table 14.7	Yonaguni intransitive/transitive verbs: derivation	446
Table 14.8	Yonaguni kaida characters	455
Table 15.1	Consonant inventory of Southern Hokkaido Ainu	465
Table 15.2	Person/number marking in the Saru and Chitose dialects of Ainu	471
Table 15.3	Subject/object indexing of transitive verbs in the Saru and Chitose dialects of Ainu	474
Table 15.4	Verb structure in the Saru dialect of Ainu	481

---

---

# CONTRIBUTORS

---

---

**John R. Bentley** is in the Department of Foreign Languages and Literatures at Northern Illinois University, USA.

**Anna Bugaeva** is at Waseda Institute for Advanced Study, Waseda University, Japan.

**Mark Irwin** is at the Faculty of Literature and Social Sciences, Yamagata University, Japan.

**Atsuko Izuyama** was formerly at Dokkyo University, Sōka, Saitama, Japan.

**Young-Key Kim-Renaud** is at the East Asian Languages and Literatures Department, The George Washington University, USA.

**Mika Kizu** is in the Department of Languages and Cultures of Japan and Korea, School of Oriental and African Studies, University of London, United Kingdom.

**Wayne P. Lawrence** is at the School of Asian Studies, The University of Auckland, New Zealand.

**Akiko Matsumori** is at the School of Humanities, Japan Women's University, Tokyo, Japan.

**Heiko Narrog** is at the Graduate School of International Cultural Studies of Tohoku University, Sendai, Japan, and is also an invited associate professor at the National Institute for Japanese Language and Linguistics (NINJAL).

**Nam Pung-hyun** is emeritus professor at the University of Dankook, Republic of Korea, and is a former president of the Society for Kugyōl Studies.

**Takuichiro Onishi** is at the National Institute for Japanese Language and Linguistics (NINJAL), Tachikawa City, Tokyo, Japan.

**Mark Rosa** is a doctoral student at Tokyo University, Japan.

**Michinori Shimoji** is at Gunma Prefectural Women's University, Japan, and is also an invited associate professor at the National Institute for Japanese Language and Linguistics (NINJAL).

**Ho-min Sohn** is at the Department of East Asian Languages and Literatures, University of Hawaii at Mānoa, USA.

**Nicolas Tranter** is at the School of East Asian Studies, University of Sheffield, United Kingdom.

**John Whitman** is at Cornell University, USA, and is also an invited scholar at the National Institute for Japanese Language and Linguistics (NINJAL).

**Jaehoon Yeon** is in the Department of Languages and Cultures of Japan and Korea, School of Oriental and African Studies, University of London, United Kingdom.

---

# **PREFACE**

---

This volume presents a series of detailed language sketches of different current and historical varieties of the languages of Japan and Korea. Although language change and etymology are discussed, and comparison with other varieties is made in various places, the language sketches are intended to be primarily stand-alone descriptions and not part of a diachronic study. Each sketch covers phonology, script (where appropriate), morphology, syntax and the lexicon. There is a greater emphasis on syntax than is typical in previous language sketches, which can treat syntax cursorily in comparison with morphology.

---

# **ACKNOWLEDGEMENTS**

---

Acknowledgements are due to Kevin Cawley, Chris Dutch, James Grayson, Ross King, Hyung Hye Kyung, and Hyangjin Lee for various help given. Particular acknowledgements are due to Sukyeon Cho for frequently acting as an intermediary, and my wife Susie Tranter for standardizing the bibliographies, general proofreading, and acceptance of the state of our house in the final nine months of editing, which was carried out on our kitchen table. Finally, thanks and apologies in equal measure are due to the various editors at Routledge who have supervised this project, namely Jonathan Price, Andrea Hartill, and Samantha Vale Noya.

---

# ABBREVIATIONS

---

## Languages

C	(Modern Mandarin) Chinese
CJ	Classical Japanese
EMC	Early Middle Chinese
EMJ	Early Middle Japanese
EMK	Early Middle Korean
EOJ	Eastern Old Japanese
EOK	Early Old Korean
(Mod)J	(Modern/standard/contemporary) Japanese
JR	Japanese-Ryukyuan (as a language family) = Japonic
(C)K	(Modern/standard/contemporary) Korean
LMC	Late Middle Chinese
LMJ	Late Middle Japanese
LMK	Late Middle Korean
LOK	Late Old Korean
MC	Middle Chinese
MK	Middle Korean (specifically Late Middle Korean unless stated otherwise)
MOK	Mid Old Korean
NJ	native Japanese (vs Sino-Japanese)
NK	native Korean (vs Sino-Korean)
N-K	North Korean
OJ	Old Japanese (specifically Western Old Japanese unless stated otherwise)
OK	Old Korean
pJ	proto-(Mainland) Japanese
pJK	proto-Japanese-Korean
pJR	proto-Japanese-Ryukyuan = proto-Japonic
pK	proto-Korean
pR	proto-Ryukyuan
SJ	Sino-Japanese
SK	Sino-Korean
S-K	South Korean
SV	Sino-Vietnamese
V	Vietnamese
WOJ	Western Old Japanese

For convenience, Modern Korean and Modern Japanese forms are cited distinctively in bold script in those chapters of Part II and Parts III–IV respectively other than Chapter 5 ‘Modern Korean’ and Chapter 10 ‘Modern Japanese’.

The following morphemic tags are used widely throughout the volume. It has been considered impractical and counterproductive to force across-the-board standardization of morpheme-glossing tags across or even within the Korean, Japanese-Ryukyuan and Ainu groups. However, the following tags are those that are used widely. Other tags, where not obvious, are explained in individual chapters. Morphological analyses are consistently represented with hyphens (-) between morphemes, though in Chapters 12 and 15 a double hyphen (=) is used to represent clitics as opposed to affixes, and inter-linear glossing is presented to match the breakdown accurately and clearly, with a dot (.) used instead of a space where a single morpheme would otherwise require more than one word as a gloss, e.g. 'say.HUM' for CJ *mawos-* (combining the category of humble + the lexical meaning 'to say').

ABL	ablative
ACC	accusative
AUX	auxiliary verb
CAUS	causative
COMP	complementizer
COP	copula
DAT	dative
EMPH	emphatic
FOC	focus
GEN	genitive
HON	honorific
HUM	humble
IMP	imperative
INF	infinitive
INST	instrumental
LOC	locative
NEG	negative
NMR	nominalizer
NOM	nominative
PASS	passive
PL	plural
PROH	prohibitive
Q	interrogative
QUOT	quotative
TOP	topic
*	reconstructed proto-form; ungrammatical

# TRANSCRIPTION SYSTEMS

The principle is that linguistic data and bibliographies are transcribed in the Yale System (Modern Korean) and the Revised Hepburn System (Modern Japanese), the latter in a slightly variant form (with *o* and *n*) in the case of linguistic data. Certain familiar place names (e.g. Kyoto, Seoul, Tokyo; plus Hokkaido), or names of scholars who write in English are presented in their most familiar romanization. Korean-language terms, names of literary texts and their authors, and other place names are transcribed according to each author's preference; Chapters 1 and 3, however, because of their introductory nature, use the formula 'McCune-Reischauer [Yale]' for clarity. Names of writing systems are written unitalicized and uncapitalized throughout. Major Japanese and Korean transcriptions are compared in [Tables 0.1](#) and [0.2](#). Linguistic data for varieties other than Modern Japanese and Modern Korean are presented in a consistent form throughout, e.g. Middle Korean is transcribed in the system chosen by the author of Chapter 4. The only exception to this is Early/Mid Old Korean data, where Chapters 2 and 3 present different treatments of the vowel system.

**TABLE 0.1 MODERN JAPANESE TRANSCRIPTION SYSTEMS WHERE DIFFERENCES OCCUR**

<i>s-</i>		<i>z-</i>		<i>t-</i>		<i>d-</i>			<i>h-</i>	
MH	KS/NS	MH	KS/NS	MH	KS/NS	MH	KS	NS	MH	KS/NS
sa	sa	za	za	ta	ta	da	da	da	ha	ha
shi	si	ji	zi	chi	ti	ji	zi	di	hi	hi
su	su	zu	zu	tsu	tu	zu	zu	du	fu	hu
se	se	ze	ze	te	te	de	de	de	he	he
so	so	zo	zo	to	to	do	do	do	ho	ho
sha	sya	ja	zya	cha	tya	...	...	...	fa	—
sho	syo	jo	zyo	cho	tyo	...	...	...	fi	—
shu	syu	ju	zyu	chu	tyu	...	...	...	fe	—
she	sye	je	zye	che	tye	...	...	...	fo	—

MH = 'Modified (or Revised) Hepburn'

KS = 'Kunreisiki'

NS = 'Nipponsiki'

**TABLE 0.2 MODERN KOREAN TRANSCRIPTION SYSTEMS (SOUTH KOREAN HAN'GŪL ALPHABETICAL ORDER)**

	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ
Yale	k	kk	n	t	tt	l	m	p	pp	s	ss
M-R	k/g	kk	n(*)	t/d	tt	l/r	m	p/b	pp	s	ss
NR	g/-k	kk	n	d/-t	tt	l/r	m	b/-p	pp	s	ss
	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㆁ	ㅍ	ㅎ	Assimilations?		
Yale	ng	c	cc	ch	kh	th	ph	h	Underlying		
M-R	ng*	ch	tch	ch'	k'	t'	p'	h	Surface		
NR	ng	j	jj	ch	k	t	p	h	Surface		
	ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅘ	ㅙ
Yale	a	ay	ya	yay	e	ey	ye	yey	o	wa	way
M-R	a	ae	ya	yae	ō	e	yō	ye	o	wa	wae
NR	a	ae	ya	yae	eo	e	yeo	ye	o	wa	wae
	ㅛ	ㅠ	ㅜ	ㅠ	ㅡ	ㅣ	ㅚ	ㅜ	ㅣ	ㅣ	ㅣ
Yale	oy	yo	wu/u*	we	wey	wi	yu	u	uy	i	i
M-R	oe	yo	u	wō	we	wi	yu	ū	ūi	i	i
NR	oe	yo	u	weo	we	wi	yu	eu	ui	i	i

\*: M-R: An apostrophe is used in *n'g* to distinguish from *ng*.

Yale: *u* is used after bilabials because *—* does not occur in native words (and only occurs in loanwords, where it is realized as though ㅜ) in this environment.

M-R = 'McCune-Reischauer'

NR = 'New (or Revised) Romanization', promulgated 2000

PART I

# **PRELIMINARIES**



# INTRODUCTION: TYPOLOGY AND AREA IN JAPAN AND KOREA

*Nicolas Tranter*

## 1.1 THE LANGUAGES OF JAPAN AND KOREA

Ainu is a language isolate, with no proven relationship to any other language. Moreover, Japanese and Korean are also frequently referred to as language ‘isolates’, albeit ones with a significant great degree of historical and regional diversity; there are, however, problems with tagging them as isolates.

### 1.1.1 Japanese-Ryukyuan

Despite the belief often found in Japanese writing, and even reflected in some major Western writing (e.g. Shibatani 1990: 89, 187ff), that Japanese is a language isolate and that the Ryukyuan varieties of Kagoshima and Okinawa prefectures are ‘dialects’ (*hōgen*) of Japanese, Ryukyuan varieties and Modern Japanese remain mutually incomprehensible, as non-Ryukyuan generally do not understand such sentences as presented in Chapters 12 to 14 (Shimoji, Lawrence, Izuyama) of this volume. If mutual incomprehensibility is taken as a measure of what constitute separate languages, then rather than a Japanese-language isolate we are dealing in Japan with a small language family, typically referred to as ‘Japonic’ or ‘Japanese-Ryukyuan’ (JR), and its modern members belong to two separate branches, ‘Japanese’ (or ‘Mainland Japanese’) and ‘Ryukyuan’. Leaving aside the language vs dialect terminology, this primary division of JR is uncontroversial.

The Ryūkyū islands are a long chain of island groups between the south of Kyūshū (the southernmost ‘mainland’ island of Japan) and within 70 miles of Taiwan. Topographically the island groups, north to south, are (i) the Ōsumi group, (ii) the Tokara group, (iii) the Amami group (Kikai, Ōshima, Tokunoshima, Okierabu, Yoron), (iv) the Okinawa group (the large island of Okinawa and a number of other islands), (v) the Miyako group (Miyako, Irabu and Tarama), (vi) the Yaeyama group (Ishigaki, Iriomote, Hateruma and a few other small islands) and (vii) Yonaguni. Okinawa island was united in the early fifteenth century and its capital Shuri remained the centre of an independent Ryūkyū kingdom that spread as far as (iii)–(vii) till the early sixteenth-century start of Japanese suzerainty (including outright ceding of (iii) to the lord of Satsuma province), though final annexation only occurred in 1879. Politically, therefore, (i)–(iii) are part of Kagoshima prefecture, most of which is on the mainland of Kyūshū and which was formed from the older Satsuma, while (iv) to (vii) are part of Okinawa prefecture. Linguistically, ‘Ryukyuan’ is defined as those varieties in (iii)–(vii), while ‘Okinawan’ tends to be reserved for Okinawa island, particularly the prestigious modern Shuri variety. The largest linguistic break

within Ryukyuan is accepted to be the largest ocean break in the Ryūkyū islands; that is, between the Okinawa and Miyako groups. North of this break is ‘North Ryukyuan’ consisting of (iii)–(iv), and south is ‘South Ryukyuan’ consisting either of (v)–(vi) or (v)–(vii). Japanese-language literature – as well as some literature in the West – tends to use ‘Amami-Okinawa’ (or similar) vs ‘Sakishima’ as respective terms instead. Beyond these divisions, classification of Ryukyuan varieties is more problematic for a range of reasons: the lack of historical data further back than the sixteenth century (when Japanese mixed script was adapted to writing ‘classical’ Okinawan poetry, especially the *Omoro sōshi* anthology), a paucity of reliable data for some varieties, the erosion of varieties through over a century’s imposition and dominance of Modern Japanese (Heinrich 2004) leading to moribundity or to interdialects (Matsumori 1995), a lack of agreement on what constitutes language vs dialect and little study of mutual incomprehensibility. Therefore a Ryukyuan ‘family tree’ varies from author to author. It is arguable that beyond the primary North–South division, Yonaguni at the southernmost end of the archipelago is the most distinctive Ryukyuan variety. Many authors include Yonaguni within South Ryukyuan, whereas others, especially since Hirayama *et al.* (1966) and reflected in the West in Thorpe’s (1983: 2–4) tentative classification, suggest Yonaguni to be a third primary division of Ryukyuan. The choice to include Yonaguni as well as Shuri (North) and Hatoma (South) Ryukyuan as a chapter in its own right in this volume reflects its uncertain classification as well as its range of unusual features.

Modern mainland Japanese is widely taken to be a single language with a range of dialects constituting a chain of mutual comprehensibility, even though the extremes of the chains may not understand each other. However, a West–East distinction is identified in the Nara period, between (Western) Old Japanese of the capital’s region (Kinki) vs Eastern Old Japanese of the provinces that were grouped as ‘Azuma’ (OJ *aduma*) provinces, the geographical boundary being a mountain range. The EOJ varieties are attested in 31-syllable ‘eastern poems’ (*adumauta*) and ‘border-guard poems’ (*sakimoriuta*) that constitute roughly a volume and a half (Books 14 and 20) of the *Man’yōshū* anthology (c. 749). A few substantial studies of the poems have been made in Japan (Fukuda 1965; Hōjō 1966; Mizushima 1984), but study in the West has been minimal. More recently, aspects of EOJ have fed into reconstructions of proto-Japanese–Ryukyuan. The dialect of the isolated island of Hachijōjima, 175 miles due south into the Pacific, has been noted to share features with EOJ, notably the verb attributive in *-o* (Hōjō 1966: 529–86), and is regarded as a modern relic of EOJ. Modern Japanese dialects are normally subdivided into the Kyūshū, western and eastern groups. Even though most EOJ features vanish completely after the Nara period, Onishi (2008) observes that the western–eastern dialect groups still appear to be sharply divided by the same mountains that marked off Azuma, based on the distribution of such features as negative formation (western *-nu/n*, etc., cf. WOJ/CJ attributive *-nu*, vs eastern *-nai*, etc., cf. EOJ *-napye/nopye*), arguing that this is explainable as a continuation of a prehistoric language break into the present. It should be noted that the de facto capital of Japan has moved across these mountains several times, from Heian-kyō (modern Kyoto) in the west to Kamakura in the east, back to Kyoto, and then to Edo (modern Tokyo) in the east. To what extent EOJ underlies modern Tokyo Japanese is debatable. However, the two-way WOJ/EOJ division of the Nara period is simplistic, not least because within the EOJ corpus we observe a geographic subdivision into those eastern provinces that display only EOJ features, those that display no EOJ features and those intermediate provinces that display a mix (Fukuda 1965: 246–54; cf. more recently Hino 2003). Furthermore, we have no reason not to suppose that even more westerly regions of Japan in the Nara period had their own distinct

varieties; for example, the otherwise distinctly EOJ word *sida* ‘time’ also occurs in a poem of Hizen province in Kyūshū (*Hizen fudoki*, Akimoto 1958: 396).

### 1.1.2 Korean

Modern varieties of Korean maintain a more or less reasonable level of mutual comprehensibility, although speakers of the standard language will struggle to understand such varieties as that of Cheju island. Historically, though, Korean appears to parallel Japanese in attesting much older sister varieties that subsequently vanish. Conventionally, the differences revealed in the ninth-century sinicization of place names in *Samguk sagi* [*Samkwuk saki*] (1145) are divided into three types based on the earlier Three Kingdoms: Paekche (roughly the western half of modern South Korea), Silla (roughly the eastern half of modern South Korea) and Koguryō (roughly modern North Korea and much further north). Whether these actually were three distinct languages is unprovable, as certain subsequently extinct vocabulary transcends the kingdoms, and Nam (this volume) argues that they were a single language. An apparently JR element in some Koguryō toponyms is intriguing.

Therefore, rather than dealing only with language isolates, this volume in Routledge’s Language Family Series can be regarded as a collection of three ‘language families’: Japanese-Ryukyuan, Korean and the isolate Ainu. However, there is evidence that JR and Korean are ultimately related. Starting from Martin’s (1966) first attempt to establish correspondence rules, the study of the Japanese–Korean genetic relationship and serious consideration of what is genetic and what is loan has resulted in a number of studies (Whitman 1985; Frellesvig and Whitman 2008; Vovin 2010), and Whitman (this volume) discusses proto-Japanese–Korean.

In the modern world, Japanese and Korean are languages of major developed nations, Japan and the Republic of Korea (South Korea); Korean is also the language of the Democratic People’s Republic of Korea (North Korea). In terms of number of speakers (and treating ‘Chinese’ and ‘Arabic’ each as a single language), Japanese is firmly within the world’s top ten, ranking at number nine in the *Ethnologue* rankings (Lewis 2009, [http://www.ethnologue.com/ethno\\_docs/distribution.asp?by=size](http://www.ethnologue.com/ethno_docs/distribution.asp?by=size)). Korean is at number 17 currently in *Ethnologue* (Lewis 2009), a drop from higher rankings in earlier editions (13 in 2005: Gordon 2005). Japanese and Korean are both in the top ten (four and ten respectively) of languages on the internet (<http://www.internetworldstats.com/stats7.htm>). Korean is, together with Tibetan, Mongolian, Uyghur and Zhuang, one of the five ‘major’ minority nationalities within China, and both Japanese and Korean have small but significant diasporas. In contrast with the robust health of Japanese, however, Ainu is by any measure severely moribund, and UNESCO’s atlas of endangered languages (Moseley 2010) identifies eight threatened languages in Japan: Ainu; the Hachijōjima dialect of Japanese; three North Ryukyuan varieties, namely Amami, Kunigami, Okinawan (Shimoji, this volume); and three South Ryukyuan varieties, namely Miyako, Yaeyama (Lawrence, this volume) and Yonaguni (Izuyama, this volume). Ainu is categorized as ‘critically endangered’, Yaeyama and Yonaguni as ‘severely endangered’, the rest ‘definitely endangered’.

## 1.2 AREAL RELATIONSHIPS AND TYPOLOGY

Similarities between languages can be explained variously, including chance coincidence. For example, in the semantic field of bodyparts we can observe chance similarities

between English and Shuri Ryukyuan or Japanese: E *shin* : S *sini* [eini] : J *sune*, E *bone* : S *huni* [ϕuni] : J *hone*, E *spine* : J *sebone*. No serious academic would argue anything other than coincidence for the English/JR similarities. (Nevertheless, there is a small pseudo-academic tradition of trying to prove Indo-European and JR genetic connections using ‘evidence’ like this.) The other explanations are typological (1.2.1), areal – including loanwords – (1.2.2) and genetic. The key views on genetic relationships in our region are: (a.) JR and Korean are unrelated; (b.) JR and Korean are two branches of a single family; (c.) JR and/or Korean are branches of an even larger pan-North Asian language family, ‘Altaic’. The Altaic hypothesis has existed for quite some time with three core branches, Mongolic, Tungusic and Turkic, and is based primarily on great typological similarities between the representative languages, all SOV, and on a significant proportion of common vocabulary. Unfortunately for the Altaic hypothesis, typology is irrelevant for proving a genetic relationship, and a language’s syntactic type is not immutable; for example the SOV of older Indo-European languages and widely accepted for proto-Indo-European versus the SVO pattern of most modern Indo-European languages, the VSO pattern of Celtic, etc., or the evidence that attested Chinese varieties, all SVO, developed from an earlier SOV pattern; and lexical similarities, such as between Turkic and Mongolic (Schönig 2003) increasingly are identifiable as loanwords. Though there have been attempts to tie JR and Korean in as fourth and fifth branches of Altaic (particularly Miller’s (1971) bullishly titled *Japanese and the Other Altaic Languages*), the very existence of Altaic is uncertain.

### 1.2.1 Typology

Typology is the classification of languages into ‘types’ based on specific features. The best known feature is syntactic. Languages can be broadly divided into six types based on the unmarked order of subject, object and verb, of which SOV and SVO are the two major patterns. As the verb in SOV languages is rightmost within the verb phrase of which it is the head (VP → O V, e.g. K *yenghwa-lul poassta*, J *eiga-o mita* ‘[I] saw a movie’), SOV languages are characterized as ‘left-branching’ and show a marked tendency for left-branching in other phrase types, e.g. a noun phrase is headed by a noun positioned rightmost, preceded by modifiers (K *nay-ka cinantal pon yenghwa*, J *watashi-ga sengetsu mita eiga* ‘a movie that I saw last week’), and a PP is headed by a *postposition* preceded by a noun (K *cwungkwuk-eyse*, J *chūgoku-de* ‘in China’). This is a strong tendency, but is subject for example to areal influences (see Dryer 2003 for a discussion of syntactic typology and its areally influenced exceptions in the Tibeto-Burman branch of Sino-Tibetan.) Another – morphological – typology is that primarily of ‘isolating’, ‘inflecting’ and ‘agglutinating’ languages, with a fourth ‘polysynthetic’ type.

JR, Korean and Ainu are all SOV languages, and indeed discussions of syntactic typology frequently cite Japanese as a canonical example. They all largely have modifiers before nouns and *postpositions* (or ‘particles’) rather than *prepositions*. The sole exception in Modern Standard Japanese and Korean is optional alternative positioning of quantifier and number + classifier phrases after the head noun, though the surface syntax differs in the positioning of particles: J Noun-Particle Quantifier vs K Noun Quantifier-Particle. Internally headed relative clauses are also found in Modern Japanese and Modern Korean (see Kuroda 1992: 114–74 on Japanese, and Kim 2004 on both), typically regarded as a substandard feature even though Middle Japanese made extensive use of them in writing.

In considering the relations of Korean or JR in wider north Asia, it is notable that Mongolic, Tungusic and Turkic are also all largely canonically SOV (as also are other

languages such as Burmese!), so the wide similarities in word order between these families and Korean and Japanese cannot be used to argue for a genetic relationship. In terms of morphology, JR and Korean, as well as Mongolic, Tungusic and Turkic, are all agglutinating, and agglutination – as would be typical of SOV languages – is overwhelmingly suffixal. Ainu is distinctive in being predominantly prefixal as also ‘incorporating’ (Bugaeva, this volume). Although Ainu is SOV, it differs from Japanese-Ryukyuan and Korean in a range of other features as well, such as its use of person-marking and ‘applicative’ structures.

Phonologically, Ainu’s minimal repertoire of vowels and consonants is notable, and the limited number of consonants is similar to that proposed for proto-JR and for proto-Korean. Similarly, accent in Ainu in Japan, Japanese, and historically for Korean (as seen in Middle Korean and certain modern dialects, but no longer in the standard language) is a pitch accent system, with each syllable/mora having high or low pitch, with a less common variation of changing pitch on a syllable/mora that seems associated with long vowels, contraction or possible segment loss. However, it is difficult to argue confidently that any phonological features in Ainu, Japanese and Korean are areal, as they are not uncommon in the languages of the world, and there are very marked phonological differences between the attested varieties of the languages.

### 1.2.2 Korean–Japanese–Ainu areal relationships

An areal explanation for similarities is based on features of one language seeping into another language with which it has contact, either as a result of imperfect bilingualism or deliberate code-switching or borrowing. Lexical borrowing between Japanese and Korean in prehistoric or early historical times is reflected in a small but definite set of similar words in the two languages, and borrowing at this early point is generally acknowledged to have been from Korean into Japanese, though the question of what is a cognate and what is a loanword is crucial to the theory of a Japanese–Korean genetic relationship. Moreover, there are small ancient loans between Ainu and Japanese, such as Ainu *sak-ipe* ‘salmon’ (lit. ‘summer’ + ‘food’) > J *sake*. There are also more recent cultural native-stratum loanwords between Korean and Japanese, e.g. K *kimchi* > J *kimuchi*, J *sashimi* > K *sasimi* (despite the existence of K *hoy*). In Korea in the second half of the twentieth century, native Japanese loans were largely purged from the language, though more recently some have returned. Most recent loans between Japanese and Korean, however, have been Sino-Japanese to Sino-Korean (see 1.2.3), where the graphic rather than the phonological nature of the loans has obscured their overtly Japanese origin, resulting in minimal purging in Korea.

There are also clear grammatical similarities beyond the typologically expected syntax. A handful of illustrative examples are presented below. On the one hand, many of these similarities are relatively recent, and some certainly derive from Japan’s annexation of Korea (1910–1945) and imposition of Japanese, or the slightly earlier period when Korean intellectuals studied in Japan. However, many areal features spread beyond Japan and Korea, appearing in Mongolic, Tungusic or Turkic, or even in Chinese, and so it is difficult to identify a clear Japan–Korea Sprachbund (speech area). Such areal grammatical features and their history and spread are in much need of research.

(a.) *Subject/Object honorification*: Modern Japanese and Korean share morphological honorification, both addressee honorification (hereafter ‘polite(ness)’ or ‘speech levels’)

and subject/non-subject honorification ('honorific' and 'humble' respectively). The system of the latter is markedly different in the modern languages – which can be described as 'egocentric' – and in the earlier languages – which can be described as 'relative-hierarchical'. In an egocentric system, 'humble' forms show respect to the referent of a grammatical object or of some other participant in the sentence by downgrading a subject that is the speaker ('I') or someone in the speaker's in-group as defined relative to the participant requiring respect ('we', 'my brother', 'my company'). Similarly, 'honorific' forms tend to be triggered by the respected referent's social status relative to the speaker. It is noteworthy that though Japanese maintains an extensive 'humble' system, Korean has almost entirely lost the category except for a couple of verbs. In a relative-hierarchical system, however, the choice of honorific or humble is determined by the relative status of the referent(s), and not only relative to the speaker. This fundamental difference means that a divinity or a ruler can use honorific forms where the subject is 'I' and humble forms where the subject is 'you' (e.g. the retired emperor saying *mono-domo maiō-te sarugaku tukamatur-e* 'People, **come** and **do** a comedy routine for me!', *Heike monogatari*; humble verb forms marked in bold). Another consequence of relative-hierarchical systems is that a verb can be marked simultaneously honorific and humble. (In (1) I replace SH and OH as used in Sohn (this volume) with HON and HUM for ready comparison.)

- (1) a. MK *apa-nim poj-zóβ-ósi-l? cej*  
 father-HON see.HUM-HUM-HON-REL time  
 Speaker = the narrator<sub>C</sub>  
 'when he<sub>B</sub> [= the prince] saw his father<sub>A</sub> [= the king]' (*Yongpi echenka* 91)
- b. CJ *fito na itaku wabisase-tatematur-ase-tamafi-so*  
 person PROH painfully upset-HUM-HON-HON-PROH  
 Speaker = old bamboo-cutter<sub>C</sub>  
 'Don't you<sub>B</sub> [= Kaguyahime] upset the people<sub>A</sub> [= princes and ministers] too much!' (*Taketori monogatari*, simplified from Tranter, this volume, Chapter 8 (22b))

In both examples, there are three people involved, with different status,  $A > B > C$ . The speaker C in both cases is the lowest position, and shows respect to the grammatical subject B by using an honorific form, but also shows respect to the object A, who is highest-ranking, by humbling B. Despite common similarities – the fact that both languages have such systems, the relative-hierarchical nature of the older system and the egocentric nature of the modern system – there are significant differences between the languages. For example, the Korean honorific is formed with a simple suffixal morpheme (K *-(u)si-* < MK *-(u)ósi-* < OK *-sβ-*), and the MK humble was similarly simple (*-zóp/-sóp*); in contrast, Japanese has always made heavy use of a system of forms that derive from independent verbs, e.g. CJ *-tamafu* (HON < 'give (HON)') and *-tatematuru* (HUM < 'give (HUM)') and J *-ni naru* (HON < 'become') and *-suru* (HUM < 'do'). This means that there is an observable difference in the markedness of honorific forms in the modern languages. Yet, the MK humble appears – not least because of its OK *idu/kugyōl* spelling 白 (for 曰 = 'say') – cognate with independent *sólp-* 'say.HUM' < OK *sβrp*. Historically another difference is the fact that Japanese appears to have entered into a cycle whereby the need to express genuine respect leads to more sophisticated forms, rendering previously used forms clichéd. We particularly notice this in the case of personal pronouns, with the result that earlier true pronouns

have become largely obsolete and pronoun-substitutes have become debased (e.g. casual ‘you’ *omae* < ‘imperial/divine presence’, intimate ‘you’ *kimi* < ‘lord’, aggressive ‘you’ *kisama* < ‘nobleman’). In contrast, though Korean makes considerable use of pronoun-substitutes, it retains its original set of pronouns. Another feature in common is the development of distinct honorific equivalents to kin-terms in both languages, e.g. J *chichi* vs *o-tō-san* (HON) and K *apeci* vs *ape-nim* (HON) ‘father’.

(b.) *Addressee honorification*: The addressee honorification system, or the system of speech levels, is characteristic of both Japanese and Korean. The system, however, shows marked differences. Korean, for example, has developed from a single MK polite level marker *-ngi-* to the modern system of six or so levels. Japanese, on the other hand, is typically treated as having just two levels, ‘plain’ and ‘polite’, though even in declarative sentences we can argue for more levels, e.g. ‘exist’ *aru* – *arimasu* – *gozaimasu*, ‘be good’ *ii* – *ii-desu* – *yoroshii-desu*. The choice of question-marking particles in J interrogative sentences, or the choice of command/request forms, displays a more complex system. The system in both languages involves a mix of inflected affixes (e.g. K *-supni-*, J *-mas-*) and ‘polite particles’ (e.g. K *-yo*, J *-desu<sub>2</sub>*). Ryukyuan varieties vary in whether they use affixes (e.g. Shuri *-yabii-*), particles (e.g. Yonaguni *juu*) or no marking. The need to use a verb ending that marks the correct speech level has in both Japanese and Korean arguably contributed to a tendency to use incomplete sentences to avoid the need to choose an appropriate style marker while also expressing a request without an explicit request form or achieving some other specific pragmatic effect. Korean has progressed further than Japanese, to the extent that historically incomplete sentences ending in *-e/-a* ‘and’ or *-(nu)ntey* ‘but’ now constitute a distinct egalitarian speech level, ‘panmal’, from which in turn, through addition of the polite particle *-yo*, a new polite style developed.

(c.) *Passivization*: The passive construction in Japanese and Korean is used not only in the typical ‘direct’ structure of Western passives (2), where the object of the active verb becomes the subject of the passive verb, but also in a so-called ‘indirect’, ‘adversative’ or ‘suffering’ structure, where someone other than the referent of the grammatical object that is affected by the action becomes the subject (3). This structure typically involves transitive verbs (but see Tranter and Kizu, this volume, 10.5.5), and the new subject of the passive is most frequently *the possessor of the referent of the grammatical object*. It is interesting that a similar distinction is possible in Chinese.

- |     |   |  |                    |                    |                           |
|-----|---|--|--------------------|--------------------|---------------------------|
| (2) | J | <i>nihon-ga</i>                            | <i>chōsen-no</i>   | <i>ryōdo-o</i>     | <i>ubaq-ta</i>            |
|     | K | <i>ilpon-i</i>                             | <i>cosen-uy</i>    | <i>yengtho-lul</i> | <i>ppayas-assta</i>       |
|     |   | Japan-NOM                                  | Korea-GEN          | territory-ACC      | seize-PAST                |
|     | C | <i>riběn</i>                               | <i>zhànlǐng-le</i> | <i>cháoxiān-de</i> | <i>lǐngtǔ</i>             |
|     |   | Japan                                      | occupied-PFV       | Korea-GEN          | territory                 |
|     |   | ‘Japan seized Korea’s territory.’          |                    |                    |                           |
|     |   |  |                    |                    |                           |
| (3) | J | <i>chōsen-ga</i>                           | <i>nihon-ni</i>    | <i>ryōdo-o</i>     | <i>ubaw-are-ta</i>        |
|     | K | <i>cosen-i</i>                             | <i>ilpon-ey</i>    | <i>yengtho-lul</i> | <i>ppayas-ky-essta</i>    |
|     |   | Korea-NOM                                  | Japan-by           | territory-ACC      | seize-PASS-PAST           |
|     | C | <i>cháoxiān</i>                            | <i>bèi</i>         | <i>riběn</i>       | <i>zhànlǐng-le lǐngtǔ</i> |
|     |   | Korea                                      | PASS               | Japan              | seize-PFV territory       |
|     |   | ‘Korea had its territory seized by Japan.’ |                    |                    |                           |

(d.) *Converbs*: Japanese and Korean share with Tungusic, Mongolic and Turkic the principle of coordinating or subordinating by means of ‘converbs’, whereby in a sequence of verbs only the final verb is marked for tense and speech level, whereas the earlier verbs are marked by an untensed ‘converb’ suffix, such as J *-eba* ‘if’, infinitive + *-nagara* ‘and’ (simultaneous), infinitive + *-ni* ‘[go] to’, *-tara* ‘if/when’, *-tari* ‘and’ (non-sequential), *-te kara* ‘after’, *-temo* ‘even if/though’, *-tewa* ‘if’, nonpast + *-to* ‘if/when’, and the infinitive ‘and’; and the richer inventory in K *-e/-a* ‘and’, *-ese/-ase* ‘and so’, *-eto/-ato* ‘even if/though’, *-eya/-aya* ‘only if’, *-(u)lye* ‘[go] to’, *-(u)mye* ‘and’, *-(u)myen* ‘if’, *-(u)myense* ‘and’ (simultaneous), *-nulako* ‘what with’, *-(nu)ntey* ‘and/but’ (background), *-taka* ‘and/but’ (interruption), *-tolok* ‘in order to’, *-(u)nikka* ‘since’, etc. Older Japanese shows even fewer true converbs than Modern Japanese: for example the ancestor of *-eba* (‘and then/so’ in OJ and CJ) markedly is preceded by tense/aspect forms.

(e.) *Phrasal structures*: Sequences of coordinating ‘converb’ + another verb are prone to become grammaticalized to create aspectual or pseudo-aspectual forms. In all the languages, the earlier verbs in the sequence are marked by specific endings with coordinating or converbal function: the J infinitive and *-te*; Korean *-e/-a* and *-ko*; Mongolian *-ž/-č*; Uyghur *-(I)p*. The (pseudo-)aspectual derivatives show different regional spreads. On the one hand, ‘V and see’ is widely found to express ‘try Ving’ from JR to eastern Turkic (4a) (and cf. even English *Eat it and see*). Ainu is reported to restrict ‘see’ to visual evidence, using ‘hear’ for other sensory inputs (4b (Tamura 2000: 185)). Similarly, benefactive constructions using verbs of giving or of receiving are found widely in East Asia (cf. Uyghur *-(I)p ber-*, Chinese *gěi* + Benefactor + VP, etc.). Other combinations are more localized to Japan and Korea, e.g. ‘V and put’ to express an action carried out in readiness for a future situation, though this same construction also appears in Uyghur (5). Completion with a sense of irrevocability and therefore typically also of regret is expressed by K *-e peli-*, J *-te shimau* and Uyghur *-(I)p qal-*, but this apparently similar construction is not necessarily formed from synonymous verbs (‘throw (away); finish’, ‘finish’ and ‘remain’ respectively), and the fact that Uyghur has a rich variety of such auxiliaries (Hahn 1991: 612–18; 1998: 390–2) and intervening languages generally do not suggests a degree of coincidence.

- (4) a. Japanese: *tabe-te mi-ta*  
 Korean: *mek-e pw-assta*  
 Yonaguni: *ha-i nn-itan*  
 Mongolian: *ide-ž üz-sön*  
 Uyghur: *ič-ip kör-di/baq-ti*  
 eat-and see-PAST
- b. Ainu *ipe wa inu*  
 eat and hear  
 ‘He tried eating it.’
- (5) Japanese: *kaq-te oi-ta*  
 Classical Japanese: *kaf-i oki-keri*  
 Hatoma: *ffa-i suk-utan*  
 Korean: *s-a twu-essta/noh-assta*  
 Uyghur: *el-ip qoy-di*  
 buy-and put-PAST  
 ‘He bought it (in readiness for a future action).’

Characteristic of recent Japanese and Korean are complex forms to express deontic modality such as ‘must, have to’, though a modern pseudo-CJ equivalent was also developed (6a), whereas in southern Ryukyuan a simpler construction that cancels out the double negative is common (6b). Even here we cannot say the construction is uniquely Japanese-Korean, as we can find some equivalence in Mongolic.

- (6) a. Japanese *ik-ana-kereba nar-anai*  
 go-NEG-if become-NEG  
*ik-ana-kute-wa nar-anai*  
 go-NEG-and-TOP become-NEG
- pseudo-CJ *yuk-an-eba nar-anu*  
 go-NEG-if become-NEG
- Korean *an ka-myen an toy-nta*  
 NEG go-if NEG become-DEC
- Shuri *ik-an-too nar-an*  
 go-NEG-if become-NEG
- Hatoma *par-an-kaa nar-anu*  
 go-NEG-if become-NEG
- Mongolian *yav-ah-güy bol bol-oh-güy*  
 go-FUT-NEG TOP become-FUT-NEG  
 ‘It does/will not become if [I] do not go.’ > ‘I have to go.’
- b. Hatoma *par-aba-ru nar-u*  
 go-if-EMPH become-ATTR
- Yonaguni *hir-ja-du nar-u*  
 go-if-EMPH become-ATTR  
 ‘It does/will become if [I] go.’ > ‘I have to go.’

It is obvious from the above that there are significant grammatical similarities between Korean and Japanese that go beyond typology, yet there is little that is *unique* to the two languages in north Asia, and some similarities appear to have developed in historical times. Moreover, there are significant differences between Japanese and Korean. For example, Japanese is often cited as a canonical example of a language with systematic differences between male and female speech. Other than a stronger tendency for women to use politer speech styles (including honorific/humble forms and ‘beautified’ nouns) and tag questions, which is in any case part of a universal tendency for female speech to be politer (Lakoff 1975) and which is also observable in Korean, Japanese male/female differences are mostly realized in the choice of pronouns (e.g. male *boku*, female *atashi* ‘I’) and sentence-final particles (e.g. male *zo*, female *wa*, exclamatory; or male retention and female deletion of copula *-da* before exclamatory *yo*). The nature of Japanese female language has been much studied (Shibamoto 1985), and its origins have been traced back to LMJ ‘court ladies’ speech’ (*nyōbo kotoba*) and Edo-period ‘play ladies’ speech’ (*yūjogo*) (Ide and Terada 1998). In contrast, Korean has few overtly male/female differences, though the choice between plain style interrogatives *-ni* and *-nya* is primarily a male vs female choice. Phonologically, Middle Korean is notable by pervasive vowel harmony (archiphonemes  $U = u \sim \acute{o}$ ,  $wU = wu \sim o$ ,  $E = e \sim \acute{a}$ ) and strict vowel harmonic co-occurrence restrictions, whereas in Japanese vowel harmony has been noticeably lacking, beyond certain OJ co-occurrence restrictions: *o* and *wo* generally do not co-occur.

### 1.2.3 Chinese

The largest areal influence on Korean and Japanese comes from neither language, but from Chinese, specifically from varieties of Middle Chinese. There is some influence in grammar, e.g. grammatical words like K *hwu*, J *go* ‘after’ (< 後 EMC *huw* = LMC *xhəw* > C *hòu*),<sup>1</sup> as in K *ku hwu-ey*, J *sono go* ‘after that’, but most importantly a large repertoire of Chinese-derived classifiers as well as a complete set of Chinese-derived numbers acquired by the two languages. However, Japanese and even more so Korean each still has a set of native classifiers, and we know that Old Japanese had a small native classifier system. Even Ainu has a simple two-way classificatory system of human vs unmarked. Therefore, it seems likely that at least Japan had a classifier system before Chinese influence, but Chinese greatly enhanced the range of classifiers available.

The most significant influences from Chinese, however, are not grammatical: the Chinese character script (hereafter CCS); and the wealth of borrowed lexical morphemes.

#### 1.2.3.1 Script

CCS is a morphosyllabic writing system developed specifically to write, and therefore tailored to, Chinese. As Chinese morphemes are almost entirely monosyllabic, there is a one morpheme = one syllable = one character correspondence. Moreover, the overwhelming majority of characters are historically of the ‘xingsheng’ category, i.e. constructed from two components, one (the ‘signific’) hinting at the semantic field and the other (the ‘phonetic’) suggesting pronunciation, e.g. 齡 *ling* ‘age’ and 鈴 *ling* ‘bell’ are analyzable as <TOOTH + *ling*> and <METAL + *ling*>. The ‘phonetics’ still provide learners of modern Chinese with an extremely useful clue.

There is no evidence of any history of indigenous writing in Korea or Japan before Chinese influence; attempts to identify *jindai moji* ‘writing from the age of the gods’ in Japan have proved forged. A famous example (see Seeley 1991: 3–4; illustration reproduced in Taylor and Taylor 1995: 296) is a crude reworking of han’gŭl. Therefore, when Chinese influence appeared in Korea and (via Korea) in Japan, CCS was essentially the only writing system available.<sup>2</sup> Moreover, CCS accompanied the adoption of sophisticated Chinese culture and a sizeable literary canon that included the philosophical basis of that culture. It was not surprising, therefore, that writing in both Korea and Japan, which we can categorize as full-fledged and associate members respectively of the Chinese civilization, was heavily written in Chinese. In Korea, civil service examinations were in Chinese, including a Chinese poetry component. The sinicization of Korea was so great that most place names (see the eighth-century place name changes that provide much Early Old Korean data, Nam, this volume) and most people’s names are Sino-Korean. In medieval Korea the simple equation that China = civilization meant that any dilution of the Chinese component of society was a move towards barbarism, encapsulated famously in the 1444 memorial written (in Classical Chinese) by Ch’oe Malli [Choy Manli] to King Sejong [Seycong] arguing against the promulgation of the Korean alphabet: ‘Only the likes of the Mongols, Tanguts, Jurchens, Japanese and Tibetans have their own graphs. But these are the matters of barbarians, and not worth talking about. [. . .] To now separately make the Vernacular Script is to abandon China and identify ourselves with barbarians’ (translation from Ledyard 1998: 141). The Korean alphabet therefore remained secondary to CCS in Korea for 450 years; for much of this time it was regarded as a system used by the poorly educated including, perhaps especially, women. That the Korean alphabet is sometimes known as *amk’ul* [*amkhul*]

‘women’s writing’ is paralleled by the term *onnade* ‘women’s hand[writing]’ in Japan for hiragana, or even by the development by women in small southern Chinese communities of a more phonographic form of CCS, which scholars have named *nǚshū* (女書) ‘women’s writing’ (Chiang 1995; Zhao 1998).

In Korea or Japan, Classical Chinese provided both a ready-made script and language, and a long literary tradition. This advantage was counteracted by the fact that those writing in Korea or Japan spoke a native language rather than Chinese, and so if a writing system could be developed to write one’s *native* language it would be easier to write. Contrastive discussion of how CCS was adapted in Korea and Japan is difficult simply because of the wide variety, so instead of a detailed discussion I present a formulaic approach. (7) presents a classification of the values and form of individual characters, namely how they are used, adapted or created. The three parameters considered are: whether the character is structurally a Chinese(-style) character (**c**) in form, even if it did not exist in China, or if it is non-Chinese(-style) (**n**) in form; whether it is – or is simplified from – an authentic Chinese character (**C**) in origin, or is a Japanese or Korean manufactured item (**M**); and whether its value in a given text is phonographic (**[P]**) or semantographic (**[S]**). This three-way classification is still crude, as other factors such as ‘multivalence’ (the use of an individual character for more than one value or type of value), ‘morpheme-script mismatch’ (the graphic division of a polymorphemic word differently from its morphemic division, e.g. J 待つ <‘wait’> + <*tsu*> for *mats-u* ‘wait-NONPAST’), ‘switched-script writing’ (writing some monomorphemic words with a mix of character types, e.g. J 後ろ <‘after’> + <*ro*> for *ushiro* ‘behind’), and other disambiguating devices, such as writing phonograms smaller than semantograms (e.g. Bentley, this volume, 7.3), or whether a phonogram derives directly from a Chinese pronunciation or from a native word via a semantographic adaptation (cf. *ūmgaja* vs *hun’gaja*, Nam, this volume) are not included. As the discussion is of how CCS was *adapted* to writing Japanese or Korean, it makes little sense to apply the above three-way distinction to the writing in CCS of actual Chinese words (e.g. loanwords) or to the actual writing of *Chinese* as opposed to Japanese or Korean, so I categorize these separately (**X**). Other features necessary to categorize characters in actual texts are types of ‘double-writing’: # for explicit double-writing of (parts of) words in semantograms (**X** or **Cc/Mc**) + phonograms; ● for the use of other reading aids in Chinese texts, such as specially positioned dots over characters and similar devices (e.g. the ‘(w)okoto-ten’ in earlier Japan or K dot-type *kugyōl*, Nam, this volume) to indicate native endings or the order to be read.

(7) Adaptation of individual characters to write Japanese or Korean

- X** Characters used to write Chinese, or to write Chinese loanwords.
- Cc[S]** Semantographic use of existing CCS:  
 e.g. MK 花 *koc* ‘flower’ cf. MC 花 *xwæ* ‘flower’  
 e.g. CJ 花 *fana* ‘flower’ cf. MC 花 *xwæ* ‘flower’
- Cc[P]** Phonographic use of existing CCS: man’yōgana, hyangch’al and idu phonograms:  
 e.g. OK 乙 (*i*)*r* ACC cf. MC 乙 (\**ʔil* <)*ʔit* ‘2nd’  
 e.g. OJ 乎 *wo* ACC cf. MC 乎 *hu* question particle
- Mc[S]** Made-in-Japan/Korea CCS:  
 e.g. MK 石 *tol* ‘stone’ analyzable as <‘stone’ + *l*>  
 e.g. J 畑 *hatake* ‘vegetable field’ analyzed as <FIRE + ‘paddy’>

**Mc[P]** ?**Cn[P]** Simplification of CCS used phonographically:

- e.g. J め hiragana *me* from grass-written 女  
(cf. OJ/CJ 女 *me* ‘woman’)
- e.g. J ロ katakana *ro* from abbreviated 呂  
(cf. modern values SJ *ro*, C *lǚ*)
- e.g. OK □ kugyōl *ko* from abbreviated 古  
(cf. modern values SK *ku*, C *gǔ*)

**Cn[S]** Simplification of CCS used semantographically:

- e.g. OK ヌ kugyōl *hə-* ‘to do’ from abbreviated 為 (< 爲)  
(cf. Literary Chinese 爲 ‘to do’)
- e.g. J 𠂔 *koto* ‘fact’ from abbreviated (grass script?) 事  
(now obsolete) (cf. Literary Chinese 事 ‘fact’)

**Mn[P]** Non-CCS derived characters: K han’gūl.**Mn[S]** Non-CCS derived characters: Yonaguni kaida.

- # Semantographic and phonographic double-writing
- Use of ‘reading aid’ dots. (Typically used to ‘interpret’ X texts.)

There are various points that arise from the above. Firstly, CCS was intimately tailored to writing Chinese, and any semantographic adaptation of it to write a different language than that to which it was tailored means that the dominant xingsheng loses its clues and the characters become more difficult to learn and to use. For example, the phonetic clue is lost in the adaptation of 齡 and 鈴 above to write J *yowai* ‘age’ and *suzu* ‘bell’ respectively. Secondly, there are certain but isolated cases of fused character, where two characters that have values independent of each other have been fused into a single form. Phonogram + phonogram examples include J 磨 *maro* ‘lad’, also written unfused as man’yōgana 麻呂 *ma + ro*, both still found in (historical) personal names; and the fused katakana writings of *toki* and *-domo* in Tranter and Kizu (this volume, 10.3.3). Semantogram + semantogram examples include J 𣎵 ‘type of tree’ < 正 *masa-* ‘true’ + 木 *ki* ‘tree’. The large part of made-in-Korea characters (**Mc[S]**) arguably fall into this category as semantogram + idu phonogram/semantogram fusions (see Sasse’s (1980) list); semantogram + phonogram cases when they write monomorphemic words, such as 𠂔, belong to what I termed ‘switched-script writing’ above, deriving from 石 <‘stone’> + 乙 </>. Thirdly, the limits of pictography mean that the **Mn[S]**-based ‘script’, Yonaguni kaida (Izuyama with Rosa, this volume; Rosa 2006), is incapable of writing full sentences, and is what DeFrancis (1989) terms ‘partial writing’. Finally, the only instance of invention of a true fully functioning script that does not derive from CCS is Korea, where King Sejong commissioned, and in 1446 promulgated, the Korean alphabet known officially at the time as *hunmin chōngūm* [hwunmin cengum] ‘correct sounds to teach the people’ and unofficially as *ōnmun* [enmun] ‘vernacular script’. In the twentieth century it was renamed han’gūl [hankul], the term that it now has. Discussed often as a unique invention, its basic consonant letters show similarities to ’phags-pa, the universal but short-lived alphabet commissioned by Qubilai Khan in the 1260s (Poppe 1957; Van der Kuijp 1996), or to the latter’s ultimate source, Tibetan *dbu can*. However, most of the consonant letters in han’gūl are derived by adding strokes to other letters (ㄱ /k/, ㄷ /t/, ㅈ /c/ > ㅋ /k<sup>h</sup>/, ㅌ /t<sup>h</sup>/, ㅊ /c<sup>h</sup>/) and the vowel letters appear to be combinations of just three ideal strokes · — |, so han’gūl is far from ‘no more than a transparent adaptation’ of ’phags-pa as Miller

(1996: 80) asserts. Han'gŭl is more strictly an alphasyllabary, in that, although it consists of discrete letters representing consonant and vowel segments, these are grouped into syllabic blocks reminiscent of the one syllable = one character principle of CCS.

To describe how the above character types are used in an extended text, (8) lists the main options for the distribution of the semantographic and phonographic values, while (9) lists the options for written 'word order'.

(8) Semantographic/phonographic distribution

- 0** Write in Chinese
- 1** Write K/J semantographically
- 2** Write K/J phonographically
- 12** Write K/J in a semantographic/phonographic mix
- 1(2)** Write K/J predominantly semantographically, but with some phonography
- 2(1)** Write K/J predominantly phonographically, but with some semantography

(9) Written 'word order'

- I** Write in Chinese order, e.g. SVO
- II** Write in K/J order, e.g. SOV
- II(I)** Write predominantly in K/J order, but with some sequences written in easily decodable Chinese order (cf. order difference in English \$5 vs *five dollars*)

Some comments are in order. Firstly, distribution **1** – writing entirely semantographically – always results in underspecification, because the adaptation of a writing system created for an isolating language that has few grammatical morphemes (Chinese) cannot readily be adapted to writing an agglutinative language that has a wide range of grammatical morphemes (Korean, Japanese). As a result, early idu essentially lacks any representation of noun-following particles and verb inflections. Secondly, order **I** – writing an SOV language in Chinese word order – may seem immensely impractical, but it was the writing device 'pseudo-Chinese' (J *hentai kanbun*, K *pyŏnch'e hanmun*) that was used to write the entire OJ prose of eighth-century *Kojiki*, and characterizes early OK idu texts. Aldridge (2000, 2001) has shown that *Kojiki* prose was encoded in (and decodable from) Chinese word order by extremely regular rules.

We can therefore characterize typical contemporary writing in Japan and Korea formulaically as in (10).

- (10) Modern Japanese:  $\mathbf{II} = \mathbf{12} = \mathbf{X}, \mathbf{Cc[S]}, \mathbf{Cn[P]} \quad \# = \mathbf{X/Cc[S]} \times \mathbf{Cn[P]}$   
 Modern N. Korean:  $\mathbf{II} = \mathbf{2} = \mathbf{Mn[P]}$   
 Modern S. Korean: Normal:  $\mathbf{II} = \mathbf{2} = \mathbf{Mn[P]}$   
 Academic:  $\mathbf{II} = \mathbf{2(1)} = \mathbf{Mn[P]} \quad \# = \mathbf{Mn[P]} \times \mathbf{X}$

(11) and (12) illustrate the great variety that has existed in Japan and Korea respectively, as exemplified in Nam (this volume), Bentley (this volume) and Tranter (this volume). For comparison, (13) suggests a classification of typical Vietnamese (chữ nôm) adaptation of CCS, as well as Mongolian's brief experimentation with CCS in *The Secret History of the Mongols*.

- (11) OJ: *Nihon shoki*: prose:  $\mathbf{0} = \mathbf{X}$   
 poetry:  $\mathbf{II} = \mathbf{2} = \mathbf{Cc[P]}$   
*Kojiki*: prose:  $\mathbf{I} = \mathbf{1(2)} = \mathbf{Cc[S]}, \mathbf{Cc[P]}$   
 poetry:  $\mathbf{II} = \mathbf{2} = \mathbf{Cc[P]}$

- Man'yōshū* poetry:  $\mathbf{II} = 2 = \mathbf{Cc[P]}$   
 $\mathbf{II(I)} = 2(1) = \mathbf{Cc[P], Cc[S]}$   
 $\mathbf{II(I)} = 12 = \mathbf{Cc[S], Cc[P]}$
- MJ: *Tosa nikki* (except dates):  $\mathbf{II} = 2 = \mathbf{Cn[P]}$   
 Heian female prose:  $\mathbf{II} = 2(1) = \mathbf{Cn[P], Cc[S]}$   
*Heike monogatari*:  $\mathbf{II} = 12 = \mathbf{Cc[S], Cn[P]} \quad \# = \mathbf{X/Cc[S]} \times \mathbf{Cn[P]}$
- (12) OK: earliest idu:  $\mathbf{II} = 1 = \mathbf{Cc[S]}$   
*hyangga*:  $\mathbf{II} = 12 = \mathbf{Cc[S], Cc[P]}$   
 kugyōl texts:  $\mathbf{0} = \mathbf{X} \quad \# = \mathbf{X} \times \mathbf{Cn[P/S]}, \bullet$
- LMK: *Yongpi echenka*:  $\mathbf{II} = 12 = \mathbf{X, Mn[P]}$   
 most han'gŭl texts:  $\mathbf{II} = 2 = \mathbf{Mn[P]}$   
 Some early 20C Korean:  $\mathbf{II} = 12 = \mathbf{X, Mn[P]}$
- (13) Vietnamese chŭr nôm:  $\mathbf{II} = 1(2) = \mathbf{X, Cc[S], Mc[S], Cc[P]}$   
*Secret History of the Mongols*:  $\mathbf{II} = 2 = \mathbf{Cc[P]}$

Ryukyuan varieties and Ainu have no 'standard' scripts. When written, Shuri Ryukyuan ('standard' Okinawan) tends to be written in mixed script, with kana and some kanji. An example of maximal use of kanji is illustrated by the opening sentences (14) of Issue 1 of *Okinawa Hōgen Shinbun* 'Okinawa Dialect Newspaper' (Okinawa Hōgen Fukyū Kyōgikai 2002), which clearly found the use of kanji with superscript phonographic double-writing helped those less familiar with the variety to understand it. (Upper vs lower case represents semantography vs phonography.) Devices used by others include Nakamatsu and Funatsu's (1988) invention of a series of new kana to write Okinawan morae that do not exist in Japanese.

- (14) 

うちなーぐち 沖繩語	ふいる 広みーる	じんみ 協議ぬ	くわ 会えー	くんどう 今度
UCINAA-GUCI	HWIRU-mii-ru	ZINMI-nu	KWAee,	KUNDU
Okinawa-language	wide-make-ADN	deliberation-GEN	society.TOP	now
- |                       |                    |               |           |
|-----------------------|--------------------|---------------|-----------|
| 沖繩方言新聞                | はじ<br>初みてい         | っん<br>出じゃする   | くとう<br>事に |
| [NEWSPAPER NAME]      | HAZImi.ti          | 'Nzasu-ru     | KUTU-ni   |
| Okinawa Hōgen Shinbun | for.the.first.time | bring.out-ADN | fact-DAT  |
- |                    |                      |            |                |
|--------------------|----------------------|------------|----------------|
| ないびたん。             | うちなーぐ<br>沖繩語ちえー      | しでー<br>次第に | わ<br>解からん      |
| na(j)-ibi-ta-n.    | UCINAA-GUcee         | SIDEE.ni   | WAKar-an       |
| become-POL-PST-IND | Okinawa-language.TOP | gradually  | understand-NEG |
- |            |             |             |            |                     |
|------------|-------------|-------------|------------|---------------------|
| なてい、       | ちか<br>使いる   | つちゅ<br>人ん   | いき<br>少らく  | なとーいびーん。            |
| na-ti,     | CIKA(j)i-ru | CCU-n       | IKIra-ku   | na-too(j)-ibii-n.   |
| become-GER | use-ADN     | person-even | be.few-INF | become-PROG-POL-IND |
- 'The Society for Promotion of Okinawan has come now to publish for the first time the Okinawa Hōgen Shinbun [lit. 'Okinawa Dialect Newspaper']. Regarding Okinawan, the number of people who understand and use it has become fewer and fewer.'

The scripts invented in Korea and Japan have not been adapted practically to languages beyond their borders. A recent exception is the promotion of han'gŭl (by a private organization dedicated to exporting han'gŭl) in 2009 in textbooks to teach Ciacia on Buton island off Sulawesi (*New York Times*, 12 September 2009).

## 1.2.3.2 Sino-Japanese and Sino-Korean

Together with Chinese civilization and Chinese script came Chinese vocabulary, and vocabulary from China or built from Chinese morphemes is known in the West as Sino-Japanese (SJ) and Sino-Korean (SK), and in Japan and Korea as *kango* and *hanjaō* [*hancae*] respectively. The influx of Chinese vocabulary has been so pervasive that it is conventional in Japan and Korea to distinguish SJ and SK from other types of loanword (J *gairaigo*, K *oeraeō* [*oylaye*]), and Western works therefore also normally use ‘loanword’ to exclude SJ and SK. However, there is also some justification in this distinction, not least because, especially in Japanese, a large proportion of everyday SJ polymorphemic words were invented in Japan rather than borrowed from China; in a great many cases, SJ coinages have been loaned into Korean and Chinese. This is primarily in the area of modern political, sociological, scientific, engineering, economic and military terminology, typically but not only nouns. The Korean, Japanese and Chinese terms within each row of (15) are all the same Chinese-based word in origin, and with one notable exception (15e) are written identically in traditional CCS. (Post-war character simplifications in Japan and the People’s Republic of China are not given in these examples.)

(15)	SK	SJ	Chinese	CCS	
a.	<i>chwulphan*</i>	<i>shuōpan*</i>	<i>chūbǎn</i>	出版	‘to publish*; publishing’
	<i>uywen</i>	<i>giin</i>	<i>yìyuán</i>	議員	‘parliamentarian’
	<i>kongep</i>	<i>kōgyō</i>	<i>gōngyè</i>	工業	‘industry’
	<i>cengchayk</i>	<i>seisaku</i>	<i>zhèngcè</i>	政策	‘(political) measure’
	<i>sangep</i>	<i>shōgyō</i>	<i>shāngyè</i>	商業	‘commerce’
	<i>nayyong</i>	<i>naiyō</i>	<i>nèiróng</i>	內容	‘content’
	<i>cengtang</i>	<i>seitō</i>	<i>zhèngdang</i>	政黨	‘(political) party’
	<i>kwukhoy</i>	<i>koōkai</i>	<i>guóhuì</i>	國會	‘parliament’
b.	<i>sahoy</i>	<i>shakai</i>	<i>shèhuì</i>	社會	‘society’
	<i>seykyey</i>	<i>sekai</i>	<i>shìjiè</i>	世界	‘world’
	<i>hyengmyeng</i>	<i>kakumei</i>	<i>gémìng</i>	革命	‘revolution’
	<i>-cwuuy</i>	<i>-shugi</i>	<i>-zhǐyì</i>	主義	‘-ism’
c.	<i>peplyul</i>	<i>hōritsu</i>	<i>fǎlǜ</i>	法律	‘law’
	<i>senke</i>	<i>senkyō</i>	<i>xuǎnjǔ</i>	選舉	‘election’
d.	<i>kongwen</i>	<i>kōen</i>	<i>gōngyuán</i>	公園	‘park’
e.	<i>notong</i>	<i>rōdō</i>	<i>láodòng</i>	J: 勞働 C/K: 勞動	‘labour’

\*: in verbal sense, J adds *-suru* and K adds *-hata* ‘do’.

Those examples in (15a) are widely accepted to be Japanese neo-classical coinages from Chinese morphemes, which were then borrowed by Korea and China. Because earlier Japanese had nativized Chinese verbs by adding the verb ‘do’ (Tranter, this volume, 8.6), neo-classical forms also included verbs. Because of Japan’s earlier modernization than China or Korea, the direction of graphic borrowing has tended to be regarded as overwhelmingly from Japan to China and Korea (e.g. Miller 1967). However, it has been shown that Japan resurrected and/or redefined some obsolete

Chinese words (Mair 1992; Masini 1993; Liu 1995) such as in (15b), sometimes with massive semantic change, e.g. ‘society’ < OC ‘festal gathering around communal altar’ (Mair 1992: 11; Liu 1995: 336), or even in the case of ‘-ism’ < OC ‘to comply with the virtue; proposal’ the change of a noun (Masini 1993: 220) into a derivational suffix.

Moreover, there were also many forms invented in China that Japan then borrowed (15c); the words may have taken root faster in Japan, and Japanese may therefore have reinforced their establishment/diffusion in China, but they are not Japanese coinages. Masini (1993: 149, 165–8) gives the example of ‘telegram’, where China coined one word, Japan another, and both loaned their words to the other as shown in (16). Korean and Vietnamese examples are added from Tranter (forthcoming).

- (16) a. C 電報  
*diànbào* ‘telegram’ > J *dēnpō* > K *cenpo*  
 > V *điện-báo*
- b. J 電信  
*dēnshin* ‘telegram/telegraphy’ > K *censin*  
 > C *diànxìn* ‘telegraphy’ > V *điện-tín*

There are many cases where it is not clear which country first created or redefined a word. Among these are examples of apparent coincidence (15d), sometimes with quite different meanings, as in the case of ‘park’ vs earlier Chinese ‘official’s garden’. However, all cases of such pan-East Asian common vocabulary for modern concepts, the origin is taken to be either Japan or China; it is accepted that Korea borrowed these terms from Japan (Sohn 1999: 104), in the same way that Vietnam – which also shares most of these terms – borrowed from China (Nguyễn 1980: 97). It is unlikely that future research will change our view of modern Korea primarily as a borrower.

The most notable feature that links all the words in (15) is the fact that they are not phonetic loans, but graphic loans. For example, the word for ‘commerce’ coined in Japan, *shōgyō*, is obviously not a phonetic loan because it lacks the /ŋ/ or /p/ that Chinese and Korean have (*shāngyè*, *sangep*), whereas what the three languages (and indeed also Vietnamese *thương-nghiệp*) have in common is the graphic form: in traditional CCS the word is written identically in all the languages. The morphemes had been borrowed from Middle Chinese (e.g. EMC 商 *syang* ‘trade’, 業 *ngjæp* ‘work’), had been adapted phonologically (cf. Late Middle Korean *sjang*, *ngep*; Japanese pre-1946 ‘Classical’ kana spellings *syau*, *gefu*), then over the centuries underwent sound changes internal to each borrowing language (e.g. Korean morpheme-initial *sj-* > *s-*, initial *ng-* >  $\emptyset$ ; Japanese *au* > *ââ* > *ō*, *efu* > (*y*)*eu* > *yō*) as well as within Chinese (e.g. the loss of morpheme-initial /ŋ/- and morpheme-final /p/ in Mandarin). As a result, not only had the morphemes evolved in Chinese since MC, but each borrowing language had its modern conventional readings for the characters: EMC *syang* > C *shāng*, V *thương*, K *sang*, J *shō*; EMC *ngjæp* > C *yè*, V *nghiệp*, K *ep*, J *gyō*. Essentially, Vietnamese, Korean and Japanese had an entire lexical stratum of potentially every Chinese morpheme at its disposal, a situation which Miller in reference to Japan refers to as ‘total availability’ (1967: 244–5).

In the nineteenth century, China, Japan and Korea faced military and economic pressure from the modern, industrialized countries of the United States and Europe, and had a history of minimizing contact with the West and its influences. For example, Japan had a deliberate policy of being an isolated nation (*sakoku*, lit. ‘chain[ed off] country’) since the start of the seventeenth century, and its international contact had exclusively

been with the Dutch, who were isolated to a single island in Nagasaki bay; Korea had a similar policy – nowadays referred to as the ‘hermit nation’; and even trading contact between China and Japan had been indirect, mediated through the Ryūkyū kingdom that fell under Japanese suzerainty. In the nineteenth century, China was particularly humiliated in its confrontation with the military and economic might of the West in 1840–1842 and 1857–1860. With the arrival of four warships on the Japanese coast in 1853, the United States initiated the economic opening up of Japan. The end of the shogunate in 1868 and the establishment of the Meiji period (1868–1912) famously saw Japan embark on large-scale Western-style modernization. As part of this, Japan needed to adapt or create vocabulary for the wealth of modern – Western – political, social, economic, military, technological and generally academic vocabulary, and the neo-classical compound, exploiting the ‘total availability’ of Chinese morphemes in their written (CCS) form, was the device of choice. In this way, two morphemes – or rather written characters – were combined to create a term for ‘commerce’. In the later nineteenth century, as Japan was modernizing, Korea and China were continuing to resist the modern world. Korean and Chinese intellectuals who wanted to study the modern disciplines went to Japan to study, and back in their own countries would translate Japanese texts. In this way a word such as ‘commerce’ entered China and Korea in its *written* form 商業 – i.e. as a graphic loan – and the component characters were read out in both countries according to the established readings associated with those characters. The influence became strongest on Korea: within a decade of the start of the Meiji period, Japan had emulated Western tactics and forced Korea to open up economically, and its influence increased till its defeat of China and installation of a pro-Japanese government in the Sino-Japanese War (1894–1895), resulting in formal annexation of Korea to the Japanese Empire in 1910 and the establishment of Japanese as the only official language of the Empire till liberation in 1945. This long period of Japanese domination of Korea facilitated the process of large-scale diffusion of Japanese-coined words into Korean.

More extreme cases of graphic borrowing are where the Japanese source form is not even read with Sino-Japanese readings, but belongs to the *native* Japanese stratum. In these cases, the Koreans still borrowed the word graphically and read it with the conventional *Sino-Korean* readings associated with each character. The result is that the Japanese and Korean forms are phonetically totally different and unrelated, despite their common written form. A similar process occurred in China. Some examples (adapted from Tanaka and Lee (1986: 128) and Liu (1995)), include:

- (17) J 組合 *kumiai* ‘[e.g. labour] union’ > K *cohap*  
 J 割引 *waribiki* ‘discount’ > K *halin*  
 J 立場 *tachiba* ‘viewpoint’ > K *ipcang* (also C *lichāng*, V *lập-trường*)

An interesting case example is ‘labour’ (15e) above. The ‘traditional’ character spelling involves a slightly different second character in Japan (働 ‘to work’) versus China and Korea (動 ‘to move’). When the word was first used in Japan, the latter spelling did also occur, but the semantics of the former fitted better with the meaning ‘labour’. The reason why the Chinese and Korean did not use 働 but either adopted the other early 動 version, or more likely just deleted the first (leftmost) two strokes of 働 to 動, is simple: 働 was a made-in-Japan character (<HUMAN + ‘to move’> in structure) that never existed in China, designed to write the native Japanese word *hataraku* ‘to work’. At some point a pseudo-Sino-Japanese reading *dō* was associated with 働 (the same as the true Sino-Japanese reading of 動). The Chinese and Korean character spellings are therefore ‘corrections’ from their viewpoint.

### 1.2.4 Anglo-Japanese and Anglo-Korean

The impact of English on South Korean and Japanese is predominantly a post-Second World War phenomenon. Both languages certainly borrowed from English in the half-century or so before, but the amount borrowed was on a much smaller scale, and other European languages also contributed loanwords, such as some medical terms from German. Since the Second World War, however, Japan and South Korea have been close allies of the United States, and a combination of great interest in American culture and the chic value placed upon the English language by Japanese and Koreans has led to an amazing influx of Western words, which can constitute a significant percentage of the lexemes in many genres of text (see Irwin (2011: 14–21) for a review of recent statistics and analysis in the case of Japanese). Some relatively new areas of technology, particularly computer technology, have particularly large quantities of English loans. In both countries populist writing and particularly advertising use more English loans. Miller (1967: 249) in connection with Japanese characterizes the role of English in Japan as a similar ‘total availability’ as applies to Chinese morphemes. This parallel to Chinese morphemes is revealed in several ways. Firstly, there is a strong tendency to combine English words or even roots into new compounds and phrases that do not exist in English. Secondly, there has been a tendency for Korean to borrow English words via Japanese, and even though post-war South Korea has purged overtly Anglo-Japanese forms and appears to borrow directly from English, there are numerous features in Korean forms that show a Japanese source, e.g. the wrong phoneme in *K nyusu* ‘news’ or *kullowucu-ep* ‘close-up’ (cf. *J nyūsu*, *kurōju-aōpu*), the nature of truncation in *pheme* ‘perm’ or *theylleypi* ‘television’ (cf. *J pāma*, *terebi*), or the fact that a compound does not exist in English as in the case of *othopai* ‘motorcycle’ or *khonseynthu* ‘electrical socket’ (cf. *J ōtobai*, *konsento*) (Tranter 1997). Thirdly, as discussed in Tranter (forthcoming), post-Second World War English loans in Japanese display such a range of spelling-influenced features and are distinct from genuine phonetic loans that English loanwords in Japanese at least can be characterized as predominantly graphic loans that, once they have entered Japanese, are adapted into Japanese script, and such loans into Korean arguably are the same in nature.

### NOTES

- 1 EMC follows the Baxter (1992) system, while LMC follows Pulleyblank (1991). Tones are not reproduced.
- 2 Arguably, Indic scripts used for Sanskrit or Pali and familiar in China at the time may also have been available; however, Indic writing was never adapted to write Chinese, so would equally have been unlikely to be adapted to writing Korean or Japanese. Indic writing is still found in Buddhist contexts in Korea and Japan, such as in connection with graves or statuary. It is notable that East Asian influence has meant that the Sanskrit/Pali written this way is normally written vertically in syllabic blocks rather than horizontally as is the case in South and South-East Asian Indic scripts. Indic alphabetical order has also been taken as the source of the syllabary order in Japanese dictionaries (*kV sV tV nV hV* (< OJ *pV*) *mV yV rV wV* from Indic *k . . . c . . . t . . . n p . . . m . . . y r/l v . . .*, and *Ca Ci Cu Ce Co* from Indic *a ā ī ī u ū e ai o au*), and is additional evidence for an affricate pronunciation of OJ *s* as Indic *s* is ordered later than *v*.

### REFERENCES

- Akimoto, Kichirō (ed.) (1958) *Iwanami Shoten Nihon Koten Bungaku Taikei 2: Fudoki*, Tokyo: Iwanami Shoten.

- Aldridge, Edith (2000) 'Principles of *hentai kanbun* word order: evidence from the *Kojiki*', in T. E. McAuley (ed.) *Language Change in East Asia*, Richmond, UK: Curzon, 207–32.
- (2001) '*Hentai kambun* perspective on short scrambling', *Journal of East Asian Linguistics*, 10: 169–200.
- Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.
- Chiang, William W. (1995) '*We Two Know the Script; We Have Become Good Friends*': *Linguistic and Social Aspects of the Women's Script Literacy in Southern Hunan, China*, Lanham, MD: University Press of America.
- DeFrancis, John (1989) *Visible Speech: The Diverse Oneness of Writing Systems*, Honolulu: University of Hawai'i Press.
- Dryer, Matthew S. (2003) 'Word order in Sino-Tibetan languages from a typological and geographical perspective', in Graham Thurgood and Randy J. LaPolla (eds) *The Sino-Tibetan Languages*, London and New York: Routledge, 43–55.
- Frellesvig, Bjarke and Whitman, John (2008) 'The Japanese–Korean vowel correspondences', in Mutsuko Endo Hudson, Sun-Ah Jun, Peter Sells, Patricia M. Clancy, Shoichi Iwasaki and Jung-Ock Sohn (eds) *Japanese/Korean Linguistics 13*, Stanford: Stanford University CSLI Publications, 15–28.
- Fukuda, Yoshisuke (1965) *Nara Jidai Tōgoku Hōgen no Kenkyū* [Research into the Azuma dialects of the Nara period], Tokyo: Kazama Shobō.
- Gordon, Raymond G. (ed.) (2005) *Ethnologue: Languages of the World*, 15th edition, Dallas: SIL International.
- Hahn, Reinhard F. (1991) *Spoken Uyghur*, Seattle and London: University of Washington Press.
- (1998) 'Uyghur', in Lars Johanson and Éva Á. Csató (eds) *The Turkic Languages*, London and New York: Routledge, 379–96.
- Heinrich, Patrick (2004) 'Language planning and language ideology in the Ryūkyū islands', *Language Policy*, 3(2): 153–79.
- Hino, Sukenari (2003) 'Nihon sogo no boin taikai – jōdai tōgoku hōgen shiryō ni yoru saikō' [A reconstruction of the proto-Japanese vowel system on the basis of the data from Eastern Old Japanese], in Alexander Vovin and Toshiki Osada (eds) *Nihongo keitōron no genzai* [Perspectives in the origins of the Japanese language], Nichibunken Japanese Studies Series 31, Kyoto: International Research Center for Japanese Studies, 187–206.
- Hirayama, Teruo, Ōshima, Ichirō and Nakamoto, Masatomo (1966) *Ryūkyū Hōgen no Sōgō-teki Kenkyū* [Comprehensive research into the Ryukyū dialects], Tokyo: Meiji Shoin.
- Hōjō, Tadao (1966) *Jōdai Tōgoku Hōgen no Kenkyū* [Research into the dialects of ancient Azuma], Tokyo: Nippon Gakujutsu Shinkōkai.
- Ide, Risako and Terada, Tomomi (1998) 'The historical origins of Japanese women's speech: from the secluded worlds of "court ladies" and "play ladies"', *International Journal of the Sociology of Language*, 129: 139–56.
- Irwin, Mark (2011) *Loanwords in Japanese*, Amsterdam and Philadelphia: John Benjamins.
- Kim, Min-joo (2004) 'Event-structure and the internally-headed relative clause construction in Korean and Japanese', unpublished doctoral thesis, University of Massachusetts Amherst.
- Kuroda, S.-Y. (1992) 'Pivot-independent relativization in Japanese', in *Japanese Syntax and Semantics: Collected Papers*, Dordrecht: Kluwer, 114–74.
- Lakoff, Robin (1975) *Language and Woman's Place*, New York: Harper and Row.
- Ledyard, Gari K. (1998) *The Korean Language Reform of 1446*, Seoul: Sinkwu Munhwasa.
- Lewis, M. Paul (ed.) (2009) *Ethnologue: Languages of the World*, 16th edition, Dallas: SIL International.

- Liu, Lydia H. (1995) *Translingual Practice: Literature, National Culture, and Translated Modernity – China, 1900–1937*, Stanford: Stanford University Press.
- Mair, Victor H. (1992) ‘East Asian round-trip words’, *Sino-Platonic Papers*, 34: 5–13.
- Martin, Samuel E. (1966) ‘Lexical evidence relating Korean to Japanese’, *Language*, 42(2): 185–251.
- Masini, Federico (1993) ‘The formation of modern Chinese lexicon and its evolution toward a national language: the period from 1840 to 1898’, *Journal of Chinese Linguistics*, monograph series no. 6, 1–295.
- Matsumori, Akiko (1995) ‘Ryūkyuan: past, present, and future’, in John C. Maher and Kyoko Yashiro (eds) *Multilingual Japan*, Clevedon, Philadelphia and Adelaide: Multilingual Matters, 19–44.
- Miller, Roy Andrew (1967) *The Japanese Language*, Chicago and London: University of Chicago Press.
- (1971) *Japanese and the Other Altaic Languages*, Chicago and London: University of Chicago Press.
- (1996) *Languages and History: Japanese, Korean, and Altaic*, Bangkok: White Orchid Press.
- Mizushima, Yoshiharu (1984) *Man'yōshū Azumauta no Kokugogaku-teki Kenkyū* [Linguistic research into the *Man'yōshū*'s eastern poems], Tokyo: Kasama Shoin.
- Moseley, Christopher (ed.) (2010) *Atlas of the World's Languages in Danger*, 3rd edition, Paris: UNESCO Publishing. (Online version *UNESCO Interactive Atlas of the World's Languages in Danger*, <http://www.unesco.org/culture/languages-atlas/>)
- Nakamatsu, Takeo and Funatsu, Yoshiaki (1988) *Utsukushii Okinawa no Kotoba* [The beautiful Okinawa language], Tokyo: Gikōsha.
- Nguyễn, Đình-Hoà (1980) ‘Standardization and purification: a look at language planning in Vietnam’, in Patricia Mỹ-Hường Nguyễn Thị (ed.) *Language in Vietnamese Society: Some Articles by Nguyễn, Đình-Hoà*, Carbondale, IL: Asia Books, 87–108.
- Okinawa Hōgen Fukyū Kyōgikai (2002) *Okinawa Hōgen Shinbun*, Issue 1. ([www4.ocn.ne.jp/~knaka/hogen/shinbun/index.html](http://www4.ocn.ne.jp/~knaka/hogen/shinbun/index.html))
- Onishi, Takuichiro (2008) ‘Proto-Japanese and the distribution of dialects’, in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese: Issues and Prospects*, Amsterdam and Philadelphia: John Benjamins, 57–78.
- Poppe, Nicholas (1957) *The Mongolian Monuments in ḥP'ags-pa Script*, 2nd edition, trans. J. R. Krueger, Wiesbaden: Otto Harrassowitz.
- Pulleyblank, Edwin G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver: UBC Press.
- Rosa, Mark (2006) ‘*Kaida* Writing in the Yaeyama Islands’, unpublished Master's thesis, Linguistics Department, University of Tokyo.
- Sasse, Walter (1980) ‘“Chinesische” Zeichen erfunden in Korea’ [“Chinese” characters discovered in Korea], *Asiatische Studien* [Asian studies], 34: 189–205.
- Schönig, Claus (2003) ‘Turko-Mongolic relations’, in Juha Janhunen (ed.) *The Mongolic Languages*, London and New York: Routledge, 403–19.
- Seeley, Christopher (1991) *A History of Writing in Japan*, Leiden: E. J. Brill.
- Shibamoto, Janet S. (1985) *Japanese Women's Language*, Orlando: Academic Press.
- Shibatani, Masayoshi (1990) *The Languages of Japan*, Cambridge: Cambridge University Press.
- Sohn, Ho-Min (1999) *The Korean Language*, Cambridge: Cambridge University Press.
- Tamura, Suzuko (2000) *The Ainu Language*. Tokyo: Sanseido.
- Tanaka, Katsuhiko and Lee, Yeonsuk (1986) ‘Aspekte der japanisch-koreanischen Sprachkontakte und Lehnbeziehungen’ [Aspects of Japanese–Korean language contact and loanwords], *International Journal of the Sociology of Language*, 58: 123–8.
- Taylor, Insūp and Taylor, M. Martin (1995) *Writing and Literacy in China, Korea and Japan*, Amsterdam and Philadelphia: John Benjamins.

- Thorpe, Maner Lawton (1983) 'Ryūkyūan Language History', unpublished thesis, University of Southern California.
- Tranter, Nicolas (1997) 'Hybrid Anglo-Japanese loans in Korean', *Linguistics*, 35: 133–66.
- (forthcoming) 'Graphic loans: East Asia and beyond', *Word*, 60.
- Van der Kuijp, Leonard W. J. (1996) 'The Tibetan script and derivatives', in Peter T. Daniels and William Bright (eds) *The World's Writing Systems*, New York and Oxford: Oxford University Press, 431–41.
- Vovin, Alexander (2010) *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*, Honolulu: University of Hawai'i Press.
- Whitman, John (1985) 'The phonological basis for the comparison of Japanese and Korean', unpublished thesis, Harvard University.
- Zhao, Liming (1998) 'Nūshu: Chinese women's characters', *International Journal of the Sociology of Language*, 129(1): 127–37.

# THE RELATIONSHIP BETWEEN JAPANESE AND KOREAN

*John Whitman*

## 2.1 INTRODUCTION

This chapter reviews the current state of Japanese-Ryukyuan (JR) and Korean internal reconstruction and applies the results of this research to the historical comparison of both families. Reconstruction within the families shows proto-Japanese-Ryukyuan (pJR) and proto-Korean (pK) to have had very similar phonological inventories, with no laryngeal contrast among consonants and a system of six or seven vowels. The main challenge for the comparativist is working through the consequences of major changes in root structure in both languages, revealed or hinted at by internal reconstruction. These include loss of coda consonants in Japanese, processes of syncope and medial consonant lenition in Korean.

The chapter then reviews a small number (50) of pJR/pK lexical comparisons in a number of lexical domains, including pronouns, numerals, and body parts. These expand on the lexical comparisons proposed by Martin (1966) and Whitman (1985), in some cases responding to the criticisms of Vovin (2010). It identifies a small set of cognates between pJR and pK, including approximately 13 items on the standard Swadesh 100 word list: ‘I’, ‘we’, ‘that’, ‘one’, ‘two’, ‘big’, ‘long’, ‘bird’, ‘tall/high’, ‘belly’, ‘moon’, ‘fire’, ‘white’ (previous research identifies several more cognates on this list). The chapter then concludes by introducing a set of cognate inflectional morphemes, including the root suffixes \*-i ‘infinitive/converb’, \*-a ‘infinitive/irrealis’, \*-or ‘adnominal/nonpast’, and \*-ko ‘gerund’.

In terms of numbers of speakers, Japanese-Ryukyuan and Korean are the largest language isolates in the world. I use the term ‘isolate’ here in the informal sense that includes both true isolates and small language families. Japanese-Ryukyuan, often misleadingly treated as the unitary language ‘Japanese’ rather than a language family, ranks ninth, with 122 million first language speakers; Korean ranks seventeenth, with 66.3 million (Lewis 2009, [http://www.ethnologue.com/ethno\\_docs/distribution.asp?by=size](http://www.ethnologue.com/ethno_docs/distribution.asp?by=size)). The next largest language or family-level isolate is Quechua at seventy-ninth, with 10.1 million speakers. Like Quechua(n) and Aymara(n), Japanese-Ryukyuan is unquestionably a language family, made up of the five Ryukyuan languages and a variety of dialect clusters in the Japanese main islands. Korean dialects show less depth of separation. Due to their size and importance, their proximity and long cultural contact, and their well-known typological similarity, scholars have debated the possibility of a genetic relation between Japanese and Korean for over two centuries. All scholars agree that a genetic affiliation between the two, if valid, is distant. At the same time, Japanese-Ryukyuan and Korean are distributed

in adjacent areas, and we know that their historical distribution overlapped. There is a consensus that at some point a relative of pJR was spoken on the Korean peninsula, although there is controversy over the relevant evidence. The evidence has been known for almost a century, beginning with Shinmura's (1916) observation that words preserved in toponymic material in the twelfth-century Korean history *Samguk sagi* show close resemblances with Japanese. The combined factors of genetic distance and geohistorical proximity resemble the case of Quechuan and Aymaran: in the case of these two language groups as well, specialists divide up into protagonists (e.g. Campbell 1995) and skeptics (e.g. Adelaar and Muysken 2004) of a genetic relationship. Campbell (1995: 158) points out such situations are excellent test cases for the efficacy of historical/comparative tools for distinguishing material diffused by contact from material inherited from a common parent.

Proposals for a relationship between Japanese and Korean are sometimes dated to Arai Hakuseki's Korean etymologies for a number of lexical items in his *Tōga* (1717). Proposals for a thoroughgoing cognate relationship were made in the nineteenth century by Aston (1879) and Shiratori (1897), and again in the twentieth century by Kanazawa (1929). Among Japanese specialists at mid-century, Kōno (1949) points out important phonological correspondences between Korean and Japanese but does not commit himself to a genetic relationship. Ohno (1975) expresses a positive view toward a genetic relationship between Japanese and Korean, as does this scholar in earlier work, mostly based on research by others. Among Korean scholars, Lee (1972a) affirms the existence of a relationship while stressing its distance. The first research to propose a systematic set of sound correspondences and reconstructions for a sizeable amount of vocabulary was Martin (1966); Whitman (1985) expands upon this inventory.

The current state of debate remains comparable to the Quechumaran controversy, with Vovin (2010) rejecting the majority of Martin's and Whitman's proposed cognates. Vovin addresses the problem of distinguishing diffused from inherited material by requiring that comparanda on the Japanese-Ryukyuan side be reconstructible to pJR. If a cognate is found only in Western Old Japanese (WOJ, the central dialect of eighth-century Japanese), Vovin rejects it as a probable loan, based on his hypothesis that WOJ or its immediate predecessor absorbed a large number of loanwords from Old Korean (OK).<sup>1</sup> This is a sound procedure and I adopt it insofar as possible in this chapter. Vovin's conclusion is that the set of firm Japanese-Korean cognates is far smaller than suggested by Martin (1966) or Whitman (1985). He identifies six "reliable cognates", as against 75 "obvious loans" (2010: 239). The winnowing out of probable (or possible) loans is a great step forward in this field, but I will argue that an improved set of correspondences, involving not just phoneme-to-phoneme correspondences but consideration of original root shape, uncover a core set of cognates in the lexical and functional vocabulary.

The structure of the chapter is as follows. In section 2.2 I set out assumptions about the reconstruction of pJR, based on recent research in internal and comparative JR reconstruction. In section 2.3 I do the same for Korean. In section 2.4 I discuss a set of correspondences between pJR and pK, focusing on vowel correspondences, and introducing relevant lexical comparisons in the course of the discussion. This section also explores the comparative consequences of stem shape change, in particular syncope and medial consonant lenition in earlier Korean. Sections 2.5 and 2.6 discuss lexical comparisons in two specific functional domains, pronouns and grammatical formatives.

Transcription of Late Middle Korean is presented in a slightly different system to that in Sohn (this volume, Chapter 4). Following Yale Romanization (Martin 1995), I use *y*, *o*, and *wo* where Sohn uses *j*, *ó*, and *o* respectively, and mark tone as ´

(high/‘accented’, ◌ in fifteenth–sixteenth-century han’gŭl texts), no mark (low/‘unaccented’, unmarked in fifteenth–sixteenth-century han’gŭl texts), and ◌̣ (rising, ◌̣ in fifteenth–sixteenth-century han’gŭl texts).

## 2.2 PROTO-JAPANESE-RYUKYUAN

I assume a six-vowel system for pJR (Table 2.1). Reconstruction of the mid vowels \*e and \*o follows Hattori (1976, 1977–1979).<sup>2</sup> The reflexes of pJR \*e and \*o are quite restricted, particularly in WOJ, but in other varieties as well. The WOJ reflexes of \*e and \*o are /e/ and /o/ in word-final position, /i/ and /u/ elsewhere.<sup>3</sup> A process of mid vowel raising is posited to explain this distribution and correspondences such as WOJ *sugus-* ‘cause to pass by’: Eastern Old Japanese (EOJ) *sugwos-*, Early Middle Japanese (EMJ) *sugos-* id. (Hayata 1998; Hino 2003; Miyake 2003; Frellesvig and Whitman 2004, 2008b), and similar correspondences involving WOJ /i/, /u/ and proto-Ryukyuan (pR) \*e, \*o. Comparison with pR is our richest source of evidence for pJR \*e and \*o in nonfinal position. Tables 2.2 and 2.3 list the examples of \*e and \*o proposed by Thorpe (1983) and Pellard (forthcoming) based on the correspondences pR \*e : WOJ /i/ and pR \*o : WOJ /u/.

TABLE 2.1 PROTO-JAPANESE-RYUKYUAN VOWELS

*i		*u
*e	*ə	*o
	*a	

TABLE 2.2 PROTO-JAPANESE-RYUKYUAN \*e

Proto-Ryukyuan	Japanese (WOJ except as indicated)	Proto-Japanese-Ryukyuan
*ero	<i>iro</i>	*erə ‘color’
*ezu(ro)	<i>idu(re)</i>	*entu(rə/o) ‘which’
*kezu	<i>kizu</i>	*kensu ‘wound’
*memezu	<i>mimizu</i> (MJ)	*memensu ‘earthworm’
*mezu	<i>midu</i>	*mentu ‘water’
*nebu-	<i>nibu-</i> (MJ)	*nenpu- ‘dull, slow’
*pejesi- ‘cold’	<i>piye-</i> (MJ) ‘get cold’	*peje- ‘get cold’
*peru	<i>piru</i>	*peru ‘garlic’
*pezi	<i>pidi</i>	*penti ‘elbow’

TABLE 2.3 PROTO-JAPANESE-RYUKYUAN \*o

Proto-Ryukyuan	Japanese (WOJ except as indicated)	Proto-Japanese-Ryukyuan
*ki ~ ko-	<i>kwi ~ ku-</i> (MJ <i>ko-</i> )	*koj ‘yellow’
*kusori	<i>kusuri</i>	*kusori ‘medicine’
*mogi	<i>mugi</i>	*monki ‘wheat’
*moko	<i>muko, moko</i>	*moko ‘bridegroom’
*omi	<i>umi</i>	*omi ‘sea’
*ori	<i>uri</i>	*ori ‘melon’
*tuki ~ tuko-	<i>tukwi ~ tuku-</i>	*tukoj ‘moon’
*tukos-	<i>tukus-</i>	*tuko-s- ‘exhaust it’

TABLE 2.4 PROTO-JAPANESE-RYUKYUAN CONSONANTS

	Bilabial	Dental/ Alveolar	Palatal	Velar
Stop:	*p	*t		*k
Nasal:	*m	*n		
Fricative:		*s		
Tap:		*r		
Approximant:	*w		*j	

The pR data add nine examples of nonfinal \*e. In all these examples \*e occurs before a sonorant, as pointed out by Whitman (1985), who suggested that they are reflexes of an original nonfront vowel, \*ə in the six-vowel system for pJR. Miyake (2003) and Pellard (forthcoming) observe that we cannot claim that \*ə fronted automatically in this environment, since pJR contains many unfronted examples of \*CəC[sonorant]. However, the relevant environment can be narrowed to the position before a coronal sonorant, a phonetically plausible environment for vowel fronting (Clements 1991; Flemming 2003), and perhaps further narrowed to the position before a coronal sonorant in the same syllable. This narrower environment would leave unaccounted for the examples of \*e before \*r and \*j, but these examples may mask an earlier, more complex syllable structure ('earthworm', where \*e occurs before \*m, probably originates from a reduplication). I therefore reserve the possibility that at least some examples of nonfinal \*e have been fronted from earlier ə\*.

The eight tokens of nonfinal \*o show a different pattern. In 'exhaust it' \*o occurs in root-final position, while in 'yellow' and 'moon' it is in root-final position before a glide. 'Bridegroom' may involve root-final position in an original compound with \*ko 'child'/diminutive, while 'medicine' may be formed from \*kus- 'smelly' + -or (adnominal) + i (nominalizer). These reconstructions would reflect a variety where the environment blocking mid vowel raising was broader than WOJ: root-final rather than word-final position. We see a similar contrast between EOJ (root-final \*o retained) and WOJ (word-final \*o retained) in the contrast between EOJ *sugo-s-* and WOJ *sugu-s-* 'cause to pass by'. In either case final position is privileged at the relevant level of prosodic analysis. The relevant generalization is preserved in a prosodically strong position, such as the head of a metrical foot (Zec 2003), whose location is subject to crosslinguistic variation.

Proto-Japanese-Ryukyuan consonants are presented in Table 2.4. There is general consensus that so-called 'dakuon' consonants (*b, d, g, z*) in all varieties of Japanese and Ryukyuan, usually realized as voiced, sometimes as prenasalized, go back to nasal-obstruent clusters. Thus pJR has no laryngeal or manner contrast for consonants.

There is also general consensus about the lexical pitch accent or word tone classes for pJR, although there is controversy about their interpretation (Ramsey 1979; de Boer 2010). I cite the reconstructed tonal classes in Martin (1987), without hypothesizing a pitch interpretation.

Summarizing this section, we have reconstructed a phonemic inventory for pJR with six vowels and nine consonants, the latter lacking any laryngeal contrast such as voicing. This inventory is supported by a high degree of agreement among specialists on the reconstruction of Japanese and Ryukyuan.

### 2.3 PROTO-KOREAN

The closest approach to an internal reconstruction of Korean is Martin (1996), which shows partial convergence with work by Lee (1972a, b, 1991) and Lee and Ramsey

**TABLE 2.5 OLD KOREAN (SILLA) VOWEL SYSTEM**

*i > LMK <i>i</i> [i]	*i > LMK <i>u</i> [i]	*u > LMK <i>wu</i> [u]
*ε > LMK <i>e</i> [ə]	*ə > LMK <i>ó</i> [ʌ]	*o > LMK <i>o</i> [o] (or [o])
	*a > LMK <i>a</i> [a]	

(2011). Lee and Ramsey derive the seven-vowel system of Late Middle Korean (LMK) from an ‘Altaic’-type system with palatal harmony, but virtually all of the assumptions behind this have been called into question. More recent analyses posit a tongue root harmony system for LMK (Kim 1993; Ko 2010), similar to the retracted tongue root [RTR] systems found in Tungusic and most Mongolic languages. Based on the distribution of the LMK vowels in Sino-Korean, Itō (2007: 267) proposes the vowel system in Table 2.5 for Old Korean (OK, the language of Silla) at the period when Sino-Korean was established, commonly assumed to be in the late Tang period, roughly the eighth–ninth century.

LMK had vowel harmony (Sohn, this volume, section 4.2.2), but as Itō points out, the system in Table 2.5 is incompatible with a palatal vowel harmony system. But it is compatible with a tongue root harmony system, with [+/- RTR] pairs formed by \*i and \*ə, \*u and \*o, and \*ε and \*a. Martin (2000) argues that vowel harmony was innovated some time prior to LMK. The OK system in Table 2.5 suggests a scenario: areal influence led to the reinterpretation of \*i/\*ə, \*u/\*o, and \*ε/\*a as [+/- RTR] harmonic pairs, and to the centralization of \*ε. Centralization of \*ε triggered backing and lowering of \*ə to the low-mid back [ʌ] articulation of this vowel in LMK.

I adopt the system in Table 2.5 for OK, but represent the mid front vowel as \*e. The distribution of the two ‘weak’ vowels \*i > LMK *u* [i] and \*ə > LMK *o* [ʌ] is restricted in LMK: /o/ does not occur at all in onset position and /u/ is very rare there. A number of scholars have assumed that /u/ and /o/ undergo syncope in the second syllable of disyllabic verb stems of shape CVC<sub>o</sub>/u (Ramsey 1978; Martin 1996), and in the original second syllable of at least some LMK monosyllabic nouns with rising tone. A further environment for syncope is proposed as a source both for LMK initial obstruent clusters and aspirated consonants (Lee 1991; Lee and Ramsey 2011). I will assume that unaccented \*i and \*ə are syncope, subject to constraints on the acceptability of the resultant clusters. I view syncope of accented final \*i and \*ə as metathesis, with compensatory lengthening of the vowel in the preceding syllable. In accented onset position or following an initial glide, \*i and \*ə may have merged with \*e, as in LMK *yetŭlp* ‘8’ < \*jətŭrp.

The proto-Korean consonants are presented in Table 2.6. The reinforced and aspirated consonants of Modern Korean are secondary, resulting from clusters produced by syncope. Thus pK, like pJR, had no laryngeal contrast among consonants. The modern reinforced consonants reflect the clusters /pC/, /sC/, and /psC/ in LMK. Gaps in the inventory of

**TABLE 2.6 PROTO-KOREAN CONSONANTS**

	Bilabial	Dental/Alveolar	Palatal	Velar/Glottal
Stop:	*p	*t	*c	*k
Nasal:	*m	*n		*ŋ
Fricative:		*s		*h
Tap:		*r		
Approximant:			*j	

**TABLE 2.7 OBSTRUENT LENITION ATTESTED IN LATE MIDDLE KOREAN**

LMK	Source	Segmental environments	Morphological environments
/W/ [β]	/p/	V_V, y_V, l_V, z_V	morpheme boundary
/l/ [r]	/t/	V_V	morpheme boundary, nativized SK
/z/	/s/	V_V, y_V, *l_V, n_V, m_V, V_W, V_G	morpheme boundary and root internally
/G/ [ɦ]	/k/	l_V, z_V, i_V	morpheme boundary

clusters suggest that some /sC/ clusters may reflect earlier \*ti/əC or \*ci/əC, while expected but unattested \*/kC/ clusters may result in aspirates. The Early Middle Korean (EMK) sources for the LMK aspirates cited by Lee (1991) have the form EMK *hi/əCV* > LMK *ChV*, e.g. EMK *hiki-n* > LMK *khú-n* ‘big-ADNOM’.

As we have seen, pJR and pK have rather simple and very similar phonemic inventories, with six or seven vowels, no laryngeal (manner) distinction among consonants, and lexical pitch accent. The functional load of pitch accent is less in pK than in pJR, but this is consistent with the fact that pK allows consonant codas in non-bound roots (that is, nouns), while pJR does not.

The prehistory of Korean is characterized by major changes in root structure. Syncope was a major such factor; a second was medial obstruent lenition (Lee 1972a, b; Martin 1996; Lee and Ramsey 2011). The results of lenition are observable, first, in instances of LMK /W/ [β], /l/ [r], /z/, and /G/ [ɦ] resulting from earlier /p/, /t/, /s/, and /k/ (Table 2.7, based on Lee and Ramsey 2011: 136–53). Martin (1996) takes the view that all instances of the LMK voiced spirants /W/, /z/, and /G/ result from /p/, /s/, /k/, either allophonically at a morpheme boundary or diachronically within roots, before the weak vowels /u/ [i] and /o/ [ʌ].<sup>4</sup> Martin (1996: 54) claims that “the group of nouns with medial lenition expanded to include, at least sporadically, nouns with nonminimal vowels in the second syllable.”

But as Vovin (2010: 14) points out, the restriction of lenition to medial position before the least sonorous vowels is odd. Furthermore, when pre-LMK sources give evidence for medial stop lenition, the LMK outcome is not an LMK voiced spirant. This is particularly clear in the case of /p/, where the Chinese *Jílín lèishì* (鷄林類事 1103–1104, or *Kyey-lim yusa* in the Korean reading of its name) transcriptions for LMK *twülh*, *twuül* ‘two’ and *swul*, *swuul* ‘wine’ are best interpreted as *tupir* and *supir*. To this we must add the fact that intervocalic /p/, /t/, /k/ before non-weak vowels are rare root-internally in LMK. These considerations lead me to hypothesize that the LMK spirantizations were the tail end of a more general process, which began with the lenitions in Table 2.8. Each of these processes is supported by well-known developments, such as: (a) the lenition of EMK /p/ in ‘two’ and ‘wine’, (b) the lenition of the indicative assertive suffix *-ta* > *-la* [ra] after the copula and certain auxiliaries, (c) the lenition of initial /k/ > Ø after vowels in the postnominal particles *kwá* ‘with’, *kwós* ‘precisely’, and *kwóm* ‘each’.

**TABLE 2.8 HYPOTHESIZED LENITIONS FOR PRE-LATE MIDDLE KOREAN**

	Pre-LMK	LMK
a.	*VpV > VV[labial]	VV[labial]
b.	*VtV > VrV	VlV
c.	*VkV > V(V) (with reduction of hiatus)	V

## 2.4 PHONOLOGICAL CORRESPONDENCES

I posit the vowel correspondences in Table 2.9 between pK and pJR. Other than mid vowel raising in Japanese and Ryukyuan, I have omitted from Table 2.9 what I consider to be secondary developments in the two proto-languages, such as assimilation to another vowel in the same root, or, in the case of pK, vowel harmonic alternations. The new proposal among these correspondences is for pK \*e, for which pJR \*ə or its equivalent have been proposed in previous research. In onset position this is unproblematic: in comparisons like pK \*ep- ‘bear on back’ :: pJR \*əp- id. the pK initial vowel may reflect earlier \*ə or \*i. But comparisons such as pK \*kes ‘thing, matter’ :: pJK \*kətə 2.3 id. require reconsideration.<sup>5</sup> As noted above, vowel outcomes are further complicated by the co-occurrence restrictions known as Arisaka’s laws in Japanese, and root-internal effects of vowel harmony in Korean.

In contrast to the vowel correspondences, the pJR :: pK consonantal correspondences are relatively straightforward, as we might expect from the minimal inventories in Tables 2.4 and 2.6. However, as we see in Table 2.6, pK has two consonants, \*h and \*c, that are not reconstructable for pJR. Any comparison of pJR and pK must account for these consonants. In the first part of this section I focus on this problem. In the second part I focus on the more complex matter of correspondences reflecting on earlier Korean syncope and lenition.

Following a suggestion of Vovin (1993: 340–1), I propose that the pJR correspondence for pK \*h is \*s before \*i and \*j, \*k elsewhere (Table 2.10). The first two lexical comparisons involve a secondary vowel correspondence, pJR \*Ce > \*Ci :: pK \*Cje. This correspondence is supported by examples such as pJR \*sima ‘island’ :: pK \*sjəm id., and pJR \*me, \*mi- ‘woman’ :: pK \*mjənir ‘daughter-in-law’. In the pJR forms, the vowel surfaces as \*e in root-final position (1), \*i elsewhere (2) including the first syllable of ‘island’. Reconstruction of \*Cjə in pK is supported by the stem alternation shown by LMK ‘white’ and ‘whiten’. Recall that the diphthong /yo/ < \*jə is disallowed in LMK. In stem-final position the diphthong is eliminated by metathesis of the glide and

TABLE 2.9 PROTO-KOREAN :: PROTO-JAPANESE-RYUKYUAN VOWEL CORRESPONDENCES

pK		pJR	
*i	::	*i	
*i	::	*ə	
*u	::	*u	
*e	::	*e	(in nonfinal environments > *i through mid vowel raising)
*ə	::	*ə	
*o	::	*o	(in nonfinal environments > *u through mid vowel raising)
*a	::	*a	

TABLE 2.10 PROTO-JAPANESE-RYUKYUAN CORRESPONDENCES FOR PROTO-KOREAN \*h

pJR		pK	Reconstruction
1. *se- A ‘do’	::	*hjə-, LMK <i>hó(y)</i> - id.	*hjə-
2. *siro B ‘white’	::	*hjə-, LMK <i>hóy</i> id., <i>syey</i> - ‘whiten’	*hjə- (+ pJR *-ro ATTR) <sup>6</sup>
3. *kasa 2.2b ‘bulk’	::	*ha-, LMK <i>há</i> - ‘many, great’	*ha- (+ pJR *-sa NMR)
4. *kəsi 2.? ‘lower back’	::	*heli, LMK <i>hell</i> - id.	*həti <sup>7</sup>

TABLE 2.11 PROTO-JAPANESE-RYUKYUAN CORRESPONDENCES FOR PROTO-KOREAN \*c

pJR	pK	Reconstruction
5. *kunsu 2.5 ‘arrowroot’	:: *hicirk, LMK <i>chulk</i> id.	*hincu (+ pK -irk?)
6. *kusi 2.3 ‘skewer’	:: *koc, LMK <i>kwoc</i> id., <i>kwos-</i> ‘insert’	*koc- (+ pJK -i NMR)
7. *puta- ?2.1 ‘two’	:: *pəcak, LMK <i>pca</i> ‘a pair’	*pəca (+ pK -k?) <sup>8</sup>
8. *mi(t)- ?1.1 ‘three’	:: *mjech, LMK <i>myé</i> ch ‘a few’	*mjec (+ pK -h/k?)

nuclear vowel. In other positions the diphthong is retained, but the vowel fronted to /e/, as in LMK *yetulp* ‘8’ < \*jətǝrp, discussed in section 3. Another possible outcome of pK \*Cjə may be shown by LMK *siki-* ‘cause to do’, *sikpu-* ‘want’, which as shown by Lee (1991) seem to be related to a root \*sik- ‘want’.

Further support for reconstructing an initial consonant distinct from \*s- in pJR ‘do’ may be supported by a paradigmatic alternation in WOJ and EOJ adjectives. As many linguists have observed, the conclusive and adnominal (attributive) forms in this paradigm appear to be a WOJ/EOJ innovation, as they are not found in Ryukyuan (or indeed most Kyūshū varieties). One idea about the source of this part of the paradigm is that it results from combining pJR \*se- ‘do’ as a ‘light’ verb with the originally uninflecting adjectival stem. The WOJ/EOJ conclusive suffix *-si* may directly reflect continuative \*s-i at a period prior to the development of the OJ conclusive *s-u*; alternatively, the ‘light’ verb ‘do’ may have grouped with *ar-i* ‘to exist’ in having a conclusive in *-i*. How then to relate the adnominal suffix, which we know to have been \*-ke on the basis of its EOJ form *-ke*, raised to *-ki* in WOJ by mid vowel raising? On the hypothesis that the adnominal involves the same ‘light’ verb ‘do’, we must posit a consonant initial that surfaces as OJ /s/ before /i/ but /k/ elsewhere. This is exactly the alternation predicted by the conditioned correspondences in Table 2.10. The suffix vowel \*-e in the adjectival adnominal suffix may reflect the original stem vowel of ‘do’, or it may involve a distinct suffix no longer recoverable. Whichever is the case, the alternation provides pJR-internal evidence for an initial distinct from \*s for pJR; \*h- is a phonetically plausible candidate.

Regarding pK \*c, Table 2.11 presents evidence that it corresponded to pJR \*s before high vowels and \*t elsewhere. The outcome of the pJR vowel in (7) ‘two’ is supported by several other instances of pK \*ə :: pJR \*u adjacent to a labial consonant:

9. pJR \*pej in EMJ *fe+saki* ?3.1 ‘prow’ (‘boat+tip’) :: pK \*pəj, LMK *póy* ‘boat’ < pJK \*pəj  
but \*pu- in pJR \*pu+naj OJ *pune* ‘boat’ (‘boat+root’)<sup>9</sup>
10. pJR \*pe(j) in EMJ *feso* 2.1 ‘navel’ :: pK \*pəj, LMK *póy* ‘stomach’ < pJK \*pəj  
but also EMJ *foso* < \*pəj + ?sə/o
11. pJR \*mej in OJ *me* ‘seaweed’, but also OJ *mo* ‘seaweed’ :: pK \*mər, LMK *mól* id. < pJK \*mər

The otherwise unusual alternations of /e/ and /o/ can be explained by the relative timing of labial assimilation of \*ə and monophthongization of \*əj. Where \*j is lost first, \*ə rounds to \*o after a labial; when this occurs nonfinally (in particular before a morpheme boundary that has become opaque, as in ‘boat’), mid vowel raising applies, \*o > u. Where diphthongization applies first, the result is /e/.

Next we proceed to comparisons involving root structure change in Korean. We saw two examples involving syncope in (5) ‘arrowroot’ and (7) ‘two, a pair’. Table 2.12 presents a few more.

TABLE 2.12 SYNCOPATED PROTO-KOREAN \*i, \*ə

pJR		pK	Reconstruction
12.	*sik- A ‘spread it’	:: *səkər-, LMK <i>skór-</i> id.	*sjək- (+ pK *-ər- CONT)
13.	*sup- A ‘suck, inhale’	:: *səpər-, LMK <i>spór-</i> id.	*səp- (+ pK *-ər- CONT)
14.	*pə(n)tə 2.2a ‘interval’	:: *pət-aj, LMK <i>pstáy-</i> ‘time’	*pə(n)tə- (+ pK *-aj LOC)
15.	*əmə- A ‘heavy’	:: *mi-kep-, LMK <i>mukép-</i> id.	*imi- (+ pK *-ka/ep- ADJ)
16.	*u/imə 2.3 ‘yam’	:: *mah, LMK <i>mah</i> id.	*(j)əmah <sup>10</sup>

TABLE 2.13 PROTO-KOREAN MEDIAL CONSONANT LENITION

pJR		pK	Reconstruction
17.	*takaj 2.1 ‘bamboo’	:: *taj, LMK <i>táy</i> id.	*takaj
18.	*taka-j 2.3 ‘height’	:: *tarák, LMK <i>talak</i> ‘loft’	*takar (+ pK -Vk LOC)
19.	*tuku/oj 2.3 ‘moon’	:: *tər, LMK <i>tól</i> id.	*tokər
20.	*naka 2.4 ‘inside’	:: *an(-)h 1.1 LMK <i>ánh</i> id.	*nakəh
21.	*ka(:)nkaj 2.5 ‘shadow’	:: *kənərth, LMK <i>kónolh</i> id.	*kankərth
22.	*nanka- B ‘long’, *nanka-r- B ‘flows’	:: *nái, LMK <i>náy</i> ‘throughout, during’	
23.	*sanki 2.1 ‘heron’, suffix in bird names	:: *sái, LMK <i>sáy</i> ‘bird’	*sanji
24.	*pitə- ?2.3, ?2.4 ‘one’	:: *piris, LMK <i>pilús</i> ‘first’	*piti (+ pK -s NMR)
25.	*kata ?2.3 ‘one of pair’	:: *hət(V)-, EMK <i>hət-ən</i> ‘one’, cf. LMK <i>holo</i> ‘one day’	*hatə
26.	*ita-r <sub>INTR</sub> /s <sub>TR</sub> - A ‘attain’	:: *iri/ə-, LMK <i>í(l)-/ilú/ó-</i> ‘arise’	*iti/ə- (+ pJK *-ar- INTRANS)
27.	*pinti 2.2b ‘elbow’	:: *pərth ‘arm’, LMK <i>pólh</i>	*pintəh
28.	*ap- B ‘meet, fit’	:: *e/api/ər-, LMK <i>ewúl-</i> ‘meet’, <i>awól-</i> ‘join it’	*ap- (+pK *-ər- CONT) <sup>11</sup>
29.	*əpə- B ‘big’	:: *ipij-, LMK <i>ewúy-</i> ‘broad, big’	*ipi- (+pK *-i- INTRANS)

Next consider comparisons based on medial consonant lenition in pK (Table 2.13). (17–19) exemplify a correspondence where pJK apophonic nouns correspond to pK nouns with a sonorant final. Further examples are:

30. pJR \*pəj 1.2 ‘fire’, OJ *po-* ~ *pwi* :: pK \*pir, LMK *púl* id. < pJK \*pir<sup>12</sup>  
 31. pJR \*muj 1.1 ‘body’, OJ *mu-* ~ *mwi* :: pK \*mom, LMK *móm* id. < pJK \*mom

Vovin (2011) argues that earlier Japanese ‘fire’ should be reconstructed as \*poj rather than \*pəj, based on an attestation of the compound form of this noun as 本 in the *Kojiki* songs. Vovin follows Mabuchi (1957, 1968) in interpreting this phonogram as <pw> in the *Kojiki* songs (and, according to Mabuchi, sound glosses). However, Mabuchi (1957: 86) explicitly rejects this example as a mistranscription. The philological rationale for interpreting the OJ compound form for ‘fire’ as *po-* < \*pə- is explained by Wenck (1954: 268–9).

Vovin (2010: 194) considers (31) to be a loan from Korean to WOJ “because cognates of WOJ *mī* ‘body’ are not found in Ryukyuan.” But reflexes of \*muj in its derived meaning ‘self, person’ occur throughout Ryukyuan (e.g. Nakijin *a-ga-mi* ‘we’, Nakasone 1983; cf. WOJ *a ga mwi* ‘myself’). Since the direction of grammaticalization is clearly ‘body’ > ‘person’ > ‘self’, Vovin’s hypothesis would require that ‘body’ was borrowed into WOJ, then the grammaticalized form was borrowed into pR.

TABLE 2.14 PROTO-JAPANESE-KOREAN NUMERALS

Gloss	pJR		pK
'1'	*pitə	::	*piris 'first'
	*kata 'one of pair'	::	*hət(V)- 'one'
'2'	*puta	::	*pcak < *pəcak 'double'
			*tupir 'two'
'3'	*mit	?::	*mjech 'how many, a few'
			*se- 'three'
'4'	*jə	?::	*ne
'5'	*itu		*tasə
'6'	*mu(t)		*jəsəs
'7'	*nana		*nilko/up
'8'	*ja	?::	*jətərp ?< *jə+tərp '4 × 2'
'9'	*kəkəna		*ahop
'10'	*təwə		*jer

In this section we have seen comparisons involving basic vocabulary, including body parts (4, 10, 27, 31), numerals (7, 8, 24–25), and basic verbs and adjectives (1–3, 15, 18, 22, 28–9). It may be appropriate to conclude the section with a comparison of the pJR and pK numeral inventories (Table 2.14), taking into account the four comparisons made here. It has long been observed (e.g. Ellis 1873: 50) that the Japanese numerals '2', '6', and '8' are the product of a doubling game, exploiting vowel alternations with the base numerals '1', '3', and '4', where \*i alternates with \*u and \*ə with \*a in the doubles. The doubles are thus less likely to show cognates in any language that does not employ a similar strategy; but the base numerals '1', '3', and '4' all show possible pK cognates, although '3' is semantically and '4' phonologically weaker. pK \*pcak < \*pəcak 'double, one of a pair' and 'eight' suggest that pK may have employed the strategy as well. pK \*jətərp 'eight' is analyzable as '4 × 2'; the alternative analysis 'two (from) ten' would not predict the first syllable vowel.

## 2.5 PRONOUNS

I present in Table 2.15 candidate cognates from the pJR and pK pronominal systems.

TABLE 2.15 PROTO-JAPANESE-KOREAN PRONOUNS

	pJR		pK	Reconstruction
32.	*a, are 2.4 1P (exclusive?)	::	*a- in kin terms <sup>13</sup>	*a
33.	*wa 1.3a, ware 2.4 1P (inclusive?)	::	*uri, LMK <i>wūli</i> 1P plural, ?*ij LMK <i>ūy</i> 1P (矣 in EMK idu texts)	?*wa
34.	*na, nare ?2.4/5 2P, pR reflexive	::	*ne, LMK <i>ne</i> 2P	*na <sup>14</sup>
35.	*kə 1.1, *kəre 2.1 'this' (proximal)	::	*ki, LMK <i>ku</i> 'that'	*ki <sup>15</sup>
36.	*sə 1.1, *səre 2.1 'that' (mesial)	::	*sə, LMK <i>so</i> nominal complementizer 'that'	*sə
37.	*e-, OJ <i>i-</i> 'which' <sup>16</sup>	::	*e-, LMK <i>e-</i> id.	*e-
38.	*məsi 2.2b 'perchance', OJ adverb introducing polar interrogatives	::	*misi(k), LMK <i>musú(k)</i> 'what'	*misi(k)

1P: first person; 2P: second person.

## 2.6 GRAMMATICAL FORMATIVES

Table 2.16 introduces verb affixes, Table 2.17 postnominal particles. The suffixes in (39–42) are noteworthy because they include all of the verbal inflectional root suffixes reconstructable for pJR. Of these *-ku* is primarily an adjectival suffix, but it attaches to the stative verb \*ar- ‘exist’ to derive adverbial \*aku in OJ *kaku* ‘thusly’ < \*kə ‘this’ + aku ‘being’. The one OJ inflectional root suffix not included in this group, conclusive *-u*, is not clearly reconstructable for pJR, as reflexes of the conclusive category in Ryukyuan show up primarily in the form of reflexes of the distinctive conclusive forms of existential \*ar-i and \*wor-i ‘exist-CONC’.

## 2.7 CONCLUSION

This chapter has presented a brief argument for a genetic affiliation between Korean and Japanese. Internal and language group-comparative reconstruction leads to the reconstruction of quite similar phonemic inventories. Earlier research has uncovered a number of potential cognates, which I have supplemented here. The main challenge for the comparativist is to account for changes in root structure, especially in Korean, where our historical and comparative information is relatively shallow.

**TABLE 2.16** PROTO-JAPANESE-KOREAN VERBAL AFFIXES AND POSTVERBAL PARTICLES

pJR	pK	Reconstruction
39. *-i infinitive/adverbial	:: *-i, LMK <i>-i</i> adverbial	*-i <sup>17</sup>
40. *-ro clausal nominalizer	:: *-i/ər, LMK <i>-ú/ól(?)</i> clausal nominalizer. LMK <i>-ú/ól</i> forms object nominalizations.	*-or
41. *-ku gerund	:: *-ko/-ku, LMK <i>-kwó/ú</i> gerund	*-ku/o <sup>18</sup>
42. *-a infinitive in irrealis conditional	:: *-é/á	*-a
43. *-Vs- (OJ) honorific	:: *-i/əsi-, LMK <i>u/osi</i> id.	*-as- (+ pK <i>-i</i> ADV)
44. *-nV- perfective	:: *-nə-, LMK <i>no-</i> processive	*-na-
45. *-i- active prefix	:: *-i, LMK <i>-i</i> nominative < ergative postnominal particle	*-i
46. *tə ‘that’ complementizer	:: *tə, LMK <i>tó</i> ‘that’ complementizer (follows nominalized clause)	*tə
47. *ka interrogative complementizer (with nominalized clause)	:: *ka, LMK <i>ká</i> interrogative complementizer (follows nominalized clause)	*ka

**TABLE 2.17** PROTO-JAPANESE-KOREAN POSTNOMINAL PARTICLES

pJR	pK	Reconstruction
48. *-nə genitive	:: *-i/ən, LMK <i>-ú/ón</i> noun modifier	*-ən
49. *-tə comitative, adverbial	:: *tə/is, LMK <i>tó/ús</i> adverbial, attaches to verb/adjective stems, nominalizations	*tə (?+ pK <i>-s</i> ADV)
50. *-pa topic	:: *pa, LMK <i>pá</i> bound noun ‘way, thing that’ (attaches to nominalizations)	*pa

## NOTES

- 1 In this chapter I use Vovin's terms WOJ and EOJ to refer to specific properties of these varieties; OJ is used to refer to properties shared by both.
- 2 Hattori himself proposed reconstructing an additional high vowel, but Whitman (1985) shows the correspondences Hattori adduces are otherwise explicable. Similarly, but for a different set of cases, Frellesvig and Whitman (2004, 2008a, 2008b) argue for reconstruction of a high central \*i on the basis of alterations of *o* with *e* rather than the expected *wi* in WOJ. However, the number of relevant alternations in WOJ is small, and Frellesvig and Whitman's argument is based in part on external evidence involving both loans from Korean and possibly inherited cognates. To avoid circularity, in this chapter I adopt the standard six-vowel hypothesis.
- 3 WOJ does attest nonfinal /e/ in examples such as *simyes-* 'indicate', *sakyeb-* 'shout', *kapyer-* 'return', *kyepu* 'today', *pyeta* 'edge, shore', *pyera* 'moldboard', as pointed out by Pellard (forthcoming), who also observes correctly that such examples are problematic for an account that predicts word-nonfinal mid vowel raising in WOJ. The first three examples may involve original root-final \*e. 'Moldboard' probably stands in a loan relationship with LMK *pyet*. 'Today' may involve contraction of the \*ki- found in *kinopu* 'yesterday' and *kiso* 'last night', although there is no clear source for a second element in \*ki+apu. Frellesvig and Whitman (2004, 2008b) reconstruct the proximal demonstrative *ko-* as \*ki- and claim that its function was originally mesial. If this is correct, 'today', 'yesterday', and 'last night' may involve \*ki- in the first syllable, with fronting of \*i before a coronal sonorant in the last two forms.
- 4 The role of accent in Martin's hypothesis is unclear. According to Martin, lenition occurs before accented /ó/, /ú/ in verbs of shape /(C)VCó/ú, but it also occurs before unaccented /u/, /o/ in nouns.
- 5 Vovin (2010: 149) rejects this comparison because he claims that the Ryukyuan reflex of \*kətə is not used in the bound noun or nominalizing function 'the one, the fact that' of WOJ *koto* and LMK *kes*. But surely this latter, grammaticalized function is derived from the basic meaning 'thing', which Ryukyuan retains. At the same time, Vovin is right to point out that the potential consonantal correspondence pK \*s# :: pJR \*tV needs further clarification (Vovin suggests that it is a loan correspondence). A better comparison involving what I believe to be the same root is LMK *kotho-* 'be the same' < pK \*kot+ho(j)- 'do' :: pJR \*nkətə + adjectival formants id., which is abundantly attested in Ryukyuan (cf. Nakijin *gutuu* < \*nkətə-ku 'be the same-GER', Nakasone 1983).
- 6 Reconstruction of pJR \*-ro is based on the hypothesis that the verbal adnominal suffix \*-ro (or \*-or) attached to some roots that have survived as adjectival stems. Examples of such \*CV roots include *siro* < \*si-ro 'white' and *kuro* < \*ku-ro 'black'; examples of \*CVC roots include WOJ *awo* 'blue/green' < \*aw-ro and *kuso* 'shit' < \*kus-ro (cf. *kusa-* 'smelly').
- 7 As noted by Vovin (1993: 340), this is a widely cited comparison. The reconstruction posited here assumes lenition of \*t to /r/ in Korean, and palatalization of \*t before \*i in Japanese (Whitman 1985).
- 8 Vovin (pers. comm.) points out that LMK *pcak* and Modern Korean *ccak* can have the meaning 'one of a pair', as indicated by the common LMK character gloss 隻 (Chinese *zhī*) 'one of a pair' and compounds like *pcak nwun*, Modern Korean *ccak nwun* 'mismatched eye'. But the 'pair, double' meaning is clear in expressions like *pcak machwo-* 'match as a pair' < *pcak* + 'match, fit together', and Modern Korean *ccak swu* 'even number' < *ccak* + 'number'. The basic meaning is 'matched pair'; synecdoche gives 'pair' > 'double' > 'a double' > 'one of a pair'.
- 9 The idea that pJR \*punaj is a compound with \*naj 'root' is due to Osada (1982). Similar compounds with *-ne* in OJ are *ki* ~ *kine* 'pestle' and *kaki* ~ *kakine* 'fence'.

- We know the pJR form is \*pəj because it is attested as OJ *pe* ‘prow’. Vovin (2010) suggests that OJ *pe* is a borrowing, but this would require that *pe* was borrowed prior to OJ, spread to non-central varieties in the compound form, and then semantically narrowed in OJ.
- 10 This comparison assumes progressive assimilation in pJR prior to labialization of the initial vowel.
  - 11 Vovin (2010: 229) rejects this comparison because the LMK [+RTR] variant *awól-* is transitive ‘join it’. But the [-RTR] variant *ewúl-* ‘meet, join together’ preserves the intransitive meaning. Vovin also objects that the function of \*-ər- (which I have glossed as CONT(inuative)) is unclear. But as pointed out in Whitman (1985), \*-ər- must be a suffix, because this verb has an OK attestation cited in a Silla toponym in *Samguk sagi* 34, 阿火 <apər>. In this attestation -ər- is the adnominal suffix, so the form must be segmented *ap-ər*. The OK form shows clearly that the original stem-final consonant was /p/, and that -ul- in the LMK form must have accreted between OK and LMK. The suffix, whatever its function, created the environment for lenition of /p/.
  - 12 I am grateful to Sven Osterkamp for bringing Wenck’s discussion to my attention.
  - 13 E.g. LMK *āki* ‘baby’, *atól* ‘son’ (cf. *stól* ‘daughter’), *apí* ‘father’ (cf. OJ *pi+kwo* ‘male’ <pi + ‘child’), *azo* ‘younger brother’ (cf. OJ *se* ‘brother’), *azóm* ‘kin, relatives’, *acapí* ‘uncle’, *acómi* ‘aunt’ (cf. *émi* ‘mother’), *ahú/óy* ‘child’.
  - 14 2P *ne* is in vowel harmonic opposition to 1P *na*.
  - 15 Frellesvig and Whitman (2004, 2008b) argue for a deictic shift in the pJR demonstrative paradigm, where an original proximal demonstrative \*i (cf. *ima* ‘now’ 2.4/5 <\*i-ma ‘this interval’) is replaced by \*kə.
  - 16 The pJR reconstruction follows Thorpe (1983).
  - 17 (37) and (38) are probably the same morpheme, but they show accentual differences in Japanese. (37) occurs mainly after adjectives in LMK, but adverbializes verbs in OK and in EMK *kugyōl* texts. (38) occurs in e.g. LMK *khúy* ‘height’ <*khú-* ‘big’ + *i*. See Martin (1992: 553–5).
  - 18 The OJ form is limited to adjectival gerunds, except the OJ form *aku* <\*ar-ku ‘be-GER’ visible in e.g. *kaku* 2.2b ‘thus’ <\*kə+ar-ku ‘this be-GER’.

## REFERENCES

- Adelaar, Willem F. H. and Muysken, Pieter (2004) *The Languages of the Andes*, Cambridge: Cambridge University Press.
- Aston, William G. (1879) ‘A comparative study of the Japanese and Korean languages’, *Journal of the Royal Asiatic Society of Great Britain and Ireland* (New Series), 11: 317–64.
- de Boer, Elisabeth M. (2010) *The Historical Development of Japanese Tone: From Proto-Japanese to the Modern Dialects. The Introduction and Adaptation of the Middle Chinese Tones in Japan*, Veröffentlichungen des Ostasien-Instituts der Ruhr-Universität Bochum 59, Wiesbaden: Harrassowitz Verlag.
- Campbell, Lyle (1995) ‘The Quechumaran hypothesis and lessons for distant genetic hypothesis’, *Diachronica*, 12(2): 157–200.
- Clements, George N. (1991) ‘Place of articulation in consonants and vowels: a unified theory’, *Working Papers of the Cornell Phonetics Laboratory*, 5: 77–123.
- Ellis, Robert (1873) *On Numerals as Signs of Primeval Unity among Mankind*, London: B. D. Tribner.
- Flemming, Edward (2003) ‘The relationship between coronal place and vowel backness’, *Phonology*, 20: 335–73.
- Frellesvig, Bjarke and Whitman, John (2004) ‘The vowels of proto-Japanese’, *Japanese Language and Literature*, 38: 281–99.

- Frellesvig, Bjarke and Whitman, John (2008a) 'The Japanese-Korean vowel correspondences', in Mutsuko Endo Hudson, Sun-Ah Jun, Peter Sells, Patricia M. Clancy, Shoichi Iwasaki, and Sung-Ock Sohn (eds) *Japanese/Korean Linguistics* 13. Stanford: Center for the Study of Language and Information, 15–29.
- (2008b) 'Evidence for seven vowels in proto-Japanese', in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese*, Amsterdam: John Benjamins, 15–41.
- Hattori, Shirō (1976) 'Jōdai nihongo no boin onso wa muttsu de atte yattsu de wa nai' [The vowel phonemes of Old Japanese are six, not eight], *Gekkan Gengo*, 5(12): 69–79.
- (1977–1979) 'Nihon sogo ni tsuite' [Regarding proto-Japonic] 1–22, *Gekkan Gengo*, 7(1)–7(3), 7(6)–8(12).
- Hayata, Teruhiro (1998) 'Jōdai nihongo no onsetsu kōzō to o-retsū kō-otsu no betsu' [The syllable structure of Old Japanese and the *o/wo* distinction], *Onsei Kenkyū*, 2: 25–33.
- Hino, Sukenari (2003) 'Nihon sogo no boin taikai – jōdai tōgoku hōgen shiryō ni yoru saikō' [A reconstruction of the proto-Japanese vowel system on the basis of the data from Eastern Old Japanese], in Alexander Vovin and Toshiki Osada (eds) *Nihongo keitōron no genzai* [Perspectives in the origins of the Japanese language], Nichibunken Japanese Studies Series 31, Kyoto: International Research Center for Japanese Studies, 187–205.
- Itō, Chiyuki (2007) *Chōsen Kanjion Kenkyū* [A study of Sino-Korean], Tokyo: Kyūko Shoin.
- Kanazawa, Shōzaburō (1929) *Nissen Dōsoron* [The Japanese–Korean common origin hypothesis], Tokyo: Tōkō shoin.
- Kim, Chuwōn (1993) *Moum Cohwa-uy Yenkwu* [A study of vowel harmony], Kyōngsan: Yōngnam University Press.
- Ko, Seongyeon (2010) 'A contrastivist view on the evolution of the Korean vowel system', in Hiroki Maezawa and Azusa Yokogoshi (eds) *MIT Working Papers in Linguistics (MITWPL) 61: Proceedings of the Sixth Workshop on Altaic Formal Linguistics (WAFSL 6)*, Cambridge, MA: MIT Department of Linguistics & Philosophy.
- Kōno, Rokurō (1949) 'Nihongo to chōsengo no ni san no ruiji' [Two or three resemblances between Japanese and Korean], in Hachi gakkai rengō (ed.) *Jinbun Kagaku no Shomondai* [Issues in social sciences], Tokyo: Seki shoin, 74–80.
- Lee, Ki-Moon (1972a) *Kwuke-sa Kaysel* [Introduction to the history of the Korean language], Seoul: Thap Chwulphansa.
- (1972b) *Kwuke Umwun-sa Yenkwu* [A study of Korean phonological history], Seoul: Thap Chwulphansa.
- (1991) *Kwuke Ehwi-sa Yenkwu* [A study of Korean lexical history] Seoul: Thap Chwulphansa.
- Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge: Cambridge University Press.
- Lewis, M. Paul (ed.) (2009) *Ethnologue: Languages of the World*, 16th edition, Dallas: SIL International.
- Mabuchi, Kazuo (1957) 'Kojiki no si, o, po no kana' [The *si, o, po* kana in *Kojiki*], *Kokugogaku*, 31: 61–90.
- (1968) *Jōdai no Kotoba* [Old Japanese], Tokyo: Shibundō.
- Martin, Samuel E. (1966) 'Lexical evidence relating Korean to Japanese', *Language*, 42(2): 185–251.
- (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.
- (1992) *A Reference Grammar of Korean*, Burlington, VT: Tuttle.
- (1996) *Consonant Lenition in Korean and the Macro-Altaic Question*, Honolulu: University of Hawai'i Press.
- (2000) 'How have Korean vowels changed through time?', *Korean Linguistics*, 10: 1–59.

- Miyake, Marc Hideo (2003) 'Philological evidence for \*e and \*o in Pre-Old Japanese', *Diachronica*, 20(1): 83–137.
- Nakasone, Seizen (1983) *Okinawa Nakijin Hōgen Jiten* [A dictionary of the Nakijin dialect of Okinawa], Tokyo: Kadokawa Shoten.
- Ohno, Susumu (ed.) (1975) *Nihongo to Chōsengo* [Japanese and Korean], Tokyo: Mainichi Shinbunsha.
- Osada, Natsuki (1982) *Genshi Nihongo Kenkyū – Nihongo Keitōron e no Kokoromi* [Studies on proto-Japanese – an essay on the genetic affiliation of Japanese], Kobe: Kobe Gakujutsu Shuppan.
- Pellard, Thomas (forthcoming) 'Ryukyuan perspectives on the proto-Japonic vowel system', in Bjarke Frellesvig and Peter Sells (eds) *Japanese/Korean Linguistics 20*, Stanford: Cambridge University Press/Center for the Study of Language and Information.
- Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects: A Descriptive and Historical Study*, Seoul: Thap Chwulphansa.
- (1979) 'The Old Kyoto dialect and the historical development of Japanese accent', *Harvard Journal of Asiatic Studies*, 39(1): 157–75.
- Shinmura, Izuru (1916) 'Kokugo oyobi chōsengo no sūshi ni tsuite' [Regarding numerals in Japanese and Korean], *Geibun*, 7(2): 4.
- Shiratori, Kurakichi (1897) 'Nihon no kogo to chōsengo to no hikaku' [Comparison of early Japanese with Korean], *Kokugakuin Zasshi*, 4: 4–12.
- Thorpe, Maner Lawton (1983) 'Ryūkyūan Language History', unpublished thesis, University of Southern California.
- Vovin, Alexander (1993) 'Notes on some Japanese-Korean phonetic correspondences', in Soonja Choi (ed.) *Japanese/Korean Linguistics 3*, Stanford: Center for the Study of Language and Information, 338–50.
- (2010) *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*, Honolulu: University of Hawai'i Press.
- (2011) 'On one more source of Japanese *i*<sub>2</sub>', *Journal of East Asian Linguistics*, 20(3): 219–28.
- Wenck, Günther (1954) *Japanische Phonetik 2: Die Phonetik der Manyo gana*, Wiesbaden: Otto Harrassowitz.
- Whitman, John (1985) 'The phonological basis for the comparison of Japanese and Korean', unpublished thesis, Harvard University.
- Zec, Draga (2003) 'Prosodic weight', in Caroline Féry and Ruben van de Vijver (eds) *The Syllable in Optimality Theory*, Cambridge: Cambridge University Press, 123–43.

PART II

**KOREAN**



# OLD KOREAN

*Nam Pung-hyun* (南豐鉉)

## 3.1 PERIODIZATION

We divide Old Korean (OK) into Early, Mid and Late Old Korean (EOK, MOK, LOK). EOK was the Korean of the Three Kingdoms period, roughly from the start of the fifth century until Silla unified the Three Kingdoms in the 660s. MOK was the Korean of the Unified Silla [Sinla] period, from the 660s until the 930s when Koryō [Kolye] re-unified the country. LOK was the language of the earlier part of the Koryō dynasty from the 930s till the mid-thirteenth century.

The reason why I treat the language of the earlier part of the Koryō period as Late *Old* Korean is because the grammar of ‘interpretive’ kugyōl, presumed to be data from the tenth to the mid-thirteenth century, and that of ‘consecutive’ kugyōl, which constitutes the Early *Middle* Korean data of the later thirteenth century, show quite significant differences. The social factor that separates the OK and MK worlds is taken to be the Mongol invasion and occupation. This resulted in large population movement, and the nomadic culture that subsequently entered Korea greatly changed the traditional culture that had continued since ancient times, heavily affecting language and culture.

The EOK period followed the establishment of the Three Kingdoms: Silla [Sinla] 新羅, Paekche [Paykcey] 百濟 and Koguryō [Kokwulye] 高句麗. As the Three Kingdoms were independent kingdoms, some believe that each must have had its own independent language, but the territory unified by Silla was roughly the territory of the early city-states (the ‘Samhan’ period, from first century BC), and the texts left by the Three Kingdoms are not so diverse that we should regard the Three Kingdoms’ varieties as separate languages. Moreover, it is preferable to treat what has been left from this period as texts in the Korean language, and the characteristics shown in them suggest dialectal differences of a single language.

## 3.2 DATA

It is a universal fact of any language that the further back we trace it into the past the scarcer the written materials are, and OK texts are in addition very crude considering that they extend over 900 years or more. I shall divide them into internal and external materials.

Internal materials are texts which are documented by borrowing Chinese characters. We call this writing ch’aja [chaca] 借字 or ‘loan character’ writing, and texts written with ch’aja writing have been divided previously into idu [itwu] 吏讀, hyangch’al [hyangchal] 鄉札 and kugyōl [kwukyel] 口訣. We can include vocabulary lists as a fourth category.

I shall start briefly with vocabulary list materials. These consist of the recording by means of Chinese characters of native Korean words. They record proper nouns

such as toponyms and names of people, and the names of official ranks or roles and names of things. There are many written inside Chinese texts (epigraphs) and idu or hyangch'al texts, and there are also collations of vocabulary such as the geographical chapters (地理志) of *Samguk sagi* [*Samkwuk saki*] 三國史記 (mid-twelfth century). These are difficult to decipher and different researchers suggest different views, but in the case of EOK for which texts are scarce we have to rely on them to obtain crucial information. The folk vocabulary recorded in the *Kyerim yusa* [*Kyeylim yusa*] 鷄林類事 (c. 1103) and the mid-thirteenth century *Hyangyak kugŭppang* [*Hyangyak kwukuppang*] 鄉藥救急方 are also lexical list materials. There are records of EOK/MOK vocabulary in the *Samguk yusa* [*Samkwuk yusa*] 三國遺事 (late thirteenth century) in addition to its *hyangga* poems, and we can also use the *Koryŏsa* [*Kolyesa*] 高麗史 (mid-fifteenth century) for snapshots of LOK vocabulary.

### 3.2.1 Idu data

'Idu' in the broad sense sometimes refers to all the ch'aja texts, but in practical terms, once we separate off hyangch'al and kugyŏl, it refers to prose in ch'aja. Most kugyŏl texts, in terms of what is extant, consist of administrative texts.

Extant EOK idu texts are in the form of texts on metal or stone, known as 'pseudo-Chinese' (*pyŏnch'e hanmun* [*pyenchey hanmun*] 變體漢文) or 'early idu' texts. Koguryŏ idu texts include the King Kwanggaet'o [Kwangkaytho] stela (廣開土大王碑銘, 414), the Sŏbong-ch'ong [Sepong-chong] silver bowl inscription (瑞鳳冢銀盒杆銘, 451?), the Chungwŏn Koguryŏ [Cwungwen Kokwurye] stela (中原高句麗碑銘 590s?) and the four P'yŏngyang Koguryŏ castle wall engravings (平壤高句麗城壁刻字). The Kwanggaet'o stela is typical Koguryŏ Chinese, but its word order and use of declarative之 display features of early idu. The Chungwŏn inscription is in word order and its use of idu a typical Koguryŏ idu text.

Paekche idu texts consist of the silver bracelet inscription (銀釧銘, c. 520) and the so-called *sukse-ga* [*swuksey-ka*] 宿世歌 ('songs to past lives') written on wooden slats excavated at the temple Puyŏrŭng-sa [Puyelung-sa]. Their word order and expressions are Korean.

The extant idu texts of Silla are the most numerous of the Three Kingdoms. Abbreviated names are given below, with a full name in Chinese characters in parentheses.

- EI-i. Naengsuri [Nayngswuli] inscription (迎日冷水里碑), 503?
- EI-ii. Pongp'yŏng [Pongphyeng] inscription (蔚珍鳳坪新羅碑), 524?
- EI-iii. Ch'ŏnjŏlli [Chencenli] first inscription (蔚州川前里書石原銘), 525?, and additional inscription, 539?
- EI-iv. Chŏksŏng [Cekseng] inscription (丹陽新羅赤城碑), 540?
- EI-v. Imsin oath inscription (壬申誓記石銘), 552?
- EI-vi. Ojak [Ocak] inscription (戊戌塢作碑), 578?
- EI-vii. Sinsŏng [Sinseng] inscriptions (南山新羅碑) 1 to 9, 591?
- EI-viii. Myŏnghwal-san [Myenghwal-san] inscription (明活山城作成碑), 611?

(EI-i–iv) mix Korean and Chinese word orders, whereas (EI-v) and (EI-viii) are arranged completely in Korean word order but use characters as ūmdokcha (see 3.3). This writing continued in later periods, when it constituted one of the styles of writing. (EI-vi–vii) are also arranged in Korean word order, but use hundokcha. There are many collocations used in these that continue in later administrative texts.

Early idu style continued into the MOK period; the following are representative texts:

- MI-i. In'yang-sa [Inyang-sa] inscription (昌寧仁陽寺碑銘), 810
- MI-ii. Chungch'o-sa [Cwungcho-sa] pillar inscription (中初寺幢竿石柱記), 827
- MI-iii. Samhwa-sa Buddha inscription (東海市三和寺鐵佛造像銘), 860

A sentence of (MI-ii) is parsed in (1):

- (1) 中初寺 東方 僧岳 一 石 分 二 得 同月  
 Chungch'o-sa east Sūngak one stone split two gain same.month  
 二十八日 二 徒 作 初 奄 九月 一日 此處 至  
 28th.day two crowd make start then 9th month 1st.day here reach  
 'At Sūngak to the east of Chungch'o-sa they split a stone into two. On the 28th of the same month two groups began shaping them, and on the 1st of the 9th month they arrived here.'

MOK texts emerged with 't'o [tho] 吐 marks' added to early idu style, displaying features of full-blown idu texts. Twenty-five or so such idu texts are known at present, of which the following are representative:

- MT-i. Kamsan Amida [Kamsan Amita] inscription (甘山寺阿彌陀佛像造成銘), 720
- MT-ii. Mujin-sa [Mucin-sa] bell inscription (无盡寺鐘銘), 745
- MT-iii. Hwaōm-sa [Hwaem-sa] sutra notes (華嚴經寫經造成記), 755
- MT-iv. The Silla [Sinla] Register (新羅帳籍), pre-758
- MT-v. Karhang-sa [Kalhang-sa] pagoda inscription (葛項寺石塔), 758
- MT-vi. Ch'ōpp'ogi [Chepphoki] (正倉院所藏毛氈貼布記) notes, mid-eighth century
- MT-vii. Vairocana-Buddha inscription (永泰2年毘盧遮那佛造成銘), 766
- MT-viii. Ch'ōngje [Chengcey] inscription (永川菁堤碑貞元銘), 798
- MT-ix. Sōllim-wōn [Senlim-wen] bell inscription (禪林院鐘銘), 804
- MT-x. Kyuhūng-sa [Kyuhung-sa] bell inscription (竅興寺鐘銘), 856

These are inscriptions on metal or stone, and old manuscripts. There are, in addition, texts on wooden slats. (MT-i) is a relatively simple record, but shows the earliest t'o. (MT-iii) is a typical idu text of the period, displaying diverse idu symbols. (MT-iv) is a Silla administrative document that displays diverse administrative terms and idu symbols. A sentence of (MT-iii) is parsed in (2):

- (2) 經之 成內 法者 楮 根中  
 sutra-GEN make-CONSID steps-TOP paper.mulberry.tree root-LOC  
 香水…… 散尔 生長令內旆  
 perfumed.water sprinkle-MEANS grow-CAUS-CONSID-LINK  
 佛 菩薩 像 筆師 走使人那  
 Buddha bodhisattva statue calligrapher errand.boy-ALT  
 菩薩戒 授令旆 齊食旆  
 commandment receive-CAUS-LINK take.ritual.meal-LINK  
 (Underlined characters are t'o.)

'The way of making the sutra is to pour perfumed water on the root of the paper mulberry tree, mature it, . . . and to give commandments to the errand boys of the calligraphers of the Buddhist statues and have them take ritual meals, and . . .'

Sixty-plus Koryŏ period idu texts are known, and 20 or so of these are LOK. The most important of these are:

- LT-i. Myŏngbong-sa [Myengpong-sa] inscription (醴泉鳴鳳寺慈寂禪師塔碑陰銘), 941  
 LT-ii. Yaksa (= Bhaisaiyagura) inscription (校里磨匡藥師坐像銘), 977  
 LT-iii. Kaesim-sa [Kaysim-sa] pagoda record (醴泉開心寺石塔記), 1010  
 LT-iv. Pulguk-sa [Pulkwuk-sa] pagoda record (佛國寺無垢淨光塔重修記), 1024  
 LT-v. Chŏngdo-sa [Cengto-sa] pagoda record (若木淨兜寺五層石塔造成形止記), 1031  
 LT-vi. Pulguk-sa [Pulkwuk-sa] west pagoda record (佛國寺西石塔重修形止記), 1038

(LT-i) is an administrative document in which the Top'yŏng-sŏng [Tophyeng-seng] 都評省, a Koryŏ central government office, authorizes the founding of the temple Myŏngbong-sa. (LT-iv) and (LT-vi) are records of repairs to the pagoda at Pulguk-sa and show various idu symbols. In particular, they use characters abbreviated as – or according to the principle of – kugyŏl abbreviated forms, and this is the same logical principle as the use of old katakana in Japan. (LT-v) has been known for a long time. Part of it is parsed in (3):

- (3) 郡百姓 光賢亦... 石塔 五層乙  
 district.commoner Kwanghyŏn-NOM stone.pagoda 5.layers-ACC  
 成是白乎 願 表 爲遣  
 make-CAUS-HUM-NMR wish express do-CONJ  
 成是 不得 爲乎 天禧 六年... 五月  
 make-CAUS can't do-REL Ch'ŏnhŭi 6th.year 5th.month  
 初七日 身病以 遷世 爲去在乙  
 7th.day illness-INST depart.this.world do-CONF-CONT-DISC  
 同生兄 副戶長 稟柔亦... 修善僧 覺由...乙  
 elder.brother deputy.hojang P'umyu-NOM good.karma.monk Kagyu-ACC  
 繼 願 成畢 爲等 勸善 爲...  
 inherit wish completion do-[NMR]-fact-ACC encouragement do[-CONJ]  
 'District commoner Kwanghyŏn expressed the desire to build a five-storey stone pagoda, and on the seventh of the fifth month of the sixth year of the Ch'ŏnhŭi era when he had not been able to build it he departed this world because of illness, but deputy *hojang* [= Koryŏ period local headman] P'umyu who was his elder brother encouraged the noble monk Kagyu to inherit this wish and to complete the construction, and ...'

### 3.2.2 Hyangch'al data

Hyangch'al was mainly used to write *hyangga* [*hyangka*] 鄉歌 songs in sentences recorded entirely in Korean. Scriptural commentaries that explained difficult passages of sutras are also included here. The following 14 *hyangga* have been preserved in *Samguk yusa*:

- H-i. *Mojukchirang-ga* [*Mocwukcilang-ka*] 慕竹旨郎歌  
 H-ii. *Hŏnhwa-ga* [*Henhwa-ka*] 獻花歌

- H-iii. *Anmin-ga* [*Anmin-ka*] 安民歌
- H-iv. *Ch'angip'arang-ga* [*Chankiphalang-ka*] 讚著婆郎歌
- H-v. *Ch'öyong-ga* [*Cheyong-ka*] 處容歌
- H-vi. *Södong-yo* [*Setong-yo*] 薯童謠
- H-vii. *Toch'önsugwanüm-ga* [*Tochenswukwanum-ka*] 禱千手觀音歌
- H-viii. *P'ung-yo* [*Phung-yo*] 風謠
- H-ix. *Wönwangaeng-ga* [*Wenwangsaying-ka*] 願往生歌
- H-x. *Tosol-ga* [*Tosol-ka*] 兜率歌
- H-xi. *Chemangmae-ga* [*Ceymangmay-ka*] 祭亡妹歌
- H-xii. *Hyesöng-ga* [*Hveyseng-ka*] 彗星歌
- H-xiii. *Wön-ga* [*Wen-ka*] 怨歌
- H-xiv. *Ujök-ka* [*Wucek-ka*] 遇賊歌

*Samguk yusa* was written in the later thirteenth century, so (H-v), (H-vi) and (H-viii) are thought, because of how they are written and their grammar, to have been recorded not long before the thirteenth century. I treat these as LOK texts. (H-ii) is deciphered in (4):

- (4) 紫布 岩乎 过希 執音 乎 手 母牛  
tólpo-y pahuy kós-öy cap-óm [ho]-n son amsyo  
 purple rock edge-LOC hold-DUR [do]-NMR hand cow-ACC
- 放教遣 吾盼 不喩 慚盼伊賜 等  
noh-kisi-ko na-lól antí puskuli-só-n tó-n  
 let.go-HON-CONJ I-ACC NEG be.ashamed-HON-NMR fact-TOP (= if)

花盼 折叱可 獻乎理音如  
koc-ól kesk-e pat-o-li-ms-ta  
 flower-ACC pick-INF give-INTENT-PROSP-ESSEN-DEC  
 ‘Beside the purple rock [of azaleas] you made me let loose the cows [because of your beauty], and if you do not feel ashamed of me I shall pick a flower and give it to you.’

This is deciphered initially in MOK grammar and in MK lexis and morphology. Underlined elements are t’o.

Another 11 songs, the *Pohyön sibwön-ga* [*Pohyen sipwen-ka*] 普賢十願歌, or ‘songs of Samantabhadra’s ten great vows’, are included in the *Kyunyöjön* [*Kyunyecen*] 均如傳:

- P-i. *Yegyöng chebul-ga* [*Yeykyeng ceypul-ka*] 禮敬諸佛歌
- P-ii. *Ch'ingch'an Yörae-ga* [*Chingchan Yeray-ka*] 稱讚如來歌
- P-iii. *Kwangsü kongyang-ga* [*Kwangswu kongyang-ka*] 廣修供養歌
- P-iv. *Ch'amhoe öpchang-ga* [*Chamhoy epcang-ka*] 懺悔業障歌
- P-v. *Suhüi kongdök-ka* [*Swuhuy kongtek-ka*] 隨喜功德歌
- P-vi. *Ch'öngjön pömyun-ga* [*Chengcen peplyun-ka*] 請轉法輪歌
- P-vii. *Ch'öngbuljuse-ga* [*Chengpulcwusey-ka*] 請佛住世歌
- P-viii. *Sangsu purhak-ka* [*Sangswu pulhak-ka*] 常隨佛學歌
- P-ix. *Hangsun chungsaeng-ga* [*Hangswun cwungsaying-ka*] 恒順衆生歌
- P-x. *Pogae hoehyang-ga* [*Pokay hoyhyang-ka*] 普皆廻向歌
- P-xi. *Ch'onggyöl mujin-ga* [*Chongkyel mucin-ka*] 總結無盡歌



(5b) reorders the text into Korean word order, while (5c) converts the kugyŏl characters in (5b) into their source Chinese characters. Underlined characters in (5c) are hundokcha (see 3.3). These (5c) characters are identical to hyangch'al texts, so we know interpretive kugyŏl formed the basis for hyangch'al. (5d) is a transcription; Chinese loans are transcribed in capitals based on their MK forms.

Dot-type interpretive kugyŏl texts are:

- i. *Hwaŏm munŭi yogyŏl* [*Hwaem munuy yokye*] 華嚴文義要訣 (estimated mid-eighth century)
- ii. *Chinbon Hwaŏm-gyŏng* [*Cinpon Hwaem-kyeng*] 晋本華嚴經 (the Chinbon text of the Garland Sutra, estimated ninth or tenth century), volume 20
- iii. *Yugasaji-ron* [*Yukasaci-lon*] 瑜伽師地論 (estimated first half of eleventh century), volumes 3, 5, 8
- iv. *Chubon Hwaŏm-gyŏng* [*Cwupon Hwaem-kyeng*] 周本華嚴經 (the Chubon text of the Garland Sutra, estimated first half of the eleventh century), volumes 6, 22, 33, 34, 36, 57
- v. *Pŏphwa-gyŏng* [*Pephwa-kyeng*] 法華經 (Lotus Sutra, estimated tenth or eleventh century), volume 1
- vi. *Happu Kŭmgwangmyŏng-gyŏng* [*Happu Kumkwangmyeng-kyeng*] 合部金光明經 (Golden Light Sutra), volume 3

The symbols used in this dot-type kugyŏl are as in Figure 1. The clause 令一切衆生悉能除滅諸障闕業 is from line 7, page 2 of volume 20 of (ii), and the three characters 生, 能 and 業 in the sutra have dot-type t'ŏ attached to them. I shall use these to explain the way of reading dot-type kugyŏl. (6) is the sentence reordered into Korean word order and with character-type kugyŏl added.

- (6) 一切衆生<sub>し</sub> 悉 能<sub>支</sub> 諸 障闕業<sub>し</sub> 除滅  
 一切衆生<sub>-ir</sub> 悉 能<sub>-ti</sub> 諸 障闕業<sub>-ir</sub> 除滅  
 all.living.beings-ACC all capable every obstacle-ACC destroy

々 々 々

*hŕi-miə*

CAUS-LINK

‘causes all living beings to be able to surmount every obstacle, and . . .’

A single dot at mid-low position on the left side outside of 生 expresses the accusative particle *-(i)r*. This shows that 一切衆生 is used contextually as the object. A single dot at top right within 能 expresses *ti* and shows that this 能 is interpreted as the adverb

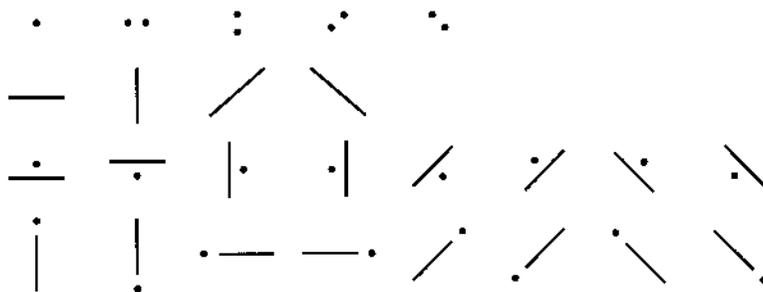


FIGURE 1 DOT-TYPE KUGYŏL

**TABLE 3.1 DOT-TYPE KUGYŎL SYSTEMS OF CHUBON HWAŎM-GYŎNG AND YUGASAJI-RON**

a. <i>Chubon Hwaŏm-gyŏng</i>				b. <i>Yugasaji-ron</i>															
<i>iə</i>		<i>l</i>	<i>tə/n</i>	<i>ti</i>		<i>na/to</i>	<i>ko</i>												
<i>tiə</i>	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td><i>čiəi</i></td><td><i>ii</i></td><td><i>ti</i></td></tr> <tr><td></td><td><i>n</i></td><td><i>ko</i></td></tr> </table>	<i>čiəi</i>	<i>ii</i>	<i>ti</i>		<i>n</i>	<i>ko</i>		<i>s</i>	<i>ti</i>	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td></td><td><i>ii</i></td><td></td></tr> <tr><td><i>ra</i></td><td><i>n</i></td><td><i>l</i></td></tr> </table>		<i>ii</i>		<i>ra</i>	<i>n</i>	<i>l</i>		<i>sa</i>
<i>čiəi</i>	<i>ii</i>	<i>ti</i>																	
	<i>n</i>	<i>ko</i>																	
	<i>ii</i>																		
<i>ra</i>	<i>n</i>	<i>l</i>																	
<i>r</i>		<i>həi</i>	<i>sa</i>	<i>l</i>		<i>s</i>	<i>lo</i>	<i>kii</i>	<i>hi</i>										
<i>miə</i>		<i>kii</i>	<i>ta</i>	<i>ta</i>		<i>kiə</i>	<i>iə</i>	<i>ha</i>	<i>miə</i>										

能-*ti*. Three dots are clustered on 業. The dot at mid-low position on the left side outside it is the *t*'o that expresses accusative *-(i)r*, as seen above. This shows that 諸障闕業 is read as an object. What looks like an exclamation mark made by a dot under a slant line at mid position on the right side outside 業 is the *t*'o expressing causative *-həi*. This corresponds to the function of 令 in the original Chinese, but one reads this *t*'o suffixed to the verbal 除滅 rather than read the 令. The dot at the bottom left outside 業 expresses *-miə*, a suffix conjoining two clauses.

There are many dot-type interpretive kugyŏl marks that are not yet deciphered. However, the single dot was used most. We know that in most cases double dots, lines or complex marks represent expanded versions of morphemes: in the *Chubon Hwaŏm-gyŏng*, for example, *.* at mid-low position on the left outside the Chinese character represents *-r*, while in the same position *..* represents *-tər*, *:* represents *-holir*, and *·|* represents *-holtər*, all of which include accusative *-r*.

Table 3.1 gives charts that show the various positions of dots relative to the imagined square occupied by a Chinese character.

In the library at the temple Tōdaiji in Nara is an abridgement of the *Hwaŏm-gyŏng*, volumes 12–20, believed to have been brought to Japan before the 740s. This is currently being studied because it has Silla-language *t*'o produced with a stylus as well as annotations and all sorts of marks. There are characters and forms in this that do not appear in Silla idu or *hyangga*.

External materials include Chinese histories: *Shiji* (史記), *Book of Han* (漢書), *Book of Later Han* (後漢書), *Records of Three Kingdoms* (三國志), *Book of Jin* (晉書), *History of the Southern Dynasties* (南史), *History of the Northern Dynasties* (北史), *Book of Sui* (隋書), *Book of Tang* (舊唐書), *New Book of Tang* (新唐書). These fragmentarily show us OK lexis. Japanese histories are *Kojiki* (712) and *Nihon shoki* (720). These also present us with data on Three Kingdoms lexis, but mostly they preserve information on Paekche OK. Also, Japanese metal and stone inscriptions and old manuscripts present data on EOK. The *Kyerim yusa* was recorded by Song dynasty scholar Sun Mu, and within it 350 or so Koryŏ words are recorded under the heading of 'dialect'.

### 3.3 ADAPTATION OF CHINESE CHARACTERS AND THEIR DEVELOPMENT

This notation developed together with the way that Chinese characters and Chinese were learnt and read. This development had already begun in the EOK period and developed into a rather precise notation by the LOK period.

In studying Chinese, Koreans read Chinese characters out as Chinese (*ŭm* [um] 音) or as semantically equivalent Korean words (*hun* [hwun] 訓). Hence, by borrowing the characters' pronunciation and meaning, a method of representing Korean came to develop.

**TABLE 3.2 ADAPTATION OF CHINESE CHARACTERS TO OLD KOREAN**

	Semantographic value (讀)	Phonographic value (假)
Sino-based (音):	ũmdokcha [umtokca] (音讀字) Sino-based semantogram = Chinese loan morpheme pronunciation	ũmgaja [umkaca] (音假字) Sino-based phonogram = phonogram based on Chinese
Koreo-based (訓):	hundokcha [hwuntokca] (訓讀字) Koreo-based semantogram = native Korean morpheme of similar meaning	hun'gaja [hwunkaca] (訓假字) Koreo-based phonogram = phonogram based on the native Korean morpheme's pronunciation

Their usage came to differ depending on whether they used a character's meaning as a semantogram or ignored the meaning and used only its pronunciation as a phonogram. Table 3.2 shows the division of ch'aja use into four types. The semantographic use is called 'read' (*tok/dok* 讀) in the sense that they are read just as they are, and the phonographic use is called 'borrowed' (*ka/ga* 假).

All ch'aja notation, i.e. the ch'aja used to notate sentences and words, corresponds to one of these categories. Sometimes only ũmdokcha and ũmgaja are written, and sometimes only hundokcha and hun'gaja are written. However, the most widespread device was the structure semantogram + phonogram.

Ūmgaja were used from the earliest period in which characters were used to represent Korean. These correspond to the *jiajie* category of character which the Chinese used to write loanwords, and proper nouns that appear in the Kwanggaet'o stela are almost only in ũmgaja. The *hun* device is a native device that Koreans developed. It is already recognizable in the metal and stone inscriptions of the Three Kingdoms period, and the Silla official title 大舍 used in the Ch'ŏnjŏlli inscription (?525) is written in eighth-century texts as 韓舍. This shows that 大 was a hundokcha read as 韓 *han*. Also, in Ch'ŏnjŏlli supplementary inscription (?539) the Silla title 波珍干支 appears, where 波珍 means *patul* 'sea' and so 珍 is a hun'gaja read as *tul*. On the other hand, the Koguryŏ minister 淵蓋蘇文 is recorded in *Nihon shoki* as 伊梨柯須彌 *irikasumi*. This shows that in Koguryŏ the character 淵 was a hundokcha for *iri*. No Paekche texts are extant, but considering the point that Old Japanese literary activity was carried out by migrants from Paekche, one must see that using the *hun* device also developed in Paekche. These facts indicate that all the Three Kingdoms used notation that used the *hun* device.

There are proper 'Chinese' texts extant from the Three Kingdoms. Koguryŏ's Tongsu epitaph (冬壽墓誌, 357) and the Kwanggaet'o stela, Paekche's King Munyŏng tomb inscription (武寧王陵誌石銘, 525) and Sat'aek Chijŏk [Sathayk Cicek] inscription (砂宅智積碑銘, ?654), and Silla's inscription of King Chinhŭng's visit (眞興王巡狩碑銘, 540s or later), etc. are proper Chinese texts. In contrast, the Chungwŏn Koguryŏ inscription of Koguryŏ mixes Chinese and Korean word order and uses ch'aja such as 之 or 節 associated with later idu. The P'yŏngyang castle wall engravings include Korean linguistic elements. The Ūnch'ŏn inscription excavated from Paekche King Munyŏng's tomb is Korean in word order, while the *sukse-ga* mix Korean and Chinese word orders. The Naengsuri, Pongp'yŏng, Ch'ŏnjŏlli and Chŏksŏng inscriptions, etc. from Silla also mix Chinese and Korean word orders. However, the Imsin oath and Myŏnghwal-san inscriptions arrange Chinese, i.e. ũmdokcha, entirely in Korean word order. These are early idu texts written only in semantograms. On the other hand, notation that arranges hundokcha entirely in Korean word order also developed in this period, as in the Ojak and Sinsŏng inscriptions.



to idu texts. Because idu had a lot of omission of its t'o we have to read it by supplying the missing t'o. When we analyze them together, we can call the development of reading devices for Chinese, or kugyōl which was the device for reading Chinese, the womb for the development of ch'aja writing.

To write t'o in kugyōl, sometimes ch'aja were used unchanged in their proper written form but in most cases unique kugyōl forms of characters were used. If it was a simply written ch'aja the whole character was used, but mostly an abbreviated character was used by missing out strokes of the original character. Typically this abbreviated form involved taking the first part or the last part of a square- or grass-script character. Its value was determined according to whether it used the *ūm* or the *hun* of the character. Kugyōl characters used in some interpretive kugyōl were generally about 55 characters, but depending on text we also find increased numbers of characters and different forms of the same character.

In Table 3.3 are presented the kugyōl characters used among the interpretive kugyōl of kugyōl texts (ii) and (vi). Following the kugyōl character is its reading, the character it is derived from, then a classification into whether it is based on square or grass script, whether it derives from the early or latter part of the full character, and whether it is *ūm* or *hun*. We can tell by this the process by which kugyōl characters came about.

Most of these kugyōl characters are phonograms, but auxiliary verbs such as *hve*, *kiə*, *šerp*, *i*, *hvi*, etc. and bound noun *čahi* are mostly used as hundokcha.

We can easily trace these abbreviated kugyōl characters to their original Chinese characters, and most of the latter are used to write idu and kugyōl t'o.

### 3.4 EARLY OLD KOREAN

I regard Early Old Korean as a single language rather than as Koguryō, Paekche and Silla being separate languages. The three-language approach is based on the Koguryō toponyms in the geographical chapters of *Samguk sagi*. However, Koguryō ruled the territory corresponding to these toponyms for not even 100 years. This territory was not greatly separate from the territorial range of the Samhan period, which preceded the division into the Three Kingdoms, and I regard the latter's language as a single 'Han'-branch variety.

If we regard the Han-branch varieties as the 'Southern branch' and, from the fact that Koguryōan succeeded the 'Gojosōn' and 'Puyō' languages in the north, we regard it as the Northern branch, we can show how the two branches were related as follows. It is extremely rare to be able to decipher with Koguryō texts but the spelling 伊梨柯須彌 *irikasumi* recorded in *Nihon shoki* for 淵蓋蘇文 cited above shows that Koguryōan was close to the Han branch. In other words, 伊梨 *iri*, which is the *hun* of 淵 (a Chinese character meaning 'abyss'), corresponds to the 乙 *əl* of 奈乙 *naəl* ('well') in Sillan and to the 於乙 *äəl* of the Koguryōan toponym 於乙買串, which in the Chinese version of its name (泉井口縣) means 'Spring Water Wellmouth county'. Also, the fact that 蓋蘇文 *kaisumon* of 淵蓋蘇文 is also found as 蓋金 shows that 'metal' (金) was 蘇 *su* in Koguryōan. 蘇 *su* is the same as Sillan 素/省 *su* 'metal' and is connected with MK *soj* 'iron'. Also the original name of 鉛城 (lit. 'lead castle'), a castle north of the Yalu river, was 乃勿忽, allowing us to obtain the reading 乃勿 *namül* for 'lead', and this matches with Old Japanese *namari* 'lead' and with 那勿 *namər* given in the statement 'lead is colloquially known as *namər*' (鉛俗云那勿) in the *Hyangyak kugūppang* which reflects the early Koryō language. I shall discuss more regarding the commonality of the Three Kingdoms' varieties in the following section.

TABLE 3.3 INTERPRETIVE KUGYŎL CHARACTERS

ハ	k(i)	只	square, bottom, <i>ũm</i>
去	kə	去	grass, whole, <i>ũm</i>
ナ	kiə	在	square, top, <i>hun</i>
口	ko	在	square, bottom, <i>ũm</i>
尔	kim	古	grass, bottom, <i>hun</i>
人	koa	彌	square, bottom, <i>ũm</i>
斤	kin	果	square, whole, <i>ũm</i>
十	kii	斤	grass, whole, <i>hun</i>
𠂇	kə	中	square, whole, <i>ũm</i>
𠂇	kəs	良	square, whole, <i>hun</i>
𠂇	(i)n	這	square, top, <i>ũm</i>
乃	na	隱	square, bottom, <i>ũm</i>
ト	nu	那	square, bottom, <i>hun</i>
ヒ	ni	臥	square, bottom, <i>ũm</i>
ヒ	nə	尼	square, top, <i>hun</i>
丨	ta	飛	grass, whole, <i>hun</i>
丁	tiə	之	square, whole, <i>ũm</i>
刀	to	丁	square, whole, <i>ũm</i>
巴	torok	斗	unknown
斗	tu	斗	square, whole, <i>ũm</i>
支	tì	支	square, whole, <i>ũm</i>
矢	tì	知	square, top, <i>ũm</i>
𠂇	tì	止	grass, whole, <i>ũm</i>
入, の	tə	入	s/g, whole, <i>hun</i>
冬	təl	入	square, whole, <i>hun?</i>
ム	təi	冬	square, top, <i>hun?</i>
シ	(i)r	矣	square, whole, <i>ũm</i>
尸	l	乙	square, whole, <i>hun?</i>
𠂇	ra	尸	square, top, <i>ũm</i>
𠂇	lo	羅	grass, whole, <i>hun</i>
𠂇	ri	以	square, top, <i>ũm</i>
𠂇	li	利	grass, whole, <i>ũm</i>
𠂇	(i)m	令	square, top, <i>ũm</i>
𠂇	ma	音	?
𠂇	miə	彌	grass, top, <i>ũm</i>
毛	mo	毛	square, whole, <i>ũm</i>
火	pə	火	square, whole, <i>hun</i>
七	s	叱	square, bottom, <i>ũm</i>
𠂇	sa	沙	square, top, <i>ũm</i>
一	sam	三	square, top, <i>ũm</i>
立	siə	立	square, whole, <i>hun</i>
𠂇	siə	徐	square, top, <i>ũm</i>
二	si	示	square, top, <i>ũm</i>
𠂇	sə	賜	grass, bottom, <i>ũm</i>
白	səp	白	square, whole, <i>hun</i>
𠂇	a	良	grass, top, <i>hun</i>
𠂇	iə	亦	s/g, top, <i>ũm</i>
𠂇	o	乎	s/g, top, <i>hun?</i>
𠂇	u	于	square, whole, <i>ũm</i>
𠂇	ii	衣	square, top, <i>ũm</i>
𠂇	i	是	grass, top, <i>hun</i>
𠂇	ii	之	square, whole, <i>hun</i>
𠂇	čahi	第	grass, whole, <i>hun</i>
𠂇	čiai	齊	grass, whole, <i>ũm</i>
下	ha	下	square, whole, <i>ũm</i>
𠂇	ho	乎	square, top, <i>ũm</i>
𠂇	hi	乎	square, whole, <i>ũm</i>
𠂇	hə	爲	grass, top, <i>hun</i>
𠂇	həi	令	grass, top, <i>hun</i>

It is still too early to discuss the phonology of EOK, and I shall adopt the romanization used for MOK.

### 3.4.1 Lexis

I shall describe EOK lexis predominantly in connection with toponyms in *Samguk sagi*, with reference also to *Samguk yusa* and Chinese and Japanese texts. I list Koguryō (9), Silla (10) and Paekche (11) EOK words, followed by romanization and their meanings.

#### (9) The Koguryō dialect:

- a. **Nouns:** 加 *ka* ‘great man, noble’; 皆 *kai* ‘king’; 琉璃/類利/孺留/累利 *#rūri* ‘world, society’; 且/吞/頓 *tən* ‘valley’; 內/奴/惱 *nui* ‘earth’; 達 *tar* ‘mountain’; 巴衣/波衣/波兮 *#paɨ* ‘rock’; 於乙 *är* ‘spring’; 勿 *#mör* ‘water’; 買 *mai* ‘water, river’; 古衣 *#kuɨ* ‘swan’; 功木 *#kumuk* ‘bear’; 首 *#siu* ‘ox’; 烏斯含 *usoham* ‘rabbit’; 夫斯/扶蘇 *pūsɔ* ‘pine tree’; 蘇/素 *#su* ‘iron’; 乃勿 *#namör* ‘lead’; 也次 *#iačho* ‘mother’; 伊伐支 *#ipärti* ‘beside’; 鄒牟/朱蒙 *čümü* ‘skilled archer’; 忽次/古次 *kučho* ‘mouth’; 斤乙 *#kör* ‘written character’; 吐 *tu* ‘dike’; 溝漣 *kürü*, 骨 *kur*, 忽 *hur* ‘district’.  
e.g. 奈兮忽 *nahiäi-hur* = 白城郡 ‘White Fort district’, 沙伏忽 *sapuk-hur* = 赤城縣 ‘Red Fort county’, 買忽 *mai-hur* = 水城郡 ‘Water Fort district’, 述爾忽縣 *süni-hur* = 峰城縣 ‘Peak Fort county’, etc.
- b. **Numbers:** 密 *mir* ‘three’; 于次 *üčho* ‘five’; 難隱 *nanön* ‘seven’; 德 *täk* ‘ten’.
- c. **Verbals:** 多勿 *tamör* ‘reclaim territory’; 於斯 *#äsɔ* ‘lean’; 伯 *paik* ‘welcome’; 伊 *i* ‘enter’; 沙非斤/沙伏 *sapuk* ‘be red’; 位 *#üi* ‘resemble’; 屈火 *#kürpür* ‘bend’; 今勿 *#kömör* ‘be black’.

#### (10) The Silla dialect:

- a. **Nouns:** 干 *kan*, 邯 *kam*, 今 *kəm* ‘leader’; 次次雄 *#čöčüing*, 慈充 *čöčüing* ‘king, shaman’; 居西干 *käsakan* ‘king’; 尼師今 *#nisököm* ‘king, elder’; 吉士 *kirsɔ* ‘14th-rank title’; 儒禮·弩禮·世理 *#nüriäi* ‘world’; 內 *nai* ‘world’; 只 *ki* ‘fortress’; 乙 *ör* ‘well’; 勿 *#mör* ‘water’; 那 *na* ‘river’; 比自/比斯 *#piso* ‘light’; 一利 *#tri* ‘star’; 素/省 *#su* ‘metal/iron’; 首 *#siu* ‘ox’; 巨老 *#käru* ‘goose’; 毛 *#mu* ‘mosquito’; 三 *#sam* ‘hemp’; 朴 *#pak* ‘gourd’; 伊史/異斯 *#isɔ* ‘moss’; 柯半 *#kapan* ‘breeches’; 麻立 *#marip* ‘stake’; 嘉俳 *#kapɨ* ‘15th of 8th lunar month, mid-autumn festival’; 吐 *tu* ‘dike’; 洗 *#siän* ‘shoes’; 尼叱/尼師 *#nis* ‘tooth’; 闕知 *#arti* ‘infant’; 都 *#tu* ‘barbarians north of Tumen river’; 牟羅/慕羅 *mura* ‘county’; 火 *#pör*, 伐 *pär* ‘district’.
- b. **Numbers:** 推 *mir* ‘three’.
- c. **Verbals:** 密 *#mir* ‘push’; 阿火 *#apör* ‘join’; 南 *#nam* ‘exceed’; 內 *an* ‘be good’; 異次/異處 *#ičho* ‘hate’; 今勿 *#kömör* ‘become shady’; 吉 *#kir* ‘be long’; 居柒 *#käčhör* ‘be rough’; 弗矩 *#pörkä* ‘be red, bright’; 赤牙 *#pörkäm* ‘be red’; 翰 *#han* ‘big’, 韓 *#han* ‘big’.

#### (11) The Paekche dialect:

- a. **Nouns:** 韃吉之 *känkirči* ‘king’; 於羅瑕 *äraha* ‘king’; 吉士/吉師 *kirsɔ* ‘noble’; 二リム *#nirim* ‘lord’; ハシカシ *#hasikasi* ‘wife’; 所非 *#supi* ‘forest’; 斯摩/斯麻 *#soma* ‘island’; 仇池/仇知 *#küti* ‘metal/iron’; 珍惡 *#turak* ‘stone’; 只/己 *ki* ‘fort’; 牟羅 *mura* ‘village’; 忽 *hur* ‘fort’; 夫里 *#püri* ‘city, village’.

- b. **Verbals:** 奈 #*na* ‘come out’; 近 #*kön* ‘big’; 沙 #*sa* ‘new’; 烏 #*u* ‘be lonely’; 所比 *supi* ‘be red’; 勿居 #*mörkä* ‘be clear’; 今勿 *kömör* ‘be black’.

Three Kingdoms lexical items we can identify are extremely few in number. We can see that the number of words in common between them is significant. In particular the number of words connected to MK is significant and this fact shows that they were all Han languages. I mark such words with # above.

The 忽 *hur* in Koguryō county and district place names was used not only in the Samhan territory but also mostly north of the Yalu river. In the ‘eastern barbarians’ chapter of the Wei volume of the *History of the Three Kingdoms*, it says that 溝瀆 denotes a Koguryō castle, and if we refer to the fact that a Koguryō city name was 紇升骨 then we can assume the development 溝瀆 *kürü* (> 骨 *kur*) > 忽 *hur*. This name was used as an administrative term meaning ‘district’ or ‘county’. This 忽 *hur* did not come from the corresponding area’s spoken language but was established uniformly in the centre. We can confirm that in the southern dialects 忽 *hur* was used in 慰禮忽國 in Silla and 伏忽 in Paekche, and so it reflected the indigenous language. 夫里 *pūri* is an administrative term used in Paekche toponyms and its distribution is relatively wide, and 只己 *ki* is also used several times in the sense of ‘fort’. The corresponding administrative term in Silla was 火 *pö*/伐 *pär*. We trace these, 夫里 *pūri* and 火 *pör*, to the 卑離 *pīri* in the Mahan language. On the other hand, 牟羅/慕羅 *mura* was widely used across Paekche, Silla and Kaya. This was used as a toponym for relatively large settlements in the south before the break-up of the Samhan. The corresponding northern word is 牟婁 *murū*, and it is assumed that Japanese *mura* ‘village’ is cognate.

Words corresponding to ‘king’ differ between the Three Kingdoms, and this is thought to derive from the different processes by which the dynasties were established. In contrast, we can see that the titles of officials or for ‘noble’ are in common.

The word for ‘rabbit’ is in common between Koguryō (*usōham*) and Japanese (*usagi*). Though 吐 *tu* ‘dike’ does not occur in MK, it is found in both Koguryō and Silla. The number 密 *mir* ‘three’ is found in Koguryō and Silla, and 沙非斤/沙伏 *sapuk* ‘be red’ in Koguryō and 所比 *supi* ‘be red’ in Paekche show the same heritage, so we know that both are Han languages. 今勿 *kömör* ‘be black’, despite semantic differences between the Three Kingdoms, is shared between them. The dialects of the Three Kingdoms show differences, but these are small compared with what they have in common.

EOK has many words in common with Japanese. The following are often discussed and have a high credibility:

- |      |             |            |                           |                           |            |                       |
|------|-------------|------------|---------------------------|---------------------------|------------|-----------------------|
| (12) | <i>tōn</i>  | ‘valley’   | OJ <i>tani</i>            | <i>mōi</i>                | ‘water’    | OJ <i>mi</i>          |
|      | <i>koma</i> | ‘bear’     | OJ <i>kuma</i>            | <i>nami</i>               | ‘lake’     | OJ <i>nami</i> ‘wave’ |
|      | <i>mair</i> | ‘garlic’   | OJ <i>mira</i> ‘leek’     | <i>usōham</i>             | ‘hare’     | EMJ <i>usagi</i>      |
|      | <i>kuōi</i> | ‘swan’     | EMJ <i>kofu</i>           | <i>kučō</i>               | ‘mouth’    | OJ <i>kuti</i>        |
|      | <i>piär</i> | ‘layer’    | OJ <i>pe</i>              | <i>ki</i>                 | ‘castle’   | OJ <i>ki</i>          |
|      | <i>sōma</i> | ‘island’   | OJ <i>sima</i>            | <i>mir</i>                | ‘three’    | OJ <i>mi</i>          |
|      | <i>üčō</i>  | ‘five’     | OJ <i>itu</i>             | <i>täk</i>                | ‘ten’      | OJ <i>towō</i>        |
|      | <i>mura</i> | ‘village’  | OJ <i>mura</i>            | <i>kōpōr</i> > <i>kur</i> | ‘district’ | OJ <i>kopori</i>      |
|      | <i>muri</i> | ‘mountain’ | OJ <i>mure</i> ‘mountain’ | <i>patah</i>              | ‘sea’      | OJ <i>wata</i>        |
|      |             |            | OJ <i>mori</i> ‘forest’   | <i>pat</i>                | ‘field’    | OJ <i>pata</i>        |
|      | <i>tiär</i> | ‘temple’   | OJ <i>tera</i>            | <i>mum</i>                | ‘body’     | OJ <i>mu-</i>         |
|      | <i>nat</i>  | ‘sickle’   | J <i>nata</i> ‘hatchet’   |                           |            |                       |

### 3.4.2 Grammar

In EOK the verb root had a very independent nature. In MK, the verb root could be made into an adverb, e.g. *kót-* ‘be like’: *kót* ‘like’, two verb roots could be combined into a compound verb, e.g. *tat-ni-* ‘run’+‘go’, verb roots could modify nouns like *pswus* ‘rub’ in *pswustol* ‘whetstone’, and verb roots could occur as nouns, e.g. *sin-* ‘put on [shoes]’: *sin* ‘shoes’ or *póy-* ‘be pregnant’: *póy* ‘belly’. We know the verb root’s independent nature was far greater in EOK. Koguryō 伯 *paik* ‘welcome’, 伊 *i* ‘enter’, 多勿 *tamur* ‘reclaim territory’, 於斯 *āsō* ‘yellow’, 今勿 *kōmōr* ‘black’; Silla 居渠 *kāčir* ‘rough’, 吉 *kir* ‘long’, 密 *mīr* ‘push’, 阿火 *apōr* ‘join’, 異次/異處 *ič* ‘hate’; and Paekche 烏 *u* ‘lonely’, 所比 *supi* ‘red’, 奈 *na* ‘go out’ indicate that this sort of syntax was found in the Three Kingdoms. 南內 *naman*, renamed in Chinese 餘善縣 (lit. ‘exceedingly good county’), shows a compound of the roots *nam* ‘leave’ and *a* ‘good’. Even though we do not have many Three Kingdoms lexemes, there is a lot of this sort of commonality.

‘Samhan’ or ‘the three Hans’ is the coverterm for the 78 states of the Mahan 馬韓 (west), Chinhan [Cinhan] 辰韓 (east) and Pyōnhan [Pyenhan] 弁韓 (south) confederacies. 韓 *han* is formed from the verb *ha* ‘great’ and the nominalizer *-n* and has come to have the meaning ‘chieftain’. This *han* probably designated the chiefs of each of the 78 tribes, and developed as a nation name when the tribes formed confederacies and became the Samhan. The 韓 *han* in Silla and the 翰 *han* of Paekche toponyms are also cognate. Another form is 干 *kan* and it is thought that it diverged from the same word and the two forms co-existed very early on. With the growth and development of tribal societies we find the development of titles for ‘king’ by adding modifiers to this morpheme or such variants as 邯 *kam* or 今 *kōm*, giving 居西干, 居瑟邯, 尼師今 or 麻立干. Also it developed the meaning ‘noble’ and was used in Silla for such official ranks as 伊伐干 *ipārkan*, 波珍干 *patorkan*, 阿干 *akan* or 一吉干 *irkirkan*, as well as in the title of nine Karak officials, 刀干 *tokan*. 近 *kōn* in the names of the Paekche kings Kūnch’ogo [Kunchoko] 近肖古, Kūngusu [Kunkwuswu] 近仇首 and Kūngaeru [Kunkaylwu] 近蓋婁 is also a variant of *kan*, with nominalizer *-n* added to the root 幹 *kō* ‘great’. The 韃 *kān* ‘great’ of 韃吉支 is also a variant. Some think that the 加 *ka* ‘noble’ used in Koguryō titles derives from the root underlying *kan*, and the 瑕 *ha* of the Paekche royal title 於羅瑕 is also of the same root.

From the above we can confirm the existence of nominalizer *-n*, and the nominalizer *-on* can be confirmed by analyzing the name of Silla’s founder 弗矩內/赫居世 as *pōrk* ‘red’ + *-on* nominalizer + *oi*.

In Three Kingdoms inscriptions the respectful suffix 智/知/支 *ti* is widely used. It is used in the Naengsuri inscription, as 智 *ti* in people’s names (斯夫智, 乃智, 斯德智, 子宿智, 爾夫智) or as 支 *ti* in titles (阿干支, 居伐干支, 壹干支). On the other hand, 智 is used in titles such as 壹今智, showing 支 and 智 were homophones. In the Pongp’yōng inscription 智 is used in personal names and 支 is used consistently in the title 干支, but 智 is used in names of low-ranking officials such as 邪足智, 小舍帝智, 吉支智, 小烏帝智. In the Sinsōng inscription 知 *ti* is used, but in the name of a local man 之 is also found. 智/支/知/之 *ti* was used in this way as a Silla EOK honorific suffix, but it is the same root as 智 *ti* in the Mahan chieftain 臣智, 支 *ti* in Paekche 捷吉支 above, and 支 *ti* in Koguryō 莫離支.

It was noted above that EOK idu texts arrange ūmdokcha in Korean word order. Koguryō’s Chungwōn inscription mixes Korean and Chinese word orders. (13) is from the Korean word-order part:

- (13) 太位 諸位 上下 衣服 來 受 教 跪  
 top.rank all.ranks high.and.low robes come receive order kneel  
 營 之  
 camp DEC  
 ‘When he commanded that all ranks come and receive their robes, they knelt before the camp.’

The sequence 上下衣服來受 here is Korean SOV word order. The declarative suffix 之 and the verb 教 are used in this inscription and are connected with Silla idu texts. These are ūmdokcha but must be viewed as underlyingly spoken Koguryō.

The inscription of the queen’s silver bracelets excavated in King Munyōng’s tumulus is typical of Paekche idu (14).

- (14) 庚子年 二月 多利 作 大夫人 分 二百三十 主 耳  
 520 2nd.month Tari make queen part 230 1/24 oz. DEC  
 ‘Tari made this in the second month of 520. It is for the king’s wife. It is 9.59 Chinese ounces.’

This is also entirely in Korean word order. The sentence-final 耳 is a lexical particle and ūmdokcha used to correspond to the Paekche declarative suffix.

The Imsin oath inscription is a typical Silla idu text, ordering all the text entirely in Korean word order. In contrast is the Sinsōng inscription (15), an idu text which uses hundokcha.

- (15) 南山新城 作 節 如 法…以 作 後 三年 崩破…者  
 \*cis- \*tiüi \*-lo \*cis- \*-n  
 Namsan.Sinsōng build time if law-INST build after 3.years destroy-NMR  
 罪 教 事 爲 聞 教令 誓 事…之  
 \*-isi- \*il \*sam- \*piäki- \*il \*-ta  
 commit.crime-HON fact make.into tell.king teachings swear fact-DEC  
 ‘When they built Namsan Sinsōng, if it is built in accordance with the regulation and if it crumbles within three years, report it to the king to be punished, and the builders accordingly took oaths.’

The text concerns those who undertook the building work on the castle swearing an oath to the king.

The grammatical morphemes that appear in idu texts of the period are identified as follows. The particles 以 *lo* (instrumental) and 者 *n* (topic) occur, and 中 is used to represent spoken *a* or *kōi*. 之 as a suffix is used to represent declarative *-ta*. 在 *kiä* is an auxiliary verb expressing continuing action, and 節 *tiüi* is a conjunctive adverb. These were inherited into later idu.

### 3.5 MID OLD KOREAN

#### 3.5.1 Phonology

T’o developed in the MOK period and there are more texts than EOK. Yet it is only after the 1446 promulgation of han’gŭl [hankul] that a full description of Korean becomes possible. OK research is based predominantly on tracing back in time from fifteenth-century MK texts, applied more vigorously with phonology. The MOK consonants had two sets of

obstruents: plain (*p t č k*) and aspirate (*?ph ?th čh kh*). However, aspirates' functional load was extremely low, and so it is dubious whether *ph th* existed. Fricatives (*s h*) and nasals (*m n ŋ*) were the same as in MK. Liquids are thought to have distinguished *r* (ㄹ) vs *l* (ㄴ). MK *z* did not exist in MOK, and the phonemic status of *β* is uncertain even in MK, and we cannot trace it back into OK. The fifteenth century had a seven-vowel system with vowel harmony, so we presume these applied more strictly in OK. Vowel harmony is not reflected in the texts, so this is hard to verify. Some believe that a systematic obeying of vowel harmony meant an eight-vowel system, but the basis for this is weak. Current research regards it as least risky to trace back and adapt the MK seven-vowel system, which I also accept here (Lee 1972: 72; Lee and Ramsey 2011: 67; for a recent alternative, see Whitman, this volume):

<i>i</i>	( <i>= MK wu</i> )	<i>u</i>	( <i>= MK o</i> )
<i>ǰ</i>	( <i>= MK u</i> )	<i>ɔ</i>	( <i>= MK ó</i> )
<i>ǎ</i>	( <i>= MK e</i> )	<i>a</i>	

We cannot say that the MK semi-vowels had developed. I represent MK semi-vowel + vowel sequences as two vowels, i.e. *ia, iā, iu, iū, ua, uā*, etc. A phonological process that we can identify in this period is monophthongization of diphthongs, e.g. the use of 令 (*liāng*) – thought to be used for *\*liā* – in its abbreviated form 令 for *li* in the twelfth century, or the development of nominative 亦 *iäk* > 弋 *ik* > ㄹ *i* (see 3.6.2.(ii.)). The writing of one king's name as 儒禮, 弩禮, 儒理 or 世里 shows monophthongization of 儒 *nüü* > 弩 *nu* and 禮 *liäi* > *liä* > 理/里 *li*. *ɔ* was weak even in MOK, so such contractions as *ɔ+i* > *i* or *ɔ+iä* > *iä* occur.

### 3.5.2 Grammar

I shall divide MOK grammar between that in the *hyangga* of *Samguk yusa* and that in idu texts. The writing of grammatical morphemes in *hyangga* is largely full, but as they were composed at a different time to when they were committed to writing, they are difficult to date accurately. There have been many decipherment attempts, but not only are there so few texts but we also do not understand OK grammar enough, so decipherments have conflicted with each other. Following the recent discovery and decipherment of interpretive kugyöl and ancient idu texts, we are now able to close in on *hyangga* grammar, but parts still remain undeciphered, so we must bear this in mind when we describe the grammar. Idu texts' dates are certain and they are accurately decipherable, but the writing of grammatical morphemes is cursory and we certainly do not have enough of these texts. We can describe the grammatical system to an extent by combining these two types of text, but gaps are inevitable. In Tables 3.4 to 3.9 I present morphemes from *hyangga* on the left and from idu on the right. Superscript <sup>a</sup> = fusion with a stem-final sound, <sup>b</sup> = fusion with the bound noun ㅈ 'fact', and <sup>c</sup> = special cases.

(i) *Noun-following particles* (Table 3.4). I shall start with case particles in *hyangga*. Nominative *-i* was used as in 雪是 snow-NOM (H-iv), 爲賜尸 知 do-HON-NMR fact.NOM = 'the fact that X does' (H-iii). Genitive *-ǰi* was used of sentient beings (心未 際叱 heart.m-GEN edge.s = 'edge of the heart' (H-iv)), *-s* of non-sentients, and *-l* of direction-nouns (東尸 汀叱 east-GEN shore.s (H.xii)). 盼 *-hir* was used as accusative. It is unclear if the *h* characteristic of MOK *hyangch'al* in the accusative and locative was actually pronounced, and does not appear in LOK. The locative has *-ǰi* and *-a* types

TABLE 3.4 MID OLD KOREAN CASE AND DISCOURSE PARTICLES

		Hyangch'al	Idu
<i>a. Case particles:</i>			
Nominative:		是 <i>-i</i> , 知 <sup>b</sup> <i>t-i</i>	
Genitive:	Sentient:	矣 <i>-ŏi</i> , 未 <sup>a</sup> <i>m-ŏi</i>	
	Non-sentient:	叱 <i>-s</i>	之 <i>-s</i>
Accusative:		盼 <i>-hir</i>	
Locative:	<i>ŏi</i> -Type:	衣/矣 <i>-ŏi</i> , 希 <i>-hŏi</i> , 中 <i>kŏi</i> , 未/米 <sup>a</sup> <i>m-ŏi</i>	中 <i>-ŏi</i>
	<i>a</i> -Type:	良 <i>-a</i> , 乃 <sup>a</sup> <i>n-a</i>	
	Combined:	惡希 <i>-a-hŏi</i> , 良中 <i>-a-kŏi</i>	
Instrumental:		以 <i>-lu</i>	以 <i>-lu</i>
Connective:		也...耶/也 <i>-ia...-ia</i>	
Vocative:	Plain:	良 <i>-a</i> , 也 <i>-ia</i>	
	Honorific:	下 <i>-ha</i>	
Complex:	LOC + GEN:	之叱 <i>-ŏi-s</i> , 阿叱 <i>-a-s</i>	
<i>b. Discourse particles:</i>			
Topic:		隱/焉 <i>-ŏn</i> 吞/等/等焉 <sup>b</sup> <i>t-ŏn</i>	者 <i>-ŏn</i>
Inclusive:		置 <i>-tu</i>	
Selective:			那 (乃) <i>-(i)na</i>
Emphatic:		沙 <i>-sa</i>	
Distributive:		爾/尔 <i>-kŏm</i>	尔 <i>-kŏm</i>

(心未 *masŏm-ŏi* heart.m-LOC, 人米 *saram-ŏi* person.m-LOC, 前乃 front-LOC), as well as combinations of *-a* + *-kŏi/hŏi*. The connective was used as in 邊也 藪耶 shore-CONN woods-CONN 'the shore and the woods' (H-xii). The writing of case particles was cursory in idu: neither nominative nor accusative are written, and the genitive only occurs in 經之 成內 法者 sutra-GEN make-CONSID[-NMR] way-TOP 'the way in which one makes the sutra' (example (2)).

Topic 隱 and rarely 焉 were used in *hyangga*, some believing the difference being vowel harmony. There was no special conditional converb in OK and the topic marker fulfilled this function. Emphatic *-sa* either directly attached to a noun or followed another particle. Distributive 爾 *-kŏm* is used as in 八切爾 eight-cut-DISTR 'eight cuts each' (H-xii), or 經心 內中 一收 舍利尔 入內如 sutra.heart inside-LOC one śārīra-DISTR insert-CONSID-DEC (MT-iii) 'inserts one śārīra inside each sutra'.

(ii) *Verb enders* (Table 3.5). Declarative *-ta* was used as in 人是 有叱多 person-NOM exist.s-DEC 'there's someone' (H-xii), 植內之 plant-CONSID-DEC 'planted them' (MT-iv), 妹者 敬信大王 孀在也 younger.sister-TOP Kyōngsin-king maternal.aunt[-COP]-CONT-DEC 'his sister is King Kyōngsin's aunt' (MT-v). 如 emerged in the mid-eighth century and was used till the end of the nineteenth, and 也, though rare, had been used since EOK. 矣 was mostly used for a converb *-tŏi*, but there are MOK examples for *-ta*: 鐘成內矣 bell[-ACC] make-CONF-DEC 'they (had) made the bell' (MT-x). *-čŏi* was used for an objective statement, but in MOK idu *-čŏi* was also used for a wish: 後代 追愛人者 此 善 助在哉 posterity cherish people-TOP this virtue assist-CONT-DEC 'may people who cherish this in later times assist in this virtue' (MT-i). Interrogative *-ku* and *-ka* were both used only as polar interrogatives, unlike the LOK polar *-ka* vs wh *-ku* distinction. 民焉 狂尸恨 阿孩古 爲賜尸 知... people-TOP foolish child-QUOT do/

TABLE 3.5 MID OLD KOREAN VERB ENDERS

	Hyangch'al	Idu
<i>a. Sentence-enders:</i>		
Declarative 1:	如/多 <i>-ta</i>	之/如/也/矣 <i>-ta</i>
Declarative 2:	齊 <i>-čiäi</i>	哉 <i>-čoi</i>
Interrogative:	古/遣/故 <i>-ku</i> , 去 <i>-ka</i>	
Exclamatory:	也/邪 <i>-ia</i> , 丁 <sup>b</sup> <i>-t-ia</i>	
	下 <i>-ha</i> , 下是 <i>-hai</i> , 多羅 <i>-tara</i>	
Admiratory:	叱多 <i>-sta</i>	
Quotative:	古 <i>-ku</i>	
Imperative:	羅 <i>-ra</i> , 羅良 <i>-raa</i>	
Request:	賜立 <i>-sa/-siä</i>	
<i>b. Converbs:</i>		
Infinitive:	良 <i>-a</i> , 可 <sup>a</sup> <i>k-a</i> , 也 <sup>a</sup> <i>i-a</i>	
Means:	惡只 <i>-ak</i>	
Conjunctive:	古/遣 <i>-ku</i>	
Means:	古只/遣只 <i>-kuk</i> , 古音 <i>-kom</i>	
Durative:	音 <i>-m</i>	
Causal:	米 <i>-m-oi</i>	
Simultaneous:	跡 <i>-miä</i>	跡 <i>-miä</i>
Adversative:	乃 <i>-na</i>	
Causal:	羅 <i>-ra</i>	
Parallel:		哉 <i>-čoi</i>
Volitional:		欲 <i>-kua</i>
Means:		尔 <i>-käm</i>
Pre-Quotative:		矣 <i>-toi</i>
<i>c. Nominalizers:</i>		
Established fact:	隱 <i>-n</i>	乎 <i>u-n</i> , 者 <i>-(š/š)n</i>
Prospective:	尸 <i>-l</i> , 乎尸 <i>u-l</i>	

say-HON-NMR fact-NOM ‘that the people are called foolish children’ (H-iii) illustrates quotative *-ko*.

A few examples of converbs are: 折叱可 *käsk-a* pick-INF ‘pick and’ (H-ii), 執音乎手 *čap-om-ho-n sön* hold-DUR-AUX-NMR hand ‘the hand holding it’ (H-ii), 此矣有阿米 here-LOC exist-CONF-CAUSE ‘because X is here’ (H-xi), 望阿乃 look.at-CONF-ADV ‘although I stare at it’ (H-xiii), 二于萬隱 吾羅 two lack-NMR I[-COP]-CAUSAL ‘because it is I, who hasn’t even two of them’ (H-vii), 若大小便爲哉…若食喫哉爲者 if defecation.urination-AUX-TWO . . . if eat.drink-PARA do-TOP ‘if X defecates and urinates or eats and drinks’ (MT-iii).

Nominalizer of an established fact, *-n*, most commonly modifies a noun, e.g 去隱 春 pass-NMR spring ‘the spring which has passed’ (H-i), but also functions like MK *-ni*: 白乎隱 花良 汝隱 say(HUM)-INTENT-NMR flower-VOC you-TOP ‘I say to you, O you flowers’ (H-x). It also occurs on the scope of the negative *anti* in nominal sentences. In idu there is no *ch’aja* that expresses the nominalizer by itself, and so we have to reinstate it according to context. The prospective nominalizer *-l* expressing factual or future meaning is mostly used to modify nouns, but in for example 愛尸 知古如 love-NMR know-ASSERT-DEC ‘they should know that he loves them’ (H-iii) it is the object of a verb.

TABLE 3.6 MID OLD KOREAN AUXILIARY ROOTS

	Hyangch'al	Idu
Honorific:	賜/史 <i>-sɔ</i>	賜 <i>-sɔ</i>
Intentional:	烏/乎/屋 <i>-u</i>	乎 <i>-u</i>
Confirmatory:	阿/惡 <i>-a</i> , 於 <i>-ä</i> , 邪 <i>-ia</i>	去 <i>-ä</i>
Assertive:	古/遣 <i>-ko</i>	
Essential:	音叱 <i>-ms</i>	
Prospective:	里/理 <i>-ri</i>	
Past tense:	如 <i>-tä</i> , 頓 <i>-tä-n</i>	
Present tense:	奴 <i>-nu</i>	飛 <i>-nɔ</i>
Honorific+Confirmatory:	立 <i>-siä</i>	
Emotive:	省 <i>-su</i>	
Volitive:	將來 <i>-riä</i>	

(iii) *Auxiliary roots* (Table 3.6). I use ‘auxiliary root’ for affixes that when attached to verb roots form new roots. They have been called ‘pre-final suffixes’, but this is not as appropriate in OK because they are also used as enders, and ‘pre-final suffixes’ are by nature uncharacteristic of agglutinative languages. In previous research the causative and passive have been regarded as auxiliary roots, but here I only deal with non-derivational forms.

Some view that the intentional mood *-u* expresses ‘the speaker’s psychological proximity’ (Kōno 1996). *-a/-ä/-ia* are used as the confirmatory mood; in idu there are examples of 去 *-kä* used word-finally: 法界有情皆佛道中到內去誓內 universe sentient.beings all Buddhism-LOC reach-CONSID-CONF (*nirɔ-a-kä*) swear-CONSID ‘he swears that all sentient beings of the universe will certainly arrive at Buddhism’ (MT-ix). This illustrates the independence of roots. *-ku* combines the speaker’s desire in the confirmatory mood: 白遣賜立 say(HUM)-ASSERT-HON-IMP ‘make sure you tell [Amitābha]’ (H-ix). Essential *-ms* expresses the sense ‘that is how it should be’. *-ri* expresses factual or conjectural meaning, and intention when the speaker is the subject: 去賜里遣 go-HON-PROSP-Q ‘[Moon,] will you go?’ (H-ix), 獻乎理音如 give-INTENT-PROSP-ESSEN-DEC (above). *-nu/-nɔ* and *-tä* express present and past tense respectively: 去奴隱處 go-PRES-NMR place ‘the place X goes to’ (H-xi), 仰頓隱面 look.at-PAST-NMR face ‘the face I used to look at’ (H-xiii). *-siä* is a fusion of honorific *-sɔ* and confirmatory *-ä*: 陪立羅良 accompany-HON-CONF-IMP ‘please accompany him’ (H-x).

(iv) *Derivative suffixes* (Table 3.7). Verb noun *-m* continues into later Korean: 誓音 swear-VN ‘oath’ (H-ix) > MK *tati-m* > **taci-m**, 岳音 *urɔ-m* climb-VN ‘mountain’ (H-xii). The adverbializer *-ül/-u* in *hyangga* and idu differs from MK in that it can have an argument: 白雲音 逐于 white cloud .m follow-ADV ‘following white clouds’ <H-iv>.

(v) *Auxiliary verbs* (Table 3.8). OK auxiliary verbs often link directly to verb roots without a linking suffix. *-hɔ* is an auxiliary verb ‘do’ that covers all verbals: 狂尸恨 foolish.l-AUX-NMR ‘foolish’ (H-iii). *-hɔ* is widely added to Chinese verbs to Koreanize them, e.g. Chinese 供養 ‘perform a memorial service’ > 供養爲 perform.memorial.service-AUX[-NMR] (MT-iii). Considerate 內 *-a* relates to performing an action towards someone senior, and accepts the doer’s action as appropriate, and came also to be used for

**TABLE 3.7 MID OLD KOREAN DERIVATIVE SUFFIXES**

	Hyangch'al	Idu
Verb noun:	音 <i>-m</i>	
Adverbializer:	于 <i>-ü</i> , 乎 <i>-u</i>	亏 (于) <i>-ü</i>

**TABLE 3.8 MID OLD KOREAN AUXILIARY VERBS**

	Hyangch'al	Idu
Auxiliary:	爲 <i>-hɔ</i> , 恨 <sup>c</sup> <i>-hɔ-n</i> , 好 <sup>c</sup> <i>-h-u</i>	爲 <i>-hɔ</i>
Considerate (Humble):	内 <i>-a</i>	内 <i>-a</i>
Superhonorific:	教 <i>-kiso</i> (只 <i>-ki</i> )	
Factual:	支/支 <i>-ti</i> , 多支 (支) <i>-tati</i>	
Causative:		令只 <i>-siki*</i>
Humble:		白 <i>-sɔrp</i>
Continuative:		在 <i>-kiä</i>

\*: There is also the view that this was read as the causative of *hɔ*, *hɔ-ki* > *hɔi*

humble mood. The source verb 内 *a* means ‘be good, appropriate’. *-kiso* is an honorific auxiliary, which developed from the meaning of a command issued by someone senior or ‘instruction’ so was much more honorific than *-sɔ*. *-ti* is assumed to derive from *tɔ* ‘fact’ and *-i*, meaning ‘to be certainly the case’. Humble *-sɔrp* appears in (MI-iii) from the 860s and began to take over the humble function of *-a*. 在 *-kiä* < *kiä* ‘exist; put’ expresses progressive and resultative. In the MK period *is-* ‘exist’ assumes this function, and *-kiä* now remains fossilized in honorific **kyesi-** ‘exist’ only.

(vi) *Copula*. The copula is mostly 是 *-i*, but in Silla *hyangga* there are also examples of 以 *-i*. In idu 是 *-i* was used with largely the same function as in Modern Korean, but differed in that it could combine with the continuative: 内 物是在之 inside thing-COP-CONT-DEC ‘it is something remaining inside’ (MT-vii). *-i* was often omitted in writing.

(vii) *Quasi-grammatical forms* (Table 3.9). There are forms that perform an intermediate function between content words and grammatical forms. They are grammaticalized content words.

**TABLE 3.9 MID OLD KOREAN QUASI-GRAMMATICAL FORMS**

	Hyangch'al	Idu
Factual:	等/冬 <i>tɔ</i> , 知 <sup>c</sup> <i>t-i</i>	等 <i>tɔ</i>
‘across (to)’:	念丁 <i>nämtiä</i>	
‘when, after’:	第 <i>tiäi</i>	
‘instruct’?:	只 <i>ki</i>	
‘make into’:		爲 <i>sam</i>
‘starting from’:		初/元 <i>pirɔs</i>
‘but’:		而 <i>mari</i>

The bound noun 等 *to* of the factual mood was widely used with the meaning of ‘certainly the case’, which is the meaning of declarative *-ta. nämtiä* (cf. MK *nem-* > **nem-** ‘go over’) occurred in 西方念丁 去賜里遣 west-across go-HON-PROSP-Q (H-ix), Chinese loan 第 *tiäi* ‘in sequence’ grammaticalized to ‘when, after’. The verb *ki*, which since vanished, is assumed to have meant ‘instruct’ from the fact that it was used in *-kisə*, but at present it is difficult to be accurate. *sam* originally meant ‘make’, but was grammaticalized as ‘for the purpose of’. *pirös* was a verb root and expressed the meaning of ‘begin’ in constructions such as ‘beginning from X till Y’: 二月 十二日 元 四月 十三日 此 間中 了治內之 2nd.month 12th.day begin 4th.month 13th.day this period-LOC complete-CONSID-DEC ‘they completed [the repairs] in the period from the 12th of the 2nd month to the 13th of the 4th’ (MT-viii).

(viii) *Negation*. As in LOK nominal negation and verbal negation differed: 不喩/安支 *anti* was the nominal negator. In 不喩 慙盼伊賜等 NEG be.ashamed-HON-NMR fact-TOP ‘if you do not feel ashamed’ (H-ii), *anti* negates the nominalizer *n*. Preverbal 不冬 *antər* was the verbal negator and 毛冬 *mutər* the potential negator.

(ix) *Conjunctions*. Converbs function to link clauses in Korean, and it is thought that the use of conjunctions in *hyangga* as a part of speech was due to Chinese influence: 但 *tamən* ‘however’ and 唯只 *očik* ‘only’ (H-xiv).

### 3.5.3 Lexis

Because ch’aja writing used semantograms to express the content part of a word, it usually does not show us its pronunciation. Just occasionally a word stem is written in phonograms, e.g. 阿孩 *ahəi* ‘child’, 於冬是 *ätəri* ‘where’, 爾處米 *kəmčhəmci* ‘because X stops’, 是史 *is* ‘exist’, 毛冬 *mutər* ‘cannot’, deciphered based on context and corroboration from other materials. However, it is extremely difficult to decipher words written only in phonograms. Words such as 阿冬音, 乃叱好支, 窟理叱 and 乃乎尸 remain hard to decipher.

It is easy to understand when the ch’aja writing is of the structure semantogram + phonogram, common in *hyangga*, and we can work out a word’s form based on the final phonogram and on its MK form. For example, because of the addition of phonogram 音 *m* to 心 ‘heart’, we can work out that 心音 is the OK cognate *məsəm* of MK *mózóm* ‘heart’. The following are OK words identified in this way.

- (16) a. Nouns: 雲音 *kürəm* ‘cloud’ (H-iv), 夜音 *pam* ‘night’ (H-i), 誓音 *tatim* ‘oath’ (H-ix), 憂音 *sirəm* ‘worry’ (H-i), 道尸 *kil* ‘road’ (H-i), 二尸 *tüpəl* ‘two’ (H-iv), 秋察尸 *kasəl* ‘autumn’ (H-xiii), 兒史 *čs* ‘looks’ (H-iv), 栢史 *čas* ‘pine’ (H-xiii), 母史 *äs* ‘mother’ (H-iii), 際叱 *kəs* ‘edge’ (H-iv), 兵物叱 *piängkas* ‘weapon’ (H-xiv), 物叱 *kas* ‘thing’ (H-xiii), 枝次 *kač* ‘branch’ (H-iv), 蓬次 *tapč* ‘mugwort’ (H-i), 一等 *hətən* ‘one’ (H-xi), 千隱 *čmən* ‘1,000’ (H-iv), 紫布 *tərpəi* ‘purple, azalea’ (H-ii), 岩乎 *pahə* ‘rock’ (H-ii), 國惡 *nara* ‘country’ (H-iii), 沙矣 *mərəi* ‘sand’ (H-xiii), 風未 *pərəm-ai* ‘wind-LOC’ (H-xi), 前乃 *čän-a* ‘front-LOC’ (H-ix)
- b. Verbs: 有叱 (H-i) *is* ‘exist’, 折叱可 *käsk-a* ‘pluck-INF’ (H-ii), 修叱如良 *task-a* ‘wash/polish-INF’ (H-viii), 待是 *kitəri* ‘wait’ (H-xi), 持以 *tini* ‘hold’ (H-iv), 使以 *pəri* ‘use’ (H-x), 遣知 *kiti* ‘leave’ (H-iv), 慕理尸 *kəri-l* ‘long.for-NMR’

(H-i), 慙昃伊 *püsköri* ‘be shy’ (H-ii), 改衣 *kätöi* ‘mend/cure’ (H-xiii), 望良 *pöra* ‘desire’ (H-xii), 哀反 *sörpö-n* ‘sad-NMR’ (H-viii), 白反 *sörpö-n* ‘say(HUM)-NMR’ (H-xii), 爾處 *kümčhiä* ‘stop’ (H-iv), 集刀 *mätö* ‘gather’ (H-ix), 直等 隱 *kätö-n* ‘straight-NMR’ (H-x)

The following words had a word-medial *r*, which fell out by MK.

- (17) 川理 *nari* ‘stream’ (H-iv) > MK *nay* > **nay**  
 倭理 *iäri* ‘comet’ (H-xii) > MK *yey*  
 世理 *nüri* ‘world’ (H-xii) > MK *nwi*, but **nuri**  
 舊理 *niäri* ‘old times’ (H-xii) > MK *nyey* > **yey**  
 邀里白 *muri-sörp* ‘to worship’ (P-i) > MK *moy-zöp-* > **mosi-**

Separate from these are examples where syllables or vowels later dropped out. MK *töl* ‘moon’ appears in *hyangga* (H-iv) as 月羅理 *törar-i* moon-NOM, the middle *ra* dropping out in MK. MK *mom* ‘body’ appears in LOK (P-i) as 身萬隱 *muma-n* body-TOP, the word-final *a* falling out in MK. On the other hand, 星利 *piäli* ‘star’ (H-xii) is the same as MK *pjel* ‘star’, the word-final *i* dropping out. Added final phonograms appear in MOK idu only in 令只 *siki*. Added final phonograms are useful in working out OK word forms, but it is still hard to work out the forms of 掌音 ‘palm of hand’ and 花判 ‘flower’ (H-iv) despite the final phonograms, as there are no cognates in MK.

Loanwords from Chinese entered by various routes, but central was importation via Confucian and Buddhist scriptures. The increase in Chinese words was so great that King Kyöngdök [Kyengetek] tried changing the names of the provinces, districts and counties of Silla to Chinese names in the mid-eighth century. Amongst Chinese loanwords are made-in-Korea Sino-Korean words. Characteristic of them is the arrangement of Chinese characters in Korean word order. 楮皮脫 <paper.mulberry + bark + peel> ‘peeler of paper mulberry bark’, 脫皮練 <peel + bark + treat> ‘person who treats the peeled bark’, 紙作人 <paper + make + person> ‘person who makes the paper’, and 苦離樂得 <suffering + avoid + joy + obtain> ‘to avoid suffering and obtain joy’ are examples, but they are coined according to the earliest idu grammar which arranges ümdokcha in Korean word order. Administrative terms such as 古有人 ‘old resident’, 烟受有畚 ‘paddy field allotted to a category of household’ are used in (MT-iv), but these too are Sino-Korean coinages with the characters arranged in Korean word order.

There are even Chinese characters invented in Korea. The 畚 ‘paddy field’ above was formed by combining 水 <water> and 田 <field>, and is a made-in-Silla character. 太 ‘bean’ is also thought to have been created at this period, writing 大 <big> and 豆 <bean> in sequence vertically, changing 豆 to grass script and then reducing it to a dot.

### 3.6 LATE OLD KOREAN

We can characterize LOK texts by the fact that they have detailed t’o writing. This means that we can describe this period’s language in more detail. T’o became detailed in both idu and *hyangch’al*, and we confirmed the large number of t’o anew through the discovery of interpretive *kugyöl*. Not only was the t’o of interpretive *kugyöl* (hereafter ‘*kugyöl*’ for short) written relatively complete, we can analyze the linguistic facts of it by comparison with the contents of the Chinese text, which allows us to describe relatively accurately the nature of OK. The *Kyerim yusa* and *Hyangyak kugüppang* are also important materials regarding LOK.

### 3.6.1 Phonology

The LOK phoneme system is identified by comparing the sounds of the Song dynasty Chinese characters in which Korean words are transcribed in *Kyerim yusa* with MK (Kang 1980). Plain consonants are *p t ĉ k* as in MOK; aspirates *ph th ĉh kh*, which still had low functional load; fricatives *s h* and now also *z*; nasals *m n ng*. I regard liquids to have distinguished *r* and *l*. *z* had a low functional load, occurring only between voiced sounds. MK *β* still did not exist. 只 *k* 隱 *n* 乙 *r* 尸 *l* 音 *m* 邑 *p* 叱 *s* were used as syllable codas. *ng* was used as a coda in Sino-Korean, but as it had no letter of its own this cannot be confirmed. There appears to have been a seven-vowel system:

<i>i</i>	<i>i</i> (= MK <i>u</i> )	<i>u</i> (= MK <i>wu</i> )
	<i>ə</i> (= MK <i>e</i> )	
	<i>v</i> (= MK <i>ó</i> )	<i>o</i> (= MK <i>o</i> )
<i>a</i>		

### 3.6.2 Grammar

In Tables 3.10 to 3.14 I present the grammatical forms of LOK *hyangga* (the eleven *Pohyŏn sibwŏn-ga*, plus the *Sŏdong-yo* and *Ch'ŏyong-ga*), idu and kugyŏl.

(i) *Noun-following particles* (Table 3.10). The nominative had *-i* and *-ik*. *-ik* is used once in *Kyunyŏjŏn*, i.e. 身靡只 *mom-ik* body-NOM (P-viii), and is rarely used even in kugyŏl. We may hypothesize a development 亦 *-iək* > 弋 *-ik* > 伊 *-i*. The genitive distinguishes sentient and non-sentient possessors, but there are exceptions. We note in kugyŏl, apart from *-ii* and *-s*, the use of *-l* as an honorific genitive, especially but not only with Buddhist deities: 菩薩<sup>尸</sup>法<sup>レ</sup> bodhisattva-GEN law-ACC ‘the law of the bodhisattva’ (K-ii). Sometimes in LOK, the accusative was used instead of the locative, and the Chinese locative preposition 於 was interpreted as the accusative: [於]法<sup>レ</sup>歸<sup>ト</sup> *law-ACC converting-AUX-NMR-INST* ‘by converting to Buddhism’ (K-iii). The *-a* locative was used generally, while the *-ii* tended to be used for sentient beings. Combining them gave *-a(k)ii* > MK *-ay/-ey*. There are examples of the locative used as a dative. *-iəkii* often followed the characters 時 ‘time’ and 第 ‘sequence’. The semantic range of the instrumental was broad, e.g. requisite, source, instrument, means or cause. The connective (kugyŏl only) is normally *-iə . . . -iə*, rarely *-koa . . . -koa*. *-koa . . . -koa* is linked to a following noun with *-s*, and *-iə . . . -iə* with *hw*: 樂<sup>ト</sup>利<sup>ト</sup>衰<sup>ト</sup> *joy-CONN profit-CONN weakness-CONN do-NMR* ‘being joyful, profiting and being weak’ (K-iii).

In idu *t-a-n* is a fusion of the bound noun *tə* + locative + topic (as also is kugyŏl *t-akii-n*), and *-ra-n* is accusative *-r* + *-a* (? locative) + topic *-n*. The topic particle was also used for the conditional as in EOK. *-sa* was emphatic, but also conjunctively for Chinese 乃 ‘then, indeed’, attached to nouns, locative particles, adverbs, and honorific auxiliary roots as *i-si-sa* COP-HON-EMPH. *-mata* ‘every’ developed from an adverb meaning ‘universally’ examples of the adverb occurring in kugyŏl. Delimitative *pəs* occurs in kugyŏl, but rarely. Distributive *-akim* and alternative *-ina* were also rare. *-iki* (idu only) expressed a temporal endpoint.

(ii) *Verb enders* (Table 3.11). LOK verb roots retained a high degree of independence. In *Kyerim yusa*, the 鋪 *po* of “讀書曰乞鋪” ‘reading books is called *kir po*’ is the root of MK *po-* ‘to see’, the 索 *so* of “射曰活索” ‘shooting is called *hoar so*’ is the root of

TABLE 3.10 LATE OLD KOREAN CASE AND DISCOURSE PARTICLES

		Hyangch'al	Idu	Kugyōl
<i>a. Case particles:</i>				
Nominative:		伊 <i>-i</i> , <i>-ik</i>	亦 <i>iə k</i> , 弋只 <i>-ik</i>	ㅁ <i>-i</i> , ㅁ <i>-ik</i>
Genitive:	Sentient:	衣 <i>-ii</i>	矣 <i>-ii</i>	ㄹ <i>-ii</i> , ㅍ <i>-l</i>
	Non-sentient:	叱 <i>-s</i>	之 <i>-s</i>	ㅅ <i>-s</i> , ㅍ <i>-l</i>
Accusative:		乙/盼 <i>-ir</i>	乙 <i>-ir</i>	ㅅ <i>-ir</i>
Locative:	<i>a</i> -Type:	良/阿/惡 <i>-a</i>	良/ㄹ 中 <i>-akii</i> , 亦中 <i>-iəkii</i>	ㅅ <i>-a</i> , ㅅ + <i>-akii</i> , ㅅ + <i>-iəkii</i>
	<i>ii</i> -Type:	衣/矣 <i>-ii</i> , 中 <i>-kii</i>	中 <i>-kii</i>	ㄹ <i>-ii</i> , ㄹ + <i>-iikii</i> , ㅅ + <i>-kii</i>
	Combined:	良衣, 阿希 <i>-aii</i> , 惡中 <i>akii</i>		
Instrumental:		留/乙留 <i>-lo</i>	以/乙以 <i>-lo</i>	ㅅㅅㅅ <i>-lo</i>
Connective:				ㅅ <i>-koa</i> ...ㅅ <i>-koa</i> , ㅅ <i>-iə</i> ...ㅅ <i>-iə</i>
Vocative:	Plain:	良 <i>-a</i>		ㅅ <i>-a</i>
	Exclam.:	也 <i>-ia</i> , 亦 <i>-iə</i>		ㅅ <i>-iə</i>
	Honorific:	下 <i>-ha</i>		下 <i>-ha</i>
Complex:		阿叱 <i>-a-s</i> , 惡之叱 <i>-avi-s</i> , 留叱 <i>-lo-s</i>		ㅅ <i>-koa</i> ...ㅅ <i>-koa-s</i>
<i>b. Discourse particles:</i>				
Topic:		隱/焉 <i>-(i/v)n</i> , 恨 <sup>a</sup> <i>hv-n</i> , 等 <sup>b</sup> <i>tə-n</i>	段 <sup>c</sup> <i>t-a-n</i> , 乙良 <sup>c</sup> <i>r-a-n</i>	ㄱ <i>-n</i> , ㅁ + ㄱ <sup>c</sup> <i>-t-akii-n</i>
Emphatic:		沙 <i>-sa</i>	沙 <i>-sa</i>	ㅅ <i>-sa</i>
Inclusive: 'every':		置 <i>-tu</i> , 刀 <i>-to</i> 每如 <i>-mata</i> , 馬洛 <i>-marak</i>		ㄴ <i>-to</i> , 斗 <i>-tu</i> ㅅㅁ <i>-mata</i>
Extent:				ㅅ <i>-ma</i>
Limiting:				火(七) <i>-pvs</i>
Terminative:			已只 <i>-iki</i>	
Distributive:				(ㅅ) 亦 <i>-akim</i>
Alternative:				乃 <i>-na</i>

MK *sso-/pso-* ‘to shoot’, the 耻 *ti* of “雪下曰嫩耻” ‘snowing is called *nun ti*’ is the root of MK *ci-* ‘to fall’, and the 打里 *tari* of “煖酒曰蘇孛打里” ‘heating rice-wine is called *sopər tari*’ is the root of MK *tali-* ‘to boil, infuse’. Such independence led to great numbers of root + root compounds, and agglutination especially with auxiliary verbs in threes, fours or more. Verb enders, however, could not be used without a verb root.

Declarative *t-iə* was a contraction of the bound noun *tə* with declarative 2 *-iə* and was emphatic. *-iə* displays wider functions in *kugyōl*, including conjunctive function, also seen with *-čīiəi*; one could regard each clause listed as its own respective sentence. In *hyangga*, *-čīiəi* is used as a hortative: 禮 爲白齊 greeting do-HUM-DEC2 ‘let’s greet’ (P-i). Plain imperative *-ra* is seen in *Kyerim yusa* in “來曰烏囉” ‘come! is called *o-ra*’. Lamentary 良羅 *-ara* and 兮 *-hiəi* express the speaker’s lament, e.g. 脚烏伊 四是良羅 leg-NOM four-COP-LAMENT ‘the legs are four in number’ (H-v), while exclamatory *-siə* seems more objective.

TABLE 3.11 LATE OLD KOREAN VERB ENDERS

	Hyangch'al	Idu	Kugyōl
<i>a. Sentence enders:</i>			
Declarative 1:	如/多 <i>-ta</i>	之/如 <i>-ta</i>	ㅅ <i>-ta</i>
Declarative 2:	也/耶 <i>-ia</i>	亦 <i>-ia</i> , 丁 <sup>b</sup> <i>-t-ia</i> , 齊 <i>-čiai</i>	ㅅ/ㄷ <i>-ia</i> , ㅅ <sup>b</sup> <i>-t-ia</i> , 齊 <i>-čiai</i>
Lamentory:	良羅 <i>-ara</i> , 兮 <i>-hiəi</i>		
Exclamatory:	立/舌 <i>-siə</i>		ㅅ <i>-siə</i>
Interrogative (polar):	去 <i>-kə</i>		ㅅ <i>-a</i>
Interrogative (wh):	古 <i>-ko</i> , 焉古 <i>-ən.ko</i> , 叱過 <i>-s.ko.a</i>		ㅅ <i>-ko</i>
Interrogative (rhetor.):	下呂/下里 <i>-hari</i>		
Hortative:	去良 <i>-kəa</i> , 齊/制 <i>-čiai</i>		ㅅ <i>-siə</i>
Imperative:	Plain: ( <i>Kyerim yusa</i> ) 良 <i>-ra</i>		ㅅ <i>-esia</i>
	Honorific:	少時/受勢 <i>siosia</i>	
<i>b. Converbs:</i>			
Infinitive:	良 <i>-a</i> , 可 <sup>a</sup> <i>k-a</i>	良/ㅅ <i>-a</i>	ㅅ <i>-a</i> , 下 <i>-ha</i>
	Means: 良只/惡只 <i>-ak</i> , 良尔 <i>-akim</i>	良尔 <i>-akim</i>	ㅅㅅ <i>-ak</i> , ㅅ尔 <i>-akim</i>
Conjunctive:	古/遣 <i>-ko</i> , 昆 <i>-ko-n</i>	遣 <i>-ko</i>	ㅅ <i>-ko</i>
Simultaneous:	脉 <i>-miə</i>	彌/脉/ <i>-miə</i>	ㅅ' <i>-miə</i>
Adversative:	乃/奈 <i>-na</i>		乃 <i>-na</i>
Concessive:	而也 <i>-mariə</i> 馬於隱 <i>-marən</i>		
Interrupted:	如可 <i>-taka</i>		
Discontinued:		乙 <i>-r</i>	ㅅ <i>-r</i> , 斤 <i>-nər</i>
Resultative:	色只/所只 <i>-torok</i>		巴 <i>-torok</i>
Volitive:		欲 <i>-koa</i>	人 <i>-koa</i>
Pre-quotative:		乎矣 <i>-oltwi</i>	ㅅㅅ <i>-ntwi</i> , ㅅㅅ <i>-rtwi</i>
Causal 1:	米 <i>-mai</i>		
Causal 2:		等以 <i>-təlo</i> , 等用 <i>-təlsia</i>	入 ㅅㅅ, <i>-təlo</i> , ㅅ <i>-ra</i>
<i>c. Nominalizers:</i>			
Established fact:	隱/焉 <i>-n</i> , 根 <sup>a</sup> <i>ko-n</i> , 斤 <sup>a</sup> <i>k-ən</i> , 反 <sup>a</sup> <i>p-ən</i> , 昆 <sup>a</sup> <i>ko-n</i> , 孫 <sup>a</sup> <i>so-n</i> , 仁 <sup>a</sup> <i>i-n</i> , 寸 <sup>a</sup> <i>cč-in</i>	乎 <sup>c</sup> <i>-o-n</i>	ㅅ <i>-n</i>
Prospective:	尸 <i>-l</i> , 乙 <i>-r</i>		尸 <i>-l</i> , ㅅ <i>-r</i>

*-a* expresses the premise for another action. In LOK idu abbreviated kugyōl was sometimes used in writing t'ŏ, shown by 僧人乙合 ㅅ monk layman-ACC bring.together-INF 'bringing together monks and laymen', hinting to us that kugyōl letters were widely used for writing Korean. Concessive 而也 *-mariə* and 馬於隱 *-marən* are not thought to have differed functionally. Interrupted 如可 *-taka* had the same function as modern **-taka**: 夜入伊遊行如可 night enter-ADV gallivant-INTERRUPT 'as I was gallivanting into the night, ...' (H-v). Discontinued *-r/-nər* were used when a situation was suspended and a new situation started following it, e.g. example (3) 遷世 爲去在乙 pass.away do-PAST-CONT-NMR-DISC 'even though he passed away'. *-ol/-r/-n + tɁ-i*

TABLE 3.12 LATE OLD KOREAN AUXILIARY ROOTS

	Hyangch'al	Idu	Kugyöl
Honorific:	賜 <i>-sɐ</i>	賜 <i>-sɐ</i>	𠄎 <i>-sɐ</i> , 二 <i>-si</i>
Intentional:	乎/屋 <i>-o</i> (留 <i>-ru</i> )	乎 <i>-o</i>	ㄱ/ㄴ <i>-o</i>
Confirmatory 1:	去 <i>-kə</i>	去 <i>-kə</i>	ㄱ <i>-kə</i>
Confirmatory 2:	良/於 <i>-a</i> , 焉 <sup>◦</sup> <i>-ə-n</i> , 下 <i>-ha</i>	良/ㄱ <i>-a</i>	ㄱ <i>-a</i>
Convictive:	遣/遣只 <i>-ko</i>		ㄱ <i>-ko</i>
Essential:	音叱 <i>-ms</i>		ㄱ 七 <i>-ms</i>
Prospective:	理 <i>-ri</i>		禾/禾 <i>-ri</i> , ㄱ <i>-li</i>
Present:	臥 <i>-nu</i> , 飛 <i>-nɐ</i>	臥 <i>-nu</i>	ト <i>-nu</i> , 又 <i>-no</i> , ㄱ <i>-nɐ</i>
Past:	如 <i>-tə</i>	如 <i>-tə</i>	ㄱ <i>-tə</i>
Admiratory:	叱 <i>-s</i> , 孫 <sup>◦</sup> <i>-s-o-n</i>	叱 <i>-s</i>	七 <i>-s</i>
Volitional?:	將來 <i>-riə</i>		呂 ?

introduces a quotation like MK *-ótój*, which dropped the nominalizer *l/n* within the form. Causal 2 *-tɐlo* consists of the bound noun *tɐ* plus instrumental *-lo*; *-tɐlsia* substitutes accusative *-l* + *si-a* ‘using’ for *-lo*, developing from Chinese text interpretive style.

The nominalizer for an established fact was *-n*, used with various *ch’aja* reflecting fusion with the final sound of (auxiliary) verb roots in *hyangga*: 見根 *po-ko-n* see-CONVIC-NMR, 明斤 *pelk-en* bright-NMR, 迷反 *ip-in* be.confused-NMR, 惡寸 *kuč-in* bad-NMR. There was no distinct letter in idu for *-n*, but 乎 *-o-n* was used for intentional *-o* + *-n*. The functions of nominalizers *-n* and *-l* were diverse in *kugyöl*, acting as nominalizations with particles (e.g. *hɐ-n-en* do-NMR-TOP) and the copula (e.g. *hɐ-n-i-miə* do-NMR-COP-LINK) as well as noun modifiers.

(iii) *Auxiliary roots* (Table 3.12). Honorific *-sɐ* was used in *kugyöl* till (K-ii), but from (K-iii) it changed to *-si*. In contrast with confirmatory 去 *-kə* which confirmed an objective fact, *-a* is thought to have involved the speaker’s will, and *ə-n* has the nominalizer added. In *kugyöl* *-kə* was not only an auxiliary root but was also widely used word-finally. *-ha* was used in rhetorical *-hari* above, e.g. 不冬 應爲賜下呂 NEG respond-AUX-HON-Q ‘might X not respond?’ (P-vii), and is thought to be honorific. Convictive 遣 *-ko* was a contraction of *-kə* with intentional *-o*. We can tell the function of essential *-ms* because it is often used in *kugyöl* to gloss Chinese 應 ‘should’: ㄱ 應七 應七 應七 *ho-ms-ta* ‘must certainly do’. There are various views on the reading and the function of volitional 將來 *riə*, and *kugyöl* 呂 only appears in (K-iv) and is an unknown form whose function is not clear.

(iv) *Derivative suffixes* (Table 3.13). *tɐr* was used to express plurality: 善陵等沙 good.deed-PL-EMPH ‘good deeds’ (P-v). We can confirm transivitizers 是 *i* and 乎 *o* from idu, e.g. 成是 *ir-i* ‘make’. Deadjectival nominalizer *vi* in 明期 *perk-vi* ‘bright’ (H-v) shows a function close to that of nominalizer *n*. As adverbializers, we find *-u/o* and *-i* (逐好/追于 *čoč-o* ‘following’, 加于 *tə-u* ‘more’, 并以 *apɐl-o* ‘together’, 具 *kač-o* ‘ready’, 夜 入伊 *tir-i* night enter-ADV ‘into the night’, 深以 *kip-i* ‘deeply’, 密只 *kisik-i* ‘secretly’, 普 ㄱ ‘widely’, 常 ㄱ ‘always’), and *-s* which was not used in MK (e.g. 然叱 爲賜隱 that-ADV do-HON-NMR ‘did it like that’ (P-viii)) and is thought to express a mental situation like the *-ehó-* does in MK *phelehóta* ‘be blue’,

TABLE 3.13 LATE OLD KOREAN DERIVATIVE SUFFIXES

	Hyangch'al	Idu	Kugyöl
Plural:	等/冬 <i>-ter</i>		
Transitivizer:	伊 <i>-i-</i>	是 <i>-i-</i> , 乎 <i>-o-</i>	<i>-i-</i> , 〇 <i>-o-</i>
Nominalizer:	期 <i>k-ii</i>		
Adverbalizer:	好 <i>h-o</i> , 以 <i>l-o</i> 伊/以 <i>-i</i> , 只 <i>k-i</i> 叱 <i>-s</i>	于 <i>-u</i> , 以 <i>l-o</i> 兮 <i>-hi</i> 叱 <i>-s</i>	〇 <i>-o</i>    <i>i</i> , 兮 <i>-hi</i> 七 <i>-s</i> ㄱ <i>-miə</i> 矣/支 <i>-ti</i> , 去 <i>-kə</i> , ハ <i>-ki</i>

TABLE 3.14 LATE OLD KOREAN AUXILIARY VERBS

	Hyangch'al	Idu	Kugyöl
Auxiliary:	爲 <i>-hə-</i> , 好 <i>-ho-</i>	爲 <i>-hə-</i>	ㄴ <i>-hə-</i> , ㄴ <i>-ho-</i>
Causative:	海 <i>-hvi-</i>	令是 <i>-hvi-</i>	ㄴ <i>-hvi-</i>
Considerate:	內/良 <i>-a-</i>	內/於 <i>-a-</i>	ㄹ <i>-a-</i>
Humble 1:	白 <i>-sərp-</i>	白 <i>-sərp-</i>	白 <i>-sərp-</i>
Humble 2:		白內 <i>-sərpa-</i>	白ㄹ <i>-sərpa-</i>
Factual:	支 <i>-ti-</i>		支 <i>-ti-</i>
Perfect:	置 <i>-tu-</i>	置 <i>-tu-</i>	
Continuative:		在 <i>-kiə-</i>	ナ <i>-kiə-</i>
Executive:		向 <i>-a-</i>	
Superhonorific:	只 <i>-ki-</i> , 教 <i>-ki.sə-</i>	教 <i>-ki.sə-</i>	ハ <i>-ki-</i> , ㄴ <i>-ki.sə-</i>

*kulehóta* 'be so'. ㄱ *-miə* occurs solely in kugyöl, used in 善 ㄱ 'excellently', 能 ㄱ 'skillfully'. It is the same as the conjunctive suffix *miə* and includes a present-tense meaning. *-ti* was used in e.g. example (6) 能矣 and expressed factual mood. ㄴ is read as *ki* or *k* and was used as in 當 ㄴ *pentə-ki* 'without fail', 今 ㄴ *iət-ki* 'now', 最 ㄴ *anči-k* 'most'.

(v) *Auxiliary verbs* (Table 3.14). *hə* was used just as in MOK. 好/乎 *ho*, a contraction of *hə* and intentional *-o*, does not appear in idu texts, probably because of the limitations of the texts. In kugyöl it is often found attached to Chinese verb roots, e.g. 減 ㄴ *ho-l* 'will decrease' (K-iii). Causative 海 *hvi* of *hyangga* is used as 海伊 *hvi* in 願海伊過 'I want to make wish' (P-xi). This is a combination of auxiliary *hə* and a causative derivative suffix *i*. Considerate 內/良/於 is used in *hyangga* as in 拜內乎隱 'I bow (in reverence)' (P-i), 直體良焉 'I set right' (P-iii), and examples occur in idu. In kugyöl it is in the 白 ㄹ *sərp-a* used in 佛 ㄴ 見 白 ㄹ ㄴ ㄴ Buddha-ACC see-HUM-CONSID-NMR fact-LOC-TOP 'when he saw the Buddha' (K-iii). Humble *sərp* occurs frequently in *hyangga*, idu and kugyöl and was ubiquitous in the tenth century. Idu 白內 *sərpa* was also written 白於 and in kugyöl as 白 ㄹ. This expressed the humble mood strongly by combining two humble auxiliary verbs. Factual 支 *ti* was used in *hyangga* and kugyöl, but it is not known in idu. There are already examples in the twelfth-century (K-iii) of it changing its rules, and in the thirteenth century it was only used as the *hun* of 如, which shows us that its duties had vanished. Perfect aspect 置 *tu* expressed something completed in the past that continues into the present, as in 嫁良置古 'had married,

and’ (H-vi) and 爲白置在教如乙 do-HUM-PERF-CONT-HON[-NMR] fact-ACC ‘the fact that I had done’ (LT-iv). There are no examples in kugyöl. Continuative *kiə* is frequent in idu and kugyöl, but is completely absent in *hyangga*. 向 *-a* was used in idu in 向事 *-a-n ir*, as in 右塔 重修向 事之 段 right.pagoda repair-EXEC[-NMR] matter-GEN fact-LOC-TOP ‘at the time of the repairs to the right-hand pagoda’ (LT-vi). I treated 只 *ki* ‘instruct’ as a quasi-grammatical form in MOK. In LOK it was used in the form 只賜 *ki-sə* in 十方叱 佛体 闕遣只賜立 ten.directions-GEN Buddha know-CONV-HON-IMV ‘please know the Buddhas of all the directions’ (P-iv). This corresponds to kugyöl *kisə/kisi*, and where *hyangga ki-sə* showed greatest respect to Buddhas, kugyöl *kisə* did so to the king. This 只賜 *kisə* corresponds to the idu superhonorific auxiliary *kisi*.

As auxiliary roots and auxiliary verbs had the independence that verb roots had, they were attached without an intermediary suffix. 安爲白置在教 ㄱ ‘once he had enshrined . . .’ is analyzed as five verbs combined in their root form: 安 ‘enshrining’ + auxiliary 爲 *hə* + humble 白 *sərp* + perfect 置 *tu* + continuative 在 *kiə* + superhonorific 教 *kisi* + converb ㅅ *miə*.

(vi) *Quasi-grammatical forms.* 等 *tw* was used as the bound noun of the factual mood in *hyangga* and in idu, and ㄸ *tw* was used consistently in kugyöl. 段 *ta-n*, frequently found in idu, was explained above under discourse particles. The bound nouns 所 *pa* and 事 *ir* were used in the sense of ‘thing, matter’. The 下 *ha* in 吾下於叱古 ‘they are mine’ and 誰支下焉古 ‘whose are they?’ (H-v) is a bound noun used in the sense of ‘possession’. ㅅ *a* is used as a sentence-nominalizing bound noun in 十四王ㅅㅅ 說ㅅㅅ 14.kings-GEN matter-ACC explain-HON-NMR ‘he explains the matter of the 14 Kings’ (旧仁 11) means matter, confirmed also in dot-type kugyöl. 爲 is read in context as *sam*, meaning either ‘for the purpose of’ or ‘because of’, used with locative *-ii* in the purpose sense. In kugyöl *sam* is indicated with 三/一/ㅅㅅ on the characters 爲, 以, 由 or 因 in the Chinese text which express a *causal* relationship, and the use of *sam* in a causal sense is considered grammatical influence from Chinese texts. Very commonly seen in LOK idu is 元 in an ablative sense, read *piris*. 念丁 *nəmtiə*, used in MOK in the sense of ‘across’ space, is used in LOK in the sense of ‘through’ time: 同年 春秋冬 念丁 ‘through spring, autumn and winter of that year’ (LT-v). 仍于 *čičiru* is common in LOK idu meaning ‘in accordance with’.

(vii) *Negation.* The distinction between nominal and verbal sentence patterns of negation is clearly seen in interpretive kugyöl (18): nominal sentence pattern (a) noun + *anti* + copula *i*; nominal sentence pattern (b) verb root + nominalizer *-n* + *anti* + copula *i*; verbal sentence pattern (c) verb root + nominalizer *-l* + *anter* + auxiliary *hə*.

- (18) a. 內 非矣ㅅ, 外 非矣ㅅ  
inside *anti*[-COP]-LINK outside *anti*[-COP]-LINK  
‘it is not the inside, it is not the outside . . .’ (K-ii, 13)
- b. 堅固ㅅㅅ 非矣ㅅ  
firm-AUX-NMR *anti*[-COP]-LINK  
‘is not firm and . . .’ (K-ii, 18)
- c. 捨ㅅㅅ 不冬ㅅㅅㅅ  
abandon-AUX-NMR *anter*-AUX-NMR-FACT  
‘does not abandon’ (K-ii, 14)

Only the verbal pattern is seen in *hyangga*: 不冬 喜好尸 ‘shall we not rejoice?’ (P-v). In *kugyōl* 不ハ *antek* occurs as the potential negator (知尸 不ハノ冬七 過失 ㄱ 知-NMR NEG-AUX-PROSP-GEN fault-COP-LINK ‘being a fault that one cannot know, . . .’ (K-vi, 13)), while in contrast it is 毛冬/毛等 *moter* that was used in *hyangga* (毛冬 長乙隱 ‘[where] we cannot cultivate’ (P-vi)). Potential negator 不得 *motsir* is only known in idu: 成是 不得爲乎 make[-REL] NEG-AUX-INTENT ‘could not bring it to fruition’ (LT-v). It was a calque from Chinese 不得, so *moter* is found in later idu texts too.

### 3.7 CONCLUSION

I personally feel that the description of OK remains problematic. Firstly, OK covers a long period of more than 900 years. Each sub-period has scarce materials, and there is much that we cannot decipher or analyze by relying on discovered texts. There are many divergent views regarding Three Kingdoms toponyms and *hyangga*, and amongst interpretive *kugyōl* there are undecipherable forms too. However, the discovery and research of newly discovered interpretive *kugyōl* and idu texts have provided new views on OK. In this chapter I have outlined based on such achievements what has been researched up to very recently. As I prepare this chapter, there is newly edited data, and I have abstained from including what I cannot yet verify concretely.

I shall briefly conclude with some words about differences between OK and MK. OK nominalizers *-n* and *-l* have more or less vanished as such by MK, remaining as relativizers. The functions of MK nominalizer *-(ó/u)m*, whose origins are obscure, were much reduced compared with OK. The independent nature of verb roots was weakened by MK, and they were used as bound roots before endings; root + root sequences were fossilized in compound verbs. Auxiliary verb 白 *sərp* changed into the auxiliary root *-sóp/cóp/zóp-*. 在 *kiə* ceded its function to *is-* and vanished. Factual mood 支 *-ti* vanished in the twelfth century, and considerate 內 *-a* and superhonorific 教 *-kisi* also vanished by the fifteenth century. The nominal vs verbal sentence distinction of negation vanished, and the *-ti* and *-ter* of *anti* and *anter* developed into adverbializers on the verb and the functional difference between them vanished. The OK conditional was expressed through the topic particle, but *-myen* and *-ketun* developed as a new grammatical category in MK. Sentence enders *-čičai* and *-iə* and auxiliary root *-ms* virtually vanished and were used only in an extremely restricted way. *-oltvī* and *-ontvī* developed into productive *-otvī* (LMK *-otój*) through loss of the component nominalizer *-l/-n*.

### NOTE

- \* This chapter was translated from the original Korean by Nicolas Tranter, with great help from a Japanese version prepared by Yun-jin Nam. Metalinguistic – and metagraphic – issues in making this chapter accessible to readers with no knowledge of Korean or of Chinese characters meant a number of concessions, particularly a severe reduction in the number of examples from primary texts.

### REFERENCES

- Kang, Sinhang (1980) *Kyeylim Yusa Kolye Pangen Yenkwu* [A study of Koryō dialect in *Kyerim yusa*], Seoul: Sungkyunkwan University Press.

Kōno, Rokurō (ed.) (1996) *Gengogaku Daijiten 6-kan: Jutsugo-hen* [Encyclopedia of linguistics: Predicates], Tokyo: Sanseido.  
 Lee, Ki-Moon (1972) *Kwukesa Kaysel* [Introduction to the history of the Korean Language], Seoul: Thap Chwulphansa.  
 Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge: Cambridge University Press.

**FURTHER READING (ALL IN KOREAN)**

權仁瀚 (2006), 武寧王陵 出土 銘文들에 대한 語學的 考察, 『口訣研究』 17, 口訣學會.  
 金永萬 (1997), 석독구결 ‘皆七’, ‘悉’와 고려향찰 ‘頓部七’, ‘盡’의 비교 고찰, 『口訣研究』 2, 口訣學會.  
 金永旭 (2003), 百濟 吏讀에 대하여, 『口訣研究』 11, 口訣學會.  
 金完鎭 (1980), 『鄉歌解讀法研究』, 서울大學校出版部.  
 김완진 (2000), 『향가와 고려가요』, 서울대학교출판부.  
 김지오 (2010), <懺悔業障歌>의 國語學的 解讀, 『口訣研究』 24, 口訣學會.  
 南廣祐編著 (1995), 『古今漢韓字典』, 仁荷大學校出版部.  
 南廣祐編著 (1999), 『古語辭典』, 敎學社.  
 남기삼·고영근 (1985), 『표준국어문법론』, 塔出版社.  
 南豐鉉 (1981), 『借字表記法研究』, 檀大出版部.  
 南豐鉉 (1999), 『國語史를 위한 口訣 研究』, 太學社.  
 南豐鉉 (2000), 『吏讀研究』, 太學社.  
 南豐鉉 (2009), 『古代韓國語研究』, 시간의 물레.  
 南豐鉉 (2010a), 獻花歌의 解讀, 『口訣研究』 24, 口訣學會.  
 南豐鉉 (2010b), 韓國語史 研究에 있어 口訣資料의 寄與에 대하여, 『口訣研究』 25, 口訣學會.  
 南豐鉉 (2011), 古代韓國語의 謙讓法 助動詞 ‘白/□’과 ‘內/아’의 發達, 『口訣研究』 26, 口訣學會.  
 노명호·이승재 (2009), 釋迦塔에서 나온 重修文書의 判讀 및 譯註, 『重修文書』, 국립중앙박물관·대한불교조계종.  
 都守熙 (2005), 『百濟語 語彙 研究』, 제이앤씨.  
 박진호 (1998), 고대 국어 문법, 『국어의 시대별 변천 연구 3』, 국립국어연구원.  
 박진호 (2008), 鄉歌解讀과 國語文法史, 『國語學』 51, 國語學會.  
 박창원 (2002), 『고대국어 음운 (1)』, 태학사.  
 백두현 (1997), 高麗時代 釋讀口訣에 나타난 선어말어미의 계열관계와 통합관계, 『口訣研究』 2, 口訣學會.  
 徐在克 (1975), 『新羅 鄉歌의 語彙 研究』, 啓明大學 韓國學研究所.  
 서종학 (2004), 借字 ‘令是-’攷, 『口訣研究』 12, 口訣學會.  
 梁柱東 (1965), 『增訂 古歌研究』, 一潮閣.  
 李基文 (1981), 『韓國語形成史』, 三星美術文化財團.  
 李基文 (1991), 『國語 語彙史 研究』, 東亞出版社.  
 李崇寧 (1961), 『中世國語文法』, 乙酉文化史.  
 李丞宰 (1992), 『高麗時代의 吏讀』, 國語學叢書 17, 國語學會, 太學社.  
 李丞宰外 (2005-2006), 『角筆口訣의 解讀과 翻譯 1-3』, 태학사.  
 장경준 (2007), 『「瑜伽師地論」 点吐釋讀口訣의 解讀 方法 研究』, 國語學叢書 58, 國語學會, 太學社.  
 장경준 (2009), 점두구결 자료의 문법 형태에 대하여, 『國語學』 56, 國語學會.  
 장윤희 (1998), 석독구결 자료의 감탄법 중결어미, 『口訣研究』 4, 口訣學會.  
 鄭在永 (2003), 百濟의 文字生活, 『口訣研究』 11, 口訣學會.  
 鄭在永 (2004), 鷄林類事의 高麗方言에 나타난 文法形態에 대한 研究, 『口訣研究』 12, 口訣學會.  
 鄭在永 외 (2003), 『韓國 角筆 符號口訣 資料와 日本 訓點 資料 연구』, 태학사.

- 황선엽 (2008), 三國遺事와 均如傳의 鄉札 表記字 비교, 『國語學』 51. 國語學會.
- 황선엽 (2010), 향가의 연결어미 '-아' 표기에 대하여, 『口訣研究』 25, 口訣學會.
- 小林芳規 (2008), 角筆による新羅語加點の華嚴經, 『南都佛教』 91. 南都佛教研究會  
東大寺.
- 小林芳規 (2010), 日本의 오코토點의 起源과 古代 韓國語의 點吐와의 關係, 『口訣  
研究』 25, 口訣學會.

# MIDDLE KOREAN

*Ho-min Sohn*

## 4.1 INTRODUCTION

### 4.1.1 Historical background

Middle Korean (MK) is usually viewed as covering the period from the tenth to the sixteenth century (e.g. K. M. Lee 1976; Y. Choy 1987; K. Kang and Hwang 2003; Ko 2007). That is, MK ranges over the Kolye dynasty period (918–1392) and the first 200 years of the Cosen dynasty.<sup>1</sup>

It is further divided into Early Middle Korean (EMK) and Late Middle Korean (LMK). EMK lasted from the beginning of Kolye until the creation of the Korean alphabet *han'gŭl* (hereafter romanized in the Yale system as *hankul*) in the early Cosen dynasty. LMK was the language represented in the *hankul* texts in the fifteenth and sixteenth centuries, specifically from King Seycong's promulgation of the *hwunmin cengum* (later termed *hankul*) in 1446 to the Japanese invasion in 1592 and the subsequent seven-year Imcin Wars.

Of all the stages of Korean language development, MK, especially LMK, is linguistically the most important, in that with the invention of the Korean alphabet in the fifteenth century detailed aspects of the Korean language were revealed for the first time. EMK data in the Kolye period were recorded in Chinese characters, whereas LMK data in the Cosen period were recorded in *hankul*. The creation of *hankul* made extensive and exact transcriptions of the Korean language possible.

With the founding of the Kolye dynasty (from which the national title *Korea* came) in the early tenth century, the capital was moved from Kyengcwu (in the Kyengsang province) to Kaykyeng (present Kayseng in the southern part of North Korea), which became the new political and cultural center until the capital moved to current Seoul (then called Hanyang) in 1392 with the establishment of the Cosen dynasty. Thus, unlike Old Korean with Silla language data as the backbone, available MK data are assumed to heavily reflect the language of the central area (Kaykyeng and Seoul). Even so, it is safe to say that MK is the direct descendant of the language of unified Silla, of which Kaykyeng and Seoul were a part.

### 4.1.2 Early Middle Korean

Although the situation is slightly better than for Old Korean, EMK is still observed in fragmentary reflexes in existing records written in Chinese characters used as phonograms. Thus, our knowledge of EMK is limited, since linguistic facts can essentially be understood only from Chinese characters by way of internal reconstruction. Such scanty written texts include 11 vernacular poems (*hyangka*) in *Kyunye-cen* (Stories about Abbot Kyunye, by Hyeklyen Ceng, 1075), *Kyeylim yusa* (Things on Korea, by Chinese Song dynasty scholar Sun Mu, 1103–4), *Samkwuk saki* (Historical Record of the Three

Kingdoms, by Kim Pusik, 1145), *Hyangyak kwukuppang* (Herbal Medicine Emergency Treatments, c. 1250), 14 *hyangka* poems in *Samkwuk yusa* (Memorabilia of the Three Kingdoms, by Monk Ilyen, 1285), *Taymyengyul cikhay* (Translation of the Chinese Ming dynasty penal code *Ta ming lii*, 1395), and *Cosenkwan yeke* (A Korean–Chinese Translated Lexicon, 1408). In addition to these books, there are a number of kwukyel (‘oral formulae’) texts, including Buddhist sutras, annotated in the Kolye period. Such texts contain extensive information about the use of grammatical endings (see Nam, this volume).

Of these, *Kyeylim yusa*, *Hyangyak kwukuppang*, and *Cosenkwan yeke* are considered the most important and valuable materials for EMK. *Kyeylim yusa* is a vocabulary list of about 360 Kolye words and phrases written with a total of 385 Chinese characters.<sup>2</sup> The *Hyangyak kwukuppang* is a pharmacological work from 1250 that lists the names of some 180 plants, animals, and minerals used in Korean herbal medicine. *Cosenkwan yeke* is a list of 596 words covering 19 areas including astronomy, geography, time, and flora and fauna. *Kyeylim yusa* is the most representative of these to overview the general linguistic profile of EMK.<sup>3</sup>

Most of the basic Kolye words appearing in *Kyeylim yusa*, which are written with Chinese characters used as phonograms, are reflected in Middle and Contemporary Korean. In the following illustrations, reconstructions of Kolye words from the Chinese phonograms (e.g. *mól* ‘horse’ from the Chinese character 末 *mal*; *onal* ‘today’ from Chinese characters 烏捺 *o-nal*; *hanól* ‘sky’ from 漢捺 *han-nal*; *kelim* ‘picture’ from 乞林 *kel-lim*; *pólam* ‘wind’ from 勃纜 *pól-lam*) are approximate, based on S. Kang (1980), B. An (1985), Chung (2004), P. Nam (2009), and others. The fifteenth-century words are from K. Nam’s (1982) MK dictionary, omitting tone marking. MK is transcribed with the Yale Romanization system.<sup>4</sup>

(1) <i>Kyeylim yusa</i>	15C MK	CK	
<i>pal</i>	<i>pal</i>	<b>pal</b>	foot
<i>iup</i>	<i>ip</i>	<b>ip</b>	mouth
<i>ni</i>	<i>ni</i>	<b>i</b>	tooth
<i>son</i>	<i>son</i>	<b>son</b>	hand
<i>kwui</i>	<i>kwuj</i>	<b>kwi</b>	ear
<i>pói</i>	<i>pój</i>	<b>pay</b>	belly
<i>mwun</i>	<i>mom</i>	<b>mom</b>	body
<i>hanapi</i>	<i>hanapi</i>	<b>halapeci</b>	grandfather
<i>nach</i>	<i>nóch</i>	<b>nach</b>	face
<i>caapi</i>	<i>api</i>	<b>api, apeci</b>	father
<i>ami</i>	<i>emi</i>	<b>emi, emeni</b>	mother
<i>p(o)tal</i>	<i>stól</i>	<b>ttal</b>	daughter
<i>atal</i>	<i>atól</i>	<b>atul</b>	son
<i>aca</i>	<i>aó</i>	<b>awu</b>	younger brother
<i>kelim</i>	<i>kulim</i>	<b>kulim</b>	picture
<i>nwun</i>	<i>nwun</i>	<b>nwun</b>	snow
<i>hanól</i>	<i>hanól</i>	<b>hanul</b>	sky
<i>tol</i>	<i>tolh</i>	<b>tol</b>	stone
<i>kwulim</i>	<i>kwulwum</i>	<b>kwulum</b>	clouds
<i>psal</i>	<i>psól</i>	<b>ssal</b>	husked rice
<i>pólam</i>	<i>pólóm</i>	<b>palam</b>	wind
<i>pól</i>	<i>pul</i>	<b>pul</b>	fire
<i>mój</i>	<i>moj</i>	<b>moy</b> (archaic)	mountain

<i>sjwupól</i>	<i>swupul</i>	<b>swul</b>	wine
<i>kahi</i>	<i>kahi</i>	<b>kay</b>	dog
<i>koki</i>	<i>koki</i>	<b>koki</b>	fish, meat
<i>hoal</i>	<i>hwal</i>	<b>hwal</b>	bow
<i>mól</i>	<i>mól</i>	<b>mal</b>	horse
<i>sjo</i>	<i>sjo</i>	<b>so</b>	cow, ox
<i>mil</i>	<i>mil</i>	<b>mil</b>	wheat
<i>kot</i>	<i>koc</i>	<b>kkoch</b>	flower
<i>omol</i>	<i>wumul</i>	<b>wumul</b>	water well
<i>mol</i>	<i>mul</i>	<b>mul</b>	water
<i>namk</i>	<i>namk-, namo</i>	<b>namu</b>	tree
<i>sam</i>	<i>sam</i>	<b>sam</b>	hemp
<i>nipul</i>	<i>nipul</i>	<b>ipul</b>	blanket, cover
<i>kachuki</i>	<i>kachi</i>	<b>kkachi</b>	magpie
<i>pittari</i>	<i>pitólki, pitwuli</i>	<b>pitwulki</b>	pigeon
<i>jangci</i>	<i>jangci</i>	<b>yangchi</b>	tooth cleaning
<i>onal</i>	<i>onól</i>	<b>onul</b>	today
<i>mosipaj</i>	<i>mosipoj</i>	<b>mosipey</b>	ramie cloth
<i>moloj</i>	<i>moloj</i>	<b>molay</b>	day after tomorrow

Some numerals appearing in *Kyeylim yusa* include the following:

(2)	<i>haton</i>	<i>hónah</i>	<b>hana</b>	1
	<i>topal</i>	<i>twulh</i>	<b>twul</b>	2
	<i>taswut</i>	<i>tasós</i>	<b>tases</b>	5
	<i>iswut</i>	<i>jesus</i>	<b>yeses</b>	6
	<i>ilkup</i>	<i>nilkwup</i>	<b>ilkop</b>	7
	<i>itap</i>	<i>jetulp</i>	<b>yetelp</b>	8
	<i>aho</i>	<i>ahop</i>	<b>ahop</b>	9
	<i>iol/iel</i>	<i>jel</i>	<b>yel</b>	10
	<i>swumol</i>	<i>sumul</i>	<b>sumul</b>	20
	<i>silhan</i>	<i>sjelhun</i>	<b>selhun</b>	30
	<i>siwun</i>	<i>sujn</i>	<b>swin</b>	50
	<i>on</i>	<i>on</i>	<b>[payk]</b>	100

*Kyeylim yusa* also reveals EMK morpho-syntax, which is essentially the same as that of LMK and Contemporary Korean, for example, nominative marking (3a), relative clause formation (3b), negation (3c), imperative constructions (3d), vocatives (3e), collocations (3f), conjunctives (3g), and questions (3h) are shared by EMK and Contemporary Korean.

(3)	<b><i>Kyeylim yusa</i></b>	<b>Contemporary Korean Glosses of MK</b>	
a.	<i>hólk-i</i>	<b>hulk i</b>	soil (nominative)
	<i>kach-i</i>	<b>kacwuk/kecwuk i</b>	skin (nominative)
	<i>namk-i</i>	<b>namu ka</b>	tree (nominative)
	<i>nó-j</i>	<b>nay ka</b> ‘I, me’	I (nominative)
	<i>nach-i tjoh-ón</i>	<b>nach i coh-un</b>	good-looking (‘face good’)
b.	<i>noph-ón</i>	<b>noph-un</b>	high
	<i>hój-n</i>	<b>huy-n</b>	white
	<i>isi-l</i>	<b>iss-ul</b>	to exist
	<i>noló-n soj</i>	<b>nola-n soy</b>	gold (‘yellow metal’)

	<i>sik-un mol</i>	<b>sik-un mul</b>	cold water
	<i>nik-un mol</i>	<b>[tew-un] mul</b>	hot water
c.	<i>nach i mot tjoh-ón</i>	<b>nach i an (not mos)</b> <b>coh-un</b>	bad-looking ('face not good')
	<i>ani masj-a</i>	<b>an masy-e</b>	not drinking
	<i>paj cha ani</i>	<b>pay cha-[ci] an[h-]</b>	not full (of one's stomach)
d.	<i>aci-kala</i>	<b>anc-kela</b>	Sit!
	<i>tó-la</i>	<b>tal-la</b>	Give!
	<i>o-la</i>	<b>o-la, w-ala</b>	Come!
	<i>sal-la</i>	<b>sal-la, sal-ala</b>	Live!
	<i>o-swusej</i>	<b>o-sole</b>	Please come! (super-polite)
	<i>pili-swusej</i>	<b>pilli-sole</b>	Borrow!, Lend!
e.	<i>ak-a</i>	<b>ak-a</b>	Child!
f.	<i>kel po-</i>	<b>kul po-</b>	read books ('see writing')
	<i>pal chi-kola</i>	<b>pal chi-kela</b>	Hang a bamboo-blind!
g.	<i>son sis-ja</i>	<b>son ssis-e</b>	wash hands and so
	<i>sjwupól masj-a</i>	<b>swul masy-e</b>	drink wine and so
h.	<i>nwu-ko</i>	<i>nwu-ko</i> (Kyengsang dialect)	Who is it?
	<i>mjech isj-e</i>	<b>myech i iss-e</b>	How many are there?

As observed in (3a) and (3f), case particles are frequently omitted. In particular, no examples of the accusative particle are given in *Kyeylim yusa*, although *itwu*, *hyangka*, and *kwukyel* data of the EMK period show abundant examples of case particles including nominative, accusative, and locative.

Even with the above *Kyeylim yusa* data alone, it is apparent that the sounds, as well as structures, of EMK were not particularly different from those of LMK. Other records also show striking similarities between Early and Late Middle Korean. For instance, in *Samkwuk saki* (1145), 'horn' was recorded with two Chinese characters, which would be pronounced as *spul*, exactly the same pronunciation in hankul (*spwul*) during the LMK period. It is thus safe to assume that fifteenth- and sixteenth-century Korean (LMK) of which abundant hankul texts are available for the first time is representative of overall MK.

There are possibilities to internally reconstruct a more comprehensive system of EMK language phenomena based on fifteenth- and sixteenth-century hankul-based written records, *itwu* and *hyangchal* transcriptions, and the traditional readings of Chinese characters, which were standardized in the Silla period (as was attempted for example in P. An 1977, 1992; P. Nam 1999, 2009; Y. Chang 2002; T. Pae 2002). Yet, the majority of Korean historical linguists have preferred to work on the abundant written data of fifteenth-century MK. Following this scholarly trend, the survey of MK in this chapter will be based largely on fifteenth- and sixteenth-century data.<sup>5</sup>

#### 4.1.3 Late Middle Korean texts

The hankul texts that appeared during the fifteenth and sixteenth centuries are impressive in terms of quantity, quality, and variability. In terms of number, Korean translations of Buddhist sutras and works of Buddhist literature are predominant. Representative texts that are frequently quoted for linguistic description include the following. Abbreviations are given in brackets.

- Hwunmin cengum haylyey* (Explanations and Examples of Hwunmin Cengum, by Ceng Inci *et al.*, 1446) [Cengum Haylyey]
- Yongpi echenka* (Songs of the Dragons Flying to Heaven, by King Seycong's royal subjects, 1447) [Yongka]
- Sekpo sangcel* (Detailed Articles on the Record of Sakyamuni, by Prince Swuyang, 1447) [Sekpo]
- Tongkwuk cengwun* (Correct Rhymes of the Eastern Country, 1447) [Tongkwuk]
- Welin chenkang ci kok* (Songs of the Moon's Imprint on a Thousand Rivers, by King Seycong, 1449) [Welin]
- Welin sekpo* (combination of the preceding two volumes, by King Seyco [= Swuyang], 1459) [Welsek]
- Hwunmin cengum enhay* (Korean translation of Hwunmin Cengum, included in Welin Sekpo, 1459) [Cengum Enhay]
- Nungemkyeng enhay* (Korean translation of Nungem Sutra, 1462) [Nungem]
- Pephwakyeng enhay* (Korean translation of Pephwa Sutra, 1463) [Pephwa]
- Kumkangkyeng enhay* (Korean translation of Kumkang Sutra, 1464) [Kumkang]
- Kumkangkyeng samkahay* (Korean translation of Three Authorities in Kumkang Sutra) [Kumkang Samka]
- Wenkakkyeng enhay* (Korean translation of Wenkak Sutra, 1465) [Wenkak]
- Kwukuppang enhay* (Korean translation of Emergency Treatments, 1466) [Kwukuppang]
- Mongsan pepe enhay* (Korean translation of Mongsan's Buddhist Teachings, 1467) [Mongsan]
- Samkang nayhwun* (Teaching for Females, 1475) [Nayhwun]
- Twusi enhay* (Korean interpretation of Du Fu's Poems, 1481) [Twusi]
- Kwanumkyeng enhay* (Korean translation of Kwanum Sutra, 1485) [Kwanum]
- Akhak kweypem* (Music Codes, 1493) [Akhak]
- Akcang kasa* (Movement Lyrics, sometime during 1506–1567) [Akcang]
- Pakthongsa enhay* (Korean translation of Interpreter Pak, by Choy Seycin, before 1517) [Pakthongsa]
- Nokeltay enhay* (Korean translation of the Old Cathayan, by Choy Seycin, before 1670) [Nokeltay]
- Hwunmong cahoy* (Chinese–Korean Glossary, by Choy Seycin, 1527) [Hwunmong]
- Samkang hayngsilto* (Depictions of Confucius Role Models, 1581) [Samkang]
- Sekpong chencamun* (Sekpong's 1,000 Characters, 1583) [Sekpong]
- Sohak enhay* (Korean translation of Primary Learning, 1588) [Sohak]
- Tayhak enhay* (Korean translation of Advanced Learning, 1590) [Tayhak]

In addition, there are many Kolye ballads that were written in hankul in the fifteenth and sixteenth centuries (Y. Choy 1987; K. Kang and Hwang 2003; Ko 2007).

## 4.2 PHONOLOGY

### 4.2.1 Background

MK succeeded the Silla language (Old Korean). K. M. Lee (1976: 93–9) assumes that the following phonological developments occurred in EMK: (a.) addition of the voiced fricative *z* to the consonant system, as in *kózól* 'autumn', *azó* 'younger brother', and *mózóm* 'mind'; (b.) addition of the voiced bilabial fricative *β*, as in *nwuβi* 'younger sister', *swuβel* 'liquor', and *twuβul* 'two'; and (c.) development of the tensed consonant series

(*pp*, *tt*, *cc*, *kk*, *ss*, *hh*), as observed in *ppul* (< *spul* < *sepul*) ‘horn’, *ttal* (< *ptól* < *potól*) ‘daughter’, and *ssal* (< *psól* < *posól*) ‘rice’.

B. An (1985: 47) identifies 13 consonants, six vowels, and a number of diphthongs in EMK based on his phonological analysis of the words that appear in *Kyeylim yusa*. The 13 consonants are: stops *p*, *t*, *c*, *k*, *ʔ*, *ch*; fricatives *z*, *s*, *h*; nasals *m*, *n*, *ng*; and liquid *l*. The six vowels are: *i*, *u*, *wu*, *o*, *a*, *ó*. An’s system shows that the tense consonant series had not developed yet and even the aspirated consonant series was not fully established. The later mid-central vowel *e* has not been identified, which makes the vowel system asymmetrical. An’s system is considerably different from what K. M. Lee assumed above. It appears that Lee’s assumption of the tensed consonant series in EMK is somewhat weak because the series was not firmly established even in LMK, as seen shortly. On the other hand, as An’s reconstructions are based only on words in *Kyeylim yusa*, which are transcribed with approximate Chinese characters, his assumed system cannot be said to be complete. This is particularly the case when we compare the system with that of LMK.

T. Kim (2002) compares seven sets of EMK consonant systems proposed by seven different linguists, from a nine-consonant system to a 15-consonant system. For instance, the 15-consonant system includes: stops *p*, *t*, *c*, *k*, *ph*, *th*, *ch*, *kh*; fricatives *s*, *h*; nasals *m*, *n*, *ng*; and liquids *l*, *r* (e.g. B. Choy 1990). None of these systems include *z* or *ʃ*. Only one includes *ʔ*.

The later period of MK (from the fifteenth century) observed a flood of literature written in hankul, most of which reflected the language used in the capital, the Kaykyeng dialect. The phonemic system of fifteenth-century Korean consisted of 22 (à la Huh 1983; Y. Choy 1987, and others) or 23 (à la K. M. Lee 1972 and others) consonants, seven vowels, and two semivowels. Unlike Huh (ibid.) and others, K. M. Lee (ibid.: 15–26), Ko (2007: 29), and a few earlier Korean linguists include *ʃ* in the MK consonant system. K. M. Lee assumes that contemporary standard words like **kasay-** (or **kawi**) ‘scissors’ and **molay** ‘sand’ derived from fifteenth-century words like *\*kózʃaj* and *\*molʃaj*, respectively, in view of the existence of forms like *kasikay* ‘scissors’ and *molokay* ‘sand’ in southern and other dialects (see 6.2.5).

#### 4.2.2 Consonants

The 23 consonant phonemes are given in Table 4.1. Examples with these phonemes are given in (4).

TABLE 4.1 LATE MIDDLE KOREAN CONSONANTS

		Bilabial	Dental	Palatal	Velar	Glottal
Stop:	Lax	p ㅍ	t ㅌ	c ㅊ	k ㅋ	
	Aspirated	ph ㅍᄀ	th ㅌᄀ	ch ㅊᄀ	kh ㅋᄀ	
	Tensed	pp ㅍㅍ	tt ㅌㅌ	cc ㅊㅊ	kk ㅋㅋ	
Fricative:	Lax-voiced	β ㅃ	z ㅅ			(f) ㅇ
	Lax-voiceless		s ㅆ		h ㅎ	
	Tensed		ss ㅆㅆ		hh ㅎㅎ	
Nasal:	m ㅁ	n ㄴ		ng ㄹ		
Lateral:		l ㄹ				

(4) <i>pul</i> ‘fire’	<i>twul</i> ‘two’	<i>cohój</i> ‘paper’	<i>koc</i> ‘flower’
<i>phól</i> ‘house fly’	<i>kothi</i> ‘cocoon’	<i>chej</i> ‘sieve’	<i>khong</i> ‘soy bean’
<i>ppwul</i> (< <i>spwul</i> ) ‘horn’	<i>tta</i> (< <i>sta</i> ) ‘earth’	<i>ccek</i> ‘time’	<i>-kka</i> (question)
<i>nwuß-</i> ‘lie down’	<i>kezuj</i> ‘almost’	<i>mol(h)aj</i> ‘sand’	
<i>sjem</i> ‘island’	<i>sso-</i> ‘shoot’	<i>him</i> ‘power’	<i>tolóhhi-</i> ‘return’
<i>skwum</i> ‘dream’	<i>nap</i> ‘monkey’	<i>cjwung</i> ‘monk’	<i>nolaj</i> ‘song’

In fifteenth-century Korean, words with a tensed stop written *pp*, *tt*, *cc*, *kk* were rare. Rather, tensed sounds resulted from sandhi with a preceding ‘tensifying’ sound such as *l* or a stop as in CK, e.g. *hol kkes* ‘single thing’ and *coch ccap-* ‘follow and catch’. The tensed series was developing as an innovation in the system, especially from clusters as discussed shortly.

King Seycong’s *Hwunmin Cengum* includes a hankul symbol (ㄷ) representing the glottal stop [ʔ]. However, most historical linguists of Korean agree that it was not a phoneme. *Hwunmin Cengum Hyeylyey* itself indicates that word-initial ʔ + vowel is similar to the same vowel without ʔ. Furthermore, the book does not provide any examples illustrating the sound.

The hankul symbol ㄷ occurred rather rarely in fifteenth-century hankul texts, e.g. *hanólʔ ptut* ‘heaven’s will’ < *hanól* ‘heaven’ + *ptut* ‘will’.

As indicated by Ko (2007: 23–4, 28), ʔ was recorded in fifteenth-century Korean only after *l*, or after the final vowel of the first noun in a noun–noun compound, and other letters such as *s*, *z*, *k*, *t*, and *p* occur in the same role: *pójs koc* ‘pear flower’, *nimkumz mal* ‘king’s words’, *myes kan.t cip* ‘house of how many rooms’, *phjeng.sajng. k ptut* ‘life-time wish’, and *salómp ptut* ‘people’s wish’. It is thus clear that the function of ʔ, as well as the other epenthetic consonants, was to indicate tensification of the following lax consonant across word boundaries. Only the so-called epenthetic *s* continues in this function.

Y. Choy (1987: 21) indicates that the Chinese 日 ‘day’ was transcribed as *ziʔ* in *Hongmu Cengwun* and as *zilʔ* in *Hwunmin Cengum Enhay*. The current Korean form is *il*, while the current Japanese form is *nichi* or *niq-* (as in *niqpon* ‘Japan’) reflecting Middle Chinese *-t*. In view of all this, the sound must have been a glottal stop or catch not having a full-fledged phonemic status but had the function of keeping the meanings apart as the epenthetic *s* in CK does.

Due to their limited distribution and low functional load, the voiced fricative consonants *β*, *hh*, and *h* disappeared between the mid-fifteenth and early sixteenth century. *β*, which generally occurred between voiced sounds and did not occur word-initially, changed primarily to *w*, as in *teβi* > *tewi* ‘heat’ and *kulβal* > *kulwal* ‘sentence, letter’, but also merged with *p* in some dialects as in the Kyengsang dialect *tepi* ‘heat’ and the Cenla/Kyengsang dialect *saypi* ‘shrimp’ from MK *safβi* (= *saywu*). The tensed *hh* occurred only in a small number of words, such as *hhje-* ‘pull’ and *chihhje-* ‘raise’ and later changed generally to *kh*: *khye-* and *chikhi-*. *h* has either dropped or remained as *k*, as in *kózhaj* ‘scissors’ > *kawi*, *kasay* or *kasikay* (Cenla, Kyengsang, etc.); *molhaj* ‘sand’ > *molay*, *molkay* (Kyengsang); *pelhej* ‘bug’ > *pelley*, *pel(k)eci* (Cenla, Kyengsang, etc.); *melhwuj* ‘wild grapes’ > *melwu*, *melkwu* (Cenla, Kyengsang, etc.); and *silhej* ‘shelf, rack’ > *sileng* (standard CK), *silkeng* (Cenla, Kyengsang, etc.).

*z* generally occurred between voiced sounds, as in *mózóm* ‘heart’, *kózól* ‘autumn’, *kjezul* ‘winter’, *mózól* ‘village’, *nózzil* ‘tomorrow’, *niz-e* ‘connect-INF’, *hanzum* ‘sigh’, *sózi* ‘space’. It persisted until the end of the sixteenth century and thereafter either (rarely) merged with *s* (especially some words in southern and other dialects) or (more

frequently) disappeared completely: **maum**, **kaul**, **kyewul**, **maul**, **nayil**, **i-e**, **hanswum**, **sai**. Forms in other dialects are generally the same, except that ‘connect-INF’ is *is-e* in Cenla, Kyengsang, and some other dialects, where ‘autumn’, ‘winter’, and ‘village’ are reflected in two forms: *kaul* and *kasil*, *kyewul* (new form) and *cesil* (old form), *maul* and *masil*. The two forms *kasil* and *masil* acquired specialized meanings: *kasil* meaning ‘autumn harvest’ and *masil* used only in *masil ka-* ‘goof around (in the village)’.

*c* and *ch* were alveo-dental [ts] and [tsʰ], respectively. When followed by *j*, they became palatal. Thus, there were contrasts between *cang* ‘cupboard’ and *cjang* ‘soy sauce’ (both **cang**) and between *cho* ‘vinegar’ and *chjo* ‘candle’ (both **cho**). *s* was pronounced as it is today, but, in many words, was followed by a *j* that has since been lost: *sjem* ‘island’ > **sem**, *sjo* ‘cow, ox’ > **so**.

*t* and *th* could occur before *i* or *j*, since palatalized to **c** and **ch**: *thjenti* ‘heaven and earth’ (**chenci**), *tjotha* ‘good’ (**cohta**), *tje* ‘that over there’ (**ce**), *tina* ‘elapsing’ (**cina**), *thita* ‘strike’ (**chita**). Palatalization of *t* and *th* before *i* or *j* occurred around the seventeenth or eighteenth century. MK also allowed *n* word initially before *i* or *j*: *ni* ‘tooth’ (**i**), *niph* ‘leaf’ (**iph**), *njelum* ‘summer’ (**yelum**). Most other sound alternations such as the epenthetic *s* phenomena, tensification of lax stops after the prospective relativizer *-(u)l*, and nasalization of non-nasal sounds before a nasal sound were, by and large, similar to parallel CK phenomena. Also as in CK, *l*-ending predicates lost their final *l* before a homorganic sound in a suffix, as in *mójngkól-noni* ‘make and so’, where *mójngkól* ‘make’ lost its final *l*.

There was a set of consonant clusters (5), illustrated in (6).

- (5) Sequence of two consonants: *pt* *pth* *ps* *pc*  
*sp* *st* *sn* *sk*  
 Sequence of three consonants: *pst* *psk*
- (6) *ptul* ‘yard’ *pthuj-* ‘jump’ *psi* ‘seed’ *pcak* ‘pair’  
*spje* ‘bone’ *sto* ‘again’ *snóhój* ‘boy’ *skoli* ‘tail’  
*pstaj* ‘time’ *pskum* ‘gap’

The values of orthographic *sp*, *st*, *sk*, *ps*, *pst*, *psk*, *pt*, *pc* are still disputed: were they pronounced as written or were they simply allographs of tensed *pp*, *tt*, *kk*, *ss*, *tt*, *kk*, *tt*, *cc*? Certainly they correspond to modern tensed consonants: *ptul* (**ttus**) ‘meaning, intention’, *psi* (**ssi**) ‘seed’, *psu* (**ssu**) ‘use’, *skwum* ‘dream’ (**kkwum**), *spul* ‘horn’ (**ppul**), *pcak* (**ccak**) ‘pair’, and *pskwul* (**kkwul**) ‘honey’. Most Korean historical linguists agree that they were pronounced as written. For instance, standard **ttek** ‘rice cake’ and **ttong** ‘dung’ are still pronounced as *sitekwu* in some northern dialects and *sitong* (in the sense ‘manure’) in some southern (e.g. Cenla) dialects. The reduction processes involved are assumed to have been \**sitekwu* > MK *stek(wu)* > **ttek** and \**sitong* > MK *stong* > **ttong**. Similarly, MK *psól* ‘hulled rice’ is currently **ssal**. However, **p** is retained in such compounds as **copssal** (**co** ‘millet’ + /p/ + **ssal**) ‘hulled millet’, **haypsal** (**hay/say** ‘new’ + /p/ + **ssal**) ‘new hulled rice’, and **chapssal** (**chal** ‘sticky’ + /p/ + **ssal**) ‘glutinous rice’, where the syllable boundary falls between **cop/hayp/chap** and **ssal**. MK *pstaj* ‘time’ is now **ttay**, but the initial *p* remains in **cepttay** (**ce** ‘that’ + /p/ + **ttay**) ‘the other day’. MK *snóhój* ‘boy’ is **sanai**; it seems that, unable to be tensified, *sn* instead acquired an epenthetic vowel.

MK observed implosivization (unrelease) of syllable-final consonants, except *s* which was, unlike in CK, not neutralized to *t* but spelled *s*, suggesting it was pronounced [s], as in *mos* [mos] ‘pond’ versus *mot* [mot] ‘unable’. Thus, all bilabial consonants were

neutralized to *p*, all dental, palatal, and glottal consonants to *t*, and all velar consonants to *k* syllable-finally, as in CK. As a result, only *p*, *t*, *k*, *m*, *n*, *ng*, *l*, *s* were actually pronounced syllable-finally.<sup>6</sup>

#### 4.2.3 Vowels and semivowels

The fifteenth-century seven vowels and two semivowels are as follows. The semivowels did not have their unique hankul symbols, but were written, in the case of *j*, by adding an additional dot or short stroke (horizontal or vertical) to a vowel symbol, as in ㅓ /jo/ or ㅑ /ja/; and, in the case of *w*, by combining two vowels in a syllable block, as in ㅓㅑ /u/ + ㅑ /a/ = /wa/.

##### (7) a. Vowels

	Front	Central	Back
High	<i>i</i> /i/ ㅓ	<i>u</i> /i/ —	<i>wu</i> /u/ ㅓ
Mid		<i>e</i> /ə/ ㅑ	<i>o</i> /o/ ㅓ
Low		<i>a</i> /a/ ㅑ	<i>ó</i> /ʌ/ ㅓ

##### b. Semivowels      *j*                      *w*

The vowel system was asymmetrical with only one front vowel. King Seycong and his linguists classified vowels into three groups: the bright or *yang* vowels *a*, *ó*, *o*; the dark or *yin* vowels *e*, *u*, *wu*; and the neutral vowel *i*. This classification explained vowel harmony patterns.

MK had rigid vowel harmony: *yang* vowels *o*, *a*, *ó* occurred only with *yang* or neutral (*i*) vowels, and *yin* vowels *e*, *u*, *wu* occurred only with *yin* or neutral vowels. *Yang* vowels are characterized as having more sonority than other vowels in that they are pronounced with retracted tongue root. Vowel harmony in MK was observed within a word (8a), across nominal and particle (8b), and across predicate stem and suffix (8c). Current Korean often no longer retains this vowel harmony.

- (8) a. *namo*      **namu**      ‘tree’  
*kamakoj*      **kamakwi**      ‘crow’  
*taló-*      **talu-**      ‘different’  
*kwulwum*      **kwulum**      ‘clouds’  
*nilkwup*      **ilkop**      ‘seven’  
*halmi*      **halmi**      ‘old woman’
- b. *son-ólo*      **son-ulo**      ‘with a hand’  
*mal-ól*      **mal-ul**      ‘horse-ACC’  
*skwum-ul*      **kkwum-ul**      ‘dream-ACC’  
*son-ój*      **son-uy**      ‘hand-GEN’  
*cip-uj*      **cip-uy**      ‘house-GEN’
- c. *mak-óni*      **mak-uni**      ‘as one blocks’  
*mek-uni*      **mek-uni**      ‘as one eats’  
*nól-kaj*      **nalkay**      ‘wing’  
*tep-kej*      **tepkay**      ‘cover’

The vowel *ó* is believed to have been a sound between *a* and *o*, primarily because Ceycwu dialect still retains it. It was lost in all other dialects, becoming *a* in the first syllable and *u* elsewhere, e.g. *mózóm* ‘heart’ > **maum**.

With *j* and *w*, 18 diphthongs and triphthongs were formed, given in (9) and illustrated in (10). Unlike in CK, MK *j* occurred as an off-glide as well. There were no *w* off-glides.

## (9) Diphthongs

On-glides:	<i>ja</i>	<i>je</i>	<i>jo</i>	<i>jwu</i>		
	<i>wa</i>	<i>we</i>				
Off-glides:	<i>aj</i>	<i>ej</i>	<i>oj</i>	<i>wuj</i>	<i>uj</i>	<i>ój</i>

## Triphthongs

On-off glides:	<i>jaj</i>	<i>jej</i>	<i>joj</i>	<i>jwu</i>
	<i>waj</i>	<i>wej</i>		

- (10) On-glides: *pójam* ‘snake’      *nje-* ‘go’      *tjoh-* ‘good’  
*sjwulwup* ‘umbrella’      *cwasi-* ‘eat’ (hon.)      *pwuhwengi* ‘owl’  
Off-glides: *haj* ‘a lot’      *sejs* ‘three’      *moj* ‘mountain’  
*swujβ-* ‘easy’      *pkuj* ‘meal’      *hój* ‘sun’  
On-off glides: *cókaj* ‘nacre’      *njej* ‘old times’      *sjojkoki* ‘beef’  
*chjwujhó-* ‘drunk’      *tasswaj* ‘5 days’      *kwej* ‘chest’

## 4.2.4 Suprasegmentals

MK was a tonal language. Early hankul used a simple way to mark tones. There were three tonemes: high pitch or “going tone” (indicated by one dot to the left of the hankul syllable block), rising pitch or “rising tone” (two dots to the left), and low pitch or “even tone” (absence of dots):

- (11) a. *·kil* ‘road’      *·son* ‘hand’      *·pal* ‘foot’  
*ka·ci* ‘variety’      *·kaci* ‘branch’      *·mal* ‘a unit of measure’  
b. *:saj* ‘bird’      *:pal* ‘the blinds’      *:mal* ‘word’  
c. *cip* ‘house’      *son* ‘guest’      *mal* ‘horse’  
d. *pwulhwuj* ‘root’ + *·i* (NOM) → *pwul:hwuj*  
*tóli* ‘bridge’ + *·i* (NOM) → *tó:li*

Tones ceased to exist from the middle of the sixteenth century on, except in parts of the Kyengsang and Hamkyeng dialects. All vowels that had a rising tone became (and still are) long, whereas vowels with a high or low tone remained short.<sup>7</sup>

## 4.3 SCRIPT

Chinese character script has long been an integral part of the writing systems of Koreans and was the only system before the creation of hankul in 1446, and literacy was restricted to male members of the ruling class, who devoted their entire lives to the study of Classical Chinese written in characters, because it was the main goal of education, the official means of government affairs, and the medium of civil service examinations.

In an effort to improve this predicament, itwu script (including hyangchal and kwukyel transcriptions) was used to record Korean by means of (a) Chinese characters borrowed for their Chinese meaning but read as the corresponding semantically equivalent Korean word (semantograms) or (b) Chinese characters borrowed for their Chinese pronunciation only (phonograms). (see Nam, this volume.)

Itwu script allowed people to record personal names, place names, and vernacular songs and poems. It was also used to clarify government documents and other books written in Chinese. However, it was extremely inadequate as a means of written communication as well as being difficult to learn. It was under these circumstances that hankul was created, a phonetic writing system completely disengaged from the Chinese script.

The indigenous phonetic alphabet now called hankul is one of the most remarkable writing systems ever devised.<sup>8</sup> For the design of the hankul alphabet, Seycong (1397–1450), the fourth king of the Cosen dynasty, and the scholars of the royal academy called Ciphyencen (lit. ‘Hall of Assembled Sages’) studied the rich Chinese linguistic tradition, such as the concepts of consonants, syllables, and tones, as well as their underlying philosophical background. The orthographical design of hankul, which was based on a rigorous analysis of Korean and Sino-Korean (SK) sound patterns, was completed in 1443. To test the new writing system, the king ordered subjects (Kwen Cey *et al.*) to write the voluminous *Yongpi Echenka* (Songs of Flying Dragons, 1447) in hankul with translations in Chinese characters, a eulogy cycle in 125 cantos comprising 248 poems, composed to celebrate the founding of the Cosen dynasty and the achievements of Seycong’s predecessors.

King Seycong promulgated hankul on October 9, 1446, as *hwunmin cengum* ‘the correct sounds to educate people’. *Hwunmin Cengum* (written by Seycong himself) was accompanied by *Hwunmin Cengum Haylyey* (Explanations and Examples of the *Hwunmin Cengum*), compiled by a group of scholars (headed by Ceng Inci). These two documents were published as a single book.

The text of *Hwunmin Cengum* consists of three parts: preface, pronunciation, and usage rules. The preface briefly summarizes Seycong’s motives for inventing hankul as (12) in English translation.

- (12) The speech sounds of Korea are distinct from those of China and thus are not communicable with Chinese characters. Hence, many people having something to put into words are unable to express their feelings. To overcome such distressing circumstances, I have newly devised 28 letters that everyone can learn with ease and use with convenience in daily life.

In the second part, the sounds of the newly devised 17 consonant (ㄱ ㅋ ㅇ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ ㆁ) and 11 vowel (ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅟ ㅠ) letters are explained in Chinese. It is also indicated that when ㄱ ㆁ ㆁ ㆁ ㆁ are doubled, they are pronounced tense: ㄱ ㄱ, ㆁ ㆁ, ㆁ ㆁ, ㆁ ㆁ, ㆁ ㆁ.

The third part of *Hwunmin Cengum* presents rules regarding the use of hankul letters and other symbols in syllable blocks. For instance, when ㅇ is placed under a bilabial sound, a bilabial fricative sound results (e.g. the bilabial fricative sound  $\beta$  is represented by placing ㅇ under ㅁ as in ㅁㅇ). It also stipulates use of tone dots.

*Hwunmin Cengum Haylyey* consists of seven parts: design of the letters, syllable-initial sounds, syllable-medial sounds, syllable-final sounds, combination of letters, examples of use, and a brief postface by Ceng Inci.

Since its creation, hankul has undergone several major ordeals in the course of its diffusion in the face of the long tradition of Chinese character use by the nobility and government officials during the Cosen dynasty. Although King Seycong’s invention of hankul was the greatest cultural achievement, a group of scholar-officials led by Choy Manli, then associate academician of Ciphyencen, vehemently opposed the king’s promulgation of hankul to the public. They presented an anti-hankul appeal to the throne in 1444. Their main argument was that Korea had long emulated Chinese ideas and institutions

and that adoption of Korea's own writing system would make it impossible to identify Korean civilization with that of China but rather identify Korea with "barbarians" such as Mongolians, Tanguts, Jurchen, Japanese, and Tibetans, who had their own scripts. This appeal had little effect on Seycong's determination.

After the death of Seycong, opposition to hankul resumed and continued until Seyco, the seventh Cosen king, propagated hankul with Buddhism. Classical Chinese and *itwu* remained predominant while hankul served as an aid for the study of Classical Chinese and was used mostly by women.

#### 4.4 MORPHOLOGY

The hankul literature of the fifteenth century is the oldest data to show the comprehensive grammatical system of Korean. Overall, the basic morphological structure of MK, such as word formation and inflection patterns, is similar to that of CK, although there are conspicuous differences in detail.

Nine word classes are recognized in MK: Noun, Pronoun, Numeral, Verb, Adjective, Copula, Determiner, Adverb, and Particle. Noun, Verb, Adjective, and Adverb are open classes. Noun, Pronoun, and Numeral function as heads of noun phrases, whereas Verb, Adjective, and Copula function as heads of predicates, inflecting in tense, aspect, modality, mood, and sentence/clause type ender. Determiner (demonstratives and specifiers) modifies nominal arguments, while Adverb modifies predicates, noun phrases, clauses, and sentences. Particle marks a wide variety of syntactic and semantic functions, including case, delimiter, conjunctive, and complementizer functions. In addition, various derivational affixes contributed to generating new words, while inflectional affixes (mostly suffixes) mark grammatical categories, relations, and functions. For initial approximation and illustration of uses of different word classes and inflectional suffixes, observe some excerpts from the *Yongpi Echenka* (1447) in (13a), *Hwunmin Cengum Enhay* (1450) in (13b), and *Welin Sekpo* (1459) in (13c).

- (13) a. *pwulhwuj kiph-un namk<sup>9</sup> ón pólóm aj ani mwuj-lssój koc*  
 root deep-REL tree TOP wind by not move-so flower  
*tjo-kho jelum ha-nó-ni*  
 good-and fruit much-IND-DEC  
 'Deep-rooted trees do not move by wind. So, flowers are good and fruits are plentiful.'
- b. *na j i lól wujhój-a ejespi nekj-e sajlo sumul.jetulp*  
 I NOM this ACC serve-for pitifully think-so newly 28  
*ca lól mójngkó-no-ni*  
 letter ACC make-IND.SS-as  
 'I felt sorry for this and newly made 28 letters. And thus [. . .]'
- c. *salóm ój mokswum i mwusanghó-n kes i-la*  
 person GEN life NOM uncertain-REL thing be-DEC  
 'People's life is something uncertain.'

The above sentences can be analyzed as in (14) in terms of word classes and inflectional suffixes.

- (14) Nouns: *pwulhwuj* 'root', *namk* 'tree', *pólóm* 'wind', *koc* 'flower', *jelum* 'fruit', *ca* 'letter', *salóm* 'person', *mokswum* 'life', *kes* 'thing, fact'

Pronouns: *na* ‘I’, *i* ‘this fact/thing’

Numerals: *sumul jetulp* ‘28’

Verbs: *mwuj-* ‘move’, *mójngkól-* ‘make’, *wujhój-* ‘(do) for’, *neki-* ‘think’

Adjectives: *kiph-* ‘deep’, *tjoh-* ‘good’, *ha-* ‘many, much’, *ejespu-* ‘pitiful’, *mwusanghó-* ‘uncertain’

Copulas: *i-* (in *i-la*) ‘be’

Adverbs: *ani* ‘not’, *sajlo* (*saj* ‘newness’ + *lo* ‘with, in’) ‘newly’, *ejespi* (*ejespu* ‘pitiful’ + *-i* ‘-ly’) ‘pitifully’

Determiners: *i* ‘this’ (when it occurs before a nominal)

Particles: *ón* ‘as for’ (topic), *aj* ‘at, on, in, to’ (locative/goal), *lól* (object or accusative), *j*, *i* (subject or nominative), *ój* (genitive)

Inflectional suffixes: *-(u)n* relativizer, *-lssój* ‘as, so, because’, *-ko* ‘and’, *-nó* indicative mood, *-ni* neutral declarative sentence ender, *-noni* ‘as, since’, *-a/-e* ‘and, so’, *-la* declarative sentence ender after a copula

#### 4.4.1 Nouns and particles

##### 4.4.1.1 Nouns

Due to the all-out influx of Chinese words into MK, the lexicon was rich with many native and SK doublets with some connotational differences. Most Chinese words were imported as nouns, regardless of their original nominal or predicative status in Chinese. Thus, there were and still are a large number of SK verbal nouns and adjectival nouns derived from verbs or adjectives in Chinese. Proper nouns were generally written in Chinese characters in MK texts.

##### Bound nouns

In MK most nouns were free and independent, but there were also a considerable number of bound nouns, which cannot stand alone without a preceding determiner or relative clause. Most carry abstract or general meanings, which become more specific with a preceding determiner or a relative clause:

(15)	Semantic class	Bound noun	Gloss
	Person	<i>pwun</i>	‘person’ (honorific)
		<i>i</i>	‘person’ (adult)
	Thing/fact	<i>kes</i>	‘thing, fact’
		<i>pa</i>	‘thing, fact, what’
		<i>i</i>	‘thing’
		<i>cwul</i>	‘way, assumed fact’
	Place	<i>tó</i>	‘assumed fact’
		<i>tó(j), taj</i>	‘place’
		<i>njek</i>	‘direction, region’
	Time, chance	<i>cek</i>	‘time’
		<i>tet</i>	‘during, interval’
		<i>cej</i>	‘time’
	Limit	<i>spwun</i>	‘only’
		<i>stólóm</i>	‘alone’
		<i>man</i>	‘only, extent’

Cause	<i>tas</i>	‘blame, cause’
	<i>ach</i>	‘reason, cause’
State, condition	<i>jang</i>	‘state’
	<i>tós</i>	‘seeming, appearance’
	<i>cahi</i>	‘intact, just as it is’
Quantity, volume	<i>pól</i>	‘the span of both arms’
	<i>móli</i>	(counter for animals)
	<i>nat</i>	‘piece, unit’

### Compounding

New nouns in MK derived most productively from compounding. Examples of co-compounds are *achóm-naco* (morning-evening) ‘morning and evening’, *ezi-atól* (parents-son) ‘parents and children’, *mó-sjo* (horse-cow) ‘horses and cows’, *cwuk-sal-i* (die-live-NMR) ‘life and death’, and *anh-pask* (inside-outside) ‘inside and outside’. Examples of sub-compounds are *amh-thalk* (*amh* ‘female’ + *talk* ‘hen, cock’) ‘hen’, *hwa-sal* (*hwal* ‘bow’ + *sal* ‘arrow’) ‘arrow’, *petu-namo* (*petul* ‘willow’ + *namo* ‘tree’) ‘willow tree’, *muls-kóz* (*mul* ‘water’ + epenthetic *s* + *kóz* ‘edge’) ‘waterside’, *muls-kjel* (*mul* ‘water’ + epenthetic *s* + *kjel* ‘texture’) ‘wave’, *sjoj-koki* (*sjo* ‘cow’ + genitive *j* + *koki* ‘meat’) ‘beef’, *nól? -cjwungsójng* (*nól* ‘fly’ + epenthetic *?* + *cjwungsójng* ‘animal’) ‘winged animal’, *hanapi* (*ha* ‘big’ + relativizer *-n* + *api* ‘father’) ‘grandfather’, and *kan-pam* (*ka* ‘go’ + relativizer *-n* + *pam* ‘night’) ‘last night’.

### Affixation

Noun formation via affixation used prefixation and suffixation. Prefixation was not able to change word classes (–WCC), while suffixation includes those which changed word classes (+WCC), e.g. noun from verb or adjective, and those that could not (–WCC).

(16)	MK affixes	MK examples
Prefixes –WCC	<i>ni-</i> ‘unglutinous’	<i>ni-psól</i> ‘unglutinous rice’
	<i>tul-</i> ‘wild’	<i>tul-pskaj</i> ‘Perilla japonica’
	<i>achón-</i> ‘small’	<i>achón-atól</i> ‘nephew’ <i>achón-stól</i> ‘niece’
Suffixes +WCC	<i>-i</i> ‘activity, state, thing’	<i>hój-tot-i</i> (sun-rise-) ‘sunrise’ <i>wuzum-wuz-i</i> ‘laughing’
	<i>-ój/-uj</i> ‘state’	<i>noph-ój</i> ‘height’ <sup>10</sup> <i>nep-uj</i> ‘width’
	<i>-(u)m</i> ‘activity, state, thing’	<i>kel-um</i> ‘walking, stepping’ <i>kuli-m</i> ‘picture’ <i>el-um</i> (freeze-) ‘ice’
	<i>-kaj/-kej</i> ‘instrument’	<i>pje-kaj</i> (rest.head-) ‘pillow’
	<i>-aj/-ej</i> after <i>l</i>	<i>nól-aj</i> (fly-) ‘wing’ <i>nol-aj</i> (play-) ‘song’ <i>wul-aj</i> (cry-) ‘thunder’
Suffixes –WCC	<i>-ki</i> ‘activity, state’	<i>mól-po-ki</i> (horse-see-) ‘easing nature’
	<i>-pal</i> ‘lines, streaks’	<i>pis-pal</i> (rain-) ‘streaks of rain’ <i>hós-pal</i> (sun-) ‘streaks of sun’
	<i>-nim</i> (honorific)	<i>acóma-nim</i> ‘aunt’ (hon.)

	<i>apa-nim</i> ‘father’ (hon.)
	<i>mótató-nim</i> ‘oldest son’ (hon.)
<i>-math</i> ‘side’	<i>meli-math</i> (head-) ‘bedside’
<i>-cil</i> ‘activity’	<i>pwulmwu-cil</i> (bellows-) ‘blowing with the bellows’
<i>-aci</i> ‘young’	<i>mó-j-aci</i> (horse-GEN-) ‘foal’
	<i>móng-aci</i> (horse-) ‘foal’
	<i>sjo-j-aci</i> (cow-GEN-) ‘calf’
	<i>sjong-aci</i> (cow-) ‘calf’
<i>-i</i> ‘animate, thing’	<i>kongcak-i</i> ‘peacock’

#### 4.4.1.2 Particles

Particles, along with inflectional suffixes, perform important grammatical and semantic functions in MK. While inflectional suffixes occur after predicates, particles occur most frequently after nouns. They also follow pronouns, numerals, nominalized clauses, complement clauses, and conjunctive clauses. Observe the following MK sentence, where particles are bold-face while inflectional suffixes are in roman letters.

- (17) *sejcon i sangtwusan aj ka-sj-a ljong kwa kwujsin kwa*  
 Sakya NOM Mt.Sangtwu to go-SH-and dragon and ghost and  
*wujhó-j-a selpephó-te-si-ta*  
 serve-for lecture-RETR-SH-DEC  
 ‘I saw Sakya going to Mt. Sangtwu and giving a Buddhist lecture for dragons  
 and ghosts.’ (Sekpo 6:1)

#### Case particles

Case particles occur after nominals such as nouns, pronouns, and numerals. In most cases, vowel harmony was strictly observed. In Table 4.2, *U* stands for *u* after a dark-vowel syllable and *ó* after a bright-vowel syllable.

The nominative *i*, typically marking a subject, was reduced to *j* after a vowel, as in *sjo j* ‘cow’ (*sjo* ‘cow’ + nominative *j*). When a nominal stem ends in *i* or *j*, the nominative particle is omitted. Thus, the subject forms of *tóli* ‘bridge’, *pwulhwuj* ‘root’, and *saj* ‘bird’ are the same as the nominal stems. The nominative particle was often used like a comitative, as in *sajngpwul i kóthó-si-mje* (living.Buddha NOM same-SH-and) ‘is the same as a living Buddha and’.

The genitive particle *uj/ój* was reduced to *i* or *j* after a vowel, as in *sinha i mal* (royal.subject GEN words) ‘royal subject’s words’, *sjo j meli* (cow GEN head) ‘cow’s head’, and *na j mom* (I GEN body) ‘my body’. Genitive particles were often used to mark a subordinate clause subject or a static locative, as in *jelaj s niló-sja-n senpep* (enlightened.one GEN say-SH-REL good.law) ‘the good Buddhist laws that the enlightened one mentioned’, *sajpjel i nacó j tot-oni* (morning.star NOM evening GEN rise-as) ‘as the morning star rises in the evening’, and *thongcwu i tatól-a* (Thongcwu GEN arrive-and) ‘after arriving at Thongcwu’. A genitive particle as a static locative marker + the inceptive particle *-sje* together behaved like the dynamic locative particle (*j*)*ejsje*.

- (18) *chwuksajng ón salóm i cip uj.sje chi-nón cjwungsójing i-la*  
 livestock TOP people NOM home at raise-REL animal be-DEC  
 ‘Livestock are animals raised at home.’ (Welsek 1:46)

TABLE 4.2 LATE MIDDLE KOREAN CASE PARTICLES

Function	MK forms
Nominative	<i>i/lj</i>
Genitive 'of, 's'	<i>s</i> (after honorific animate nominal or inanimate nominal) <i>Uj</i> (after nonhonorific animate nominal)
Accusative	<i>lUl</i> (after open syllable) <i>Ul</i> (after closed syllable)
Goal/dative/static locative 'to, at, in'	Animate: <i>skej/skuj</i> (honorific) <i>Ujkej/jkej</i> (neutral) Inanimate: <i>ej</i> (after closed dark-vowel syllable) <i>jej</i> (after open dark-vowel syllable) <i>aj</i> (after bright-vowel syllable)
Source/dynamic locative 'from, at, in'	Animate: <i>skejsje/skujkje</i> (honorific) <i>Ujkejsje</i> (neutral) <i>sontój</i> (neutral) Inanimate: <i>ejsje</i> (after closed dark-vowel syllable) <i>jejsje</i> (after open dark-vowel syllable) <i>ajsje</i> (after bright-vowel syllable)
Instrument/directional	<i>Ulo</i> (after closed syllable) <i>lo</i> (after open syllable)
Comitative 'with'	<i>kwa</i> (after closed syllable) <i>wa</i> (after open syllable and after <i>l</i> )
Connective 'and'	<i>kwa</i> (after closed syllable) <i>wa</i> (after open syllable and after <i>l</i> ) <i>imje</i> (after closed syllable) <i>mje</i> (after open syllable) <i>hóko</i> <i>ilang</i> (after closed syllable) <i>lang</i> (after open syllable)
Comparative 'than'	<i>lawa</i> <i>twukon, twukwun, tokon</i>
Ablative 'from'	<i>puthe</i> <i>Ulosje</i> (after closed syllable) <i>losje</i> (after open syllable)
Vocative	<i>ha</i> (honorific) <i>a</i> (plain; after closed syllable) <i>ja</i> (plain, after open syllable) <i>ije</i> (exclamatory; after closed syllable) <i>je</i> (exclamatory; after open syllable)

The accusative particle, typically marking an object, was frequently reduced to *l* after an open (vowel-final) syllable, as in *pójam i kachi l mul-e* (snake NOM magpie ACC bite-and; Yongka 7) 'A snake bit a magpie and'.

The locative/goal particle was often used to mark a comitative, as in *nala s malssóm i tjwungkwujk ej tal-a* (country GEN language NOM China LOC different-and.so) 'the national language is different from (that of) China, so'.

MK dative honorific *skej/skuj* was sometimes used to mark a comitative, as in *estjej sejcon skuj kót-cóo-lio* (how Sakya DAT same-OH-PROS) 'How could (he) be the same as Sakya?'

All MK source or dynamic locative particles except *sontój* 'from' derived from goal/static locative particles with the suffixation of the inceptive particle *-sje*.

- (19) a. *namcong-ljwukco skujsje na-lssój namin i-la hó-ni-la*  
 Namcong's-six.ancestors from born-as Namin be-DEC call-DF-DEC  
 'As (he) was born from Namcong's six ancestors, he was titled Namin.'  
 (Wenkak se 7)
- b. *sjo-jkejsje cjes na-ko cjec ejsje lak na-ko*  
 cow-from milk come.out-and milk from dairy.food come.out-and  
 'From cows comes milk, and from milk comes dairy food, and [. . .]' (Pepwha 5:155)
- (20) *motón acómój sontój tul-oni*  
 all aunt from hear-as  
 'As (I) hear from all aunts [. . .]' (Nayhwun 2:27)

The instrument/directional particle often replaced the accusative or comitative particle. Thus, in *nolaj lo pulu-zóp-nónila* (song with sing-OH-DEC) 'sang a song' (Welsek 2:17), the instrument particle *lo* was used instead of the accusative *lól*, and in *kumkang ólo kacólpi-si-ko* (diamond with compare-SH-and) 'compared with diamond' (Kumkang se 7), *ólo* was used instead of the comitative (*k*)*wa*.

The MK comitative/connective *kwa* was reduced to *wa* after vowels and *l*. It was generally attached to the last connected nominal as well.

- (21) *hój wa tól wa pjel wa-j ta pólk-ti ani hó-ja*  
 sun and moon and star and-NOM all bright-NMR not be-so  
 'The sun, the moon, and stars were not bright, so [. . .]' (Welsek 2:15)

Comparative particles *lawa* 'than' and *twukon/tokon* 'than' [< *twu-ko nun* (place-and TOP)] are illustrated in (22).

- (22) b. *sekun sjenpój lawa nu-to-ta*  
 rotten scholar than better-EXC-DEC  
 '(He) is better than a rotten scholar.' (Twusi 6:40)
- c. *etwum tokon na-uli-la*  
 darkness than better-PROS-DEC  
 '(It) will be better than darkness.' (Nokeltay sang 39)

*puthe* 'from' developed from *puth-* 'stick to, depend on' + the verbal connective suffix *-e* 'and'. It appears right after a nominal, as in *chezem puthe* 'from the beginning', after an accusative particle, as in *mózóm ól puthe* 'from one's heart', or after a directional particle, as in *hwiin ulo puthe* 'from the empress', or after the inceptive suffix *-s(j)e*, as in *elje-se puthe* 'from childhood'.

The compound particle *ulo-sje* (directional-inceptive) was used to denote source 'from', as in *wangsaseng ulosje* 'from Wangsa castle'.

A new pseudo-case particle, *tólje* 'to (a person)', developed from *tóli* 'accompany' + the verbal connective suffix *-e*. In *posal tólje niló-sja-tój* (Buddhist.saint to say-SH-that; Welsek 21:139) 'said to a Buddhist saint that', *tólj-e* lost its original meaning and obtained a pseudo-case meaning.

Case particles occurred in sequence, as in (*k*)*wa j* (connective + nominative/subject), *ulo puthe* (directional + ablative), and *lul puthe* (accusative/object + ablative); similarly a case particle with a delimiter, as in *ej-sje* (locative + inceptive), *uj-sje* (genitive + inceptive), *ulo-sje* (directional + inceptive), and *ej to* (locative + 'also').

- (23) *saks kwa saks kwa j psi lól puthé na-ko*  
 bud and bud and NOM seed ACC from come.out-and  
 ‘Buds come out from seeds and [. . .]’ (Wenkak sang 1.2:14)

*Delimiter particles*

Some delimiter particles were used after nominals and predicates, while others were used only with nominals. The former set was much more productive. Some frequently used ones are presented below.

a. Topic (TOP) particle *nón* (after an open bright-vowel syllable), *nun* (after an open dark-vowel syllable), *ón* (after a closed bright-vowel syllable), *un* (after a closed dark-vowel syllable). The first two were optionally reduced to *n*. The topic particle occurred after a nominal (24a), after a conjunctive suffix (24b), and after another non-topic particle.

- (24) a. *na nón icej silum i kiph-e*  
 I TOP now worry NOM deep-and  
 ‘I am full of worries now, and [. . .]’ (Welsek 2:5)
- b. *meli isj-e n po-zóp-ko kaskai w-a n mot*  
 far stay-by TOP see-OH-and near come-by TOP unable  
*po-zóβ-óli-le-la*  
 see-OH-PROS-RETR-DEC  
 ‘(We) can see (him) by staying far but are unable to see (him) by coming close.’ (Welsek 7:55)

b. The ‘toleration’ delimiter (*i*)za/(*i*)sa (derived from copula *i*/zero + verbal suffix *-za/-sa* ‘only if’) ‘only if it be, only for’: *iza/lisa* occurs after a consonant and *za/sa* after a vowel. It occurred after a nominal as in *mokswum iza* ‘as for only one’s life’. After a predicate stem, the form *-aza/-esa* occurred, where *-a/-e* is an infinitive (connective) suffix, as in *ka-sj-asa* (go-SH-INF.sa) ‘if only (he) goes’.

c. The ‘inceptive’ particle *sje* occurred after an overt or omitted case particle to give it an inceptive or dynamic meaning, as in *kózól tól alaj (aj)sje* (autumn moon bottom at) ‘under the autumnal moon’. Preceded by an infinitive suffix *-a/-e*, it functions as a predicate conjunctive suffix denoting cause or temporal sequence ‘as, so, and then’, as in *tajtongkang nepu-n ti moll-a.sje* (Taytong.River wide-REL fact unaware-so) ‘as (one) does not know how wide the Taytong River is’ (Akcang).

d. The particle (*k*)om ‘each’: Its use after a nominal is illustrated in *un-ton hón nat kom* (silver-money one unit each) ‘each one of silver coins’, *hón atól om* ‘one son each’, and *hón pen kom* (one time each) ‘each one time’. After an adverbial, it indicated emphasis, as in *noph-i kom* (high-ly EMP) ‘highly’ and *salóm ólo hój-e kom* (person with cause-by EMP) ‘causing people (to do)’.

e. The ‘limitation’ particles *pós* ‘only’ and (*k*)os ‘only’:

- (25) a. *skwum pós ani-mjen enu kilh ej tasi po-zóβ-li*  
 dream only not-if some road on again see-OH-PROS  
 ‘If it is not a dream, (we) will meet again on some road.’ (Welsek 8:82)

- b. *ocik mawang kos ce j cwa aj phjenanhi mot anc-wa*  
solely Satan only he GEN seat at comfortably cannot sit-and  
‘Only the Satan could not comfortably sit in his seat, and [. . .]’ (Welsek 2:42)
- c. *puthje os cwuksali l jehuj-sj-a*  
Buddha only life.death ACC lose-SH-and  
‘Only Buddha lost life and death, and [. . .]’ (Welsek 1:21)

f. The ‘inclusion’ particle *to* ‘also’, ‘alternative’ particle (*i)na* ‘or (something), rather’, and ‘focus/transference’ particle *taka* ‘while; and then’ (‘focus’ after the goal particle *ej* as in *ejtaka*; ‘transference’ after a predicate as in *noni-taka* ‘after strolling’) were all used after a nominal and a predicate in the same way as in CK.

The following delimiter particles are examples that were used only after a nominal.

g. The particles (*u)lan* ‘as for’ and (*u)llang* ‘and so on’ were used in the following contexts.

- (26) a. *tjoh-ón il lan na jkej ponaj-o*  
good-REL thing as.for me to send-IMP  
‘As for good things, send them to me.’ (Kumkang 21)
- b. *ing mwut-un cangku llang kaci-ko*  
moss stick-REL plough etc. carry-and  
‘carrying a moss-covered plough and so on, and [. . .]’ (Chengsan Pyelkok in *Akcang Kasa*)

h. Delimiter particles such as *cocha* ‘even, as well’, as in *pulhwuj cocha psu-ni-la* (root even use-DF-DEC) ‘Even the roots are used’ (Kumkang Samka 2:50), *man* ‘only’, and *mata* ‘every, each’, *kócang/skócang* ‘until, up to, even, to the extent of, to the utmost of’ (derived from *kócang* ‘extremity, edge’):

- (27) a. *stut skócang kólómh hanól hjanghó-no-la*  
mind/will up.to river sky face-IND-DEC  
‘Even (up to) the mind is looking out on the rivers and the sky.’ (Twusi 10:13)
- b. *mózóm skócang kongjangh-kej hó-si-ni*  
mind up.to offer.food.to.Buddha-to cause-SH-DEC  
‘(He) was caused to offer food to Buddha to his utmost.’ (Pulceng ha 9)

#### 4.4.2 Predicates

MK predicates comprised verbs, adjectives, and the copula, which shared many morphological properties, more so in MK than in CK. Some morphological factors that distinguish adjectives from verbs in CK were not as strict in MK, as seen in 4.4.2.2.

##### 4.4.2.1 Verbs

MK verbs can be subclassified into (a) transitive and intransitive, (b) independent and auxiliary, and (c) active, passive, and causative. (c) will be dealt with in 4.5.8.

Transitive verbs take an object, implicit or explicit. Examples are *pota* ‘see’, *makta* ‘block’, *kitulita* ‘wait’, *kilita* ‘praise’, *kojta* ‘love’, *kólóchita* ‘teach, point to’, *katota* ‘collect’, *mushta* ‘tie’, *mechwuta* ‘stop’, *alta* ‘know’, *molóta* ‘do not know’, *mastita* ‘deposit, give (a thing) to (a person’s) keeping’, *ilpezta* ‘steal’, *kacólpita* ‘compare’,

*keskta* ‘break, snap’, *kucihata* ‘limit’, *kujjihóta* ‘promise’, *paskota* ‘exchange’, *salanghóta* ‘think’, *sjehta* ‘erect’, *kwucjonghóta* ‘scold’, *hójta* ‘cause, order’, and *cwasita* ‘eat’ (honorific).

- (28) a. *pjengma lól mechwu-esi-ni*  
military.horse ACC stop-SH-DEC  
‘(He) stopped arms and horses.’ (Yongka 54)
- b. *na nón kutój lól al-enul*  
I TOP you (hon.) ACC know-now.that  
‘Now that I know you, [...]’ (Sohak 10:5)

Verbs unable to take an implicit or explicit object are intransitive verbs. Examples are *kata*, *njeta* ‘go’, *nazakata* ‘proceed’, *ancta* ‘sit’, *motta* ‘come together, swarm’, *mólóta* ‘dry up’, *kaskolta* ‘fall head first’, *kuchuta* ‘stop’, *nathota* ‘appear’, *njelumcista* ‘do farming’, *niluta* ‘say’, *tóojta*, *tóβójta* ‘become’, *tólita* ‘run’, *sulhta* ‘feel sad’, *pajta* ‘go to ruin’, *sjeta* ‘stand’, *sólota*, *sólpta* ‘tell, inform’ (to a senior), *cehta* ‘be scared’, *ptetita* ‘fall’, and *thóta* ‘burn’.

- (29) a. *kómól aj ani kuchu-lssój*  
drought by not stop-PROS  
‘(It) won’t stop by drought.’ (Yongka 2)
- b. *nulk-un hanapi tóoj-jejs-tota*  
old-REL grandpa become-PAST-EXC  
‘(He) became an old grandfather.’ (Twusi 21:31)

A small number of verbs are both transitive and intransitive, for instance, *kólita* ‘be sorted out; sort out’ (vs. CK *kalita* ‘sort out’), as in *seng i kólita* ‘sexes are distinguished’ and *li lól kalita* ‘sort out reasons’. Similarly, *tuliwuta* ‘hang down; hang something down’ (*tuliwuta* ‘hang something down’), and *tita* ‘fall down; cause to fall down’ (*cita* ‘fall down’).

There were a relatively small number of auxiliary verbs. Most auxiliary verbs derived from independent verbs, e.g. auxiliary *pota* ‘try, experience’ < independent *pota* ‘see, look at’. Independent and auxiliary uses coexisted in most cases, as they do in CK. Auxiliary verbs appeared after an independent verb, with or without an intervening connective suffix such as an infinitive suffix (-*a*/*e*/*ja*/*je*) or a nominalizer suffix -*ti*/*ke*.

- (30) *pota* ‘try, experience’: *hjej-je pota* ‘try to think over’, *ka pota* ‘try going’  
*pólita* ‘finish up, completely’: *nic-je pólita* ‘completely forgot’  
*kjesita* ‘stay, have’ (hon.): *wang i tóoj-ja kjesita* ‘became a king’, cf. *jetulp wangca lól twu kjesita* ‘has eight princes’  
*nata* ‘come out’: *pes-e nata* ‘get out of, escape’  
*najta* ‘take out’: *ciz-e najta* ‘make out, produce’  
*tita* ‘get to, become’: *sul-e tita* ‘disappear’  
*cita* ‘wish’: *kwukjenghó-ja cita* ‘wish to watch’, *ni-ke cita* ‘wish to go’  
*isita* ‘stay, have’: *nóm ój son ój cuj-je isita* ‘be dominated by others’, *hón mal to mot hó-ja is-tesita* ‘stayed without saying anything’  
*twuta* ‘keep, for later use’: *móngkól-a twuta* ‘make for later use’  
*kata* ‘continue, gradually’: *hój ti-je kata* ‘the sun is setting’  
*malta* ‘stop’: *i ptut ul nis-ti mal-ósjosje* ‘don’t forget this intent’  
*ani-hóta* ‘be/do not’: *tini-ti ani-hóta* ‘do not carry’

New verbs were formed through compounding or affixation. Compounds included *cwuk-salta* (die-live) ‘live and die’, *cwuk-pajta* (die-perish) ‘die and perish’, *oló-nólita* (rise-descend) ‘go up and down, rise and fall, fluctuate’, *kams-tolta* (wind-turn) ‘circle around’, *nje-kata* (go-go) ‘drop in’, *tik-mekta* (dip-eat) ‘dip and eat’, *pte-tita* (crack-become) ‘burst’, *nil-e-sjeta* (rise-INF-stand) ‘stand up’, *tjangka-tulta* (wife’s.house-enter) ‘marry, take a wife’, *alph-sjeta* (front-stand) ‘go before’, and *tat-najta* (separately-put.out) ‘separate’. Affixation was not very productive, some examples being *pi-wusta* ‘laugh scornfully’, *es-makta* (aslant-block) ‘block askew’, *nep-nolta* (wide-play) ‘stroll around’, and *naj-thita* ‘throw away’. Zero derivation from nouns includes *kómólta* ‘be arid’ from *kómól* ‘drought’, *kutta* ‘cut’ from *kut* ‘end’, *nechwulta* ‘dangle down’ from *nechwul* ‘vine’, *stujta* ‘wear a belt’ from *stuj* ‘belt’, and *pójta* ‘conceive’ from *pój* ‘belly’.

#### 4.4.2.2 Adjectives and the copula

Examples of adjectives are *kapójjafta* ‘light’, *kantafta* ‘simple’, *kóskós(hó)ta* ‘clean’, *koj.ojhóta* ‘quiet’, *kakókta* ‘urgent’, *tes-kwucta* ‘worse’, *ojta* ‘wrong’, *cephuta* ‘scared’, *kazómjeta* ‘well-off’, *cjeta* ‘small (size, quantity)’, *spospohóta* ‘pointed’, *phelehóta* ‘blue, green’, *wuzfta* ‘funny’, and *ejkwutta* ‘very bad’. Some words had meanings different from corresponding CK ones, as in MK *ejesputa* ‘pitiful’ > *yeypputa* ‘pretty’.

In CK, the surest morphological criterion to distinguish adjectives and verbs is to see their non-past indicative form. Verbs have the indicative suffix **-n** or **-nun**, as in **ka-n-ta** ‘goes’ and **mek-nun-ta** ‘eats’, whereas adjectives have zero, as in **coh-ta** ‘is good’ and **huj-ta** ‘is white’. MK adjectives were more like verbs than CK adjectives are in that many (not all) MK adjectives inflected exactly the same way as verbs in their indicative forms. Thus, many adjectives were followed by the indicative suffix *-no* like verbs:

- (31) a. *koc tjoh-ko jelum ha-nó-ni*  
 flower good-and fruit much-IND-DEC  
 ‘Flowers are good and fruits are plentiful.’ (Yongka 2)
- b. *pólóm to ep-nó-ni-la*  
 wind also not.exist-IND-DF-DEC  
 ‘There was even no wind.’ (Welsek 1:50)
- c. *kwulwum i etwup-nó-ta*  
 cloud NOM dark-IND-DEC  
 ‘Clouds are getting dark.’ (Nungem 4:44)

Unlike in standard CK, some adjectives had the same indicative relative suffix as that of verbs, as in (32a). This practice is still observed in contemporary Kyengsang dialect. Also, unlike in standard CK, the indicative relative suffix of the existential adjective *epsta* ‘not exist’ was *-un*, as in (32b; see also 31b).

- (32) a. *etuketukhó-nón naco ój*  
 dusky-REL evening in  
 ‘in a dusky evening’ (Twusi 1:28)
- b. *kuci eps-un mimjohó-n ptut*  
 end lack-REL subtle-REL meaning  
 ‘endless subtle meanings’ (Welsek se 8)

Another feature is that causativization of MK adjectives into transitive verbs by means of causative suffixes *-i*, *-hi*, *-ki*, *-o*, *-wu*, etc. was much more productive than in CK. Thus, while causativized adjectives such as *pólkita* ‘brighten’ (**palkhita**), *nuchita* ‘delay’ (**nucchwuta**), *kwuthita* ‘harden’ (**kwuthita**), and *epsiwuta* ‘eliminate’ (**epsayta**) have been inherited to CK, words such as *njethota* ‘make it shallow’, *kiphita* ‘deepen’, *kiskita* ‘make happy’, and *cehita* ‘terrify’ are expressed only with periphrastic causative constructions in CK.

The copula root was *i*. It always took *-la* as its declarative ender in MK. It reduced to *j*- or zero after a vowel (33b, c). The copula was frequently omitted before the question ender *-ko* (question word question) or *-ka* (polar question) (34).

- (33) a. *kujk ón nalah i-la*  
country TOP country is-DEC  
‘*Kujk* is (means) a country.’ (Cengum Enhay)
- b. *swujko nón sajng kwa lo wa pjeng kwa sa*  
suffering TOP life and oldness and sickness and death  
*wa-j-la*  
and-be-DEC  
‘*Swujko* is (means) living, being old, being sick, and death.’ (Sekpo 6:4)
- c. *chi nón ni la*  
tooth TOP tooth be.DEC  
‘*Chi* is (means) a tooth.’ (Cengum Enhay)
- (34) a. *hjenljang ón sto mjes salóm ko*  
wise.person TOP also how.many person Q  
‘Also, how many wise people were there?’ (Twusi 19:10)
- b. *ku nil-on kes un musu ko*  
that arise-REL thing TOP what Q  
‘What was that which happened?’ (Wenkak se 12)
- c. *i twu salóm i cinsillo ne j hangkes ka*  
this two person NOM truly you GEN master Q  
‘Are these two persons truly your masters?’ (Welsek 8:94)

The verb *tóojta* ‘become’ took both the subject and complement in the nominative case.

- (35) *ku o-pajk salóm i cejca j tóoj-aci-ngi-ta*  
that 500 people NOM disciple NOM become-wish-AH-DEC  
‘Those 500 people wished to become his disciples.’ (Welsek 1:9)

#### 4.4.2.3 Inflection

Verbs, adjectives, and the copula function as sentential or clausal predicates. All predicates inflected with suffixes, which consist of prefinal suffixes that denote tense, aspect, modality, and honorification (subject, object, addressee), and final suffixes that denote clause or sentence type. While prefinal suffixes are optional in occurrence, clause- or sentence-type suffixes, which occur at the end of a clause or sentence and are thus termed enders, are obligatory. As in CK, sentence enders were of four major

types: declarative, interrogative, imperative, and propositive. In addition, promissive, exclamatory, and desiderative enders were also available. Clause enders are sub-classified into conjunctive, relative (or adnominal), complement (or adverbial), nominalized, and quotative. Let us briefly observe only tense, aspect, and modality elements and the declarative enders below just for illustration of the complexities involved, relegating detailed discussions of these and other inflectional elements to appropriate sections in syntax.

Tense, aspect, and modality (mood and modal) were not always clearly demarcated in form but often portmanteaued in the same forms. Thus, linguists do not always agree with each other in categorizing these suffixes, naming, for example, the suffix *-nó* a present tense marker or an indicative mood marker (e.g. Huh 1983; K. Kang and Hwang 2003; Ko 2007). The following classification is made in this chapter. The tense category consisted of past and non-past. The past tense was marked by *-ais/-eis*, *-jajs/-jejs*, *-ajs/-ejs*. As these forms had diverged and grammaticalized from the resultative compound *-a/-e isi* (infinitive + ‘exist’), they retained resultative or perfective aspect meaning as well. Past time was also indicated by the confirmative modal *-ke/-kwa/-ka*. Non-past, indicating present and future time, was unmarked as in CK.

The mood category, which often indicates tense function as well, consisted of (a) indicative (often referring to present events) *-nó/-ni/-nón/-no* (*-no* = *-nó* + speaker subject *-o*; *-n* changed to *-l* after prospective *-li*), (b) retrospective (referring to the speaker’s past observation or experience) *-tó/-ta/-le/-la*, (c) prospective (frequently referring to future events) *-(u)li/-(u)lh*, and (d) prospective-retrospective *-(u)li-le*. Modal suffixes, which appeared immediately before the sentence ender, include the definitive (DF) *-n(i)* (as in *-nila*), exclamatory *-(t)o(s)* (as in *-tota*), confirmative *-ke/-kwa/-ka*, and speaker subject marker *-o/-wu* (Huh 1983: 436–7).

The declarative enders were: (a) *-ta/-la* and (b) *-ni-la* (DF *-ni* + DEC *-la*). Compared to *-ta/-la*, *-ni-la* was used in a more definitive, authoritative, and conservative manner. In addition, there were neutral-level declarative/interrogative enders *-ni* (< subordinator *-ni* (?)) and *-li* (< prospective *-li* (?)). These enders were used either as they were or preceded by other prefinal suffixes such as the indicative, retrospective, prospective, past, confirmative, and speaker-subject marker. For instance, the predicate *hóta* ‘do, be’ inflected as *hó-ta*, *hó-nó-ta*, *hó-te-la*, *hó-li-la*, *hó-jajs-ta*, *hó-ke-ta*, and *hó-la*. Notice that *-ta* changed to *-la* after *-te*, *-li*, and *-o*. Similarly, the definitive declarative ender *-ni-la* was used right after a predicate root or preceded by a prefinal suffix, as in *hó-ni-la*, *hó-nó-ni-la*, *hó-te-ni-la*, and *hó-li-ni-la*.

#### 4.4.3 Numerals

There were two numeral systems in MK: the native system and the SK system. Numerals in the native system were available up to 1,000. Above that number, SK numerals were used. As SK numerals were similar to those in CK, suffice it to overview native ones. MK native cardinal numbers are as follows: *hónah* ‘1’, *twulh* (< *twuβulh*) ‘2’, *sejh* ‘3’, *nejh* ‘4’, *tasós* ‘5’, *jesus* ‘6’, *nilkwup* ‘7’, *jetulp* ‘8’, *ahop* ‘9’, *jelh* ‘10’, *sumulh* ‘20’, *sjelhun* ‘30’, *mazón* ‘40’, *sujn* ‘50’, *jesjwujn* ‘60’, *nilhun* ‘70’, *jetun* ‘80’, *ahón* ‘90’, *on* ‘100’, *cumun* ‘1,000’, *jeleh* ‘many’. As determiners, *hónah* ‘1’, *twulh* ‘2’, *sejh* ‘3’, and *nejh* ‘4’ were reduced to *hón*, *twu*, *se(j)/sek*, and *ne(j)/nek*, respectively: *hón sal* ‘one year old’, *twu kulwel* ‘two sentences’, *sej sal* ‘three years old’, *se mal* ‘three mals (unit of measure)’, *sek tól* ‘three months’, *nej salóm* ‘four persons’, *ne mal* ‘four mals’, and *nek tól* ‘four months’.

- (36) a. *sajlo sumul-jetulp ca lól mójngkó-noni*  
 newly 28 letter ACC make-as  
 ‘As (I) have made newly 20–8 letters, [. . .]’ (Cengum Enhay)
- b. *tól i cumun kólóm aj pichujjom i kóth-ónila*  
 moon NOM 1,000 river at shining NOM same-DEC  
 ‘The moon seems to be shining in 1,000 rivers.’ (Welsek 1:1)

Approximate numbers were expressed by compounding neighboring numerals. Examples of numbers up to ten are: *hón-twu* ‘one or two’, *twu-ze* ‘two or three’, *se-ne* ‘three or four’, *ta-jes* ‘five or six’, *jej-nilkwup* ‘six or seven’, *nil-jetulp* ‘seven or eight’, and *twu-e-jel* ‘scores’. Ten or more approximate numbers were expressed by suffixing the phrase *namón* (< *nam-* ‘remain’ + relativizer *-ón*), as in *je-namón* ‘about ten’, *sumu-namón* ‘about 20’, *sjelhu-namón* ‘about 30’, *mazó-namón* ‘about 40’, and *swuj-namón* ‘about 50’. Also used were multiplied compound numerals, as in *twu-nej* ‘8’, *twu-jesus* ‘12’, *twu-nilkwup* ‘14’, *nej-nilkwup* ‘28’.

Ordinal numbers consisted of a cardinal number + *chahi* (also, *chaj*, *cha*, *cahi*, *caj*, *ca*, etc.): *hóna-chahi* ‘first’, *twul-chahi* ‘second’, *sej-chahi* ‘third’, *jel-chahi* ‘tenth’, *sumul hóna chahi* ‘twenty-first’, and *mazón-chahi* ‘fortieth’.

## 4.5 SYNTAX

MK syntax was largely the same as CK syntax. It had a predicate-final, post-positional, and head-final structure. Sentence and clause types were essentially the same in MK as in CK. Some quantative differences are: (a) in MK, determiners were not well developed; (b) nominal constructions were used much more extensively in MK. Thus, many predicative constructions including interrogatives consisted of nominalized forms; (c) compound or complex sentences were much more extensively used in MK; (d) in direct and indirect quotations, quotative particles (CK *ko* or *lako*) were not used in MK.

### 4.5.1 Noun phrase structure

Noun phrases (NPs) functioned as major (subject or object) or minor (oblique cases and adjuncts) nominal arguments of sentences and clauses. As such, they took a wide variety of particles that indicated the grammatical and semantic relations of the NPs vis-à-vis other parts of the sentences or clauses. NPs were of the structure:

- (37) (Relative Clause) (Genitive NP) (Determiner) HEAD (Plural) (Particles)

The head, the only obligatory element, is occupied by a simple or compound nominal (a noun, a pronoun, a numeral, or a nominalized word). A relative clause, a genitive NP, and a determiner are all optional and their order may be scrambled for stylistic or pragmatic purposes if they co-occur.

As indicated in 4.4.1.2, two different particles, *ój/uj* and *s*, were used to mark the genitive. *ój* (after a bright vowel) or *uj* (after a dark vowel) was used with nonhonorific animate nouns as in (38a, b). The particle *-s* followed honorific nouns (39a) or inanimate nouns (39b) (cf. I. Lee and Ramsey 2000: 290–1). The bracketed constructions are NPs, in which bold-faced parts are genitive constructions. A genitive could be modified by another genitive recursively (40).

- (38) a. [salóm ój ptut ul] ‘people’s wishes’ (object)  
 person GEN wish ACC  
 b. [kepwup uj theli wa] ‘turtle’s hair and’  
 turtle GEN hair and
- (39) a. [sejcon s sinljek ulo] ‘with Sakyamuni’s divine power’  
 Sakyā GEN divine.power with  
 b. [nala s malssóm i] ‘the country’s language’ (subject)  
 country GEN language NOM
- (40) ku točók i [posal s cen-sejsajng s wenswu]  
 that thief NOM Buddhist.saint GEN former-life GEN enemy  
 -j-le-la  
 be-RETR-DEC  
 ‘That thief was the Buddhist saint’s former life’s enemy.’ (Welsek 1:6)

Determiners were the least developed word class in MK. CK pure determiners were still generally nouns, pronouns, relativized verbs or adjectives in MK, which *also* syntactically functioned like determiners. For example, CK demonstrative determiners were demonstrative *pronouns* in MK and were also used like determiners, e.g. *i kes* ‘this thing’ and *ce kes* ‘that thing over there’. *saj* was an independent noun ‘new thing’, yet functioned like a determiner ‘new’ in *saj kwusul* ‘new beads’. Indefinite *enu* was a pronoun ‘which one’ and took case particles. It was used like a determiner in constructions such as *enu nala* ‘which country’. The determiners *oón* ‘all the’ (*oón nala* ‘all the country’), *motón* ‘every’ (*motón hakca* ‘every scholar’), and *ojn* ‘wrong’ (*ojn jak* ‘wrong medicine’) were relativized forms of the adjective *oólta* ‘be all’, the verb *motta* ‘gather’, and the adjective *ojta* ‘be wrong’, respectively. Thus, only a handful of fossilized determiners included *cjokomas* ‘a small bit’ (**cokoman**), *njenamón* ‘some ten odd’ (**jenamun**), and *ptón* ‘different’ (**ttan**). All numerals also functioned like determiners when they preceded nouns, as they are today.

Plain (nonhonorific) nouns and pronouns were optionally followed by the plural suffix *-tólh*, as in *i salóm-tólh* ‘these people’, *cwungsajng-tólh* ‘human beings’, *wuli-tólh* ‘we’, and *nehuj-tólh* ‘you all’. Honorific nouns and pronouns were pluralized by the plural suffix *-naj* (CK **-nay** means ‘the family of’), as in *emanim-naj* ‘mothers’, *nwuynim-naj* ‘sisters’, *kutuj-naj* ‘you all’ (hon.), *cókjas-naj* ‘they’ (hon.), *pwuin-naj* ‘ladies’.

#### 4.5.2 Pronouns and anaphora

MK pronouns are classified into personal, demonstrative, indefinite-interrogative, and reflexive. Among personal pronouns, unlike in CK, there was no separate humble form for the first person. Thus, *na* covered both CK **na** ‘I’ (plain) and **ce** ‘I’ (humble) as in (41), where *na* is used even to the king.

- (41) tajwang ha estjej na lol moló-si-nó-ngi-sko  
 Great.King VOC why me ACC not.know-SH-IND-AH-Q  
 ‘Oh the Great King, why don’t you know me?’ (Welsek 8:92)

The second person had plain *ne* and honorific *kuti/kutuj* ‘you’, also used to an adult equal. The third person had plain *tje/ce* ‘he, she’, also used as a demonstrative pronoun (‘that over there’), and honorific *cókja*, also used as a reflexive pronoun ‘self’. This

honorific form developed from Chinese 自家 ‘self’s house’, but degraded to the plain level after the sixteenth century (cf. K. Kang and Hwang 2003: 69) and then disappeared. The third person *ce* was also used as a reflexive pronoun. Plural forms were *wuli(-tólh)* ‘we’, *njehuj(-tólh)* ‘you (pl.)’, and *tje-tólh* ‘they’.

Demonstrative pronouns were as in Table 4.3. Demonstratives were both deictic and anaphoric, with *ku* and *ku* compounds used most frequently as anaphors (42).

- (42) a. *na j i lól wujhój-a ejespi nekj-e*  
I NOM this ACC for-INF pitifully think-and  
‘As I felt pitiful about this, [. . .]’ (Cengum Enhay 2)
- b. *ku nón kóccang swuj.wu-ni-la*  
that TOP most easy-DF-DEC  
‘That is the easiest.’ (Pakthongsa sang 48)

Indefinite-interrogative pronouns covered persons, things, quantity, place, and time (Table 4.4 and (43)). Unlike *mjes* ‘how many, several’, which was used as both a pronoun and a determiner, *hjen* was used only as a determiner (44). *enu* (and its allomorphs) were used not only as a determiner (45a) but also as an adverb in the sense of ‘how’ (45b).

- (43) a. *sa j mwul-otój pwuthje j nwu-ko*  
teacher NOM ask-that Buddha NOM who-Q  
‘The teacher asked, “Who is Buddha?”’ (Welsek 21:195)
- b. *myes kan-t cip uj saló-si-li-ngi-sko*  
how.many room-GEN house in live-SH-PROS-AH-Q  
‘A house of how many rooms would you live in?’ (Yongka 110)
- c. *enma-s pok ól tukhó-li-ngi-sko*  
how.much-GEN blessing ACC get-PROS-AH-Q  
‘How much blessing is (he) going to get?’ (Welsek 17:44)

TABLE 4.3 LATE MIDDLE KOREAN DEMONSTRATIVE PRONOUNS

	Proximate to speaker	Proximate to hearer	Distal
	<i>i</i> ‘this’	<i>ku</i> ‘that near you; that we know’	<i>tje</i> ‘that over there’
Thing (dem + bound noun <i>kes</i> ‘thing’)	<i>i-kes</i>	<i>ku-kes</i>	<i>tje-kes</i>
Place (dem + noun <i>ekuj</i> ‘place’) <sup>11</sup>	<i>i-ngekuj</i> (> <b>yeki</b> )	<i>ku-ngekuj</i> (> <b>keki</b> )	<i>tje-ngekuj</i> (> <b>ceki</b> )
Time (dem + bound noun <i>cjej</i> ‘time’)	<i>i-cjej</i> ‘now’	<i>ku-cjej</i> ‘that time’	<i>tje-cjej</i> ‘that time’

TABLE 4.4 LATE MIDDLE KOREAN INDEFINITE-INTERROGATIVE PRONOUNS

Referents	Forms
Persons	<i>nwu/nwuko/nwukwu</i> ‘who, someone’, <i>amo</i> ‘any’
Things	<i>musu/muzus/musu/musuk/musum</i> ‘what, some’, <i>amo</i> ‘any’, <i>enu/enwu/enó</i> ‘which, certain’
Quantity	<i>mjes</i> ‘how many, several’, <i>hjen</i> ‘how many, several’, <i>enma</i> ‘how much, some amount’
Places	<i>etuj</i> ‘where, somewhere’, <i>etumej</i> ‘where, somewhere’, <i>etule</i> ‘to where, to somewhere’
Time	<i>enu-cjej</i> ‘when, some time’

- d. *i póli-ko etuj puth-uli-o*  
 this abandon-and where stick-PROS-Q  
 ‘Where is (he) going to stick to, while abandoning this?’ (Welsek se 15)
- e. *et-non yak i musu kes-ko*  
 get-REL medicine NOM what thing-Q  
 ‘What medicine are you getting?’ (Welsek 21:215)
- f. *musus il ól kjes-ko o-lje hó-nón-ko*  
 what thing ACC experience-and come-to do-IND-Q  
 ‘What kind of things are you going to do and come back?’ (Sekpo 6:27)
- g. *musuk i eljeβ-ulj-o*  
 what NOM difficult-PROS-Q  
 ‘What would be difficult?’ (Sekpo 11:19)
- h. *musum liik isi-li-o*  
 what advantage exist-PROS-Q  
 ‘What kind of advantage will there be?’ (Kumkang 64)
- (44) a. *sa-co j phjenanhi mot kje-sj-a hjen kot ól*  
 4-ancestor NOM peacefully unable stay-SH-and how.many place ACC  
*olma-si-njo*  
 move-SH-Q  
 ‘How many places did the four ancestors move around without staying peacefully?’ (Yongka 110)
- b. *hjen pen pthujwu-ntól nómi i oló-li-ngi-ska*  
 several time jump-even.if others NOM climb-PROS-AH-Q  
 ‘Would others go up even if they jump several times?’ (Yongka 48)
- (45) a. *enó nal aj pi kaj-ja*  
 certain day on rain clear-so  
 ‘As the rain was over one day, [. . .]’ (Twusi 12:34)
- b. *kujkin-s ptut ul enu ta sólβ-óli*  
 people-GEN wish ACC how all say-PROS.Q  
 ‘How can (I) say all the wishes of the people!’ (Yongka 118)

#### 4.5.3 Quantification

Quantification was expressed by numerals, native or SK. First, numerals directly precede countable nominals as determiners: *sek ca* ‘three inches’, *nej salóm* ‘four persons’, *nek tól* ‘four months’, *sumul jetulp ca* ‘28 letters’, *nilhun sal* ‘70 years old’, *sjelhu-namón hój* ‘three odd years’, *cumun kólóm* ‘1,000 rivers’, *o-pajk sólóm* ‘500 people’, and *il-chen nalah* ‘1,000 countries’. Nominals were also quantified by means of the construction numeral + classifier + nominal as in [*se mal swul*] *ul mek-ko* ‘drinking three *mals* of liquid’ (Twusi 15:40) or nominal + numeral + classifier as in [*swul se mal*] ‘three *mals* of liquid’. Various quantifiers were used as determiners to quantify the following nominals, as in *motón hakca* ‘all scholars’, *oón nalah* ‘all the country’. Interrogative-indefinite quantifiers such as *mjes* ‘how many, several’, *enma* ‘how much, how many’, *hjen* ‘how many, several’ were used as determiners, as in *mjes kan-t cip* ‘house of how many rooms’, *enma-s pok* ‘how much blessing’, *hjen pen* ‘how many times’. The particle *mata* suffixed to a countable nominal indicates each and every member of the set denoted by the nominal.

- (46) *nip alaj mata hanól s tongnam i is-nó-ngi-ta*  
 leaf under every heaven GEN baby.boy NOM exist-IND-AH-DEC  
 ‘Under every leaf is a heavenly baby boy.’ (Sekpo 11:32)

Finally, days are counted by special native forms: *hóló* ‘1 day’, *ithul* ‘2 days’, *saól* ‘3 days’, *naól* ‘4 days’, *tassswaj* ‘5 days’, *jessswaj* ‘6 days’, *niłwej* ‘7 days’, *jetólaj* ‘8 days’, *ahólaj* ‘9 days’, *jólhul* ‘10 days’.

#### 4.5.4 Sentence word order

All MK sentences are canonically predicate-final. Thus, all other sentential constituents are pre-predicate. The typical order is subject-initial. The subject may be followed by one or more nominals indicating various cases, copular complement, or adjuncts. Time and place nominals frequently precede subjects. Adverbials may be appropriately inserted before or after these nominal elements. Word order was not affected by sentence types such as interrogatives.

- (47) a. [subject + copular complement]  
*i nón wuli hemul i-la*  
 this TOP our fault be-DEC  
 ‘This is our fault.’ (Pepwha 2:5)
- b. [subject + object]  
*na j thajca lól sjemki-zóβó-tój*  
 I NOM prince ACC serve-OH-but  
 ‘Although I serve for the prince, [. . .]’ (Sekpo 6:4)
- c. [subject + indirect object + direct object]  
*na j hóma cjwungsajng ójkungej culkeβ-un kes ul*  
 I NOM already public to happy-REL thing ACC  
*phosiho-tój*  
 disseminate-as  
 ‘I already disseminated happy things to the general public, such as [. . .]’  
 (Sekpo 19:3)
- d. [time + subject + object]  
*hwu ej półóm i mul wuh ul pwul-e*  
 later at wind NOM water top ACC blow-and  
 ‘At a later time, the wind blew the water surface, and [. . .]’ (Welsek 1:39)
- e. [instrument + object] (interrogative)  
*musu kes ulo to lól sam-alj-o*  
 what thing with doctrine ACC make-PROS-Q  
 ‘With what can we make doctrines?’ (Welsek 9:22)

The typical order of these major nominal elements is scrambled rather freely for stylistic or focusing or topicalizing purposes.

- (48) a. [object + subject]  
*pwuthje s kujjel ul cisin i nił-enul*  
 Buddha GEN news ACC Earth.God NOM tell-now.that  
 ‘Now that the Earth God is telling the news about Buddha, [. . .]’ (Welin  
 ki 82)

- b. [direct object + indirect/dative object]  
*tjohón il lan na jkej ponaj-o*  
 good thing as.for me to send-IMP  
 ‘As for good things, send them to me.’ (Kumkang 21)

Sentential elements are freely omitted when recoverable from sentential or situational contexts. Thus, a great many sentential utterances are without a subject, object, or other major argument. Moreover, if no ambiguity ensues, nominal particles are also omitted extensively, leaving the co-occurring arguments alone.

As in CK, double or multiple subject constructions extensively occur in MK, where a higher subject and a lower subject appear in that sequence.

- (49) a. *mól i pjeng i kiph-e*  
 horse NOM sickness NOM deep-so  
 ‘As the horse is seriously ill, [. . .]’ (Yongka 109)
- b. *salóm i sa-hajng i kong i ilu-si-mjen*  
 person NOM four-deeds NOM merit NOM achieved-SH-if  
 ‘If the merit of the people is achieved in four moral deeds, [. . .]’ (Pephwa 5:22)

#### 4.5.5 Sentence types

There are four major sentence types: declarative, interrogative, imperative, and propositive. In addition, some less canonical sentence types such as promissive, exclamatory, and desiderative occur. All these sentence types are marked by respective sentence enders.

##### 4.5.5.1 Declarative sentences

Declarative sentence enders are: (a) *-ta/-la* and (b) *-ni-la* (definitive). As indicated in 4.4.2.3, these enders are used either as they are or preceded by other prefinal suffixes such as the indicative *-nó*, retrospective *-te*, prospective *-li*, past *-jajs/-jejs*, confirmative *-ke/-kwa/-ka*, speaker-subject marker *-o*, addressee honorific *-ngi*, etc. Observe the declarative sentence examples in (50), which are uttered in interactive situations.

- (50) a. *nilkwup hój nemwu ola-ta*  
 7 year too long-DEC  
 ‘Seven years is too long.’ (Welsek 7:2)
- b. *na j icej ne j em(i) uj ka-n stah*  
 I NOM now you GEN mom GEN go-REL place  
*ól poi-o-li-la*  
 ACC show-SS-PROS-DEC  
 ‘Now I will show you the place your mom went to.’ (Welsek 21:21)
- c. *pulhwuj cocha psu-ni-la*  
 root even use-DF-DEC  
 ‘Even the roots are used.’ (Kumkang Samka 2:50)
- d. *na j kulen ptut ul moll-a hó-ta-jng-ta*  
 I NOM such meaning ACC not.know-INF do-RETR-AH-DEC  
 ‘I did not know such a meaning.’ (Sekpo 24:32)

- e. *nip alaj mata hanól s tongnam i is-nó-**ngi-ta***  
 leaf under every heaven GEN baby.boy NOM exist-IND-AH-DEC  
 ‘Under every leaf is a heavenly baby boy.’ (Sekpo 11:32)
- f. *sengca j na-sj-a cengkak ilwu-si-li*  
 holy.son NOM born-SH-and enlightenment achieve-SH-PROS  
 ‘A holy son will appear and achieve enlightenment.’ (Welin ki 15)

Sentences (50a, b, c) are of the plain level, (50d) is a low-polite level, (50e) is a deferential level, and (50f) is a half-talk style without a declarative ender.

#### 4.5.5.2 Interrogative sentences

Interrogative sentence enders are: (a) *-ko* (*-sko* after the addressee honorific suffix *-ngi*), (b) *-ka* (*-ska* after the addressee honorific suffix *-ngi*), and (c) *-ta*. *-(s)ko* occurs in question-word question sentences (51), and *-(s)ka* in yes-no (polar) questions (52).

- (51) a. *ezi atól ipjel i este-n-ko*  
 parent son separation NOM how-IND-Q  
 ‘How is the parting of parents and children from each other?’ (Welin 144)
- b. *sejcon i hóló mjes li lól nje-si-nóni-**ngi-sko***  
 Sakya NOM a.day how.many ri ACC go-SH-IND-AH-Q  
 ‘How many *ri*’s does Sakya go a day?’ (Sekpo 1:39)
- (52) a. *pil-on pap ól estjej cwa-si-nón-ka*  
 beg-REL rice ACC how eat-SH-IND-Q  
 ‘How can you eat the rice you begged?’ (Welin 122)
- b. *hanapi l mit-uni-**ngi-ska***  
 grandfather ACC believe-IND-AH-Q  
 ‘Do you believe the grandfather?’ (Yongka 125)

Both *-ko* and *-ka* are also attached directly to a copula complement, in which case the enders may be considered particles rather than suffixes. In this case, both the copula and its mood suffix are omitted. Sometimes *-ko* and *-ka* are reduced to *-o* and *-a*, respectively, especially after a vowel or *l*.

- (53) a. *hjenljang ón sto mjes salóm ko*  
 wise.person TOP also how.many person Q  
 ‘Also, how many wise people were there?’ (Twusi 19:10)
- b. *i twu salóm i cinsillo nej hangkes ka*  
 this 2 person NOM truly your master Q  
 ‘Are these two persons truly your masters?’ (Welsek 8:94)
- c. *i nón sang ka pel a*  
 this TOP prize Q punishment Q  
 ‘Is this a prize or a punishment?’ (Mongsan 53)

After the mood suffixes *-ni* (indicative) and *-li* (prospective), the enders *-ko* and *-ka* lost the consonant *k*, and became new question enders *-nio/-njo* and *-nia/-nja/-nje* on the one hand and *-lio/-ljo* and *-lia/-lja/-lje* on the other.

- (54) a. *nuj hóollo sulh-e hó-nó-nio*  
 who alone sad-INF do-IND-Q  
 ‘Who alone feels sad?’ (Twusi 19:40)
- b. *sto eljep-ti ani hó-nia*  
 again difficult-NMR not be-Q  
 ‘Isn’t it difficult again?’ (Twusi 10:42)
- c. *i tajsicwu uj tukhó-n kongtek i ha-nje cjek-unje*  
 this great.offerer GEN obtain-REL merit NOM big-Q small-Q  
 ‘Is the Buddhist merit that this great offerer has achieved big or small?’  
 (Welsek 17:48)
- d. *hemwul eps-uni etuli najthi-ljo*  
 flaw lack-because where.to desert-Q  
 ‘As (he) is flawless, where can (you) desert (him)?’ (Welsek 2:6)
- e. *toj uj mokswum i ku nunghi ola-lia*  
 barbarian GEN life NOM that possibly long-Q  
 ‘Would the barbarian’s life be possibly long?’ (Twusi 1:8)

The question ender *ta* is frequently used in a sentence whose overt or understood subject is a second person. Before *ta*, a relativizer suffix may appear as if *ta* were a bound noun, hence the space before it in (55). In this respect, it is distinguished from the declarative ender *-ta*. An emphatic counterpart of this ender is *són.ta* (87b).

- (55) a. *sjang i niló-sj-atój towen-kwun ani ta*  
 king NOM tell-SH-that Town-prince not be.Q  
 ‘The king asked, “Aren’t you Prince Town?”’ (Kumkang sasil 3)
- b. *kutuj nón enu cek-uj tolao-l ta*  
 you TOP which time-at return-PROS Q  
 ‘When will you return?’ (Twusi 22:30)
- c. *ne j esti amh ól naj-ja cwu-ti ani hó-nó-n ta*  
 you NOM why female ACC take.out-INF give-NMR not do-IND-REL Q  
 ‘Why don’t you give out the woman?’ (Welsek 7:17)

The lower-polite level examples, in which the addressee honorific suffix is *-ni*, are few. The two examples in (56) show that there is no distinction between the question word and the yes–no question enders, both appearing as *-ska* (Ko 2007: 339–40).

- (56) a. *cwuín i musum chapán ól sonzo tótñj-e mǒngkó-nó-ni-ska*  
 owner NOM what food ACC in.person go-and make-IND-AH-Q  
 ‘As owner, what food are you making yourself by moving around?’  
 (Sekpo 6:16)
- b. *kutuj s apenim i is-nó-ni-ska*  
 you GEN father NOM exist-IND-AH-Q  
 ‘Do you have father?’ (Sekpo 6:14)

The half-talk (neutral) style occurs extensively in writing. This style does not have a question ender.

- (57) a. *nimkum s malssóm i ku j ani olh-ósi-ni*  
king GEN words NOM it NOM not right-SH-IND.Q  
'Are the king's words not right?' (Yongka 39)
- b. *icej estjej wenswu lól nic-ósi-nó-ni*  
now how enemy ACC forget-SH-IND-AH.Q  
'Now how can you forget the enemy?' (Sekpo 11:34)
- c. *sengin sinljek ul enó ta sólfó-li*  
saint divine.power ACC which.one all narrate-PROS.Q  
'Which divine powers can (we) narrate all?' (Yongka 87)

#### 4.5.5.3 Imperative sentences

Imperative sentence enders are *-(a/e)la* (plain-level), *-(a/e)ssje* (low-polite level), *-(ó/u)šjosje* (deferential level), and *-kola/-kolje* (half-talk style) (58).

- (58) a. *wu-nón sengnje ije sulh-e mal-la*  
cry-REL saintess VOC sad-INF stop-IMP  
'Oh you crying saintess! Don't be sad!' (Welsek 21:21)
- b. *ne j wen-tahi hój-ala*  
you GEN wish-as do-IMP  
'Do as you wish.' (Sekpo 24:14)
- c. *"na j po-acje hó-nó-ta" sólf-assje*  
I NOM see-PROP do-IND-DEC tell-IMP  
'Please tell (him) that I would like to see (him).'
- d. *nimkum ha al-ósjosje*  
king VOC know-IMP  
'Oh King, please be aware of this!' (Yongka 125)
- e. *sajngsajng aj na j wen ul ilh-ti ani-h-kej hó-kola*  
life in I GEN wish ACC forget-NMR not-do-to do-IMP  
'Be sure not to forget my wishes in your life.' (Welsek 1:13)
- f. *na j aki wujhój-a et-e po-kolje*  
I GEN child for-INF get-INF try-IMP  
'Try to obtain (it) for my child.' (Sekpo 6:13)

#### 4.5.5.4 Propositive sentences

Not many data on propositive sentences exist. Ko (2007: 344–6) lists about a dozen examples. Only plain and deferential sentences are found. The plain ender is *-cje* (59a, b). If this ender is preceded by the suffix *a* (usually called a 'confirmative' suffix), as in *-acje*, its meaning is strengthened with the sense of 'certainly, surely' (59c). The deferential propositive ender is *-sangita* (60).

- (59) a. *i kep ilhwum ulan hjenkep i-la hó-cje*  
this eternity name TOP wisdom-eternity be-DEC call-PROP  
'As for the name of this eternity, let us call it "wisdom-eternity".' (Welsek 1:40)

- b. *twuj-s najh aj mól siski-la ka-cje*  
back river to horse wash-to go-PROP  
'Let's go to the back river to wash the horse.' (Pakthongsa sang:21)
- c. "*na j po-acje hó-nó-ta*" *sólβ-assje*  
I NOM see-PROP do-IND-DEC tell-IMP  
'Please tell (him) that I would surely like to see (him).'
- (60) a. *cajngkethjen i thajca skuj sólβ-utój "ka-sangita"*  
(name) NOM prince to tell-that, go-PROP  
'Cajngkethjen told the prince, "Let's go."' (Sekpo 3:26)
- b. *cajngtho aj hótój ka na-sangita*  
clean.land to together go go.out-PROP  
'Let's go out to a clean land together.' (Welsek 8:100)

#### 4.5.5.5 Minor sentence types

There are several minor sentence types. The promissive is expressed by the ender *-oma* (61a), "oughtness" by *-(u)lʔ-ti-ni-la* (61b), and "exclamation" by *-(ó/u)lssje, -tota/-twuta/-lota/-lwuta*, etc. (61c, d).

- (61) a. *na j ne tólje kólóchj-oma*  
I NOM you to teach-PROM  
'I promise to teach you.' (Pakthongsa sang:10)
- b. *sto pantóki casejhi mózóm ól psu-lʔ-ti-ni-la*  
again surely in.detail mind ACC use-PROS-NMR-DF-DEC  
'Again you should pay close attention (to it).'
- c. *tajwang i chanthanhó-sja-tój tjuh-ólssje tjuh-ólssje*  
great.king NOM admire-SH-that good-EXC good-EXC  
'The Great King admired exclaiming, "Good! Good!"' (Sekpo 11:27)
- d. *cinsillo sulphu-tota*  
truly sad-EXC  
'I am truly sad!' (Kumkang hwuse 11)

#### 4.5.6 Negation

There are several ways to express negation. First, the generic negative adverb *ani* 'not' + *hóta* is used after a clause that ends in the nominalizer ender *-ti*.

- (62) *sejcon ha, estjej posal i [poktek swuh-ti] ani*  
Buddha VOC why saint NOM blessing receive-NMR not  
*ho-m i-ni-ngi-sko*  
do-NMR be-IND-AH-Q  
'Oh Buddha! Why is the Buddhist saint not willing to receive the blessing?'  
(Kumkang 142)

Secondly, the 'inability' negative adverb *mot* + *hóta* is used after a clause that ends in *-ti*.

- (63) [cjokomas pój thó-kocje s ptut ul nis-tí] **mot**  
 small ship ride-intending GEN wish ACC forget-NMR cannot  
*hó-li-lota*  
 do-PROS-EXC  
 ‘(He) will not be able to forget the wish to ride a small boat.’ (Twusi 15:55)

Both *ani* and *mos* appear in both a short form in pre-predicate position (64a) and a long form in post-predicate position (64b).

- (64) a. *nwu j ani sólanghó-zóβ-óli*  
 who NOM not think-AH-PROS  
 ‘Who won’t think about it?’ (Yongka 78)
- b. *nwu j hójjati-tí ani hó-nó-njo*  
 who NOM wear.out-NMR not do-IND-Q  
 ‘Who is free from (his/her things) being worn out?’ (Nungem 4:80)

A third type of negation is the use of the prohibitive verb *malta* ‘stop doing’. This verb is preceded by the infinitive suffix *-a/-e* (after an adjective) or by the nominalizer *-ti* (after a verb).

- (65) a. *wu-nón sengnje ije sulh-e mal-la*  
 cry-REL saintess VOC sad-INF stop-IMP  
 ‘Oh, you, crying saintess! Don’t be sad!’ (Welsek 21:21)
- b. *kwucwung ej tulu-sj-a thajphjeng ul nwuli-si-l cej*  
 palace to enter-SH-and peace ACC enjoy-SH-PROS time  
*i ptut ul nis-tí mal-ósjosje*  
 this idea ACC forget-NMR stop-IMP  
 ‘When you enter the palace to enjoy a peaceful reign, please don’t forget this idea.’ (Yongka 110)

#### 4.5.7 Topic, focus, and emphasis

The typical MK topic particle is (*n*)ón or its allomorph (*n*)un. Nominals with this suffix are usually placed sentence-initially (66). Topic is also expressed by *lan* ‘as for, if it be’ (67).

- (66) a. *i nóñ wuli hemul i-la*  
 this TOP our fault be-DEC  
 ‘This is our fault.’ (Pephwa 2:5)
- b. *ok óñ salóm kato-nón stah i-la*  
 jail TOP people confine-REL place be-DEC  
 ‘A jail is the place where people are confined.’ (Sekpo 9:8)
- c. *i tongsan óñ namk i tjoh-ólssój noni-non stah i-la*  
 this hill TOP tree NOM good-so walk.around-REL place be-DEC  
 ‘As for this hill, there are good trees and thus it is a place to walk around.’ (Sekpo 6:24)
- (67) *tjoh-ón il lan na jkej ponaj-o*  
 good-REL thing as.for me to send-and  
 ‘As for good things, send them to me, and [. . .]’ (Kumkang 21)

It appears that there is no specific morphological marker to uniquely indicate focus or emphasis, except perhaps for prosodic means of ‘stress’ and insertion of an epenthetic *s*. One syntactic means is scrambling of sentential constituents. Placing a constituent in sentence-initial position is often related to the speaker’s or writer’s intention to give focus or emphasis to that initial constituent. For instance, in (68), the object is fronted apparently for focus.

- (68) *pwutje s kujjel ul cisin i nil-enul*  
 Buddha GEN news ACC Earth.God NOM tell-now.that  
 ‘Now that the Earth God is telling the news about Buddha [. . .],’ (Welin ki 82)

#### 4.5.8 Causative and passive

Causativization with a suffix is more productive in MK than in CK. Intransitive and transitive verbs and adjectives are made into causative verbs by means of a causative suffix. Thus, intransitive verbs and adjectives become transitive verbs and transitive verbs become ditransitive verbs. There are allomorphic variants in the suffix: *-hi*, *-i*, *-ki*, *-o/-wu*, *-ho/-hwu*, *-ó*, *-iwu*. Occurrence of some of these variants is phonologically conditioned, and occurrence of others is lexically conditioned and thus unpredictable.

The three variants *-hi*, *-i*, and *-ki* are most productive. The suffix *-hi* occurs after *t*, *p*, *c*, and *k*, as in *mwutta* ‘be stained’ (Vi) vs. *mwuthita* ‘stain’; *ancta* ‘sit’ (Vi) vs. *anchita* ‘make sit, place’; *macta* ‘match’ (Vi) vs. *machita* ‘hit’; *nepta* ‘wide’ (ADJ) vs. *nephita* ‘widen’; *patta* ‘receive’ (Vt) vs. *pathita* ‘give, offer’; *nipta* (Vt) ‘put on’ vs. *niphita* ‘clothe, cause to put on’. The suffix *-i* (or *-j*) occurs after a vowel, after an aspirated consonant, and after *k* or *l(u’ó)*, as in *nata* ‘come out’ (Vi) vs. *najta* ‘take out, put out’; *sjeta* ‘stand’ (Vi) vs. *sjejta* ‘stand’; *nokta* ‘melt’ (Vi) vs. *nokita* ‘melt (something)’; *nikta* ‘ripe’ (Vi) vs. *nikita* ‘get used to’; *salta* ‘live’ (Vi) vs. *salita* ‘bring to life’; *olóta* ‘climb’ (Vi) vs. *olita* ‘raise, lift up’; *kuchuta* ‘stop’ (Vi) vs. *kuchita* ‘stop, cut (something)’; *kiphta* ‘deep’ (ADJ) vs. *kiphita* ‘deepen’; *telepta* ‘dirty’ (ADJ) vs. *telefita* ‘make dirty, blemish’; *etwupta* ‘dark’ (ADJ) vs. *etwuita* ‘darken’; *mólkta* ‘clean’ (ADJ) vs. *mólkita* ‘make clean, purify’; *mekta* ‘eat’ (Vt) vs. *mekita* ‘feed’; *malta* ‘stop doing’ (Vt) vs. *malita* ‘make stop, dissuade’. The suffix *-li* (instead of *-i*) occurs after *l*, as in *huluta* ‘flow’ (Vi) vs. *hullita* ‘shed, make flow’. The suffix *-ki* occurs after *s*, *m*, as in *swumta* ‘hide’ (Vi) vs. *swumkita* ‘hide (something)’; *olmta* ‘move (to)’ (Vi) vs. *olmkita* ‘move (something)’; *sista* ‘wash’ (Vt) vs. *siskita* ‘make wash’; *pasta* ‘take off’ (Vt) vs. *paskita* ‘unclothe, strip’.

The suffixes *-o/-wu*, *-ho/-hwu*, *-ó* are by and large vowel harmony sensitive. Examples include *totta* ‘rise’ (Vi) vs. *totota* ‘make higher’; *motta* ‘gather’ (Vi) vs. *motota* ‘collect’; *oólta* ‘sound, sincere (of mind)’ (ADJ) vs. *oólota* ‘make (mind) sound/sincere’; *ilta* ‘rise up’ (Vi) vs. *ilwuta* ‘accomplish’; *mwujta* ‘move’ (Vi) vs. *mwujwuta* ‘move (something)’; *elta* ‘be frozen’ (Vi) vs. *elwuta* ‘freeze’; *njethta* ‘shallow’ (ADJ) vs. *njethota* (instead of *njethwuta*) ‘make shallow’; *kócta* ‘be equipped’ (Vi) vs. *kóchota* ‘have, prepare’; *mecta* ‘stop’ (Vi) vs. *mechwuta* ‘stop (something)’; *tolta* ‘turn around’ (Vi) vs. *tolóta* ‘turn (something) around’. The compound suffix *-iwu/-jwu* occurs in words such as *psuta* ‘use’ (Vt) vs. *psuj.wuta* ‘make use’; *ptuta* ‘float’ (Vi) vs. *ptuj.wuta* ‘set (a ship) afloat’, as in CK.

- (69) a. *mom aj moncój mwuthi-si-ko*  
 body to dust stain-SH-and  
 ‘(He) covered his body with dust, and [. . .]’ (Welsek 21:219)

- b. *sekpjek ej mól oli-sj-a*  
 rock.wall at horse lift.up-SH-and  
 '(He) had the horse climb up the rock wall, and [. . .]' (Yongka 48)
- c. *na j icej ne j em(i) uj ka-n stah ól*  
 I NOM now you GEN mom GEN go-REL place ACC  
*poi-o-li-la*  
 show-SS-PROS-DEC  
 'Now I will show you the place your mom went to.' (Welsek 21:21)
- d. *mólk-ón mos aj kahi pój lól ptuj.w-e nol-li-lo-la*  
 clear-REL pond in easily boat ACC float-INF play-PROS-IND-DEC  
 '(We) will easily set a boat afloat in a clear pond and enjoy ourselves.'  
 (Twusi 1:14)

Periphrastic causative constructions with the sequence *-kuj/-kej hó-* (so.that do/make) 'cause' are also extensively used. In general, while suffixal causatives express more direct causation, periphrastic causative constructions tend to express indirect causation.

- (70) a. *sjwuko lól ta pesena-kuj ho-li-la*  
 suffering ACC all escape-so.that do-PROS-DEC  
 '(We) will have (them) escape from all suffering.' (Welsek 9:15)
- b. *sajngsajng aj na j wen ul ilh-ti ani-h-kej*  
 life in I GEN wish ACC forget-NMR not-do-so.that  
*hó-kola*  
 do-IMP  
 'Be sure not to forget my wishes in your life.' (Welsek 1:13)

MK passives are expressed by attaching a passive suffix to transitive verbs. In terms of forms, passive suffixes are a subset of causative suffixes, in that only the same three productive forms *-hi*, *-i*, and *-ki* are used. This fact, together with the functional relatedness between causative and passive, supports a hypothesis that passive suffixes developed from causative suffixes by way of functional shift before the MK period.

The passive suffix *-hi* appears in words such as *tatta* 'close' vs. *tathita* 'be closed'; *mekta* 'eat' vs. *mekhita* 'be eaten'; *capta* 'grasp' vs. *caphita* 'be grasped'; and *jencta* 'put on' vs. *jenchita* 'be put on'. Examples of the occurrence of the suffix *-i* are: *cochta* 'follow, chase' vs. *cochita* 'be chased'; *twuphta* 'cover' vs. *twuphita* 'be covered'; *jelta* 'open' vs. *jelita* 'be opened'; *asta* 'take away' vs. *asita* 'be taken away'; and *katota* 'collect' vs. *katoita* 'be collected'. The suffix *-ki* occurs in transitive verbs such as *tamta* 'put in' vs. *tamkita* 'be put in'; *cómta* 'soak (something)' vs. *cómkita* 'be soaked'; and *sólmta* 'boil' vs. *sólmkita* 'be boiled'.

- (71) a. *tongmwun i tolo tathi-ko*  
 East.Gate NOM again closed-and  
 'The East Gate was shut again, and [. . .]' (Welsek 23:80)
- b. *hón pwuchej lul tat-óni hón pwuchej yeli-kom hó-lssój*  
 one fan ACC close-as one fan opened-each do-as  
 'As each time (someone) closed one fan, another fan was opened, and so [. . .]' (Welsek 7:9)

#### 4.5.9 Speech levels and honorifics

*A. Address-reference terms:* In MK, unlike in CK, professional titles are used without any honorific title. Professional titles alone are sufficient for honorification: e.g. *sensajng* ‘teacher’, *nimkum* ‘king’, *wang* ‘king’, *pwuthje* ‘Buddha’, *miluk* ‘Maitreya’, *posal* ‘Buddhist saint’, and *zjelaj* ‘Enlightened One’. The honorific title *-nim* (from the noun *nim* ‘lord, king; beloved’) was attached to kinship terms: *apanim* ‘father’, *emanim* ‘mother’, *acópanim* ‘uncle’, *acómanim* ‘aunt’, *atónim* ‘(other’s) son’, and *hjeng-nim* ‘older brother’. No other suffixal title is attested.

*B. Pronouns:* See 4.5.2.

*C. Content words:* Most CK honorific words have their origins in MK, while many MK honorific words are obsolete in CK. Some MK examples are *cinci* (**cinci**) ‘meal’, *pwun* (**pun**) ‘person’, *moj* (-) ‘rice’, *macóbjój* (-) ‘oldest child’, and *sulha* (-) ‘king’s meal’. Honorific predicates include *cwasita* (**capswusita**) ‘eat’, *tukjesita* (-) ‘place (something)’, *kólota* (-) ‘say’, *kjesita* (**kyeysita**) ‘stay, exist’, *mojsita* (**mosita**) ‘accompany’, *pojta* (**poyppta**) ‘see (a superior)’, *cehta* (-) ‘fear’, and *jetcópta* (**yeccwuta**) ‘ask (a superior)’, *sólpta/sólota* (**aloyta**) ‘talk (to a superior)’, and *tulita* (**tulita**) ‘give (to a superior)’.

*D. Case particles:* The MK and CK case particles relevant to the honorific system are summarized in (72) and discussed in 4.4.1.2 and 4.5.1.

(72)			MK	CK
Genitive	Honorific		<i>s</i>	-
	Neutral		<i>ój/uj</i>	<b>uy</b>
Dative/locative	Honorific		<i>skej/skuj</i>	<b>kkey</b>
	Neutral		<i>ejkej</i>	<b>eykey</b>
	Inanimate		<i>ej</i>	<b>ey</b>
Vocative	Honorific		<i>ha, (i)je</i>	-
	Plain		<i>a</i>	<b>(i), (y)a</b>
Subject	Honorific		-	<b>kkeyse</b>
	Neutral		<i>i</i>	<b>i/ka</b>

Textual evidence and earlier analyses indicate that the dative/locative three-way distinction was innovated in the MK period. The following processes took place for the derivation of the honorific *skej/kkej* and the neutral *ejkej*. In MK, to indicate the dative/locative of an animate noun, the noun was followed by the anaphoric demonstrative pronoun *ku* ‘that’ + dative/locative particle *ngej*<sup>11</sup>.

- (73) a. MK honorific genitive *s + ku + ngej* > *skej/skuj* > *kkej* > **kkey**  
 b. MK neutral genitive *ój/uj + ku + ngej* > *ujkej* > **eykey**

Thus, for example, *puthje s ku ngej* ‘to the Buddha’ developed to *puthje skej* and both forms are attested. Similarly, *kjecip uj ku ngej* ‘to the girl’ developed to *kjecip ujkej*.

Three particles are used to mark the vocative case in MK: *ha*, *(i)je*, and *a*. The particle *ha* is to call a superior, as in *sejcon ha* ‘O Sakyamuni!’ The particle *(i)je* (defunct copula *i + je*; *ije* after a consonant and *je* after a vowel) is used with an exclamatory effect (I. Lee and Ramsey 2000: 292). The particle *a* is to call an equal or someone lower in rank. While *ha* is no longer used in CK, *(i)je* is still used only as a poetic or

specialized form, as in **cwu ye** ‘O Lord!’ or **im iye!** ‘O my beloved!’ The MK plain vocative *a* corresponds to CK (**y**)**a**, which is used only to children. In CK, (**i**) is used to adolescents, as in **hankon i** ‘Hankon!’ and **yongho** ‘Yongho!’

*E. Subject honorific (SH) suffix:* MK has three prefinal honorific suffixes: subject honorific, object honorific, and addressee honorific. The suffix *-(u)si* (contracted to *-(u)sj* or *-(u)sja*) marks the subject honorific.

- (74) a. *hajtong ljwuk-ljong i nóló-sj-a*  
Korea 6-dragon NOM fly-SH-and  
‘Six dragons flew in Korea and [. . .]’ (Yongka 1)
- b. *na nón puthje s sólanghó-si-non az i-la*  
I TOP Buddha GEN love-SH-REL brother be-DEC  
‘I am the younger brother whom Buddha loves.’
- c. *puthje s mokswum i kuc i eps-usja-m ul*  
Buddha GEN life NOM end NOM lack-SH-NMR ACC  
‘the fact that Buddha’s life is endless’ (Y. Choy 1987: 191)

*F. Addressee honorific (AH) suffix:* The prefinal suffix *-ngi* marks addressee honorification. This suffix indicates the highest deferential level, as in (75a). Its contracted forms *-ng* and zero form are used, respectively, in the declarative and interrogative of the next higher speech level (see H). Historically, the suffix *-ngi* was gradually reduced to *-i* in all positions by the early CK period as illustrated in (75b). Then, the suffix *-i* was simply absorbed by the indicative mood suffix *-ni* when the addressee honorific function was taken over by **-sup** during the early CK period, as in (75c) (see G).

- (75) a. MK: *kilu-sóß-ósi-ni-**ngi**-ta*  
raise-OH-SH-IND-AH-DEC  
‘(X) is raising (Y).’ (Werin 10:19; Huh 1972: 34)
- b. early CK: *kilu-si-**op**-na-**i**-ta*  
raise-SH-AH-IND-AH-DEC
- c. CK: *kilu-si-**p**-ni-**ta***  
raise-SH-AH-IND-DEC

Notice that MK *-sóß* precedes the subject honorific *-ósi*, but occurs after the subject honorific suffix in early CK and CK. This is because the MK object honorific suffix underwent a functional shift to the addressee honorific in early CK, driving out *-ngi* (cf. Sohn 1998).

*G. Object honorific (OH) suffix:* Unlike in CK, MK (and OK) marks object (direct and indirect) honorification by way of the prefinal inflectional suffix *-sóß*.<sup>12</sup> This suffix is used as productively as the subject honorific suffix. It functions to express the subject referent’s (and speaker’s) respect to the higher person affected by the action of the inferior subject referent (76).

- (76) a. *pilok zjelaj s sengsilhón mal ól tut-**caß**-ato*  
though Enlightened.One GEN sincere name ACC hear-OH-but  
‘Though we hear the sincere words of the Enlightened One, [. . .]’ (Welsek 21:15)

TABLE 4.5 LATE MIDDLE KOREAN SPEECH-LEVEL SENTENCE ENDERS

Speech level	Declarative	Interrogative	Imperative	Propositive
Plain	-ta/-la	-nje/-njo, -nta	-((k)e)la	-ce (rarely -cela)
Neutral	-ni/-noj	-ni	-kola/-kolje	(not attested)
Moderate respect	-ng-ta	-nó-ni-ska/sko	-esje/-asje	(not attested)
Deferential	-ngi-ta	-nó/ni-ngi-ska/sko	-sjosje	-sa-ngi-ta

- b. *na nón cungca skuj tut-cóp-ko cungca nón pwuca skuj*  
 I TOP Cungca from hear-OH-and Cungca TOP Puca from  
*tut-cóo-si-ni*  
 hear-OH-SH-DEC  
 ‘I heard it from Cungca and Cungca heard it from Puca.’ (Sohak 4:18)

In (76a), the respected person, the Enlightened One, is the possessor of the object referent (‘name’). The speaker/writer is the understood subject. In (76b), the subject (the speaker/writer) used *-cóp* in the first clause to show respect to Cungca, and used *-cóp* in the second clause to indicate the subject referent Cungca’s humility to his mentor Puca. The use of the subject honorific *-si* in the second clause denotes the speaker/writer’s respect to the subject (Cungca).

*H. Speech-level enders:* Sentence-type enders interact with speech levels in MK. Korean linguists establish three or four speech levels for MK, but they are far from agreeing with one another in terms of the number, kind, and interpretations of suffixes in some categories such as interrogative sentences. Table 4.5 is based on the four-level system (e.g. Ko 2007: 303–7). Texts published after the fifteenth century show some other sentence enders. For example, sixteenth-century data contain the imperative ender *-so* or *-(z)o*, which was used between adult equals or to one’s older brother (Y. O. Kim 1995: 118). This new imperative form is assumed to have developed from the deferential level *-sjosje* with the final syllable deletion.

#### 4.5.10 Adverbs

Adverbs, constituting a relatively open word class, modify predicates or entire clauses. MK adverbs mark, variously, time, place, degree, manner, negation, modal, conjunction, and discourse. There is a relatively small set of inherent (monomorphemic) adverbs, while the majority of adverbs have derived from other words (mostly predicates and some nouns) by means of the productive adverbial suffix *-i/-j*, some other idiosyncratic suffixes such as *-o*, *-aj*, and *-cahi*, and particles such as *naj* and *zo*. Examples include *khi* ‘greatly’ (from *khu-* ‘big’ + *-i*), *niki* ‘well, thoroughly’ (from *nik-* ‘ripe’ + *-i*), *haj* ‘a lot’ (from *ha-* ‘big’ + *-j*), *nunghi* ‘ably’ (from *nungha-* ‘able’ + *-i*), *kojojhi* ‘quietly’ (from *kojojha-* ‘quiet’ + *-i*), *oj.o* ‘wrongly’ (from *oj-* ‘wrong’ + *-o*), *itaj* ‘nicely’ (from *it-* ‘good’ + *-aj*), *cukcahi* ‘immediately’ (from *cuk* ‘immediacy’ + *-cahi*), *móchómnaj* ‘finally’ (from *móchóm* ‘end’ + *-naj*), and *sonzo* ‘in person’ (from *son* ‘hand’ + *zo*).

Some verb and adjective roots were also used as adverbs without any suffix, e.g. the verb *sómósta* ‘to go through, communicate’ vs. the adverb *sómós* ‘thoroughly’; the adjective *palóta* ‘be right, straight’ vs. the adverb *paló* ‘rightly, straightly’; the adjective *iluta* ‘be early’ vs. the adverb *il* ‘early’; and the adjective *hata* ‘be much,

big' vs. the adverb *ha* 'much, a lot'. Frequently used adverbs are: (a) Time: *mencje* 'before, first of all'; *pólsje* 'already, yet'; *hóma* 'already'; *sjangnej* 'always'; *nójlil* 'tomorrow'; *moloj* 'the day after tomorrow'; *eceskuj* 'yesterday'. (b) Place: *ili* 'this way'; *kuli* 'that way'; *tjeli* 'that way over there'; *ingekuj* 'here'; *kungekuj* 'there'. (c) Degree: *kóccang* 'most'; *keuj* 'almost'; *mós* 'most'; *kós* 'barely'; *cjekomato* '(not) at all'; *elwu* 'nearly, almost, approximately'. (d) Manner: *phele* 'in deep blue/green'; *zel-zel* 'gently'; *kwumul-kwumul* 'wigglingly'; *kót* 'like'; *selu* 'each other'; *kulu* 'wrongly'; *tamós* 'together with'; *ptólo* 'separately'; *tetuj* 'slowly, late, behind the time'; *nankjes* 'hurriedly'; *ilikom* 'like this'; *kantajlo* 'at random'. (e) Negative: *ani* 'not'; *mot* 'unable, cannot'. (f) Modal: *hótaka* 'if'; *sihok* 'by any change'; *hómólmje* 'still more, not to speak of'; *amato* 'perhaps'; *sóncoj* 'rather'; *motój* 'surely'; *pantóki* 'surely'; *molomaj* 'surely, certainly'; *etuli*, *estjej*, *maj*, *enu* 'how, why'. (g) Conjunctive: *kulena* 'but'; *kulemjén* 'then'; *kulelssój* 'therefore'; *sto* 'also'. (h) Discoursal: *ój* 'oh'; *ejng* 'yes'; *aso* 'Good gracious!'. Some sentence examples with adverbs are given in (77).

- (77) a. *nilkwup hój nemwu ola-ta*  
 7 year too long-DEC  
 'Seven years is too long.' (Welsek 7:2)
- b. *i twu salóm i cinsillo ne j hangkes-ka*  
 this 2 person NOM truly you GEN master-Q  
 'Are these two persons truly your masters?' (Welsek 8:94)
- c. *nwu j hóollo sulh-e hó-nó-nio*  
 who NOM alone sad-INF do-IND-Q  
 'Who alone feels sad?' (Twusi 19:40)
- d. *hanca lo mencje kul ul môngkól-o*  
 Chinese.character with first writing ACC make-and  
 'made the writing with Chinese characters first, and [. . .]' (Sekpo se 5)
- e. *cajngtho aj hótój ka na-sangita*  
 clean.land to together go out.go-PROP  
 'Let's go out to a clean land together.' (Welsek 8:100)
- f. *sto pantóki casejhi mózóm ól psu-l?-ti-ni-la*  
 again surely in.detail mind ACC use-PROS-NMR-DF-DEC  
 'Again you should pay close attention (to it).' (Mongsan 39)

#### 4.5.11 Subordinate clauses

Subordinate clauses are of five types: (a) relative, (b) predicative complement, (c) nominalized, (d) conjunctive, and (e) quotative. All these different clause types are so marked by pertinent inflectional clause enders. Frequently, several of these clause types occur together in a long sentence, as illustrated in (78).

- (78) *salóm mata hój-e](d) swuþi níkj-e](d) nallo psuw-m](c) ej*  
 person each cause-and easily learn-and daily use-NMR for  
*phjenanh-kuj](b) hó-kocje](b) hó-l](a) stólóm i-ni-la*  
 comfortable-so.that make-wishing do-PROS only be-DF-DEC  
 'My only wish is to have every person learn and use it easily and conveniently on a daily basis.' (Cengum Enhay)

TABLE 4.6 LATE MIDDLE KOREAN RELATIVIZER ENDERS

Relativizers	MK
Non-past	- <i>nón</i> , - <i>non</i> (after verb)
Past/perfective	-( <i>U</i> ) <i>n</i> (after adjective/copula)
Prospective	-( <i>U</i> ) <i>l</i>

#### 4.5.11.1 Relative clauses

A relative clause modifies a following head nominal by means of a relativizer ender (Table 4.6). Typical relativizers denote tense and aspect, in addition to the connective function.

Relative constructions consist of a relative clause followed by its head, as in [*hó-nón*] ***il*** (do-REL work) ‘the work (one) does’, [*noph-ón*] ***moj*** (high-REL mountain) ‘high mountain’, [*i cip uj sa-nón*] ***elwun*** (this house at live-REL adult) ‘the adult person who lives in this house’, [*pwuín i nah-ósi-l*] ***tól*** (lady NOM deliver (baby)-SH-REL month) ‘the month when the lady will give birth to a baby’, where the bracketed parts are relative clauses and the bold parts are their respective relativizers and head nouns. More sentence examples with a relative construction are given in (79).

- (79) a. [*wu-nón*] ***sengnje*** *ije* *sulh-e* *mal-la*  
cry-REL saintess VOC feel.sad-INF stop-IMP  
‘You crying saintess! Don’t be sad!’ (Welsek 21: 21)
- b. [*pwulhuj kiph-un*] ***namk*** *ón* *pólóm* *aj* *ani* *mwuj-lssój*  
root deep-REL tree TOP wind by not move-as  
‘As trees whose roots are deep will not move by wind,’ (Yongka 2)
- c. *en un* [*nilu-l*] ***ssó*** *j-la*  
*en* TOP say-PROS fact be-DEC  
‘*En* is a word meaning “to say”.’ (Cengum Enhay 2)
- d. [*eli-n*] ***pójksjeng*** *i* [*nilu-kocje hó-l*] ***pa*** *j*  
poor-REL people NOM report-wishing do-PROS thing NOM  
*isj-eto*  
have-though  
‘Even though simple people have things to report (to the government), [. . .]’  
(Cengum Enhay)

One MK pattern which is not shared with CK is the genitive particle *s* ‘of’ used as a special kind of (quotative) relativizer, which is attached to a complete sentence. Notice in (80) that the bracketed relative clause is a complete sentence followed by the genitive particle as a relativizer.

- (80) a. [*nepi kwangmjeng i pichwuj-ta s*] ***ptut*** *i-o*  
widely light NOM shine-DEC GEN meaning be-DEC  
‘(It) is the meaning that the light shines widely.’ (Welsek 2:9)
- b. [*nunghi ta mot hó-li-la s*] ***ptut*** *ul* *kjelhó-ja*  
possibly all unable do-PROS-DEC GEN mind ACC make-and

*natho-si-ni-la*

appear-SH-DF-DEC

‘(He) will appear with the decision that he cannot possibly do it all.’ (Pepwha 3:47)

- c. [*tinake-n mwuljangkep ej swuhajng i nik-usilssój mot*  
past-REL eternity in discipline NOM ripe-since cannot  
*ilwu-olʔ-ka s] ujsim i eps-usi-na*  
achieve-PROS-Q GEN doubt NOM lack-SH-but  
‘As (he) has been eternally disciplined in the past, there is no doubt that (he)  
would not succeed, but’ (Welin ki 53)

#### 4.5.11.2 Predicative complement clauses

This construction consists of a complement clause and a complement ender, which is followed immediately by its head predicate. The predicate and the preceding ender form a semantic unit (bold-faced in (81)) expressing desiderative, (periphrastic) causative, habitual, and so on, as in CK. Complement enders are the same as conjunctive enders in form, the difference being that the former are semantically merged with the head predicates, which was not the case with the latter. Frequently used enders include *-kocje/-acje* ‘wishing to, intending to’, *-kej/-kuj* ‘so that’, and *-tós* ‘as if’. The most frequently used head predicate following these complement enders is *hóta* ‘do’.

- (81) a. *eli-n pójksjeng i [nilu-kocje] hó-l pa j*  
simple people NOM report-wishing do-REL thing NOM  
*isj-eto*  
have-though  
‘Even though simple people have things that they wish to report, [. . .]’  
(Cengum Enhay)
- b. [*na j po-acje] hó-nó-ta sóβ-assje*  
I NOM see-intending do-IND-DEC tell-IMP  
‘Please tell (him) that I would like to see (him).’ (Sekpo 6:14)
- c. [*umsik ólo pój pulu-kuj] hó-koza*  
food with stomach full-so.that make-in.order.to  
‘in order to make (people) have a hearty meal with food’ (Welsek 9:25)
- d. [*pep i nepi phje-a ka-m i swulwujpti kuwu-tós]*  
law NOM widely spread-INF go-NMR NOM wheel roll-like  
*hó-lssój*  
do-as  
‘As the way the laws were spreading was as if a wheel rolls, [. . .]’ (Sekpo 13:4)

A subtype of predicative complement clauses is that of auxiliary or serial predicate constructions. In (82a), the auxiliary verb *pota* ‘try’ has the preceding clause as its complement. In (82b), the main verb *kjesita* ‘stay’ (honorific) has the preceding clause as its complement.

- (82) a. [*na j aki wujhó-j-a et-e*] *po-kolje*  
 I GEN child serve-INF get-INF try-IMP  
 ‘Try to obtain (it) for my child!’ (Sekpo 6:13)
- b. [*kos nilkwup cwulki lól kacj-e*] *kjesj-atój*  
 flower 7 stalk ACC hold-INF stay-and  
 ‘(She) was holding seven stalks of flowers, and [. . .]’ (Welsek 1:9)

#### 4.5.11.3 Nominalized clauses

A nominalized clause is composed of a clause followed by a nominalizer ender: *-(o)m*, *-ki*, or *-ti*. The most productive nominalizer suffix is *-(o)m*, which has allomorphs *-jom* and *-wum*.

- (83) a. [*tjoh-ón pep task-om*] *ól mot hó-ja*  
 good-REL law polish-NMR ACC unable do-so  
 ‘Being unable to polish up good laws, [. . .]’ (Sekpo 9:14)
- b. [*hó-non il eps-wum*] *i pilok pwuj-na*  
 do-REL work lack-NMR NOM though vacant-but  
 ‘Although doing nothing is boring, [. . .]’ (Welsek 8:31)
- c. *sejcon ha, estjej posal i [poktek swuh-ti ani*  
 Buddha VOC why saint NOM blessing receive-NMR not  
*ho-m]* *i-ni-ngi-sko*  
 do-NMR be-IND-AH-Q  
 ‘Oh Buddha! Why is the Buddhist saint not willing to receive the blessing?’  
 (Kumkang 142)

*-ti* and *-ki* are less productively used. *-ti* corresponds to CK positive **-ki** and negative **-ci**. The nominalizer *-ki*, which is assumed to have developed from *-ti*, is also used, as in CK.

- (84) a. [*tewujcap-ti*] *eljep-tota*  
 grasp-NMR difficult-EXC  
 ‘It is hard to hold on to it.’ (Kumkang Samka 2:48)
- b. *mózólh i mel-mjen [kelsikhó-ti] eljep-ko*  
 village NOM far-if beg.for.food-NMR difficult-and  
 ‘If the village is far, it is difficult to beg for food, [. . .]’ (Sekpo 6:23)
- c. *estjej posal i [poktek swuh-ti] ani ho-m*  
 why saint NOM blessing receive-NMR not do-NMR  
*i-ni-ngi-sko*  
 be-IND-AH-Q  
 ‘Why is the Buddhist saint not willing to receive the blessing?’ (Kumkang 142)
- c. *longtamhó-ja [namcin elu-ki] lól hó-mje*  
 joke-and male play.with-NMR ACC do-and  
 ‘(She) did playing with males by making jokes, and [. . .]’ (Welsek 1:44)

4.5.11.4 *Conjunctive clauses*

Conjunctive clauses, which precede their main clauses, are marked by a conjunctive clause ender. There are scores of conjunctive enders in MK, as in CK, such as: (a.) Combination: *-ko, -o, -kok* (for emphasis) ‘and’; *-mje* ‘and’; *-taka* ‘and then’; *-tój, -tójn* (for emphasis) ‘and that’; *-keniwa* ‘as well as’; *-tíβi* ‘and, though, but’; *-esje/-asje* ‘and then, and so’; *-e/-a/-ja/-je* ‘and’; *-ntój, ntójn* (for emphasis) ‘given that’. (b.) Opposition: *-na* ‘but’; *-(ke)nmalón* ‘though, despite’; *-ato/-eto* ‘but, yet’; *-sje* ‘and (then, so)’. (c.) Sequentiality: *-n tamata* ‘as soon as’; *-ni* ‘as, when’. (d.) Cause, reason: *-ssój* ‘as, since’; *-(k)enul/-(k)enól* ‘now that, since, as’; *-ketun* ‘as, because’; *-kwantój* ‘so ~ that, such ~ that’; *-ntaj, -ntajn* (for emphasis) ‘since, because’; *-i-la* ‘as/because it is’; *-m-aj* ‘as, since’. (e.) Intention, goal: *-la* ‘in order to’; *-kocje, -kejsko, -kujsko* ‘in order to, wishing to’; *-kwatje, -kwatja* ‘in order to, wishing to’; *-lje, -lja* ‘in order to’. (f.) Conditional: *-mjen* ‘if’; *-tun* ‘if’; *-lantój* ‘if, in case, provided that’. (g.) Concessional: *-lmantjeng* ‘even though, nevertheless, but’; *-lsjentjeng* ‘rather, sooner than’; *-lspwuntjeng* ‘rather, sooner than’; *-ntól* ‘even if’. (h.) Manner, degree: *-kej, -kuj* ‘so that’; *-tólok* ‘so that, to the extent that’; *-tios* ‘the more ~ the more’; *-tós(i)* ‘like, as if’; *-i* (adverbial ‘-ly’). Some sentence examples with some of the above conjunctive enders are given in (85).

- (85) a. *hanca lo mence kul ul móngkól-o*  
 Chinese.character with first writing ACC make-and  
 ‘Made the writing with Chinese characters first, and [. . .]’ (Sekpo se 5)
- b. *ing mwut-un cangku llang kaci-ko*  
 moss stick-REL plough etc. carry-and  
 ‘carrying a moss-covered plough and so on [. . .]’ (Chengsan Pyelkok in *Akchang Kasa*)
- c. *pójam i kachi l mul-e cumkejs kac(i) aj jencó-ni*  
 snake NOM magpie ACC bite-INF big.tree branch on put-as  
 ‘As the snake bit a magpie and put it on a branch of a big tree, [. . .]’ (Yongka 7)
- d. [*i tongsan ón namk i tjoh-ólssój*] *noni-non*  
 this hill TOP tree NOM good-so walk.around-REL  
*stah i-la*  
 place be-DEC  
 ‘As for this hill, there are good trees and thus it is a place to walk around.’ (Sekpo 6:24)
- e. [[*cwuk-un-tósi*] *ca-taka*] [*hen kas os twuph-esje*] *noll-a*  
 die-REL-as.if sleep-then old leather clothes cover-and play-to  
*o-la*  
 come-IMP  
 ‘Sleep as if you are dead, and then cover yourself with old leather clothes and come to play.’ (Twusi 22:1)
- f. [*na j thajca lól sjemki-zóβ-ótój*]  
 I NOM prince ACC serve-OH-but  
 ‘Although I serve for the prince, [. . .]’ (Sekpo 6:4)
- g. *iwuscip pul un [pam i kip-tólok] pólk-ajs-tota*  
 neighbor light TOP night NOM deep-until brighten-PAST-EXC  
 ‘The neighbor’s light was brightened until the night got deep.’ (Twusi 7:6)

## 4.5.11.5 Quotative clauses

The CK quotative particles **ko** and **lako/hako** had not developed in MK. The typical quotative construction was made of two clauses in sequence. Three types are observed. The first of them is the quoted sentence first and the saying verb second, as in CK, but without any quotative morpheme. This type appears to have functioned as indirect quotation. In (86a), for example, *-(a)cje* is a propositive ender, which is directly followed by the verb *hó* ‘do, say’ and furthermore the whole clause up to the declarative ender *-ta* is directly followed by the saying verb *sólβ-*. The sentence is equivalent to CK **nay ka po-ca ko ha-n-ta ko malssumtuli-sio**.

- (86) a. [[*na j po-acje*] *hó-nó-ta*] *sólβ-assje*  
 I NOM see-PROP say-IND-DEC tell-IMP  
 ‘Please tell (him that) I would like to see (him).’ (Sekpo 6:14)
- b. *zjelaj sangnjej wuli lol [atól i-la] niló-si-ni-ngi-ta*  
 Buddha always us ACC son be-DEC say-SH-IND-AH-DEC  
 ‘Buddha always calls us his sons.’ (Welsek 13:32)

The second, more productive type is that the first clause is a ‘saying’ clause with the conjunctive suffix *-atój/-otój/-ótój* ‘that, but’ and the directly quoted sentence comes next. The quoted sentence is not followed by any marker, except occasionally by a verb of saying such as *hóta* ‘say, do’. This type appears to have functioned as direct quotation.

- (87) a. *wang i hwangtanghi neki-sj-a nilu-sj-atój [i esten*  
 king NOM fabulously think-SH-and say-SH-that this what  
*kwangmjeng-ko]*  
 light-Q  
 ‘The king thought it fabulous and said that, “What kind of light is this!”’  
 (Welsek 10:7)
- b. *tocók i mwul-otój [nehuj-tólh i musuk ul po-nó-són.ta]*  
 thief NOM ask-that you-PL NOM what ACC see-IND-Q  
 ‘The thief asked, “What are you guys seeing?”’ (Welsek 10:28)

The third type is using the genitive *s* as a quotative relativizer (4.5.11.1).

- (88) a. [*cjokomas pój thó-ko cje s*] *ptut ul nis-ti mot*  
 small ship ride-to wish GEN wish ACC forget-NMR cannot  
*hó-li-lota*  
 do-PROS-EXC  
 ‘(He) will not be able to forget the wish to ride a small boat.’ (Twusi 15:55)
- b. [*tjwungsajng cejtóhó-no-la s*] *mózóm i isi-mjen*  
 people control-IND-DEC GEN mind NOM exist-if  
 ‘If (he) has the intention to control the general people, [. . .]’ (Kumkang 2:13)

In (88a) the form *-ko cje* was used not as a desiderative conjunctive ender *-kocje* ‘wishing’ but as a predicate complex equivalent to CK **-ko siph-eyo** ‘wish to do’.

#### 4.6 LEXICON

The MK period may be characterized by, among other things, the influx of a huge number of Chinese words into Korean. Before this period, Chinese words were limited, in general, to the names of places, people, and government ranks. But starting with the Kolye dynasty, Chinese words pervaded the spoken language.

When Chinese words and Chinese characters were first introduced into Korea remains obscure. The only deduction we can make is that they must have been used in Korea as early as the first century BC, when Han China colonized the western and northern parts of the Korean peninsula and established its four commanderies there. Two Japanese sources show that a Paykcheyan named Wang In (written in *Nihon shoki* as 王仁, in *Kojiki* as 和邇, OJ *wani*) went to Japan with many Chinese books around AD 400, suggesting that Chinese culture and characters had achieved considerable popularity in the Korea of the Three Kingdoms period. Earlier, the name of the country and the title of the king were changed from native terms to SK terms in the fourth year of King Cicung (503) of Silla. With the unification of the Korean peninsula by the Silla dynasty in 677, the use of Chinese characters gained more popularity in Korea in that Silla's unification was achieved with Tang China's military support and, subsequently, contact between the two countries became very frequent. In the second year of King Simmun (682) of unified Silla, Kwukhak (lit. 'national studies'), a government organization in charge of national education, was established and many Chinese classics began to be taught. In the sixteenth year of King Kyengtek (757), native place names were changed to two-character SK, and in his eighteenth year, all official titles were also Sino-Koreanized. Personal names of upper-class people also began to be Sino-Koreanized during the Silla period.

Native words began to be overwhelmed by SK words in the Kolye dynasty. This was particularly the case after King Kwangco adopted, in 958, the Chinese system of civil-service examinations based on the Chinese classics. The Chosen dynasty observed all-out infiltration of Chinese words and characters into every facet of Korean culture and society, chiefly because of the dynasty's adoption of Confucianism as a national religion and, as a result, the popular admiration of everything Chinese.

Due to the massive importation of words from this prestigious language during the MK period, Korean subsequently underwent a wide range of changes, not only in its vocabulary but also to a lesser degree in the sound system and word structure. There are several consequences of the all-out importation of Chinese words. First of all, Korean vocabulary was enormously enriched with various cultural words and abstract terms from Chinese. Thus, MK texts are full of content words written in Chinese characters, as illustrated in (89) where SK words are underlined.

- (89) sinku*n*    *i*            ljek    *ul*            tukhó-mjen    iltjenghi    *tinj*-e    ujsim  
trust.root    NOM            power    ACC            get-if            surely    carry-so    doubt

*ani*    hó-ko  
not    do-and

'If one's root of trust gets power, one constantly carries it and does not doubt, and [. . .]' (Welsek 7:45)

Due to the inundation of SK words, many native Korean words, even basic vocabulary items, have gradually been wiped out as a result of the battle for survival in doublet situations. Many of those native words which existed during the MK period have subsequently given in to SK, e.g. (36b), where native words *cumun* '1,000' and *kólóm* 'river, lake' are no longer used in CK, replaced by SK equivalents **chen** and **kang**.

Examples of such extinct nouns include *kasi* ‘wife’, *kuze* ‘shelter’, *kuwuj* ‘government office’, *kilóma* ‘saddle’, *kip* ‘silk’, *naco* ‘evening’, *naks* ‘taxes’, *nap* ‘monkey’, *njenu* ‘other person’, *nuc* ‘symptoms’, *tiwuj* ‘border’, *moj* ‘mountain’, *mwulwuj* ‘hail’, *mecun mal* ‘curse’, *mójh* ‘field’, *pijwuk* ‘chick’, *sjaong* ‘husband’, *swulis nal* ‘Tano Day’, *swum thón kes* ‘living thing’, *sjwulwup* ‘umbrella’, *sjeks* ‘reins’, *cas* ‘castle’, *cjenchó* ‘reason’, *cjwuljen* ‘female’s towel, handkerchief’, *cumkej* ‘trees and shrubs’, *cumkej* ‘limit, end’, *chól* ‘source, origin’, *hungcjeng-pachi* ‘merchant’, *azóm* ‘kindred, relative’, *achón atól* ‘nephew, niece’, *jak* ‘a kind of turtle’, *ezi* ‘parents’, and *ezi atól* ‘parents and children’. On a smaller scale, verbs, adjectives, adverbs, and other word classes have also lost native words in favor of SK counterparts. Furthermore, many SK words have been nativized with radical sound changes to the extent that general native speakers in CK no longer recognized them as derived from SK (e.g. Shim 1983). Examples include **salang** ‘love’ (from SK *saljang* ‘think’), **kakey** (from SK *kaka* ‘temporary house’), **kimchi** (from SK *chimchaj* ‘soaked vegetable’), and **paychwu** (from SK *pajkchaj* ‘white vegetable’).

## ABBREVIATIONS

ACC	Object (accusative) particle
AH	Addressee honorific suffix
CK	Contemporary Korean
CLASS	Classifier
DAT	Dative particle
DEC	Declarative ender
DF	Definitive
EMK	Early Middle Korean
EMP	Emphasis
EXC	Exclamatory ender
GEN	Genitive particle
IMP	Imperative ender
IND	Indicative mood suffix
INF	Infinitive suffix
LMK	Late Middle Korean
LOC	Locative/goal particle
MK	Middle Korean
NMR	Nominalizer
NOM	Subject (nominative) particle
OH	Object honorific suffix
PAST	Past suffix
POL	Polite ender
PROM	Promissive ender
PROP	Propositive ender
PROS	Prospective suffix or ender
Q	Question (interrogative) ender
REL	Relativizer suffix
RETR	Retrospective mood suffix
SH	Subject honorific suffix
SK	Sino-Korean
SS	Speaker-subject suffix
TOP	Topic particle
VOC	Vocative particle

## NOTES

- 1 P. Nam (2009, and this volume) claims that Old Korean extends to the middle of the thirteenth century. His argument is based on the marked difference he observed between the grammar of *sektok kwukyel* available until the middle of the thirteenth century and the grammar of *swuntok kwukyel* used thereafter.
- 2 Scholars are not in agreement as to the exact number of vocabulary items in *Kyeylim yusa*: 353 (à la B. An 1985: 25 and others' counts), 355 (S. Kang 1980), 356 (Kwuklip Kwuke Yenkwuwen 1999), 359 (Chin 1974), and 361 (T. Kim 2002).
- 3 A number of studies on the Kolye dialect (EMK) written in *Kyeylim yusa* are available, including monographs such as Chin (1974), S. Kang (1980), and B. An (1985). A recent grammatical summary is found in Chung (2004).
- 4 There are several sounds in MK which are not shared in CK. For these sounds, new symbols are introduced in addition to the Yale system:  $\beta$  (뵐),  $z$  (△),  $h$  (○),  $\rho$  (⊖),  $ó$  (·), and  $j$ . The semivowel  $j$  is used here in MK instead of CK  $y$ , because the MK semivowel functioned as both on-glide and off-glide, whereas the CK counterpart functions only as on-glide and the original off-glide has merged with the preceding vowel, together generating a separate vowel phoneme: e.g. MK *saj* [saj] vs. CK *say* [sɛ] 'bird'.
- 5 For detailed studies of overall MK structure, see P. An and Lee (1990); B. Choy (1990); Y. Choy (1987); Huh (1992, 1975); K. Kang and Hwang (2003); D. Kim (2003); M. Kim (1957); Ko (2007); Ko and Nam (1999); Kwon (1998); H. Lee (1994); K. M. Lee (1976); Lyu and Kim (2005); and Pak (1989), among other works. For studies on MK phonology, see N. Chang (1982); S. Kang (1990); W. Kim (1971); K. M. Lee (1972), and Moon (1974), among other studies.
- 6 K. M. Lee (1976) presumes that this implosivization was responsible for  $r$  and  $l$  becoming a single phoneme in MK, [l] occurring in the syllable-final position and [r] elsewhere.
- 7 For studies on MK tones, see Jeong (1971, 1972); Huh (1972); W. Kim (1973); S. O. Lee (1979); Martin (1992: 60–86); and Ramsey (1978), among other works.
- 8 This alphabet was named hankul (meaning 'Great Writing') by Cwu Sikyeng (1876–1914), a pioneer linguist of Korean. Before, it was popularly called *enmun* 'vernacular writing, vulgar script'. Works in English on the study of hankul include: Ledyard (1966); C. Lee (1972), Kim-Cho (2000); and Kim-Renaud (1997). South Koreans celebrate October 9 as Hankul Day, which is the day hankul was promulgated by King Seycong in 1446; North Koreans erase Seycong and celebrate January 15.
- 9 A small number of nouns that end in a vowel in their citation form replace this with an underlying  $k$  before vowel-initial particles, e.g. *namo* (< \**namok*) vs. *namk ón*.
- 10 Note that this word was distinguished from the adverb *noph-i* 'highly' in MK. The two words have merged as homonyms in CK.
- 11 Notice that *ekuj* was nasalized after a demonstrative. Similarly, the locative *ej* 'at, in' became nasalized after a demonstrative, as in *i-ngej*, *ku-ngej*, and *tje-ngej*.
- 12 It has a number of phonologically conditioned allomorphic shapes. Its initial consonant  $s$  changes to  $z$  (e.g. *-zóp*) after a voiced sound and to  $c$  after  $t$ ,  $c$ , or  $ch$  (e.g. *-cóp*). Its final consonant  $p$  changes to  $\beta$  before a vowel (e.g. *-sóβ*, *-zóβ*, *-cóβ*).

## REFERENCES AND FURTHER READING

- An, Byongho (1985) *Kyeylim Yusa wa Kolye Siki Cosene* [*Kyeylim yusa* and the Korean language in the Kolye period], China: Heilongjiang Chaoxian Minzu Chubanshe.
- An, Pyonghui (1977) *Cwunsey Kwuke Kwukyel uy Yenkwu* [A study of Middle Korean *kwukyel*], Seoul: Ilcisa.

- (1992) *Kwukesa Yenkwu* [A study of Korean linguistic history], Seoul: Munhak kwa Cisengsa.
- An, Pyonghui and Lee, Kwangho (1990) *Cwungsey Kwuke Munpeplon* [A grammar of Middle Korean], Seoul: Hakkwusa.
- Chang, Namgui (1982) ‘Phonological Variation in 15<sup>th</sup> Century Korean’, *Journal of Chinese Linguistics* Monograph 1.
- Chang, Yunhui (2002) *Cwungsey Kwuke Congkyelemi Yenkwu*. [A study of Middle Korean sentence enders], Seoul: Thayhaksa.
- Chin, Taeha (1974) *Kyeylim Yusa Yenkwu* [A study on *Kyeylim yusa*], Seoul: Thap Chwulphansa.
- Choy, Beomhun (1990) *Hankwuke Paltalsa* [A developmental history of the Korean language], Seoul: Kyengwun Chwulphansa.
- Choy, Yunkap (1987) *Cwungsey Cosene Munpep* [A Middle Korean grammar], Yanji, China: Yanbian University Press.
- Chung, Jaeyoung (2004) ‘*Kyeylim Yusa* ey nathanan munpep hyengthay ey tayhan yenkwu’ [A study of grammatical forms in *Kyeylim yusa*], *Kwukyel Yenkwu*, 12: 99–132.
- Han, Jaeyoung (1996) *16 Seyki Kwuke Kwumun Yenkwu* [A study of sixteenth-century Korean], Seoul: Sinkwumunhwasa.
- Huh, Woong (1972) *Cwungsey Kwuke Yenkwu* [A study on Middle Korean], Seoul: Cengumsa.
- (1975) *Wuli Yeys Malpon* [A historical grammar of Korean], Seoul: Saym Munhwasa.
- (1983) *Kwukehak* [Korean linguistics], Seoul: Saym Munhwasa.
- (1992) *15-16 Seyki Wuli Yeys Malpon uy Yeksa* [History of fifteenth and sixteenth-century Korean], Seoul: Thap Chwulphansa.
- Jeong, Yeonchan (1971) ‘Cwungsey sengco wa kyengsangto pangen sengco uy pikyo’ [A comparison between Middle Korean and Kyengsang tones], *Hankul Hakhoy 50<sup>th</sup> Anniversary Anthology*, Seoul: Korean Language Society, 29–46.
- (1972) ‘Cwungsey kwuke sengco uy pyentong kwa kiponhyeng’ [Variations in Middle Korean tones and their basic forms], *Han-geul* 150.
- Kang, Kyuseon and Hwang, Kyeongsu (2003) *Cwungsey Kwuke Munpeplon* [Middle Korean grammar], Seoul: Chengwun.
- Kang, Sinhang (1980) *Kyeylim Yusa Kolye Pangen Yenkwu* [A study of Kolye dialect in *Kyeylim yusa*], Seoul: Sungkyunkwan University Press.
- (1990) *Hwunmin Cengum Yenkwu* [A study of *Hwunmin Cengum*], Seoul: Sungkyunkwan University Press.
- (1995) *Cosenkwan Yeke Yenkwu* [A study on *Cosenkwan Yeke* ‘Korean-Chinese Translated Lexicon’], Seoul: Sungkyunkwan University Press.
- Kim, Dongso (2003) *Cwungsey Hankwuke Kaysel* [An introduction to Middle Korean], Seoul: Hankwukmunhwasa.
- Kim, Minsu (1957) *Cwuhay Hwunmin Cengum* [Annotated *Hwunmin Cengum*], Seoul: Thongmunkwan.
- Kim, Taewan (2002) ‘Kolye-e uy “kyey, thwu, pang” mo punsek ul thonghan *Kyeylim yusa* yukium kochal’ [A study of aspirates in *Kyeylim yusa* through an analysis of ‘kyey, thwu, pang’ sounds], *Cwunge Cwungmunhak*, 30: 55–68.
- Kim, Wanjin (1971) *Kwuke Umun Cheykyey uy Yenkwu* (A study of the Korean phonological system), Seoul: Ilcokak.
- (1973) *Cwubgsey Kwuke Sengco Yenkwu* [A study of Middle Korean tones], Seoul: Thap Chwulphansa.
- Kim, Y.O. [Yeng Wuk] (1995) *Munpep Hyengthay uy Yeksacek Yenkwu* [A historical study of grammatical forms], Seoul: Pakiceng.
- Kim-Cho, Sek Yen (2000) *The Korean Alphabet of 1446: The Orthophonic Alphabet for the Instruction of the People*, Amherst, NY: Humanities Books.

- Kim-Renaud, Young-Key (ed.) (1997) *The Korean Alphabet*, Honolulu: University of Hawai'i Press.
- Ko, Younggeun (1998) *Cwungsey Kwuke uy Sisang kwa Sepep* [Tense, aspect, and modality in Middle Korean], Seoul: Thap Chwulphansa.
- (2007) *Phyocwun Cwungsey Kwuke Munpeplon, Katcengphan* [Standard Middle Korean grammar – revised edition], Seoul: Cipmuntang.
- Ko, Younggeun and Nam, Kishim (eds) (1999) *Cwungsey.e Calyo Kanghay* [Annotations of Middle Korean data], Seoul: Cipmuntang.
- Kwon, Jaeil (1998) *Hankwuke Munpepsa* [A grammatical history of Korean], Seoul: Pakiceng.
- Kwu, Pongwan (1998) *15 Seyki Kwuke Phasayngpep ey tayhan Yenkwu* [A study of Middle Korean derivational morphology], Seoul: Thayhaksa.
- Kwuklip Kwuke Yenkwuwun (KKY) (1999) *Phyocwun Kwuke Taysacen* [A standard dictionary of Korean], Seoul: KKY.
- Ledyard, Gari K. (1966) 'The Korean language reform of 1446: the origin, background, and early history of the Korean alphabet,' unpublished PhD dissertation: University of California at Berkeley.
- Lee, Chongho (1972) *Haysel Yekcwu Hwunmin Cengum* [Annotated and translated *Hwunmin Cengum*], Seoul: Korean Library Science Research Institute.
- Lee, Hyonhi (1994) *Cwungsey Kwuke Kwumun Yenkwu* [A study of Middle Korean constructions], Seoul: Sinkwumunhwasa.
- Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, Albany: State University of New York Press.
- Lee, Ki-Moon (1972) *Kwuke Umunsa Yenkwu* [A study of Korean phonological history], Seoul: Thap Chwulphansa.
- (1976) *Kayceng Kwukesa Kaysel* [A revised introduction to the history of Korean], Seoul: Mincwung Sekwan.
- Lee, Sang-Oak (1979) *Middle Korean Tonology*, Seoul: Hansin.
- Lyu, Lyol and Kim, Inho (2005) *Cosenhacense 4: Cosenmal Lyeksa* [A history of Korean] vols 1, 2, and 3, Pyongyang: Sahoykwahak Chwulphansa.
- Martin, Samuel E. (1992) *A Reference Grammar of Korean*, Rutland, VT and Tokyo: Tuttle.
- Moon, Yangsoo (1974) 'A phonological history of Korean,' unpublished PhD dissertation: University of Texas at Austin.
- Nam, Kwangu (1982) *Pocung Ko-e Sacen* [A revised dictionary of Old Korean words], Seoul: Ilcokak.
- Nam, Punghyon (1999) *Kwukesa lul wihan Kwukyel Yenkwu* [Kwukyel studies for Korean language history], Seoul: Thayhaksa.
- (2009) *Kotay Hankwuke Yenkwu* [A study of Old Korean], Seoul: Sikan uy Mulley.
- Pae, Taeon (2002) *Itwu Munpepsa uy Thongsicek Yenkwu* [Diachronic study of grammatical elements in Itwu], Masan: Kyongsang University Press.
- Pak, Byongchae (1989) *Hankwuke Paltalsa* [History of Korean language development], Seoul: Seyyengsa.
- Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects: A Descriptive and Historical Study*, Seoul: Thap Chwulphansa.
- Shim, Jaegi (1983) *Kwuke Ehwilon* [Korean lexicology], Seoul: Cipmuntang.
- Sohn, Ho-min (1998) 'Evolution of addressee honorifics in Korean' in S. O. Lee and D. S. Park (eds) *Perspectives on Korea*, Sydney: Wild Peony.
- (2001) *The Korean Language*, Cambridge: Cambridge University Press.

# MODERN KOREAN<sup>1</sup>

*Young-Key Kim-Renaud*

## 5.1 INTRODUCTION

By Modern Korean we refer to contemporary Korean, called *Hyentay kwuke* in Korean.<sup>2</sup> The period spans from the Kabo Reform (1894), the cradle of Korea's modernization and enlightenment movement, through the Japanese occupation (1910–1945) and the Korean War (1950–1953), to today, when the Korean peninsula remains divided into North Korea and South Korea.

One of the issues in Modern Korean has been what it means to 'write in Korean'. King Sejong (r. 1418–1450), who invented the Korean alphabet in the mid-fifteenth century, and most people until recently have considered writing the Korean alphabet mixed with Chinese characters just like writing only in Korean, even if the texts looked all Chinese only with Korean endings (K.-M. Lee 1961/72: 224). Korean is rich in Chinese-based vocabulary and the etymology of a word is apparent when written in Chinese characters.

On the verge of losing the country to the Japanese in the late nineteenth century, educated Korean patriots 'discovered' and took the Korean alphabet as a salient symbol of national identity. They renamed it 'hankul' (the great/unique/Korean writing; alternatively romanized as han'gŭl). The nationalist, independence fighters insisted on writing everything only in hankul. The first and best known example of such writing was *Toklip Sinmun* (Independence Newspaper), which had a simultaneously published English version called *The Independent* (1896–1899). By 1897 there were about nine other newspapers publishing solely using hankul. And yet, between 1898 and 1910, there were more than 30 that continued printing with mixed script, some of which even included Japanese *kana* (Hong *et al.* 2000: 18–19).

Koreans, like the Japanese, kept making new words based on Sino-Korean or Chinese character-based roots (*hanca-e*), and writing in mixed script was considered clearer and more formal. This practice prevailed until recently in South Korea. However, in North Korea, with the introduction of the *Cwuchey* (Self-reliance) ideology, the language purification movement started in the 1960s, long before South Korea, whose policy on mixed script shifted off and on throughout the second half of the twentieth century. Today, even South Korean publications, except for a small percentage of academic papers and ritual documents, are all in hankul. However, in the South, this is generally more for pragmatic and aesthetic reasons, such as its simplicity and practicality for typing, etc., than a nationalistic one.

One of the most significant achievements by early scholars was to establish modern orthography, culminating in the *Hankul Macchwumpep Thongil An* (Proposal for a Unified Spelling System) in 1933 by the Hankul Society. Some contemporary linguists lament the effect of this spelling reform, which narrowed down dialectal differences and caused various dialects to lose their original characteristics (K.-M. Lee 1961/72: 224–5). There are still at least six distinct regional dialects in Korea, but there are progressively

fewer dialectal speakers due to a standardization effort through education, media, and social pressure. Periodic spelling reforms help hankul to maintain its efficacy by keeping the ‘linguistic fit’ between spoken and written language, in spite of language change. The most recent reform took place in 1988 in South Korea.

The Japanese occupation of Korea for 35 years has left some vestige in Korean in the form of loanwords from Japanese, Sino-Japanese neologism, and Western loanwords in Japanese. However, in both North and South Korea, purification movements put central focus on eradicating traces of Japanese ‘infiltration’ into Korean, and only a few, non-obvious ones remain.

Some clear regional diversification of language has evolved through more than half a century’s separation between the North and the South, not only through geographical distance, but perhaps more importantly, through ideological and political pressures imposed on how people ought to speak and write. Both North and South Korea have undergone various standardization reforms, which also reflect linguistic theories such as how ‘abstract’ the writing should be. For example, North Korean spelling makes it mandatory to write underlying initial consonants, which are not realized or changed into another distinctive class of sound in South Korean spoken language. North Korean spoken language has been reinforced through this orthographic convention, and even in spoken language these consonants are more clearly pronounced in the careful speech of North Koreans. In this light, efforts by some linguists including Samuel E. Martin can be appreciated, who use superscript *n* or *r* to mark words in transcription that have initial *n*- or *r*- in the North that is lacking in the South, e.g. *ˈilkop* ‘seven’ or *ˈnotong* ‘labor’. Although standardization reforms have evolved independently of each other, there has been no serious language divergence between the two Koreas other than some political or ideological vocabulary, and both sides basically continue to speak the same language.

One salient characteristic of the language spoken in South Korea since the end of the Korean War is the enormous amount of borrowing that has taken place from foreign languages – predominantly from English, but also from other European languages such as French.

In the last decade or so, South Korea has emerged as an IT country, and the internet has changed how Koreans conduct business, live their lives, and communicate with each other. The impact of web-based life on language is quite substantial, but since it is an ever-evolving process, its enduring effect on the structure of Korean is not yet clear.

The language analyzed in this chapter is a variety referred to as ‘Standard Korean’, defined usually as ‘the language of middle class Seoulites’, but in fact an old elite class variety spoken in Seoul that has enjoyed prestige for a long time. In 1988, the Republic of Korea Ministry of Culture and Education redefined it as ‘the language of modern educated Seoulites’ but there are plenty of educated Koreans living in Seoul who speak other regional dialects.<sup>3</sup>

## 5.2 PHONOLOGY

Korean is a syllable-timed language, rich in its sound inventory, sound symbolism, and allophonic and morphophonemic alternations.

There is a close fit between the Korean alphabet and the phonemes of Korean, and hankul letters can be used for phonetic notation (Kim-Renaud 1997: ix). Korean has 19 consonants, ten vowels, and two glides or semivowels /w/ and /y/. Syllables and other linguistic units as well as their boundaries play a significant role in Korean phonology.

Many phonological alternations follow universal phenomena such as assimilation, but the following characteristics are crucial for understanding various Korean phonological alternations, including the various sandhi phenomena:

- Sound symbolism is prevalent and systematic. Consonantal strength and vowel harmony play significant roles in sound symbolic words.
- Syllable-final consonants are unreleased before another consonant and at major grammatical boundaries.
- The intervocalic position is conducive to consonantal weakening.

**5.2.1 Consonants**

Korean has 19 consonants as shown in Table 5.1. Middle Korean phonemes represented by the letters ㅃ, ㅆ, ㅉ, ㅊ have vanished from the inventory.

Korean has no voiced obstruents but shows a three-way distinction in voiceless stops: lax (*p, t, c, k*), heavily aspirated (*ph, th, ch, kh*), and tense (*pp, tt, cc, kk*). Table 5.2 shows some examples. Table 5.3 shows the consonantal strength scale, significant in sound symbolism. The hankul letter shapes reflect the consonantal strength scale: the lax consonants at the bottom carry the least marked forms of the pertinent group, and letter forms for stronger sounds are made by adding more strokes.

**TABLE 5.1 CONSONANTS OF MODERN KOREAN**

	Bilabial	Dental/ alveolar	Palatal	Velar	Glottal
Plosive/Affricate:	p ㅍ	t ㅌ	c ㅊ	k ㅋ	
Aspirated:	ph ㅍᄀ	th ㅌᄀ	ch ㅊᄀ	kh ㅋᄀ	
Tense:	pp ㅍㅍ	tt ㅌㅌ	cc ㅊㅊ	kk ㅋㅋ	
Fricative:		s ㅅ			h ㅎ
Tense:		ss ㅆ			
Nasal:	m ㅁ	n ㄴ		ng ㅇ	
Tap/Lateral:		l ㄹ			

**TABLE 5.2 THREE-WAY CONTRAST AMONG KOREAN VOICELESS STOPS**

Lax/Plain C	Aspirated Ch	Tense CC
<i>pul</i> ‘fire/light’	<i>phul</i> ‘grass’	<i>ppul</i> ‘horn’
<i>tal</i> ‘moon’	<i>thal</i> ‘mask’	<i>ttal</i> ‘daughter’
<i>ca</i> ‘sleeps’	<i>cha</i> ‘is cold’	<i>cca</i> ‘is salty’
<i>kum</i> ‘gold’	<i>khum</i> ‘being big’	<i>kkum</i> ‘pulling’

**TABLE 5.3 KOREAN CONSONANTAL GROUPS ACCORDING TO THEIR STRENGTH**

Stronger ↑	Tense	ㅍㅍ pp	ㅌㅌ tt	ㅆㅆ ss	ㅊㅊ cc	ㅋㅋ kk	
	Aspirated	ㅍᄀ ph	ㅌᄀ th		ㅊᄀ ch	ㅋᄀ kh	ㅎ h
	Plain	ㅍ p	ㅌ t	ㅅ s	ㅊ c	ㅋ k	
	Sonorant	ㅁ m	ㄴ n		ㄹ l	ㅇ ng	

### 5.2.2 Vowels

Modern Korean has the vowel inventory given in Table 5.4. The unrounded, high back vowel /u/ is the weakest and most unstable Korean vowel. It is obligatorily deleted whenever it meets another vowel across a grammatical boundary, as in *sa-* ‘buy’ + *-umyen* ‘if’ → *samyen* ‘If (someone) buys’. /u/ is also often deleted in a weak or non-initial position in casual or fast speech, e.g. [kusl] for /kwusul/ ‘bead’. When a vowel is needed to break an impermissible consonant cluster in pronouncing a loanword, /u/ is inserted, as in /aisukhylim/ for ‘ice-cream’.

Modern Korean has complex syllable nuclei (Table 5.5), but unlike in Middle Korean, where off-glides were abundant, all glides are on-glides with the one exception of the *y* off-glide that follows the vowel /u/ as in *uy* ‘meaning’. This diphthong, written <-|>, is composed of /u/ and /i/, graphically containing both vowels, <-> and <|>. It is pronounced [ui] or [wi] in a word-initial syllable, which is phonologically strong. In a non-initial syllable the weak *u* is commonly deleted and it is the second part *i* that gets fully pronounced. So, we get [ui̯sa] ~ [usa] for /uy-sa/ ‘intention’ but [koui] ~ [koi] for /ko-uy/ ‘on purpose’.

The semivowel *y* can precede any vowel except unrounded high vowels. Thus there are ㅑ *ya*, ㅓ *ye*, ㅛ *yo*, and ㅠ *yu*, but no \*[ji] and \*[ju]. The semivowel /w/ can precede any vowel except /o/, /u/, and /wu/. So, the diphthongs \*[wo], \*[wu], and \*[wu] are not allowed. In both cases, the constraint may be viewed as the effect

**TABLE 5.4 VOWEL INVENTORY OF MODERN KOREAN (‘BRIGHT’ VOWELS ARE SHADED)**

Front-Backness Lip-Rounding	Nonback		Back	
	Unrounded	Rounded	Unrounded	Rounded
High	/i/	ㅜ   /wi/	ㅡ /u/	ㅜ /wu/
Mid	ㅝ   /ey/	ㅝ   /oy/	ㅝ   /e/	ㅝ   /o/
Low	ㅘ   /ay/		ㅘ   /a/	

NB: In Yale romanization /ay/ represents IPA [ɛ], /wi/ IPA [y], /oy/ IPA [ø], /u/ IPA [u], and /e/ IPA [ɛ]. Yale uses /u/ for [u] after labials, as [w] does not occur morpheme-internally after labials.

**TABLE 5.5 KOREAN DIPHTHONGS (‘BRIGHT’ VOWELS ARE SHADED)**

Front-Backness Tongue Height	With Nonback Vowels		With Back Vowels	
	High	ㅝ   /uy/ [ui̯]	ㅜ   /wi/	
Mid	ㅝ   /wey/ [we]	ㅝ   /oy/, [ø] or [we]	ㅜ   /we/[wɔ̯], ㅝ   /ye/[jɔ̯]	ㅝ   /yo/
Low	ㅘ   /way/ [wɛ], ㅘ   /yay/ [jɛ]		ㅘ   /wa/, ㅘ   /ya/	

of a particular type of the Obligatory Contour Principle, i.e. to avoid two similar sounds in sequence.

Most graphically complex vowels such as  $\text{ㅓ}$  /ey/,  $\text{ㅕ}$  /ay/,  $\text{ㅗ}$  /wi/, and  $\text{ㅛ}$  /oy/ were complex syllable nuclei in Middle Korean but are now single phonemes. Rounded front vowels,  $\text{ㅜ}$  /wi/ and  $\text{ㅠ}$  /oe/, pronounced [y] and [ø] respectively, are being re-diphthongized in Standard Korean as [wi] and [we], especially when they are not preceded by a consonant; they are marked vowels, and lip-rounding of the front vowel is being replaced by *w*, an on-glide with lip-rounding.

### 5.2.3 Suprasegmentals

#### 5.2.3.1 Vowel length

For a very few older speakers of Seoul Standard Korean, vowel length is linguistically significant. Today, however, vowel length as a distinctive feature is almost completely gone. Early Middle Korean was a pitch-accent language, and the newly created alphabet employed three tone marks, and the rising tone mark <: > also indicated vowel length. Soon the marks stopped being written. This orthographic change, which probably reflected language variation and ongoing change, represented and even contributed to the disappearance of distinctive vowel length. Today, vowel length is used mainly in exaggerated or emphatic expressions. Emphatic vowel lengthening is predictable and typically occurs in the first syllable of a word:

- (1) Emphatic vowel lengthening
- |                  |                   |  |
|------------------|-------------------|--|
| <i>acwu</i>      | <i>a:cwu</i>      | ‘very’   |
| <i>khukey</i>    | <i>khu:ke:y</i>   | ‘in a big/huge way (loud, in big letters, etc.)’ |
| <i>ccokkuman</i> | <i>cco:kkuman</i> | ‘tiny’   |

#### 5.2.3.2 Accentual patterns

Modern Standard Korean is neither a tone language like Chinese, a pitch-accent language like Japanese, nor a stress language like English. However, all languages have *prosody*, which determines the sentential meaning. In Standard Korean, a common intonation pattern of a sentence is a sequence of rising tone for each phrase except for the last word. The tonal pattern can differ across dialects. For example, in the southwestern dialect, each phrase has a falling tone.

The boundary tone, especially the one on the last syllable of the sentence, indicates the type of sentence, e.g. a statement or a question. For example, a statement or a request has a falling tone, while a WH-question has a rising–falling contour. A yes–no question has a rising contour. Table 5.6 shows how the meaning can change depending on the tone

**TABLE 5.6 MEANING ACCORDING TO THE INTONATION CONTOUR IN KOREAN**

Intonation Contour on the final syllable	Sentence Type	Pronunciation	Meaning
Low (↓):	Statement	<i>ka</i> ↓	‘(x) goes.’
	Request	<i>ka</i> ↓	‘Go!’
Falling (↘):	WH-question	<i>way ka</i> ↘	‘Why are you going?’
High (↑):	Yes–no question	<i>ka</i> ↑	‘Are you going?’

of the last syllable of a word. The boundary tone may also convey the emotion or the attitude of the speaker. For example, a rising–falling sentence-final tone can convey insistence. Therefore, in *ike kunyang pata* ‘Just take it!’, if the last syllable *ta* is produced with a low pitch it conveys ordering someone to take something without refusing, but if produced in a rising–falling pitch with a lengthened last syllable it shows that the speaker is *urging* the interlocutor to take it, whereas if the last syllable shows a falling–rising–falling contour with further lengthening of the syllable (*pataaa*), it would express *frustration*.

## 5.2.4 Morphophonology

### 5.2.4.1 Vowel harmony

Korean vowels are categorized in two contrasting groups. One group, called *yang* or ‘bright’ vowels, includes nonhigh vowels /a/, /o/, /ay/, /oy/, while the other group, called *um* or ‘dark’ vowels, includes nonlow vowels /i/, /u/, /e/, /wu/, /ey/, /wi/. Within a sound symbolism word (see 5.6.1.1), vowels must be all bright or all dark except the vowels *i* and *u*, which are neutral or ‘invisible’ to the Vowel Harmony (VH) process in a weak non-initial syllable and as such may co-occur with either dark or bright vowels.

VH is applied systematically in certain verb conjugations. When the final syllable of a verb has the yang vowel *o* or *a*, the initial *e-* of a suffix changes to its yang counterpart *a-*, as shown in the alternation of the past tense forms [mag-at<sup>ˀ</sup>] (← /mak-ass/) for /mak-ess/ ‘blocked’ and [mɔŋɔt<sup>ˀ</sup>] (← /mek-ess/) for /mek-ess/ ‘ate’, for the two verb stems *mak-* and *mek-*, respectively. However, VH applies only to an *e*-initial suffix when it immediately follows a verb stem. For example, the word /cop-ess-eyo/ ‘was narrow’ is changed to [cob-ass-ɔjo], where VH occurs only in the first suffix, and the second *-e* does not change to *-a* because it does not follow directly after a verb stem. Sequences of identical vowels merge, e.g. *ka-ss-eyo* for \**ka-ass-eyo* ‘went’ or *se-ss-eyo* for \**se-ess-eyo* ‘stopped’, vs. *po-ass-eyo* ‘saw’.

The problem with Korean VH is that the harmonic groups do not form phonetically natural classes in contemporary Korean, as can be seen in Tables 5.4 and 5.5. This is due to historical change and none of the numerous efforts to find a synchronically appropriate featural distinction of the two groups has proven convincing. Furthermore, various ongoing language changes, especially *o*-raising (to *wu*) in non-initial syllables, are rapidly bringing down VH as a productive process.

### 5.2.4.2 Unreleasing of syllable/word-final consonants

One of the most salient characteristics of Korean phonology is that syllable/word-final consonants retain oral contact after articulation, unless they are followed by a vowel or a diphthong. The unreleased consonants are prone to further changes and get easily assimilated to a following consonant. At the same time the heightened air pressure coming from unreleasing in the case of obstruents causes strengthening of a following consonant. Indeed, the unreleasing requirement provides a simple explanation for a large part of seemingly complex morphophonemic alternations in Korean.

Syllable-final sonorants are also unreleased, but its effect is minimal, as the air escapes either through the nose or the sides of the tongue despite the nonrelease, as can be observed in such examples as /ttam/ [ttam<sup>ˀ</sup>] ‘sweat’, /son/ [son<sup>ˀ</sup>] ‘hand’, and /cong/ [coŋ<sup>ˀ</sup>] ‘bell’.

The vowel is particularly important in distinguishing between the three-way contrast in stops. When there is no vowel, the important cues for such crucial features as degree of aspiration and tenseness are lost, as there is no observable Voice Onset Time. When a word ends in a consonant, as many as nine different phonemes (*t*, *th*, *tt*, *c*, *ch*, *cc*, *s*, *ss*, *h*) are neutralized as [t̚].<sup>4</sup>

Because unreleasing requires oral contact, fricatives (*s*, *ss*, *h*), sounds produced through a narrow opening in the vocal tract, take the closest, most unmarked stop [t̚] to achieve unreleasing, as in /nos/ [not̚] ‘brass’.

All the affricates (*c*, *ch*, *cc*) become unreleased as a dental stop, as in /nach/ [nat̚] ‘face’. An affricate begins as a stop and then releases as a fricative with or without strong aspiration. Naturally, when there is no release, the fricative part can no longer be pronounced, and dental stops and palatal stops are both pronounced the same as [t̚], and the underlying distinction is neutralized in a pre-boundary position.

The approximant *h*, having no point of contact in the oral cavity, is ‘unreleased’ by taking the most unmarked dental stop again, as in /hiuh/ [hiut̚] ‘the letter *h*’.

#### 5.2.4.3 Law of Initials

South Koreans avoid pronouncing the consonant *l* in word-initial position. Most *l*-initial morphemes are Sino-Korean, and the lateral is usually changed to [n]. However, word-initially, a sequence of *ni* or *ny* is also not felicitous for South Korean speakers, who drop the *n* of the sequence in a word-initial position. Koreans call this the *Twuum pepchik* or ‘Law of Initials’. That is why the family name with the plain [i] sound in Standard Korean today is romanized as *Lee*, *Rhee*, *Ree*, *Ri*, *Lie*, *Leigh*, *Li*, etc. The Law of Initials does not apply in North Korean dialects, including their standard, *Munhwa-e* (‘culture language’). Some Korean family names are now written with an initial lateral even in South Korea, too.

Korean speakers are well aware of this positional variation in pronunciation, and South Korean orthography, unlike in North Korea, dictates writing the actual pronunciation, i.e. ㅇ <zero> or ㄴ <n> in that position, as in:

- (2) 이승만 /li-sungman/ [isun̚man] ‘Syngman Rhee’  
 노무현 /lo-muhyen/ [nomuh̥jɛn] ‘Roh Moo-hyun’

#### 5.2.4.4 Assimilation

In Korean, assimilation between neighboring sounds is common. Assimilation may occur in the manner and/or the place of articulation.

##### *Intervocalic voicing and other consonantal weakening*

When surrounded by sonorants, lax consonants assimilate to the sonorousness and weaken their consonantal strength.

(a.) *Intervocalic voicing*. Lax stops are voiced when surrounded by sonorants, e.g. /chil-po/ [c<sup>h</sup>ilbo] ‘seven jewels’, /catwu/ [cadu] ‘plum’, and /samkak/ [samgak] ‘triangle’.

Strongly aspirated stops and tense stops are too strong to weaken even when they occur intervocalically. As the lax sibilant is gently aspirated, which may explain why

there is no strongly aspirated sibilant and there is only a two-way lax (*s*) and tense (*ss*) contrast, intervocalic voicing also does not apply to the lax sibilant.

(*b.*) *h-weakening*. The approximant *h* is further weakened to a very gently articulated [h] or deleted in a sonorous environment. For example, *annyenghaseyyo* ‘Hello’ (lit. ‘Are you well?’) is pronounced [annjɤhasejo] in careful speech but [annjɤ.asejo] in fast speech, where the [h] is either completely deleted or shows its trace as a gentle glottal stop.

Verb-final [h] is deleted obligatorily when followed by a vowel, regardless of the kind of speech style or speed, as in /co**h**-ayo/ [coajo] ‘That’s good’.

(*c.*) *l-weakening*. Intervocalic weakening also applies to /l/, which is pronounced [ɾ], a flap with much less firm oral contact than [l], as in /ssa**l**/ [ssaɭ] ‘rice’ vs. /ssa**l**-i/ [ssari] ‘rice-NOM’. The *l*-weakening occurs even if an [h] intervenes, as in /ssa**l**-hako/ [ssar(**h**)ago] ‘with rice’. In this example, the *h*, which often undergoes weakening in its degree of aspiration, is deleted, but only after making the surrounding sounds [ɾ] and [a] slightly devoiced.

#### *Nasal assimilation*

As a syllable-final consonant is unreleased it is nasalized when the following syllable begins with a nasal. For example, the final consonant /s/ of the first morpheme in /peses-naysay/ ‘mushroom smell’ is unreleased as [t̚], which undergoes nasalization in assimilation to the following /n/, and the word is pronounced [pɤsɤnnɛmɕe].

#### *Palatalization*

Dental stops become palatal in front of the vowel /i/ or glide /y/ across a grammatical boundary. For example, when /kath-/ ‘be the same’ is followed by an adverbial suffix *-i*, it is pronounced [kac<sup>h</sup>i]. Note that in a monomorphemic word such as *mati* ‘node’ there is no palatalization.

However, in the case of other dental consonants, i.e. nasal *n*, lateral *l*, and fricative *s* or *ss*, palatalization applies even within a morpheme, e.g. [kɯnjɤm] for /kinyem/ ‘celebration’, [huɭɭjunha] for /hwullyungha/ ‘be excellent’, [ji] for /si/ ‘poem’, and [ʃji] for /ssi/ ‘seed’.<sup>5</sup>

#### *Lateral assimilation*

The sonorants *n* and *l* together in any order are pronounced as a double [ll], as in /sin-la/ [jilla] ‘Silla (kingdom)’ and /khal-nal/ [k<sup>h</sup>alla] ‘knife blade’.

#### *Consonant-cluster simplification*

Modern Korean prohibits multiple consonants in word-initial position, but underlyingly a syllable may have up to two consonants in the final position. When consonant clusters are followed by a vowel, they are split into two different syllables by the re-syllabification rule, which allows both consonants to be pronounced. However, when no vowel follows the consonant cluster, only one of them gets pronounced. In short, an ‘unattached’ consonant has to be dropped, and the question is which one becomes

unattached. Which one survives the simplification process is largely the function of unreleasing of the entire syllable coda.

- When the first member of a consonant cluster has a clear point of contact in the oral cavity, it will be pronounced, and the second consonant will be deleted, as in /kaps-to/ [kap<sup>ˀ</sup>to] (→ [kap<sup>ˀ</sup>tto]) ‘price also’ and /anc-ko/ [an<sup>ˀ</sup>ko] (→ [an<sup>ˀ</sup>kko]) ‘sit down and’.
- A consonant following /l/ may or may not be pronounced, because the lateral opening makes near co-articulation of /l/ with the non-homorganic second consonant possible/audible. However, the cluster is still not a preferred form and Korean speakers will still seek the stop consonant that permits complete oral closure. As a consequence, final clusters such as *-lk*, *-lp*, and *-lm* each have variant forms, e.g. [kalkttaɡa] as well as [kakttaɡa] for /kalk-taka/ ‘scratch and then’, [ttɘlpttɘra] ~ [ttɘlpttɘra] for /ttelp-tela/ ‘it [e.g. a persimmon] was astringent!’ There is dialectal variation: in some southern dialects it is the unreleased [p] that survives, as in [ccaptta] for /ccalp-ta/ ‘it is short’.
- However, if the second consonant has the same point of articulation as the [l] sound, there is no way a speaker can articulate another sound in the same place after unreleasing the first, and there is no optional co-articulation there. Only the first unreleased part is pronounced and the second one is dropped, as shown in the case of /hwulth-ko/ [hul<sup>ˀ</sup>ko] → [hul<sup>ˀ</sup>kko] ‘scan through and’.

Consonant-cluster simplification is thus a direct consequence of unreleasing the syllable coda. While unreleasing can cause weakening of the syllable-final consonant, it makes the following position stronger, as shown in the next section.

#### *Tensification of plain stops*

When a consonant follows a stop, the heightened oral pressure built up due to non-release of the first consonant causes tensing of the following consonant, as can be seen in /cip-sok/ [cip<sup>ˀ</sup>ssok<sup>ˀ</sup>] ‘inside of a house’, /cokca/ [cok<sup>ˀ</sup>cca] ‘scroll’, and /nach-cam/ [nat<sup>ˀ</sup>ccam] ‘nap’ (lit. ‘day sleep’).

#### *Tensification in compounds*

In compound words, the second morpheme’s first consonant would have normally weakened due to the weak nature of non-initial syllables. Korean speakers, subconsciously aware of this, try to make the boundary between the two components clear so that the initial of the second word is not weakened. For example, when the first word ends with a sonorant, the following lax consonant would have become voiced as the whole compound is pronounced without a pause between the subcomponents. In order to prevent such voicing and other forms of weakening, there is an effort to block the usual liaison, through unreleasing of the final sound of the first component word.

One of the efforts to make the boundary clear is to unrelease the final consonant of the first word, if there is one, as if it were a word boundary. However since the two components are pronounced with no break between them, the intervocalic voicing rule still applies if the second word starts with a vowel. Note, however, that unreleasing applies first and then intervocalic voicing. Therefore, the resulting pronunciation is rather removed from the underlying forms. For example, /swuh-oli/ ‘male duck’ is

pronounced as [sudori], as the final consonant /h/ of the first morpheme is first unreleased as [t̚], which in turn becomes voiced before the following vowel.

### *Epenthetic /t/*

When the first word of a compound ends in a sonorant, however, an artificial boundary is made by resorting to insertion of the unmarked stop /t/ to make the break clear, as in the word meaning ‘seaside’ containing two nouns, /pata/ ‘sea’ and /ka/ ‘border’, which is pronounced [padakka] < /pata-t-ka/. Traditionally linguists have believed that the tensing of the lax consonant is due to an inserted /s/ sound, which originates from a possessive marker, and for that reason it is called an ‘epenthetic-s’ (*saisiot*) or ‘Bindungs-s’. The inserted consonant is hypothesized to be *s*, because the unreleased [t̚] does often originate from an *s*. However *t*-insertion is a more plausible analysis partly because /t/ surfaces in the form of a voiced dental stop in such forms as [udot] < /wu-t-os/ for a compound word made of /wu-/ ‘top’ + /os/ ‘clothes’.

### *Epenthetic n*

A nasal [n] is inserted when the first noun ends with a consonant and the second noun starts with an *-i* or *-y* sound in compounds. This is a kind of backward derivation of the Law of Initials. The ‘deleted’ sound is put back: As the second word occupies a non-initial position, reinserting the ‘lost initial *n*’ is perfect in making the first syllable of the second component stronger. Note that the inserted [n] undergoes or causes phonological change depending on what it follows (e.g. *ln* → [ll], *kn* → [ŋn], *tn* → [nn]):

- |     |                          |   |                  |
|-----|--------------------------|---|------------------|
| (3) | /puekh-il/ [puɣŋnil]     | ← <i>puekh</i> ‘kitchen’ + <i>n</i> + <i>il</i> ‘work’    | ‘kitchen work’   |
|     | /kkoch-iph/ [kkonnip]    | ← <i>kkoch</i> ‘flower’ + <i>n</i> + <i>iph</i> ‘leaf’    | ‘flower petal’   |
|     | /yelum-yong/ [jɣrumnjoŋ] | ← <i>yelum</i> ‘summer’ + <i>n</i> + <i>yong</i> ‘use’    | ‘summer use’     |
|     | /kwul-yoli/ [kuɻɻjori]   | ← <i>kwul</i> ‘oyster’ + <i>n</i> + <i>yoli</i> ‘cuisine’ | ‘oyster cuisine’ |

### *Aspiration*

A stop adjacent to the approximant *h* is heavily aspirated, as in /neh-ko/ [nɛk<sup>h</sup>o] ‘insert and’, and /pap-hako/ [pap<sup>h</sup>ago] ‘with rice’, where the two consonants just coalesce.

## 5.3 ORTHOGRAPHY

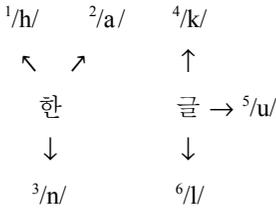
Until recently, Korean was written in a mixed script, where Chinese characters (*hanca*) are written for Sino-Korean lexical items, especially in formal and academic writing. Today, however, even these are almost completely written only in hankul.

The alphabetic symbols used today are essentially the same as those of *Hwunmin ceng’um*, except the older distinction between ◯ [‘Ø’ or dummy initial ‘consonant’] and phonemic ◊ *ng* is no longer maintained, as *ng* only occurs syllable-finally in Modern Korean.

Hankul is unusual in that the letters are not written linearly like in most other alphabets but assembled in syllable blocks, each letter occupying a particular recognizable space within a syllable. Writing in syllables made it possible to write horizontally or vertically. Modern Korean, especially today, is written horizontally from left to right. For example,

the two-syllable word 한글 ‘hankul’ is written left to right, each letter in the syllable written from top left, as follows:

(4) Writing 한글 ‘hankul’



Two phonological phenomena have governed Korean orthography. One is *liaison* of the final consonants, called ‘pachim’ in Korean, with the first vowel nucleus of the following syllable, as in /pam.i/ > [pa.mi] ‘night-NOM’ and /san.ak/ > [sa.nak] ‘mountain + peak’ = ‘mountains’. The other is unreleasing of syllable-final consonants.

The syllable-final consonant clusters are split into two different syllables, when followed by a vowel-initial syllable, by *liaison*, which causes *re-syllabification*, as shown below:

(5) Korean liaison rule

$$V (C) C.(G)V \rightarrow V (C).C(G)V$$

The following examples show how the final consonant is carried over to the next syllable:

(6) Morphophonemic Writing Phonemic Writing (re-syllabified/pronounced)

몸이 /mom.i/	as 몸미 [mo.mi]	‘body-NOM’
핥아 /halth.a/	as 핥타 [hal.tʰa]	‘licks’
없이 /eps.i/	as 업시 [ɤp.ji]	‘without’ (lit. ‘lack-ADV’)

Key principles governing Modern Korean orthography are as follows:

1. A syllable must at least have a CV structure. When there is no initial consonant, use a null/zero sign represented by a small circle ○ ‘Ø’. The reason for the use of this symbol is both aesthetic and linguistic. Each block needs to be of more or less the same size. Since the position of the vowel and that of a vowel within a syllable are interrelated, at least one consonant and one vowel are necessary to write a syllable. Most crucially, this writing is based on a theory that an optimal syllable is a CV one.
2. Hankul is written morphophonemically. In other words, each morpheme is written consistently regardless of how it is actually pronounced, depending on the other grammatical units it occurs with. Thus, the final consonant of the morpheme, 꽃 <kkoch> ‘flower’, is written ㅈ <ch>, although it is pronounced in many different ways, depending on the sound that follows it, as shown below:

(7)	꽃 /kkoch/	[kkot <sup>ˀ</sup> ]	‘flower’
	꽃이 /kkoch-i/	[kkoc <sup>h</sup> i]	‘-NOM’
	꽃하고 /kkoch-hako/	[kkot <sup>h</sup> hago]	‘-INST’
	꽃잎 /kkoch+ip/	[kkoch-n-ip] > [kkot <sup>ˀ</sup> -n-ip] > [kkonnip]	‘flower petal’
	꽃내 /kkoch+nay/	[kkot <sup>ˀ</sup> -ne] > [kkonne]	‘flower fragrance’

Hankul began to be written morphophonemically only around the turn of the twentieth century led by Cwu Si-kyeng (1876–1914). Although King Sejong himself wrote morphophonemically in his own writings (K.-M. Lee 1997: 20–3), Koreans up to then wrote

phonemically showing the forms where the liaison applied or the syllable coda with a single consonant which is the result of cluster simplification and/or unreleasing.

3. Certain surface forms Korean speakers are conscious of, such as a result of vowel harmony and the Law of Initials, are written as pronounced and not the underlying forms, e.g. 좁아요 <cop-ayo> for /cop-eyo/ 'is narrow' and 노동 <no-tong> for /lo-tong/ 'labor'. In the case of *t*-insertion in compounding, Koreans choose to represent it as <s>, applying *s*-unreleasing reversely, as can be seen in such examples as 파닷가 <patas-ka> for /pata-ka/ 'seashore'.

## 5.4 MORPHOLOGY

Nouns and verbs are primary lexical categories. Other parts of speech include adverbs that modify the predicates of the sentence and determiners that express the references of nominals (nouns and noun-like words). There is a limited set of bound nouns (<sup>B</sup>), which must be preceded by a modifier, e.g. <sup>B</sup>*cwul* 'knowledge, ability', <sup>B</sup>*i* 'person', <sup>B</sup>*ke(s)* 'thing', <sup>B</sup>*li* 'reason, rationale', <sup>B</sup>*pun* 'person (HON)', <sup>B</sup>*swu* 'means, possibility', <sup>B</sup>*tey* 'place, ground', and classifiers.

### 5.4.1 Nouns and particles

Nouns occupy the biggest portion of Korean vocabulary, because the noun category is the most open and expanding. Most loanwords, regardless of their parts of speech in their original languages, are imported as nouns, just as most of the neologisms, which keep being created on the basis of mainly Sino-Korean roots but now other roots as well, are nouns. Most verbs are nominalized through a very productive process.

Pronoun-like expressions are not limited in number in Korean as those in European languages. In the case of third-person, they are actually nouns or noun phrases consisting of the demonstratives, *i* 'this', *ku* 'that', and *ce* 'that over there', followed by nouns denoting a regular noun such as 'thing', 'person', or other animate objects with their appropriate honorific variants.

Nouns and pronouns do not display allomorphy, except for a limited number of pronouns, which show different forms when combined with the subject marker *-ka*, e.g. /na-ka/ > /nay-ka/ 'I', /ne-ka/ > /ney-ka/ 'you', /ce-ka/ > /cey-ka/ 's/he, I (HUM)', /nwukwu-ka/ > /nwu-ka/ 'who'.<sup>6</sup>

Nouns can have modifiers, including determiners, relative clauses, and genitives, which must precede them. However, Korean has no articles such as *the* or *a* in English.

Korean has a plural marker *-tul*, but plural marking is not required, unless some extra meaning calls for it. For example, in

- (8) *haksayng-(tul)-i manh-ayo*  
 student-(PL)-NOM abound-DECL.POL  
 'There are many students.'

*-tul* is not required for expressing more than one item of the noun but the explicit (or 'intrinsic') plural marking carries an extra sense of there being individual members in the group. Even in such cases, plural marking is more common with human nouns than nonhuman or inanimate ones (S. Song 1975: 542–3; Lukoff 1982: 361). Plural marking is required in the case of a demonstrative pronoun or noun phrase (NP), because of the definite reference. In (9) the demonstratives make it clear that a particular person

caught something. If more specific people than one person did it, then *-tul* is required, as *ku( ai)-tul-i*.

- (9) a. *ku ai-ka cap-ass-eyo*  
 that child-NOM catch-PAST-DECL.POL  
 ‘That child caught it.’
- b. *ku-ka cap-ass-eyo*  
 PRON-NOM catch-PAST-DECL.POL  
 ‘S/he caught it.’

Korean exhibits another, unique kind of plural marking, called ‘Plural Copying’ or ‘Extrinsic Plural Marking’, usually limited to informal and intimate situations: *-tul* can be affixed to such unlikely categories as verbs, adverbs, complementizers, and even case markers. In example (10) all three instances of *-tul* ‘copy’ the plurality of the understood subject.

- (10) *ese-tul ppalli-tul o-sey-yo-tul!*  
 without.hesitating-PL quickly-PL come-HON-IMP.POL-PL  
 ‘Please hurry, you all!’

The ‘intrinsic’ plural meaning is emphasized through its multiple appearances in other parts of the sentence, although this is most common in the case of adverbs. In this particular sentence, having just one *-tul* at the end of an adverb will express the kind welcome of the speaker, and having two may exaggerate that intent a bit. Having it only at the end of the verb/sentence will indicate impatience on the part of the speaker. It is not surprising then that this so-called ‘extrinsic’ plural marking is often pragmatically conditioned (J. Song 1997).

#### 5.4.1.1 Noun particles

Noun phrases are typically followed by particles, which mark their function in a sentence or delimit their meaning.

Only a very limited set of noun particles show honorific variants, e.g. *-kkeyse* (HON) vs. *-i/-ka* subject and *-kkey* (HON) vs. *-ey/-eykey* dative. The majority of Korean particles display allomorphy, depending on whether they follow a vowel or a consonant (Table 5.7). In general the allomorphic variants are a result of an effort to avoid hiatus or two vowels occurring consecutively in adjacent syllables (Y. Lee 2008). The most common rule at work here is that the particle initial *i-* or *u-* is deleted when following a vowel, e.g. */yenghwa.ina/ > [jʌŋhwana]* ‘movie or something’ and */pata.ulo/ > [padaro]* ‘to the sea’. Another rule at work is that the weak vowel *u-* is squeezed out when surrounded by two identical sonorants in the casual-speech topic and object markers, as in */na-nun/ > [nanun] ~ [nan]* ‘me-TOP’ and */na-lul/ > [narul] ~ [nal]* ‘me-ACC’. This change is obligatory in particles beginning with *-ulo*, e.g. */kil-ulo/ > [killo]* ‘to the road’.

Most particles belong to one of two types: case marking postpositions, and delimiters providing special semantic information. A limited number of particles are pragmatic or conjunctive.

Case particles can be broadly divided into two different types: (1) structural ones, which can be dropped when the context makes them clear and retrievable, e.g. *-i/-ka* subject; *-(l)ul* object; *-uy* genitive ‘of, ’s’; *-(y)a* (CNDES)/-i/Ø (PAN)/-(i)ye (POL)/-(i)sie (SUPERDEF) vocative; (2) meaning-carrying ones that cannot be

TABLE 5.7 ALLOMORPHY IN KOREAN NOUN PARTICLES

	Consonant_	Vowel_
NOM	-i	-ka
ACC	-ul	-lul; -l
TOP	-un	-nun; -n
COM	-kwa	-wa
VOC.CNDES	-a	-ya
VOC.PAN	-i	Ø
VOC.POL	-ie	-ye
VOC.SUPERDEF	-isie	-sie
'said/called'	-ilako	-lako
COM(Colloquial)	-ilang	-lang
concession	-ilato	-lato
enumerative	-imye	-mye
'or (the like)'	-ina	-na
'if only'	-inama	-nama
'of course'	-iya	-ya
'to'	-ulo*	-lo
'as, being'	-ulo(se)*	-lo(se)
INST	-ulo(sse)*	-lo(sse)

\*-lo(se/sse) if C = l.

easily dropped, e.g. -ey (inanimate)/-eykey (animate)/-hanthey (animate)/-kkey (animate HON) indirect object, agent of passive, static locative, direction 'to, by'; -tele/-poko '(pointing) to'; -eyse dynamic locative; -eyse (inanimate)/-eykey(se) (animate)/-hanthey(se) (animate)/-kkey(se) (animate HON) 'from'; -puthe 'from, starting at/with'; -kkaci 'until, as far as'; -(u)lo 'toward'; -(u)lo(se) 'as, in the role of'; -(u)lo(sse) instrumental 'with'; -(k)wa (written)/(i)lang (informal)/-hako comitative 'with, and'; -(i)mye/-hamye enumerative 'and'; -ey/-ta/-eytaka 'in addition to, on top of'; -pota comparative 'than'; equative -chelem 'like', -mankhum 'as much as'; -kathi 'in the same way'.

Delimiter particles can be attached to other sentence units as well as nouns, e.g. -(n)un contrastive topic 'as for'; -cengto/-ccum 'about, around, approximately'; -cocha/-kkaci/-mace 'as far as, up to, even, to the point of'; -kwulye 'I reckon'; -to emphatic addition to the list 'also, too, surely'; -man/-ppun/-pakkey + NEG 'only'; -(i)lato 'just, in the absence of anything better'; -mata 'every'; -(i)nama 'if only, in the absence of good things'; -(i)ya 'if it be, then for sure, as expected'; -(i)yamallo 'for sure, certainly'; -khenyeng 'let alone, far from, on the contrary'; -puthe 'beginning with, led by'; -tul plurality; -yo POL (sentence-ender).

Pragmatic particles connote discourse and extralinguistic factors and can be attached to other grammatical units than noun phrases, e.g. -kkaci 'even (unexpectedly)'; -man '(unexpectedly) none other than'; -(n)un, topic 'as for'; -to 'surprisingly'; -yo (POL).

Conjunctive particles connect nouns with other nouns or clauses, e.g. -e 'upon'; -e(ta[ka]) 'on top of'; -hako 'and'; -(i)na 'or (something)'; -(i)nka 'or (not sure)'; -(k)wa 'and'; -(i)tunci 'or (something)'; -ttaymun-ey 'because of'; -tung 'and the like'.

A limited set of particles can belong to more than one category, e.g. -puthe 'from, beginning with'; -kkaci '(up) to, even'; and -man 'only, none other than'. However, they are not homonyms but semantically closely related to each other.

Nouns can have more than one particle:

- (11) *sewul-eyse-puthe-chelem-mankhum-man-ina-tul hay-la* (Lim 1996)  
 Seoul-in-from-like-as much-only-or something-PL do-IMP  
 ‘(You all) do at least as much, if that much, as you have done since you started in Seoul.’

The subject marker *-i/-ka* and the object marker *-(l)ul* cannot co-occur with the topic marker, *-(n)un* or the emphatic *-to* ‘also, too, surely’, so there are no such forms as *\*(l)ul-un* or *\*-ka-to*.

Particles generally follow a certain ‘fixed order’ (Cho and Sells 1995; Yoon 1995; Chang 1996). However, variations in ordering nominal particles exist, best accounted for by non-morphological (linguistic and pragmatic) explanations (Kuno and Kim-Renaud 1987/2004; K. Lee 1993; Yoon 2005). For example, the different positioning of *-man* ‘only’ in (12) (from Kuno and Kim-Renaud (1987/2004: 19)) receives different interpretations due to their disparate syntactic structures, as shown by the bracketing.

- (12) a. [*sokum-man-ulo kimchi-lul tamku*]-*l-swu iss-ta*  
 salt-only-with kimchi-ACC make-FUT.ADN-possibility exist-DECL  
 ‘One can make kimchi with salt alone (even if other ingredients are not available).’  
 b. *sokum-ulo-man* [*kimchi-lul tamku*]-*l-swu iss-ta*  
 ‘One can make kimchi only with salt (and with nothing else).’

*Two types of topic marker (n)un: non-contrastive and contrastive*

Korean clearly marks what a sentence is about with the topic marker *-(n)un*, so it is called a ‘topic language’. In a Korean sentence, the main subject is the default topic. If some other unit than the main subject is topicalized, it precedes the subject. If the topic marked noun is not moved to the front, then it usually has the contrastive meaning. For example in (13) the second topic marker is contrastive:

- (13) *swua-nun pulkoki-nun mek-nun-ta*  
 Sua-TOP pulgogi-TOP(Contrast) eat-PRES-DECL  
 ‘Sua eats pulgogi.’ (lit. ‘As for Sua, she eats pulgogi but no other food like it.’)

Having multiple topics is impossible in reality, because one cannot talk ‘about’ more than one thing at a time. However, multiple topics in a given sentence are allowed if the second and succeeding topics have contrastive meaning as in (13). Since a relative clause is about the noun it modifies, the only topic marker it can have in it is one with a contrastive meaning, for example, in (14), ‘Sua’ gets only a contrastive meaning, because the embedded sentence is ‘about the secret’ and cannot have another non-contrastive topic marker.

- (14) *kukes-un* [*swua-nun al-nun*] *pimil-i-ta*  
 that-TOP Sua-TOP(Contrast) know-PRES.ADN secret-COP-DECL  
 ‘It is a secret Sua (and no one else) knows.’

A contrastive topic particle usually receives a high pitch (C. Lee 1996). For this reason, Choe (1995) calls this kind ‘contrastive focus’.

A non-contrastive topic marker refers to old information, shared by both the speaker and the hearer. For this reason, a nominal with a non-contrastive topic marker such as *swua-nun* in (15b) is often deleted:

- (15) a. *swua-ka eti ka-ss-eyo?*  
 Sua-NOM where go-PAST-DECL.POL  
 ‘Where is Sua?’ (lit. ‘Where did Sua go?’)
- b. (*swua-nun pang-ey iss-eyo*)  
 Sua-TOP room-in exist-DECL.POL  
 ‘She is in her room.’ (lit. ‘Sua is in her room.’)

#### 5.4.2 Action verbs and stative verbs

Korean verbs are of four types: *action*, *stative*, *copulative*, and *existential*. (Identifying) copula *i-* ‘be’ or *ani-* ‘not be’ and existential verbs *iss-* ‘exist’ or *eps-* ‘not exist’ form special subsets of stative verbs.

Action verbs are dynamic verbs meaning ‘do X’ or ‘X happens’, e.g. *ssu-* ‘write’ and *o-* ‘come’. They are also called ‘processive’ and include verbs denoting mental processes such as ‘think’ and ‘love’. Stative verbs, e.g. *chakha-* ‘be good’ and *kop-* ‘be pure’, simply describe situations or conditions, and some linguists call them ‘adjectives’. It is helpful to view this group of words as verbs rather than adjectives, as stative verbs have inflections just as other verbs. Many grammatical constructions often depend on whether a verb is active or stative. For example, only an action verb can have an imperative ending, progressive *-ko iss-* ‘be . . . ing’, or the present tense marker *-nun*. However, such constraints are a natural consequence of their inherent meaning.

All verbs are bound forms in that they cannot stand by themselves. A verb consists of the stem plus one or more inflectional suffixes ‘glued’ one after the other. Inflectional endings include ① *voice* (passive or causative), ② *subject honorification*, ③ *tense* (past, present, or future) and *aspect* (complete, experienced, repeated, or continuing), ④–⑤ *modality*, and ⑥ clause-final *conjunctives* or *sentence enders* with the appropriate speech style toward the listener and encoding speech acts such as interrogative, declarative, imperative, and propositive. There are more than 40 basic verb suffixes (I. Lee and Ramsey 2000: 173), but the number can be ‘well over 400’ if we count the possible inter-combinations of these suffixes (Martin 1992: 244). Verb suffixes are used only when needed except for the sentence ender, which is obligatory. Some affixes such as honorific *-(u)si*, retrospective *-te*, and volitional future *-keyss* are called ‘pre-final’, and sometimes erroneously called ‘infixes’, because they themselves must be followed by another ending.

(16) is an example of a verb with suffixes in all six slots:

- (16) *ppop* ① *-hi* ② *-si* ③ *-ess* ④ *-keyss* ⑤ *-te* ⑥ *-kwun*  
 elect PASS HON PAST SUP RET DECL.PAN  
 ‘(From what I observed, unexpectedly) it was likely that (the honorable) he or she would have been elected.’

‘Stem-expanding’ suffixes, such as passive, causative, and subject honorification, are put close to the stem, and those indicating modality and addressee honorification, which by definition consider the broader scope of the speaker’s attitude in discourse, are found at the end. Others that are time-related, such as tense and aspect, are placed in between.

Verb suffixes display a similar kind of allomorphy as noun particles, where most variants involve deletion of the suffix-initial vowel, in this case mostly *u* [u], for the optimal CV sequence (Table 5.8).

TABLE 5.8 ALLOMORPHY IN KOREAN VERB SUFFIXES

	Consonant_	Vowel_
DECL.DEF	-supnita	-pnita
EXH.DEF	-upsita	-psita
Q.DEF	-supnikka	-pnikka
DECL/Q.EQ	-so; -uo	-o
PROM.CNDES	-uma	-ma
FUT.ADN	-ul	-l
PROM: 'I'll ... straight away'	-ulkkey(yo)	-lkkey(yo)
Premontive 'in case ...'	-ulla	-lla
Intention	-ullay(yo)	-llay(yo)
Inquisitive 'I wonder ...'	-ulkka	-lkka
Goal 'in order to ...'	-ulyeko; -ule	-lyeko; -le
Permissive 'You may ...'	-ulyem(una)	-lyem(una)
PROM.SUPERDEF	-uolita	-olita
'if ...'	-umyen(un)	-myen(un)
'while ...'	-umyense	-myense
NMR	-um	-m
(PAST.)ADN	-un	-n
'but ...'	-una	-na
'since, as ...'	-unikka	-nikka
HON	-usi	-si

The verbalizing suffix *-ha* from *ha-* 'do' is attached to a noun to convert it into a verb when the noun in question is an action noun, and the resulting verb is an action verb.

- (17) a. *swua-ka kongpu-lul ha-n-ta*  
 Sua-NOM study-ACC do-PRES-DECL  
 'Sua does the studying.'
- b. *swua-ka swuhak-ul [[kongpu]<sub>NP</sub>-ha]<sub>V</sub>-n-ta*  
 Sua-NOM math-ACC study-do-PRES-DECL  
 'Sua studies math.'

When a descriptive noun(-like form) is verbalized, the result is a stative verb, e.g. *[[cengcik]<sub>N</sub>-ha]<sub>V</sub>* 'be honest' < *cengcik* 'honesty'. The suffix *-ha* may also be attached to certain adverbs, especially the sound-symbolic type, e.g. *[[swuktek swuktek]-ha]<sub>V</sub>* 'tittle-tattle' < *swuktek swuktek* 'whisperingly, gossiping'. Loanwords, always analyzed as nouns regardless of their part of speech in their original language, are verbalized by *-ha*, too, e.g. *[[eylleykanthu]<sub>N</sub>-ha]<sub>V</sub>* 'be elegant' < *elegant*.

More than one verb can be combined into a longer verb by putting a connective vowel *-e* between them. These are usually put in chronological order, the final item becoming the head.

- (18) *[[[[ttut-e]<sub>V</sub> kochy-e]<sub>V</sub> ssu-e]<sub>V</sub> ponay]<sub>V</sub>-*  
 tear (apart)-e change-e write-e send  
 'totally rewrite and send'

5.4.2.1 Irregular verbs

Some groups of verbs are called 'irregular' because they do not follow the general phonological rules when inflected. However, these are not random exceptions to the grammar, but often have different underlying representations from 'regular' verbs.

*p*-, *t*-, and *s*-irregular verbs are among the most common types. A group of verbs ending in *p*, *t*, or *s* do not voice their final consonants in voiced surroundings like other regular verbs ending with the same consonants, but undergo further weakening when followed by a vowel-initial suffix. It is because historically most of these verbs had vowel length, and the extra weakening comes from the increased sonorance in the environment. For example, in the case of *p*-irregular verbs, the final *-p* is pronounced [w] (and vocalized as [u] when it cannot form a nucleus with the following vowel), rather than [b] as in regular verbs (19). The final *-t* in *t*-irregular verbs becomes [l] and then [r] unlike *t*-regular verbs, where the final /t/ becomes a [d] (20). The final *-s* in *s*-irregular verbs is deleted instead of staying unchanged as in *s*-regular verbs. The deleted *s* leaves its trace by lengthening the preceding vowel (21). The final *-lu* in *lu*-irregular verbs loses its weak vowel *u* as it meets another vowel and the *l* is doubled (22).

- |   |   |
|---|---|
| (19) <b><i>p</i>-regular verbs</b><br>/kwup-e/ → [kubɻ] ‘X bends’                         | <b><i>p</i>-irregular verbs</b><br>/kwup-e/ → [kuwɻ] ‘X bakes’<br>/kwup-una/ → [kuuna] ‘X bakes, but’ |
| (20) <b><i>t</i>-regular verbs</b><br>/mit-uni/ → [miduuni]<br>‘as X believes’            | <b><i>t</i>-irregular verbs</b><br>/mut-uni/ → [mul-umi] → [murumi]<br>‘as X asks’                    |
| (21) <b><i>s</i>-regular verbs</b><br>/sos-ato/ → [sosado]<br>‘even if one rises (above)’ | <b><i>s</i>-irregular verbs</b><br>/cis-eto/ → [ci:ɻdo]<br>‘even if one builds’                       |
| (22) <b><i>lu</i>-regular verbs</b><br>/chilu-e/ → [c <sup>h</sup> irɻ] ‘Pay back!’       | <b><i>lu</i>-irregular verbs</b><br>/pulu-e/ → [pullɻ] ‘Call!’  |

A verb-stem-final *l* is deleted in front of a suffix beginning with *u*-, *o*-, *n*-, and *s*-, e.g. /al-sey-yo/ → [a:sejo] ‘does X (HON) know?’, /nol-na/ → [no:na] ‘(I wonder) does X play?’. All *l*-final verbs are called ‘irregular’ because /l/ does not get deleted in these idiosyncratic environments elsewhere in the grammar, cf. /silswu/ → [silsu] ‘mistake’ and /selnal/ → [sɻllal] ‘New Year’s Day’.

*ha*-, an action verb meaning ‘do, be’, which is also definitely the most common verbalizing suffix in Korean, has a slightly irregular conjugation: the effort to avoid a hiatus situation when followed by an *e*-initial suffix is realized in two different strategies: (i) merge two vowels into one as in /ha-e/ → [hɛ] ‘Do!’; (ii) insert a semivowel *y* between the two vowels, as in /ha-e/ → [hajɻ] ‘after doing’.<sup>7</sup> The two coexisting forms are not completely interchangeable: (i) is much more commonly used, especially in spoken Korean today; (ii) sounds a bit more formal and old-fashioned.

A group of verb stems, which derive from old compounds with *ha*- – including all basic color words such as *ppalkah*- ‘red’, *phalah*- ‘blue’, *nolah*- ‘yellow’, and *kkamah*- ‘black’, and some other descriptive verbs such as *etteh*- ‘how is it?’ and *kuleh*- ‘it is so’ – undergo a similar change when they meet an *e*-initial affix. Although these verbs’ stems manifest only the *h* part of their historical origin *ha*- ‘do’, the vowel’s trace is shown in conjugation: when *e* follows this *h*, it changes to *ay* and *h* is deleted in the process, e.g. /nolah-e/ → [noreɻ] ‘X is yellow’.

The verb *toy*- ‘become’, which often functions as an honorific replacement for copula *i*- ‘be’ and is also used in passivization, shows similar conjugation to *ha*-. Thus /toy-e/ → [twɛ] ‘it’s OK’ (lit. ‘it’s becoming’).

### 5.4.3 Numerals and classifiers

Numerals can be used as nouns, but are more commonly used as pronominals modifying a noun or a noun classifier, e.g. [*selhun*]<sub>NP</sub> ‘30’ or [[*selhun*]<sub>NP=Modifier</sub> [*sal*]<sub>NP</sub>]<sub>NP</sub> ‘30 years old’.

#### 5.4.3.1 Cardinal numerals

Koreans use two parallel sets of numbers up to 99: one of native origin (*hana* ‘1’, *twul* ‘2’, *seys* ‘3’, *neys* ‘4’, *tases* ‘5’, *yeses* ‘6’, *ilkop* ‘7’, *yetelp* ‘8’, *ahop* ‘9’, and *yel* ‘ten’, *yelhana* ‘11’, *sumul* ‘20’, *selhun* ‘30’, *mahun* ‘40’, *swin* ‘50’, *yeswun* ‘60’, *ilhun* ‘70’, *yetun* ‘80’, *ahun* ‘90’) and the other Sino-Korean (SK) (*il* ‘1’, *i* ‘2’, *sam* ‘3’, *sa* ‘4’, *o* ‘5’, (*lyuk* ‘6’, *chil* ‘7’, *phal* ‘8’, *kwu* ‘9’, *sip* ‘10’, *sipil* ‘11’, *sipi* ‘12’, *isip* ‘20’, *samsip* ‘30’, *kwusip* ‘90’). Beyond 100, only Sino-Korean numbers are used (*payk* ‘100’, *chen* ‘1,000’, *man* ‘10,000’, *ek* ‘100,000,000’, *co* ‘1,000,000,000,000’). Some native numbers are shortened when used as modifiers, that is, *han* ‘one . . .’, *twu* ‘two . . .’, *sey/sek* ‘three . . .’, *ney/nek* ‘four . . .’, and *sumu* ‘twenty . . .’ instead of the forms shown above.

SK numbers are used with certain classifiers, especially when a series goes up to a larger number. Mathematical formulae and other precision counting, such as bank account numbers, account balances, loanword measures, telephone numbers, room numbers, tickets, commemorative days, and ID numbers, use SK numbers, e.g. *sampun-uy il* ‘one third (lit. ‘one of three parts’), *il-mithe-chilsipo-senchi* ‘1m 75cm’, *chil-yuk-i-sam-kong-kwu-sam* ‘762-3093’, *o-ho-sil* ‘Rm 5’, *kwu-il-il* ‘9/11’ (lit. ‘nine one one’).

#### 5.4.3.2 Ordinal numerals

For native ordinal numbers, the suffix *-penccay* is added to the pronominal forms of cardinals, e.g. *twupenccay* ‘the second’, *neypenccay* ‘the fourth’. For ‘the first’, *chespenccay*, where *ches* means ‘the beginning’, is used rather than *\*hanpenccay*.

SK ordinals use *cey* in front and are followed by a classifier such as *-hoy*, *-cha*, *-kup*, *-tung*, *-chak*, *-kwa* (‘Xth number in a series of . . .’, ‘Xth anniversary’, ‘Xth grade’, ‘Xth ranking’, ‘Xth arrived’, and ‘Xth lesson’, respectively), e.g. *cey chil kwa* ‘Lesson Seven’. However, in most cases, *cey* is commonly dropped, although usually kept in formal speech.

There are several number-related pronominals, e.g. *cen* ‘entire’, *chong* ‘total’, *motun* ‘all’, *mus* ‘many, various, all kinds’, *myech* ‘a few, several, how many’, *pan* ‘half’, *swu-* ‘several’ (e.g. *swumanhun* ‘myriad’), *yele* ‘various’, and *yang* ‘both, the two . . .’.

#### 5.4.3.3 Noun classifiers

Many classifiers are bound forms attached to the end of numerals. (23) lists four possible structures, the first being the most common:

(23) Measuring nouns (‘five teachers’):

- |    |                    |                  |                    |
|----|--------------------|------------------|--------------------|
| a. | N                  | Nr-CL            |                    |
|    | <i>sensayngnim</i> | <i>tases-pun</i> |                    |
|    | teacher            | five-person.HON  |                    |
| b. | Nr-CL              | -GEN             | N                  |
|    | <i>tases-pun</i>   | <i>-uy</i>       | <i>sensayngnim</i> |

- c. Nr N  
*tases sensayngnim*
- d. N Nr  
*sensayngnim tases*

Noun and classifier must agree. There is a commonly used generic classifier for inanimates, *kay*. (24) lists some important classifiers:

(24) <u>Semantic categories of nouns</u>	<u>Classifiers</u>
human beings	<i>-in, -myeng, -nom</i> 'male' [Derogatory], <i>-nyen</i> 'female' [Derogatory], <i>-pun</i> [HON], <i>-salam</i> 'person'
animals	<i>-mali</i>
boats	<i>-chek</i>
bottles	<i>-pyeng</i>
bowls	<i>-kulus, -kongki</i>
buildings	<i>-chay</i>
bundles	<i>-tan</i>
characters/letters (of alphabet)	<i>-ca</i>
cigarettes	<i>-tay</i>
clusters, bunches, flowers	<i>-songi</i>
crowds	<i>-ttey</i>
days of the month	<i>-il</i>
eggs/beads/beans/grains	<i>-al</i>
games	<i>-phan</i>
glasses/cups	<i>-can</i>
grains/nuts	<i>-thol</i>
gang	<i>-phay</i>
hand-/footwear	<i>-khyelley</i>
inanimate objects	<i>-kay</i>
meals	<i>-kki</i>
mechanical objects/vehicles	<i>-tay</i>
money	<i>-phun</i>
month names (January–December)	<i>-wel</i>
months (counting)	<i>-kaywel</i>
packs	<i>-kap</i>
packages	<i>-kkwulemi</i>
sheets of paper/letters/documents	<i>-cang</i>
paintings	<i>-phok</i>
pieces of music	<i>-kok</i>
plant heads (e.g. cabbages)	<i>-phoki</i>
pieces of poetry	<i>-swu</i>
sets	<i>-pel</i>
sheets	<i>-cang</i>
stick-shaped (long, thin, rigid)	<i>-calwu</i>
telephones	<i>-thong</i>
trees	<i>-kulwu</i>
volumes	<i>-kwen</i>
years	<i>-nyen</i>



TABLE 5.9 KOREAN SENTENTIAL MODIFIER ENDINGS

	Action Verb	Stative Verb
Present	-nun (< -n + -un)	-(u)n
Past	-(u)n	N/A
Retrospective	-ten (< -te + -un)	-ten (< -te + -un)
Future	-(u)l	-(u)l

- (26) [[*chotunghakhyo ilhaknyen-puthe ha-nun*]<sub>S=>ADN</sub> *kes*]<sub>NP</sub>  
 primary.school 1st.grade-from do-PRES thing/that  
 ‘what you do from the first grade of the primary school’ (lit. ‘that one does from . . .’)

English ‘can’ has at least two different meanings: competence or possibility. The two meanings are represented by two different relative-clause derived forms in Korean: *-(u)l*<sup>B</sup>*cwul al-* (lit. ‘to have the knowhow/knowledge’) vs. *-(u)l*<sup>B</sup>*swu iss-* (lit. ‘it is possible for X to . . .’), as can be seen in (27):

- (27) a. *na-nun* [[*sukhiha-l*]<sub>S=>ADN</sub> [*cwul*]<sub>N</sub>]<sub>NP</sub> *al-ayo*  
 ‘I can ski.’  
 b. *yelum-ey-to* [[*sukhiha-l*]<sub>S=>ADN</sub> [*swu*]<sub>N</sub>]<sub>NP</sub> *iss-eyo*  
 ‘One can ski in the summer, too.’

### 5.5.2 Pronouns and anaphora

The Korean pronoun system is defective and its inventory is limited to a few neutral forms such as *na* ‘I’, *ne* ‘you’, and *uli* ‘we’. Korean lacks pronouns such as ‘it’, ‘he’, ‘she’, or ‘they’. In most cases, an NP consisting of a demonstrative, *i* ‘this’, *ku* ‘that’, or *ce* ‘that (over there)’, followed by a noun is used as a quasi-pronoun. The only exception may be that *ku* is used as a third-person (mainly masculine) singular pronoun in fiction writing. The pronoun *ce*, literally meaning ‘that (over there)’ is used to refer to the first-person in humble expressions. The extra self-deprecating expression, *sosaeng*, literally meaning ‘a small person’, may still be used, but perhaps only when humor is intended.

Kinship terms such as ‘elder sister’, ‘elder brother’, ‘uncle’, ‘aunt’, and lately even ‘mother’, ‘father’ for unrelated people as well as real kin are used as quasi-pronouns, including in the second-person. A person who is referred to or addressed as such senses the speaker’s intimacy and respect. Titles with or more commonly without the name also serve a comparable function.

The three demonstratives are often used to refer to statements that have been made, in addition to specific objects or ideas. When *i* ‘this’ is used, the empathy of the speaker is expressed, while *ce* ‘that (over there)’ is for a case from which the speaker is detached. The most commonly used is *ku* ‘that’ and words beginning with *ku*, because this demonstrative assumes certain information to be within the hearer’s cognition. It is for this reason that *ku* forms the basis of anaphoric verbs *kuleh-* ‘to be so’ and *kule-* ‘to do so’ and a large number of sentence-initial conjunctives derived from those verbs, e.g. *kuliko* ‘and in addition’, *kulentey* ‘but’, and *kulayse* ‘and then/so’, each of which carries an implied meaning of ‘the situation being such’.

Interrogatives and indefinite pronouns are expressed by the same words in Korean. For example, Korean ‘who’, ‘someone’, and ‘anyone’ are all expressed by the same

word, *nwukwu*, because they all share the meaning of ‘some person that is not specific or definite’. The same principle applies to other indefinite pronouns, which include *elma* ‘how much/some~any quantity’, *encey* ‘when/some time’, *eti* ‘where/somewhere’, *mues* ‘what/something~anything’. *mues* has an informal variant *mwe* and a casual variant *mo*.

Korean speakers disambiguate these two types of indefinite pronouns by the context and the accentual pattern. An interrogative gets an accent and the sentence containing it has a falling pitch at the end. In a yes–no question the indefinite pronoun gets no accent and the sentence ends in a rising pitch. The word meaning ‘any X’ in a negative sentence gets the accent and the sentence has a falling pitch.

The indefinite noun *amu* ‘anyone, any person’ can also act as a prenoun as in *amu kes* ‘anything’ or *amu ttay* ‘any time’. The *amu X* expression is almost always followed by the emphatic particle *-to* and commonly occurs in a negative sentence, as in *amu kes-to an po-ass-e* ‘I haven’t seen anything’.

### 5.5.3 The basic sentence

The basic sentence in Korean has the structure of Subject-Object-Verb. However, noun particles allow subjects and objects to appear in different orders within a sentence without changing the basic meaning of the sentence. When context makes it clear, nouns can be deleted rather than replaced by pronouns. One strict requirement is that a sentence always has a verb, and it should appear at the end. Nevertheless, the usual order of phrases within a sentence is as shown in (28):

- (28) time adverbial—place adverbial—subject noun phrase—dative noun phrase—manner adverbial—object noun phrase—verb

Some verbs require two objects, in which case the indirect object usually precedes the direct object (29):

- (29) *swua-ka halmeni-kkey si-lul sse-tuly-ess-ta*  
 Sua-NOM grandmother-DAT.HON poem-ACC write-dedicate-PAST-DECL  
 ‘Sua composed a poem for her grandmother.’

Some stative verbs require two subject-marked NPs, which are often in the part–whole relation (30):

- (30) *ai-ka meli-ka pisangha-ta*  
 child-NOM brain-NOM be.extraordinary-DECL  
 ‘The child is extraordinarily smart.’ (lit. ‘The child, her brain is extraordinary.’)

Here ‘child’ is the whole and ‘brain’ is part of that child. The first subject is commonly topicalized by *-nun*.

Affirmative copula *-i* ‘be’ must be attached to an NP, so that they form a single, verb-like unit:

- (31) *thokki-ka sensayng-i-ta*  
 rabbit-NOM teacher-COP-DECL  
 ‘The rabbit is the teacher.’

In a negative copula sentence with *ani-* (< etym. *an* + *i-*) ‘not be’, both the subject and the complement of the negative copula sentence are marked by subject markers, and this is another type of double-subject sentence, as shown in (32):

- (32) *cha-ka cwungko-ka ani-ta*  
 car-NOM used.one-NOM NEG.COP-DECL  
 ‘The car is not a used one.’

Existential verbs, affirmative *iss-*, and negative *eps-* are used in two different constructions: (i) existence/absence of an NP; (ii) possessive. The possessive construction is something like ‘to X, Y exists’.

- (33) *swuyengcang-i iss-ta*  
 pool-NOM exist-DECL  
 ‘(X = the house, I, etc.) has a pool.’

The possessor is not specified if context makes it clear, but an explicit possessor can be either a dative or a topicalized NP.

Korean is also a topic-prominent language, and any part of the sentence can be topicalized by adding the ending *-(n)un*. Topicalized elements appear first in the sentence.

There are five major sentence types, distinguished by sentence-end particles and also by intonation: declarative (*-ta*), interrogative (*-ni/-nya*), imperative (*-ela*), exhortative (*-ca*), and promissive (*-uma*). The last three can be grouped into one type called ‘jussive’ (Pak *et al.* 2008). Korean is rich in this type of particle, because each sentence type has multiple honorific variants required by speech protocol, depending on the formality, intimacy, and relative social-status hierarchy between the speaker, hearer, and the subject referent (Pak 2008).

### 5.5.3.1 Declarative

Declaratives are pronounced with a falling pitch. The unmarked, plain declarative sentence particle is *-ta*, but there are honorific variants such as *-ney* (respectfully condescending), *-e* (*panmal*), *-eyo* (informal polite), *-uo* (formal equal), *-(su)pnita* (formal deferential), *-naita* (hyper-deferential).

Many sentence-concluding endings are *panmal* (lit. ‘half speech’). *Panmal* originates from unfinished sentences, and various *panmal* forms are the result of politeness strategies that stop enunciating the rest of the ending but just imply it, in order to mitigate or avoid the expression of too hierarchy-conscious relationships. This is part of the general trend of the disappearance of certain forms of speech styles that imply inequality in power status other than nonthreatening inequalities such as seniority and kinship hierarchy.

These ‘cut-off’ speech forms have been grammaticalized from subordinate-clause-final endings to become addressee-honorific endings. However, *panmal* could be considered not respectful enough by virtue of not explicitly expressing deference, and the ending *-yo* (from the copula *i-* + the respectful *-uo* ending, together meaning ‘it is that’) is attached to the *panmal* ending to express respect. The result is a form conveying both respect and informality. This is why Sung (1970/1984) and others call *-eyo* the ‘*panmal*-elevator’.

Declarative *-ta* pronounced with stress can express the speaker’s spontaneous reaction, or inform the addressee of noteworthy information that needs immediate reaction in informal discourse, as in *cap-ass-ta!* ‘Got you!’ (H. Lee 1994: 528).

### 5.5.3.2 Tense/Aspect

Both tense and aspect are time-related, and they work closely together. Aspect specifies whether an event is ongoing, completed, experienced, reminisced, having an effect, etc.,

rather than simply locating an event or a state vis-à-vis a certain reference point in time. Aspects are most often expressed by suffixation but also by complex verbs which include inherently aspectual verbs. Some suffixes represent both tense and aspect at the same time. For example, the present often denotes the aspectual meaning of repetition or habit.

The unmarked tense is present, which expresses the current state or ongoing action, a common, repeated event, and also the past or future event in the proper context. The present tense of stative verbs has zero mark for tense and aspect, e.g. *yeyswul-un kil-Ø-ta* 'art is long'. The present tense of action verbs is marked by a processive or action marker *-n*, as in *pi-ka o-n-ta* 'It rains/is raining' (lit. 'the rain comes'). When the processive marker follows a consonant, an unfelicitous situation of a three-consonant cluster arises. There are two strategies to avoid the cluster. When the verb ends in *-l*, it is deleted with compensatory lengthening, e.g. */col-n-ta/ > co.nta* 'X dozes'. With other consonants, the strategy is to copy the *n* and insert the neutral vowel *u* to form an independent syllable *nun*, as in *sim-nun-ta* 'X plants'.

The past tense marker *-ess* denotes what occurred before the referent time (e.g. the utterance time), and also the fact that the event has been completed. The VH rule applies, as in *et-ess-ta* 'X gained' vs. *ssot-ass-ta* 'X poured'.

The past-perfect marker *-ess-ess* either indicates that an event happened before another in the past, or has an aspectual meaning of past experience that has no effect on today. It can also indicate a habitual action before a reference point in the past, e.g. *ku cen-ey-nun swuyoil-mata kyohoy-ey ka-ss-ess-ta* 'Before, I used to go to church every Wednesday'.

Strictly speaking, future marking is not really time-related, but predicting a future event or state (i) by *-keyss*, a volitional suffix expressing strong prediction or intention, or (ii) by the use of *-(u)l ke-ta*, 'it will be that . . .', which predicts what will happen without conviction. Thus, when one says *nayil tto o-keyss-supnita* 'I will come again tomorrow' it will certainly happen, because the speaker has a strong intention and conviction. That is why only the first-person subject can use *-keyss* to mark the future; for a third person, one must use *-(u)l ke-ta* as a kind of conjecture, as one cannot go into the third person's psyche but only guess it. Exceptional cases include a weather forecaster's *nayil pi-ka o-keyss-supnita* 'It will rain tomorrow', implying a strong, confident prediction, or an MC at a wedding announcing the bridal march, *cikum-puthe sinpuiipchang-i iss-keyss-supnita* 'Now the bride will march in' (lit. 'from now the bride's entrance will happen/exist'). *-keyss* can be used to express the speaker's worrisome prediction for a third person's behavior surely to happen, *nayil amhayngesa-ka o-keyss-ta* 'Tomorrow (I am pretty sure and afraid that) the secret royal inspector will arrive'. Because of the weak prediction force, the *-ul ke-ta* construction was originally used as a politeness strategy, but it is generalizing to become an all-future marker, especially among younger speakers.

The progressive is formed by attaching *-ko + iss-* to an action verb, as in *ssu-ko iss-ta* 'X is writing'. Unlike other present tense forms, the present progressive cannot mean a future or habitual action. The existential *iss-* conjugates to express past and future, e.g. *ssu-ko iss-ess-ta* 'X was writing' and *ssu-ko iss-ul ke-ta* 'X will be writing', and to mark deference as in *ssu-ko kyesi-ta* 'X (HON) is writing'. Adnominal *-nun + <sup>B</sup>cwung* 'middle' + copula also expresses a progressive, e.g. *cip-ul sa-nun cwung-i-ta* 'X is buying a house' (lit. 'X is in the middle of buying a house').

The resultative, sometimes called 'perfective', expresses the condition resulting from completion of an action. The most common ending is *-e iss-*, as in *chinkwu-ka*

*phulangsu-ey ka iss-ta* ‘My friend is in France (as a result of having gone there)’. However, a limited number of verbs, e.g. those belonging to the ‘wearing’ or ‘possessing’ class, take *-ko* instead of *-e*, thereby creating ambiguity with the present progressive. For example, *hanpok-ul ip-ko iss-ta*, can mean (i) ‘X is in the process of putting on a Korean dress’ or (ii) ‘X is dressed in a Korean outfit (as a result of having put it on)’. Verbs of knowing and being can only have resultative meaning, as in *al/molu-ko iss-ta* ‘X knows/does not know (as a result of (not) having understood/been informed)’.

*-(u)l + <sup>B</sup>cham* ‘crevice’ + copula expresses immediate future, as in *cip-ulo ka-l cham-i-ta* ‘I am about to leave for home’. Such a sentence, however, cannot be negated.

Adnominal + <sup>B</sup>*cek* ‘time’ + subject *-i + iss-* expresses experiential aspect, e.g. *X-ey ka-n cek-i iss-ta* ‘I have been to X’, *X-ey ka-nun cek-i iss-ta* ‘I sometimes go to X’. Similar meanings can be had by replacing *-cek* with the independent noun *ttay* ‘time’ or *il* ‘work, happening’. Experiential constructions can be negated, e.g. *X-ey ka-n cek-i eps-ta* ‘I have never been to X’, *X-ey ka-nun cek-i eps-ta* ‘I never go to X’.

When the speaker reports what s/he actually witnessed, retrospective *-te* is used. For example, in *ku cip mantwu cengmal mas-iss-te-la* ‘The dumpling of the house was really delicious!’, the speaker reports his/her personal experience of tasting the dumpling and how good it was. Declarative *-ta* weakens to *-la* following *-te*. *-te* is combined with the past modifier *-un* to express habitual past or some ongoing situation that has been discontinued, e.g. *nay-ka tani-ten hakkyo* ‘the school I used to attend’, *chakha-ten ai* ‘the child who used to be good’.

Some auxiliary verbs derived from regular verbs contribute aspectual meaning to the verbal complex. For example, *-e twu-* ‘do X in preparation/for keepsake’ is derived from *twu-* ‘keep, set aside’ and indicates an action done in readiness for another event. Thus *ton-ul cip-ey twu-ess-ta* means ‘I kept my money at home’ but *achim-ul mek-e twu-ess-ta* means ‘I had breakfast (to be ready for something else, to save time, etc.)’. Likewise, *pe-li-* means ‘throw away’, and the auxiliary derived from it means ‘do something as if throwing it away’, i.e. ‘do something completely and be done with it’, as in *ic-e peli-ess-ta* ‘I (completely) forgot it’. Because of the sentiment of finality, it can express feeling relieved or regret, depending on context. *-e cwu-* from *cwu-* ‘to give’ expresses an action that is done as a favor. In this case there exists an object-deferential form *tuli-* for *cwu-*, as in *halmeni-kkey say khemphyuthe-lul poi-e tuli-ess-ta* ‘I showed my new computer to my grandmother’. *-e po-* ‘try’ is derived from *po-* ‘to see’, as in *thaykwento-lul payw-e po-keyss-ta* ‘I will try learning Taekwondo’.

### 5.5.3.3 Negation

There are four types of negative expression in Korean: (a.) Short-form negation, putting before the verb the negative adverb *an* (from MK *ani*, which still exists as an archaic variant), e.g. *ca-n-ta* ‘X sleeps’ is negated as *an ca-n-ta* ‘X does not sleep’. (b.) Long-form negation, putting *-ci* + the negative auxiliary verb *anh-* (from MK *ani-hó-* ‘not do’) to the end of the verb stem, literally ‘not do X (X = a verb phrase or a sentence)’. For example, *ilponin-un wul-ci anh-nun-ta* ‘The Japanese don’t cry’, the negation of *ilponin-un wu-n-ta* ‘The Japanese cry’, can mean either ‘The Japanese don’t cry’ or ‘It is not that the Japanese cry’. (c.) Morphologically derived negative nouns by negative prefixes, the majority of which are Sino-Korean, such as *mi-* 未, *mol-* 沒, *mu-* 無, *pi-* 非, and *pul-* (*pu-* before a coronal consonant) 不, e.g. *mu-pica* ‘no visa’ vs. *pica* ‘visa’. (d.) Lexically/inherently negative verbs such as *eps-* ‘not exist’ and *molu-* ‘not know’.

The short form has narrow scope and negates just the verb, while the long form has a wide scope and negates the whole verb phrase or sentence. *an* may be replaced by another negative adverb *mos*, which expresses inability or impossibility of reaching an ideal, rather than simply negating the event or state. For example, *tayhak-ey mos ka-ss-ta* ‘X couldn’t go to college’ implies ‘X wanted to go and it would have been better if X did’. In the case of negative imperative and propositive sentences, another negative auxiliary verb, *-ci mal-* (final *l* is dropped before *n*, *s*, *o*, and *u*) is chosen, as in *onul-un kyohoy-ey ka-ci mal-ca* ‘Let’s not go to church today’. The *mal-* form is often used as a negative pro-verb when there is a choice between affirmative and negative, e.g. *sa-lkka ma-lkka* ‘Shall I buy or not?’. *-ci mal-* + connective *-ko* means ‘do not do . . . but do . . .’, but (*-ci mal-*) *ko* is also grammaticalized as a postposition meaning ‘instead of . . . ing’, which either follows a whole sentence (with *-ci*) or just a noun, as *yangpok (ip-ci) mal-ko hanpok ip-keyss-ta* ‘I will wear the Korean dress instead of (wearing) the suit’. Other grammaticalized endings include *V-ko mal-ko(-yo = POL)* ‘V, of course’ (lit. ‘no question whether an event or state is so or not’), *V-kena mal-kena* ‘(I don’t care) whether X does . . . or not’, *po-na ma-na* ‘no doubt’ (lit. ‘whether you see for yourself or not’), *ma-psosa* ‘God forbid!’ (lit. ‘May you (SUPERDEF) not do it!’).

Korean has a fair number of negative polarity items, most of which are adverbs, that must co-occur with a negative verb, e.g. *acik* ‘yet’, *celtay(lo)* ‘absolutely’, *cenhye* ‘at all’, *comchelem* ‘hardly’, *comcheylo* ‘unbudgingly’, *kyelkho* ‘by any means’, *pakkey* ‘but . . .’, *pyello* ‘particularly’, *tasi(nun)* ‘again’, *tomuci* ‘at any cost’, *thong* ‘completely’, and *yekan* ‘commonly’. Most take both types of negation, but *yekan* takes only the long form. <sup>B</sup>*li* takes only the negative existential verb, *eps-*, as in *ku haksayng-i kecismal-ul ha-l li-ka eps-ta* ‘It’s unthinkable that the student would lie’.

#### 5.5.3.4 Modality

There are broadly three types of modality in Korean: (a.) Epistemic: knowledge-based, e.g. ‘seems’, ‘appears’, ‘probably’, ‘perhaps’, hearsay. (b.) Deontic: external constraints, e.g. ‘should’, ‘must’, ‘may’, ‘need to’, and their negatives. (c.) Dynamic: internal psychology-based, e.g. ‘intend to . . .’, ‘want to . . .’. To express these, Koreans use different strategies.

(a.) A nominalized sentence may be applied to perception verbs such as *kath-* ‘looks like’ (lit. ‘equals’) to form epistemic constructions, e.g. a sentence nominalized by <sup>B</sup>*kes* ‘thing, that’ in a construction such as adnominal + <sup>B</sup>*kes-kath-* means ‘it looks as if’, e.g. *pi-ka o-l kes-kath-ta* ‘it may rain/it looks like rain’. The <sup>B</sup>*tus* + *ha-* construction preceded by an adnominal sentence is also a common structure, e.g. *nwun-i o-l tus-ha-ta* ‘It looks like snow’. The verb *po-* ‘see’ preceded by a yes–no question sentence (plain style) also forms an epistemic construction, as in *tto sa-na po-ta* ‘X seems to buy it again’.

(b.) A noun expressing ability/possibility (<sup>B</sup>*swu*), necessity (*philyo*), or reasonable expectedness (<sup>B</sup>*li*) preceded by an adnominal sentence is followed by the existential verb *iss-* or negative *eps-*. *philyo* may take nominative case, as in *cikum ka-l philyo(-ka) eps-ta* ‘X does not need to go now’. The form taken is ‘If . . . , it would be good/bad’. A conditional sentence is followed by an evaluative verb such as *coh-* ‘is good’, *kwoyunchanh-* ‘is all right’, *nas-* ‘is better’, and *toy-* ‘will do’ (lit. will become) and their negative counterparts, e.g. *swul masi-ko wuncenha-myen an-toy-n-ta* ‘X must not drive after drinking’.

(c.) In order to express a strong will or a definite intention, the ‘strong-prediction future’ suffix *-keyss* is attached to the verb stem, as in *nayil tto o-keyss-ta* ‘I will come again tomorrow’. To express one’s desire or want, an auxiliary verb such as *-ko siph-* ‘want to’ is used, as in *pulkoki-ka mek-ko siph-ta* ‘I want to eat pulgogi’ (lit. ‘pulgogi eating is desirable/longed for’). These expressions are limited to the first person, or to the second person in a question sentence, as Korean speakers view another’s inner psychological state basically inaccessible to them, unless the speaker is the creator of the third person such as would be the case of an author of a fiction or a master of ceremonies in a formal occasion.

There are co-occurrence relationships between items involving modal expressions. For example, <sup>B</sup>*cwul* denotes ability or knowledge of some fact and occurs only with cognitive verbs such as ‘know’/‘not know’, e.g. *kule-l cwul al-ass-ta* (\**sayngkakhayss-ta*) ‘I knew (\*thought) that would happen’. Multiply ambiguous <sup>B</sup>*swu* ‘ability, the possibility, means/way’ occurs only with an existential *-iss*. For example, *pule-lul ha-l swu iss-ta* ‘X can speak French’ is triply ambiguous and depending on context can mean X’s ability to speak French, the circumstance that allows X to speak French, or the way to learn French. The <sup>B</sup>*swu* meaning ‘means/way’ can act like a real noun and be preceded by its own modifier such as *ssunami-lul pangciha-l (coh-un) swu-ka eps-ulka?* ‘Isn’t there a (good) way (\*ability, \*possibility) to prevent tsunamis?’.

Modality in Korean is an important device for encoding the speaker’s affect and is extensively used in various politeness strategies in expressing wishes/hopes/intention, evaluation of the event/state, self-deprecation, nonthreatening attitude, etc. For example, in *na-nun ppalkansayk-i te coh-un kes kath-ta* (lit. ‘to me the red looks better’), the assertive force of the real meaning, ‘I think the red is better’, is mitigated by the use of the modal expression pretending it is just an impression rather than a strong opinion.

### 5.5.3.5 Non-declarative sentence types

As stated in 5.5.3, there are at least four different clause/sentence types in Korean other than declaratives, represented by sentence-final particles. However, sentence-final particles have different honorific forms chosen according to Korean language protocol (see 5.5.6). The sentence types and their particles in various speech styles are given in Table 5.10.

As shown earlier in Table 5.6, different sentence types are distinguished also by different intonation contours. Questions to oneself that do not expect a response from a listener or rhetorical questions use *-(n)unka(yo)* or *-na(yo)*. There is no change in word order, and no fronting in question-word questions. Yes–no questions are answered with *yey/ney* (DEF), *ung* (PLN) or *kulem(-yo)*, ‘yes/that’s right’ or *ani(-yo)* ‘no’ + repeated verbal of the original. Affirmative-leading questions use a negative verb, often with a tag-like ending *-ci*, as in *phyenci an ss-ess-ci?* ‘You didn’t write the letter, right?’ to which the response is most likely to be *ani-yo, ss-ess-eyo* ‘No (what you say is not true), I wrote (it)’. Non-leading questions are posed as alternative questions, the first with the verb in the affirmative, the second with it in the negative, as in *canton iss-ni? eps-ni?* ‘Do you have change or not?’. Responses usually give only the unrecoverable items such as *eps-ta* ‘[I] don’t have’ to the above choice question.

Command is imposing on others, and is often expressed indirectly rather than by jussives, which are considered blunt, unless the hearer is a beneficiary as in *hayngpok hay-la* ‘Be happy!’ and *annyenghi ka-sey-yo!* ‘Bye!’ (lit. ‘Go in peace!’). Other structures

**TABLE 5.10** CLAUSE/SENTENCE TYPING PARTICLES IN KOREAN

Clause types	Sentence-final particles by style (bold = HON)						
	PLN	CNDES	PAN	POL	EQ	DEF	SUPERDEF
Declaratives	<i>-ta</i>	<i>-ney</i>	<i>-e</i>	<i>-eyo</i>	<i>-uo/-so</i>	<i>-(su)pnita</i>	<i>-uopnita/ -saopnita, -naita/ -saopnaita</i>
Interrogatives	<i>-ni</i> <i>-nya</i>	<i>-na</i>	<i>-e</i>	<i>-eyo</i>	<i>-(n)(u)nka(yo), -((u)si)na(yo)</i>	<i>-(su)pnikka</i>	<i>-((u)si)naikka</i>
Imperatives	<i>-ela,</i> <i>-ulyem(tuna)</i>	<i>-key</i>	<i>-e</i>	<i>-eyo,</i> <i>-(u)seyyo/ -(u)siyo</i>	<i>-(u)sikeyna</i>	<i>-((u)si)psio</i>	<i>-((u)si)psose</i>
Exhortatives	<i>-ca</i>	<i>-sey(na~yo)</i>	<i>-e</i>	<i>-eyo</i>	<i>-sey</i>	<i>-((u)si)psita, -((u)si)ciyo</i>	<i>-saita</i>
Promissives	<i>-(u)ma</i>	<i>-keyssney</i>	<i>-e,</i> <i>-(u)lkkey</i>	<i>-(u)lkkeyyo</i>	<i>-umsey, -lita, -keyssso</i>	<i>-keysssupnita</i>	<i>-(u)olita</i>

include interrogatives, e.g. *icey ka-si-keyss-supnikka?* ‘Would you like to go now?’, or declaratives expressing wishes, e.g. *ney-ka ka-ss-umyen coh-keyss-ta* ‘It would be nice if you could go’.

Certain endings, especially some panmal forms, connote extra pragmatic meaning from the speaker’s point of view. What K. Lee (1993: 7–31) calls ‘attitude markers’ include *-ci* [expecting agreement by the listener, like a tag question], *-e* [seeking uninterested hearer’s attention], *-kwun(a/yo)* [discovery], *-ney(-yo)* [surprise (contrary to expectations)], and *-ta* [conveying information to the objective hearer].

**5.5.4 Topic, focus, and emphasis**

Korean is a ‘topic language’, which clearly marks what a sentence is about with the topic marker (TOP) *-(n)un*. *-(i)lan*, is a kind of TOP, having a more specific intention of defining something, e.g. *phansoli-lan* ‘(I will explain what) *pansori* means/is . . .’. TOP can be attached to any unbound item, but mostly NPs and adverbials.

TOP deemphasizes old/commonly known information, and for that reason the TOP-marked NP is often deleted. When TOP is attached to a subject or object NP, the subject or object particle is deleted. TOP is incompatible with either the subject or object marker because NOM and ACC are focus markers, while TOP has almost the opposite function. TOP never co-occurs with a genitive either because the genitive combines two NPs into a larger NP, and a part of a NP cannot be topicalized.

In addition to setting the scene, TOP is used to mark contrastive/exclusive meaning. Word order plays a role in topic marking. In Korean, the main subject, which usually is the first item in a sentence, is the default topic. If another item is topicalized, it precedes the subject. If the topicalized unit is not moved to the front, it has contrastive meaning. A sentence can have only one sentential topic, as one cannot talk ‘about’ more than one thing at a time. This is why an embedded sentence can only have a topic marker with contrastive meaning, e.g. *ku siin-un* [[*caki-ka ssu-n*]<sub>S</sub> *si-nun an ulph-nun*]<sub>S</sub> *koyin-i-ta*, but not *\*ku siin-un* [[*caki-ka ssu-n*]<sub>S</sub> *si-nun an ulph-nun*]<sub>S</sub> *koyin-i-ta* for

‘That poet (TOP) is an eccentric [who does not recite the poems he wrote himself [but recites all the others’]’.

However, a sentence can have any number of contrastive topics in addition to the sentential topic, e.g. *na-nun* (TOP) [*hankwuksalam-un* (TOP/contrastive) *cohaha-nun*] *ccikay-nun* (TOP/contrastive) *mos mek-nun-ta*. ‘The only thing I cannot eat is (spicy) stew [that only Koreans like]’ (lit. ‘As for me, I cannot eat Korean stew [that Koreans (and no one else) like [, but can eat all other Korean foods.]’).

Other ways to express focus and emphasis include the use of cleft sentences, which have the structure of nominalized sentences (by *kes* or other pertinent noun) + TOP + focused NP + copula *i-*, e.g. *hankul-ul changceyha-n salam-un seycong-i-ta* ‘It is **Sejong** who invented *hankul*’ or *seycong-i changceyha-n kes-un hankul-i-ta* ‘What Sejong invented is **hankul**’ instead of *seycong-i hankul-ul changceyhay-ess-ta* ‘Sejong invented *hankul*’. This structure is still rather uncommon and may have originated from translation of English texts.

### 5.5.5 Passive and causative

Korean has two kinds of passivization: morphological, with passive suffixes, and the use of the inchoative auxiliary verb *-e ci-* ‘become’.

In a sentence with passive voice, the grammatical subject of the sentence is the recipient of the action expressed by the verb, not the agent as the subject is in an active voice sentence. The actual agent phrase takes a dative form, *-ey* (inanimate), *-eykey* (animate), *-hanthey* (animate, informal), *-kkey* (HON), or the phrase *-ey uyhaye(se)* ‘by (the will/power of) . . .’, but is not necessarily expressed. The verb takes a passive form, with the structure:

- (34) Passive construction  

$$[\text{NP}_{2\text{-Subj}} \quad (\text{NP}_{1\text{-Dat}}) \quad \text{V-PASS}]_{\text{S-Pass}}$$
 AgentPhrase

For example, compare the passive *kwunswu-ka* (NOM) *amhayngesa-eykey* (DAT) *cap-hy-ess-ta* ‘The magistrate got caught by the Secret Royal Inspector’, with the active *amhayngesa-ka* (NOM) *kwunswu-lul* (ACC) *cap-ass-ta* ‘The Secret Royal Inspector caught the magistrate’.

Only transitive verbs can be passivized by suffixation. The passive suffix has four allomorphs: *-i* following a vowel (*po-i-* ‘is seen’), *-li* after *l* (*phal-li-* ‘is sold’), *-ki* after a nasal (*an-ki-* ‘is embraced/held in bosom’), and *-hi* (*nok-hi-* ‘is melted down’) elsewhere. Denominalized verbs of the N + *ha-* form, e.g. *phamyen.ha-* ‘fire’ are passivized by replacing *ha-* with *toy-* ‘become’, as *phamyen.toy-*. However, to emphasize the adversary situation, *tang-ha-* ‘suffer’ can be used as *phamyen.tanghay-ss-ta* ‘X got fired’.

There are many passive sentences with no corresponding active ones, e.g. *nalssi-ka phul-ly-ess-ta* ‘The weather turned warm’ (lit. ‘The weather got thawed’), where there is no clear agent. *-e ci-* is part of a productive process expressing change of state, and can apply not only in a passive construction, *kaci-ka kkekk-e ci-ess-ta* ‘The branch got broken’, but also to stative verbs, *nalssi-ka coh-a ci-ess-ta* ‘The weather has become fine’.

Passive constructions in general are used very commonly now, but older generations find passive sentences unnatural. Certainly familiarity with foreign languages that have an extensive use of passive sentences, especially in the academic literature, is a factor in the increased use of passive constructions in Korean.

Causative constructions turn the subject of the basic sentence into a direct or (if from a transitive verb) indirect object, e.g. *aki-lul ca-yw-ess-ta* ‘I put the baby to sleep’, vs. *aki-ka ca-ss-ta* ‘The baby slept’, *na-nun ttal-eykey phyenci-lul ssu-key hay-ss-ta* ‘I made my daughter write a letter’. There are two kinds of causative: short form and long form. Short-form causatives are either by suppletion (e.g. *sikhi-* ‘order’ for *ha-* ‘do’ and *ponay-* ‘send’ for *ka-* ‘go’), or by suffixes: *-chwu* after *c(h)*, as in (*nuc-chwu-* ‘delay/make X late’); *-ywu/-iwu* or *-i* after a vowel (*khiwu-* ‘raise’ [*< khu-* ‘grow’ + *iwu*]; *ca-ywu-* ‘put X to sleep’; *po-i-* ‘show’); *-i* after *k* (*nok-i-* ‘melt’); *-li* after *l* (*tol-li-* ‘turn’); *-ki* after a nasal ( $\pm h$ ) or *s* (*wus-ki-* ‘make X laugh’); *-hi* elsewhere (*nelp-hi-* ‘widen’). The productive construction is the long form, *-key* + *ha-*, e.g. *yak-ul mek-key hay-ess-ta* ‘X made Y take the medicine’.

The two types of causative constructions are syntactically and semantically quite different from each other. While suffixal causativization has exclusive ‘causative’ meaning, the long-form construction has various degrees of indirect causation (Sohn 1999: 376–7).

### 5.5.6 Speech levels and respect

Korean is an ‘honorific’ language in that the speaker’s attitude toward the hearer, the referent, and the speaker himself/herself is linguistically encoded. The extensive system of ‘honorifics’ or ‘linguistic protocol’ reflects Korean culture, which puts heavy emphasis on maintaining proper human relationships.

Two important axes determining the choice of appropriate ‘speech level’ are power and solidarity (Brown and Gilman 1960). Two focal points in Korean honorifics are reference and address. By the choice of the ‘style’ of language, the speaker’s attitude toward the referent or the addressee is explicitly conveyed.

#### 5.5.6.1 Honorific vocabulary

Korean has a limited set of inherently honorific lexical items, in nouns (e.g. *yakcwu* ‘liquor’, *tayk* ‘house’, *olapeni* ‘elder brother’, *sepangnim* ‘brother-in-law’), verbs (e.g. *kyeysi-* existential, *cwumusi-* ‘sleep’, *tusi-* ‘drink/eat’), particles (e.g. *-kkey* ‘to’, *-kkeyse* ‘from/NOM’). Humble terms are found in nouns (e.g. *ce* ‘I’, *cehuy* ‘we’, *soin* (archaic) ‘I’, *aypi* ‘father’).

#### 5.5.6.2 Honorific titles

Koreans use given names only when talking to or about children, junior relatives, close childhood friends, classmates, etc. Otherwise, Koreans generally talk to or about people by their names followed by or only by titles that refer to their occupations, places of residence, kinship, and other identifying terms. A name with a title can be followed by the generic honorific suffix *-nim* ‘a respected/beloved one’. This suffix can also be attached directly to the name without any title, in which case it carries the basic meaning of respect and affection. The honorific suffix *-ssi* has traditionally been used in a less intimate, or formal and neutral context, but today this form is frequently used among young colleagues or couples to show respect and intimacy. *-nim* can be attached to any name, titled or not, but *-ssi* can be directly attached only to the full or given name, and not to any other title. For an all-purpose title of respect *sensayng(nim)* ‘teacher’ is used. A clearly younger and junior male is addressed by name + *-kwun* and the female by

name + *-yang*. An older woman with a respectable social position can be given the title *-yesa* (lit. ‘a lady scholar/writer’). One can also address as *senpay* a senior from one’s own school or occupational environment, to show respect.

All titles including kinship terms are used as quasi-pronouns and carry the added meaning of respect, love, and care the particular position is expected to receive. The Korean pronoun system, therefore, looks quite complex because of various honorific variants. However, the pronoun system is rather deficient: true pronouns are limited to a few neutral forms such as *na* ‘I’, *ne* ‘you’, and *wuli* ‘we’, and their deferential or humble counterparts such as *ce* ‘I (HUM)’, *tangsin* ‘you (Neutral), s/he (HON)’, and *cehuy* ‘we (HUM)’. However, even these pronouns are far from being used as frequently as in other languages, as they can be dropped as long as context makes them clear. An important reason for pronoun avoidance is that even if choosing the proper terms of reference and address was not that difficult, a constant reminding of that relationship may go against good interpersonal relationships.

### 5.5.6.3 Grammatical honorification

There are broadly two types of grammatical honorification: referent honorification (RH) and addressee honorification (AH). RH includes subject honorification (SH) and object or non-subject honorification (OH). Today OH is not a productive grammatical process but is expressed mainly through a limited set of special vocabulary.

In order to show deference to the subject of the sentence (SH), the suffix *-(u)si* (HON) is attached to the verb stem, forming an honorific verbal together. Therefore, it comes after the other stem-expanding suffixes such as passives and causatives but before suffixes such as tense-aspect, sentence types, and addressee honorification. In general, the deferential subject particle *-kkeyse* instead of the usual *-i/-ka* is used.

Object or non-subject honorification (OH) is a historical residue of an older suffix that used to show the speaker’s humility and deference toward the object of a sentence, but today it is observable mainly in a few special vocabulary items such as *tuli-* ‘give (to someone honorable)’, *mosi-* ‘accompany (someone honored)’, *poyp-* ‘see (someone honored)’.

The speaker expresses his/her attitude toward the hearer (AH) by choosing the proper honorific form for the final verb of a sentence. Various forms of AH for all types of sentences were given in Table 5.10: the ‘plain’ style is ‘speaking down’ and therefore intimate or dominating, the ‘condescending’ style respects a lower-status hearer, the ‘equal honorific’ style respects an equal, the ‘deferential’ is formal and respects a senior or a stranger, and there is also ‘superdeferential’. Most AH forms are based on the assessment of the comparative power of the speaker and the hearer (J. Song 2005: 123–9). However, as modern language protocol dictates that speakers avoid power-based speech styles, two AH forms, ‘panmal’ and *-eyo* or ‘polite’ endings, have emerged as the most convenient forms to use. Because they originate from unfinished sentences, they denote no clear scale of deference. They could in fact be considered respectful in certain contexts and very rude in others. For example, if used toward a person for whom the speaker would have normally used a ‘speaking-down’ style (‘plain’), such as a mother to her child or a schoolteacher to a pupil, panmal would sound respectful – as long as it is not used sarcastically – because it mitigates the speaking-down aspect by not making the AH form too explicit. If panmal were spoken to a person for whom a clearly deferential style would be appropriate, as in the case of a student to his/her teacher or between strangers, it would sound extremely rude, again by not explicitly marking a deferential AH.

The *-eyo* style, because it usually derives from *panmal* + polite particle *-yo*, is like *panmal* in that its choice is not based on power relationship. The *-eyo* style denotes respect for human dignity, because it originates from an embedded structure of ‘It is that . . .’ as a politeness strategy (Kim-Renaud 1999). *Panmal* and the *-eyo* style thus occupy new, clearly definable spaces within the whole honorific system, and that is exactly why they are the most commonly used AH forms today. Strategies for mitigating the impression of aggression are employed much more frequently by women than men. This is why both *panmal* and the *-eyo* style sound intimate and are considered ‘feminine’.

Over recent years the Korean honorific system has undergone notable changes reflecting rapid transformations in Korean society. The repertoire of the AH forms is being simplified, as overly status-conscious forms are increasingly avoided, and only the nonthreatening parameters such as age and kinship hierarchy, and the formal/informal situations seem to allow differentiation. Even factors like kinship and age, which are supposed to be more or less constant (Chang 1973: 40), may no longer be clearly represented in a power hierarchy as they were formerly.

### 5.5.7 Adverbials

Adverbs usually modify verbs and precede the verbs they modify. However, they can also modify pre-nouns and other adverbs, as in *taytanhi ppalli kkuthnay-ss-ta* ‘X finished very fast’.

There are two types of adverbs: inherent adverbs and ones derived by adding an adverbializing suffix.

Some of the most common adverbs are inherent and are not derived from other parts of speech. Adverbs do not inflect and can stand by themselves. Lexical adverbs cover a wide range of semantic categories, including: degree (*acwu* ‘very’, *cham* ‘really’, *kkwoy* ‘quite, pretty much’, *phok* ‘quite a lot’); the negative (*an/ani* ‘not’ and *mos* ‘cannot’); negative polarity adverbs (*celtaylo* ‘absolutely’, *cokumto* ‘even a little’, *kyelkho* ‘resolutely’); provisional (*ama* ‘perhaps’, *manyak* ‘if’, *selsa* ‘even if’); repetitive (*cakkwu* ‘repeatedly’, *tasi* ‘again’, *tto* ‘again’); temporal (*camkkan* ‘for a little while’, *kot* ‘soon’, *ilccik* ‘early’).

One important category of inherent adverbs is sound-symbolic (SS) vocabulary. One different characteristic of SS adverbs is that the vowels and consonants in SS vocabulary can be and are manipulated for appropriately nuanced expressions (see 5.6.1.1).

Some nouns act as adverbs: *cengmal* ‘truth/really’, *chammal* ‘truth/truly’, *onul* ‘today’, *kucekkey* ‘the day before yesterday’, *molay* ‘the day after tomorrow’, *taum* ‘next’, *mucoken* ‘no condition/unconditionally’.

However, the majority of adverbs are derived from other linguistic items, including other adverbs, nouns, verbs, or determiners, through a highly productive process of adverbialization by suffixes. The most common suffix used to turn any verb or a sentence into an adverb is *-key*, e.g. *wuaha-key* ‘elegantly’, *pappu-key* ‘busily’, *swip-key* ‘easily’, *nuc-ci anh-key setwulu-seyo* ‘Please hurry **in order not to** be late’. Other adverbialization processes by suffixation include: stative V + *-i* (e.g. *kkaykkus-i* ‘cleanly’, *pantus-i* ‘straight’, *katuk-i* ‘fully’, *putuk-i* ‘as a last resort’; NB: *-ha* + *-i* > *-hi*, e.g. *swunswunhi* ‘smoothly’, *sinsokhi* ‘fast’, *ssalssalhi* ‘coldly’).

Various postpositive phrases (PPs) often form adverbials. The most common particles used include *-(u)lo* ‘for’ (e.g. *cengchayksang-ulo* ‘for policy reasons’, *hyengsikcek-ulo* ‘formally’, *macimak-ulo* ‘finally’); *-ey* ‘to/at/in/on (static)’ (e.g. *ilpon-ey* ‘in Japan’, *chayksang-ey* ‘on the desk’); *-eyse* ‘from/in (dynamic)’ (e.g. *ilpon-eyse* ‘(do X) in Japan,

from Japan’); *-puthe* ‘from’ (e.g. *twusi-puthe* ‘from 2 o’clock’, *hakkyo-puthe* ‘from school’); *-kkaci* ‘up to’ (e.g. *paksa-kkaci* ‘up to PhD’). Some PPs contain complex particles consisting of a noun + a locative particle in order to be more specific, e.g. *chayksang(-uy) alay/an/mith/pakk/wi/yeph-ey* ‘at the bottom/inside/below/outside/on/at the side of the desk’.

Some nouns, including bound ones that have a modifier (noun or sentence) with or without another postposition, form an adverbial phrase together with the modifier, e.g. *cheymyen-sang* ‘in consideration of face’, *mikwuk cheycay-si* ‘at the time of the American sojourn’, *cwungkwuk-ey ka-ss-ul ttay(-ey)* ‘when X went to China’ (NB: the past is *-ess* + adnominal *-(u)l* rather than the usual past adnominal *-(u)n*), *kecismal ha-nun tey-ey* ‘at (the circumstance of) X’s lying . . .’.

### 5.5.8 Coordination

Coordinate clauses are conjoined by suffixes of three types: (i) the ‘and’ type just enumerating a sequence of events or state (*-e* ‘and’, *-keniwa* ‘X is one thing and then’, *-ko* ‘and’, *-ko(se)* ‘and (then)’, *-ta(ka)* ‘X was doing . . . and then’, *-(u)mye(nse)* ‘while doing . . .’, *-yo* ‘doing . . . and also’); (ii) the ‘but’ type contrasting the later clause to the previous one (*-ciman* ‘but’, *-teni* ‘X did . . . , and then’, *-ulppun(man)anila* ‘not only did X do . . . but also’, *-(u)na* ‘X is true but’, *-(u)nkahamyen* ‘sometimes looks . . . and then’); and (iii) the ‘or’ type expressing choice (*-kena* ‘or’, *-tunci* ‘or’). Optionally, conjunctive *kuliko* ‘and’, disjunctive *kulena* or *manun* ‘but’, or *animyen* ‘if not’, may be used as sentence adverbs.

In coordinate sentences, only the final verb is marked explicitly for tense, modality, and sentence-typing, etc., e.g. *ecey ilccik mokyokha-ko* (vs. *\*mokyokhay-ss-supnita-ko*) *ca-ss-supnita* ‘Yesterday, I took a bath and went to bed early’, and *swua-nun chayk-ul sa-ko mina-nun kong-ul sa-ss-eyo* ‘Sua bought a book and Mina bought a ball’.

Another method of coordination is juxtaposing contrastive words, or the affirmative with the negative verb *mal-*, repeating the same conjunctive suffixes after each:

- (35) a. *ton-i iss-tunci eps-tunci muncey-ka an*  
 money-NOM have-tunci not.have-tunci problem-NOM NEG  
*toy-ss-ta*  
 become-PAST-DECL
- b. *ton-i iss-tunci mal-tunci muncey-ka an toy-ss-ta*  
 NEG  
 ‘Having or not having any money didn’t matter (lit. didn’t become a problem).’

Another common way of coordination is actually splitting the clauses into independent sentences and juxtaposing them, e.g. *cip-ul sa-sy-ess-eyo? pilla-lul sa-sy-ess-eyo?* ‘Did you buy a house or a condominium?’ (lit. ‘Did you buy a house? Did you buy a condo?’).

Colloquially, a special coordinate structure using volitional *-keyss* + *-ta* is commonly used:

- (36) *meli-to coh-keyss-ta, ton-to iss-keyss-ta, cipan-to coh-keyss-ta,*  
 brain-too be.good- money-too have- family-too be.good-  
*mues-i te philyoha-ni?*  
 what-NOM more need-Q  
 ‘X has got brains, money, and a good family background. What more do you want?’

### 5.5.9 Subordinate clauses

A complex sentence has one main clause and one or more embedded subordinate clauses. Embedded clauses are of five different types: conjunctive (5.5.9.1), relative (5.4.1.1), complement (5.5.9.2), nominalized (5.5.9.3), and quotative (5.5.9.4). Subordinate clauses can function as adverbs, noun modifiers, or nouns.

#### 5.5.9.1 Adverbial conjunctive clauses

The adverbial clause always comes before the main clause. The verbals and conjunctive endings (e.g. *-ese* ‘and so’, *-eto* ‘even though’, *-eya* ‘only if’, *-nulako* ‘what with doing . . .’, *-(n)untey* ‘but, and . . .’, *-taka(nun)* ‘do . . . and (do something else)’, *-tamyen/-lamyen* ‘if that’s the case’, *-telamyen* ‘if it were the case that . . .’, *-tolok* ‘so that . . . may’, *-(u)lye(ko)* (*-(u)le* after a verb of coming and going) ‘in order to . . .’, *-(u)myen* ‘if’, and *-(u)ni(kka)* ‘since’), which may be called ‘converbs’, together form a word. Examples with *-tolok* are *sa-tolok* ‘in order to buy, so that X buys’, *ta al-tolok* ‘to tell everyone/so that all should know’, *cwuk-tolok* ‘to death’ (lit. ‘until X dies’).

Conditional clauses are made with various endings, including *-(u)myen*, *-eya*, *-ketun*, *-takanun*, *-ta/la myen*, *-telamyen*, but the degree of hypothesis or *irrealis* information varies, the latter two forms being the most counterfactual. However, as Bak (1994) and others have observed, even the prototypical conditional *-(u)myen* is far from being categorically *irrealis*, as we can see in sentences such as *enni-myen ceyil-iya?* [*enni* and copula *-i-* fused] ‘So what if you are the elder sister’ (lit. ‘if you are the elder sister, does it make you the best?’) = ‘Don’t be so bossy!’

The use of adverbs such as *kalyeng*, *manyak*, or *manil* in the clause-initial position or a form like *cip-ey iss-ess-umyen cenhwa-lul pat-ass-ul theynte*, where *-(u)myen* follows the past tense and the main clause has the hypothetical ‘would have’, makes the *irrealis* meaning a lot stronger. The suffix *-ese* can never be preceded by a past tense marker, e.g. *yaksok-i iss-ese* (*\*iss-ess-ese*) *phathi-ey mos ka-ss-ta* ‘I could not go to the party, because I had an appointment’.

Postpositive adverbial phrases include forms such as *V-ki-ey* and *V-ki ttaymun-ey* ‘because . . .’ (see 5.5.9.3), adnominal (ADN) *cwung-ey* ‘while doing . . .’, ADN *han* ‘as long as’, ADN *kos-ey* ‘where . . .’, ADN *palam-ey* ‘in the midst of doing . . .’, ADN *ttay(ey)* ‘when . . .’, e.g. *sonnim-i o-nun palam-ey swuep-ey mos ka-ss-ta* ‘What with having a guest, I could not go to class’.

#### 5.5.9.2 Complement clauses

Complementation shares certain characteristics with the typical relativization process, such as turning the sentence-concluding ending into an adnominal suffix. The main difference between them is that complementation introduces a head noun which is not in the embedded sentence. A commonly used form for this purpose is either some common reporting or evaluating noun, or bound nouns such as <sup>B</sup>*kes* ‘thing, fact, assumed reality’, as in *mina-nun [appa-ka nolayha-nun] kes-ul tul-ess-ta* ‘Mina heard [her daddy singing]’, whose embedded sentence *appa-ka nolayha-n-ta* ‘Daddy sings’ has no relation to the head it modifies. Other nouns used include *cheji* ‘(current, unfavorable) status’, <sup>B</sup>*chek/chey* ‘pretense’, *hyengphyen* ‘situation’, *il* ‘fact, business’, *kanungseng* ‘possibility’, *kkatalk* ‘reason’, *kyengwu* ‘case’, <sup>B</sup>*li* ‘(good) sense, reason’, *mukkim* ‘feeling’, <sup>B</sup>*pa* ‘backdrop, ground, place’, *pep* ‘principle, (natural) law’, *saceng* ‘(individual) situation’, <sup>B</sup>*seng*

'quality, characteristics', <sup>B</sup>*swu* 'possibility', <sup>B</sup>*tus* 'appearance', <sup>B</sup>*ttus* 'vision', *yaksok* 'appointment', <sup>B</sup>*yang* 'as if'.

A reporting or 'hearsay' suffix *-tanun* (< *ta-ko ha-nun* 'they say (Adnominal)') is used, especially when the noun has some kind of opinionated meaning, such as *cwucang* 'opinion, claim', *iyaki* 'story', *kopayk* 'confession', *kyenhay* 'opinion, personal idea', *mal* 'words, report', *poko(se)* 'report (document)', *poto* 'media coverage', *sasil* 'true happening', *senipkyen* 'preconception', *somun* 'rumor', *sosik* 'news', *uykyen* 'opinion'. In this case, the only modifier ending used is *-tanun*, which is in the present tense, and the embedded sentence carries tense and modal information, e.g. [*weltukhep-ey iky-ess-tanun*] *sosik-ey hungpunhay-ss-ta* 'I got excited at the news [that X won the World Cup]'.

Some nouns, especially bound ones, require specific verbs with which they form complex predicates with a kind of selectional restriction between the noun and the verb. For example, <sup>B</sup>*kes* combines with *kath-* 'is the same' to form the complex meaning 'It looks like . . .', as in *caymi iss-ul kes kath-ta* 'It looks interesting'. Some other complex predicates which are almost grammaticalized units are given in (37):

- (37)
- |                                       |   |
|---------------------------------------|---|
| <i>-(u)l kes-i-</i>                   | 'it will probably happen . . .'                     |
| <i>-ten cham-i-</i>                   | 'I was about to . . .'                              |
| <i>-(u)n kyek-i-</i>                  | 'it is practically like . . .'                      |
| <i>-(u)n seym-i-</i>                  | 'it amounts to/equals . . .'                        |
| <i>-na/-nunka po-</i>                 | 'it seems that . . .'                               |
| <i>-(u)l seng siph-</i>               | 'it looks like . . .'                               |
| <i>-(u)l tus ha-</i>                  | 'it looks as if it could . . .'                     |
| <i>-(u)l cwul al/molu-</i>            | 'X knows how to . . ., has the competence to . . .' |
| <i>-nun/(u)n/(u)l cek-i iss/eps-</i>  | 'have (not) the experience'                         |
| <i>-nun/(u)n/(u)l il(-i) iss/eps-</i> | 'it makes no sense that . . .'                      |
| <i>-(u)l kkataalk-i iss/eps-</i>      | 'there is no reason that . . .'                     |
| <i>-(u)l li(-ka) iss/eps-</i>         | 'it makes no sense that . . .'                      |
| <i>-(u)l swu iss/eps-</i>             | 'X can, it is possible to . . .'                    |

### 5.5.9.3 Nominalization

Verbals can become nouns by adding nominalizing suffixes *-ki* or *-(u)m*.

The nominalizing suffix *-ki* allows the new noun to retain the original verbal meaning strongly, as in [*na-nun [chwum chwu]<sub>S</sub>-ki-lul*]<sub>S=NP</sub> *cohaha-n-ta* 'I love (the act of) dancing' and *enceyna yeppu-si-ki(-lul) kiwenha-pnita* 'I wish you will always stay pretty'. *ki*-nominalization appears in many grammaticalized endings, but the original sense of the verb, be it action or state, is strongly present. Some key conjunctives have the *ki* + postposition structure, e.g. *X-ki-nun khenyeng* 'far from X being the case', *-ki ttaymun-ey* 'because of the fact . . .', *-ki wiha-y/ye(se)* 'for the sake of X'. In all of these cases, the *ki*-nominalized clause can be replaced by a noun (phrase). The *ki*-nominalization + a stative verb (e.g. *swip-* 'be easy', *elyep-* 'be difficult') structure is very commonly used in Korean to express the meaning, 'it is easy/difficult to . . .'. Other examples of common stative verbs used in this structure are *coh-* 'be good/agreeable', *ilssu-* 'be customary/expected/routine', *kwich'anh-* 'be cumbersome', *nappu-* 'be bad', *silh-* 'be detestable/disagreeable/unpleasing', and *kantanha-* 'be simple'.

*ki*-nominalization expresses the speaker's affect (by admission, recognition, conviction, etc.) about the truthfulness of the sentence:

- *-ki* + topic *-nun* + the verb *ha-* recognizes the truth value of the particular meaning of the nominalized clause, e.g. *yosay nalssi-ka coh-ki-nun hay-yo* ‘The weather is good these days, **I admit** (but I have no time to go out . . . or something of the kind)’. *ha-* here can be omitted in panmal style. This expression can contradict compliments, e.g. *coh-ki-nun(-yo)!* ‘(You really think) it is good?’
- *-ki-to* (emphatic) *ha-*: *cham noph-ki-to ha-ta!* ‘It **sure** is high!’
- *-ki-man* (‘only’) *ha-*: *o-si-ki-man ha-sey-yo!* ‘You **just** come (and we will take care of you/the rest well)!’

Nominalization by *-(u)m* adds the meaning of ‘the fact of acting/being . . .’, as in *[[mina-ka imi uysa-ka toy-ess]<sub>S-um</sub>]<sub>N</sub>-ul moll-ass-ta* ‘I didn’t know Mina had already become a physician’.

*(u)m*-nominalization usually happens with a cognitive verb such as a verb of ‘knowing’, e.g. *al/molu-* ‘know/not know’, *alatat-* ‘understand’, *alli-* ‘inform, disseminate’, *cen-ha-* ‘transmit’, *cimcakha-* ‘guess’, *hwakinha-* ‘verify’, *ihayha-* ‘understand’, *kkyatat-* ‘awake to, deeply understand’, *musiha-* ‘ignore’, *nwunchichay-* ‘guess’.

5.5.9.4 Quotative clauses

When a directly quoted sentence is embedded within a sentence, one of the variants *-hako* or *-lako* connects it with the main sentence, as in *enni-ka “pelsse wa-ss-ni?”-lako/hako malhay-ss-ta* ‘My elder sister said, “You have already come?”’

When a sentence is indirectly quoted within a sentence, *-ko* is attached to the plain style of the final verb of the quoted sentence, i.e. *-ta* for declarative, *-nya* for interrogative, *-la* for imperative, and *-ca* for exhortative:

- (38) a. Declarative:  
*[mina-ka [swua-ka ka-ss-ta]<sub>S</sub>-ko malhay-ss-ta]<sub>S</sub>*  
 Mina-NOM Sua-NOM go-PAST-DECL-ko say-PAST-DECL  
 ‘Mina said that Sua had gone.’
- b. Interrogative:  
*[mina-ka [swua-ka ka-ss-nya]<sub>S</sub>-ko malhay-ss-ta]<sub>S</sub>*  
 -Q  
 ‘Mina asked if Sua had gone.’
- c. Imperative:  
*[mina-ka swua-eykey ka-la]<sub>S</sub>-ko malhay-ss-ta]<sub>S</sub>*  
 -to -IMP  
 ‘Mina told that Sua should go/told Sua to go.’
- d. Propositive:  
*[mina-ka swua-eykey ka-ca]<sub>S</sub>-ko malhay-ss-ta]<sub>S</sub>*  
 -to -EXH  
 ‘Mina suggested to Sua to go with her.’

As can be seen in (38) the quoting verb *malha-* can be used with all types of quoted sentence. Quoting verbs are broadly categorized as: (i) declarative, e.g. *cek-* ‘write’, *kule-* ‘say (so)’, *(mal)ha-* ‘say’, *nonha-* ‘discuss’, *sayngkakha-* ‘think’, *soksaki-* ‘whisper’, *ssu-* ‘write’, *woychi-* ‘shout out’; (ii) jussive, e.g. *aywenha-* ‘plead’, *ceyanha-* ‘propose’, *kwenha-* ‘suggest’, *myenglyengha-* ‘order’, *yochengha-* ‘request’; (iii) interrogative,

e.g. *mut-* ‘ask’, *cilmunha-* ‘question’, *chwukwungha-* ‘press hard upon, question persistently’.

## 5.6 LEXICON

Korea’s long history of contact with various foreign cultures and languages – especially through their long use of and familiarity with written, classical Chinese – as well as Korean-internal language change has resulted in four distinct lexical strata, two native and two foreign.

### 5.6.1 Native Korean vocabulary

#### 5.6.1.1 Sound-symbolic (SS) vocabulary

One of the most salient features of Korean is sound symbolism. SS words, also called ‘expressives’, ‘ideophones’, ‘mimetics’, and ‘phonaesthetics’, include but are not limited to onomatopoeia, and manifest a whole range of expressions from auditory, visual, and other sensory images to the inner feelings they evoke. Most SS vocabulary is completely or partially reduplicated.

In Korean both vowels and consonants can be varied intentionally to produce different connotations, travelling through the consonantal-strength scale and shifting between ‘bright’ (small, shallow, bright) and ‘dark’ (large, deep, dark) vowel-harmony. For example, when describing the flow of water, it can be expressed as gentle when it is described as *colcol*, but like an uncontrollable flow when the vowel is changed to *cwulcwul* (also used to express the flawless reciting of a poem).

(39)

Bright vowels	Dark vowels	Meaning
<i>colcol</i> 줄줄	<i>cwulcwul</i> 줄줄	‘flowingly’
<i>palpal</i> 발발	<i>pelpel</i> 벌벌	‘trembling’
<i>phongtang</i> 풍당	<i>phungteng</i> 풍덩	‘splashing, falling plop (into water)’
<i>taltal</i> 달달	<i>teltel</i> 덜덜	‘rumbling, rattling’
<i>talkulak</i> 달그락	<i>telkulek</i> 덜그럭	‘clattering’
<i>congalcongal</i> 종알종알	<i>cwungelcwungel</i> 중얼중얼	‘mumbling’
<i>sokunsokun</i> 소근소근	<i>swukunswukun</i> 수근수근	‘whispering’
<i>kkomcilak</i> 꿈지락	<i>kkwumcilek</i> 꿈지럭	‘dragging, taking time’
<i>kholkhol</i> 콜콜	<i>khwulkhwul</i> 쿵쿨	‘snoring’

Consonant strength has direct semantic correlates, and the stronger consonants connote more intense meaning. For example, being dark can have three different degrees of intensity: (i) neutral, e.g. *kamkam* ‘dark’; (ii) intensive, e.g. *kkamkkam* ‘desolately dark and spooky’; (iii) paraintensive *khamkham* ‘pitch dark’.

(40)

Neutral	Intensive	Paraintensive	Meaning
<i>paltak</i> 발탁	<i>ppalttak</i> 빨딱	<i>phalthak</i> 팔탁	‘jerking’
<i>pingping</i> 빙빙	<i>ppingpping</i> 땡땡	<i>phingphing</i> 핑핑	‘round and round in circles’
<i>taltal</i> 달달	<i>ttalttal</i> 딸딸	<i>thalthal</i> 탈탈	‘rumbling, rattling’
<i>teltek</i> 덜걱	<i>telkkek</i> 덜꺅	<i>thelkhek</i> 털꺅	‘clattering’

<i>songalsongal</i>	<i>ssongalssongal</i>	—		‘in small clusters’
송알송알	송알송알			
<i>swukunswukun</i>	<i>sswukunsswukun</i>	—		‘whispering’
수근수근	수근수근			
<i>colcol</i>	<i>ccolccol</i>	<i>cholchol</i>	촘촘	‘flowingly’
졸졸	졸졸	촘촘		
<i>congcong</i>	<i>ccongccong</i>	<i>chongchong</i>	충충	‘in small bits’
종종	종종	충충		
<i>kolkol</i>	<i>kkolkkol</i>	<i>kholkhol</i>	꿀꿀	‘trickling, gurgling’
꿀꿀	꿀꿀	꿀꿀		

Korean speakers are conscious of different consonantal strength and the VH principle, both of which are clearly indicated in hankul, as shown above.

However, various phonological changes have caused the VH principle not to operate on clear phonological axes, as it used to in earlier Korean. Consonantal strength also shows signs of change, and interesting variations exist in different dialects.

5.6.1.2 *Non-sound-symbolic native Korean (NK) vocabulary*

Practically all grammatical markers including all noun, verb, and sentence-concluding suffixes are native Korean. In fact, these are what make Korean Korean. When a sentence is filled with loanwords, if it has Korean grammatical markers, the sentence definitely is understood to be Korean, as the following example shows (grammatical markers italicized):

- (41) *khelle-wa*    *khethu-lul*    *cwungsim-ulo*    *kak*    *sallong-eyse*  
 color-COM    cut-ACC    focus-as    each    salon-from  
 LW-NK    LW-NK    SK-NK    SK    LW-NK
- chwuchen.ha-nun*    *heyesuthail-ul*    *han-pen*    *salphye.po-ca*  
 recommend-PRES.ADN    hairstyle-ACC    one-time    review-EXH  
 SK.NK-NK    LW-NK    NK-SK    SK.NK-NK
- ‘Let’s examine the hairstyle that each hair salon recommends, focusing on color and cut.’

Koreans think their native vocabulary gives an intimate, informal, pure, or childlike impression, compared to Sino-Korean words. In a trendy revival of traditional values and customs that started in the 1980s, South Koreans have been giving their children names made up of Korean root words such as *hannwuli* ‘one (or big) world’, *hanul* ‘sky’, *sulki* ‘wisdom’. Even large banks, chain stores, product names, and even political parties are now choosing names that are more native sounding, such as *hanalum* ‘an armful,’ *wuli unhayng* ‘Our Bank’, *cheumchelem* ‘like the beginning,’ and *yellin wulitang* ‘Uri Party’ (lit. ‘Our Open Party’). However, neologisms based on pure Korean roots are still a novelty, as most new names are created based on Sino-Korean roots.

5.6.2 **Vocabulary of foreign origin**

5.6.2.1 *Sino-Korean (SK) vocabulary*

Chinese characters are not used under ordinary circumstances today. However, a great proportion of the Korean vocabulary originated from classical literary Chinese, because Koreans used and internalized it for well over a thousand years until the invention of the Korean alphabet and continued to do so well into the twentieth century. Even today, neologisms are most often based on Sino-Korean roots.

Sino-Korean words (SK) are different from typical loanwords in that their sound values do not originate from the Korean effort to Koreanize contemporary Chinese sounds but from a system codified in earlier riming dictionaries and rime tables, which developed to conform with Korean sound patterns. Koreans thus have appropriated literary Chinese, called 漢字 (*hanca*), as their own in spite of their clearly Chinese origin. Although the absolute majority of Korean words are SK, only about ten percent of the so-called 'basic vocabulary' is thought to have come from Chinese (Martin 1992: 94).

SK roots have become so nativized that even those Korean speakers who do not know any Chinese character understand SK in common usage. Thus, there is no Korean speaker who does not know the word *sensayng* 'teacher', written 先生 in Chinese. The same is true for words such as *cangso* 'place', written 場所 in Chinese. Indeed, the most common Korean proper names, including almost all personal names and names of organizations, and even kinship terms, are Sino-Korean based, e.g. *haksaeng* 'student', and *pumo* 'parents', written in Chinese as 學生 and 父母, respectively.

Koreans prefer creating new vocabulary based on SK roots because SK roots generally consist of one syllable with a clear, independent meaning and the process is highly economical compared to using native words which tend to be polysyllabic.

However, there are also many true loanwords from Chinese in Korean, through Koreans' long study and admiration of the great Chinese classics. These words, borrowed centuries ago, have undergone Korean sound changes, and often semantic changes as well. Some examples that give the impression of being native Korean (NK) but are actually Chinese borrowings include *pus* ← 筆 (SK *phil*) 'writing brush' and *mek* ← 墨 (SK *muk*) 'ink (stick)'. Some other words traditionally thought to be completely NK but now strongly hypothesized to be Chinese loanwords include *kul* ← 契 (SK *kye*) 'writing' and *munuy* ← 文 (SK *mun*) 'pattern'. There are also some examples of SK words which have been coined by Koreans based on Chinese characters but are now considered native: *sanyang* ← 山行 (SK *sanhayng* 'going to the mountain') 'hunting'; *puche* ← 佛體 (SK *pulche* 'Buddha's body') 'Buddha'; *sselmay* ← 雪馬 (SK *selma* 'snow horse') 'sled'.

SK words belong to their own stratum, because some phonological rules apply only to this group of vocabulary. For example, in a specific environment where two SK roots combine to make a word, if the last consonant of the first root is /l/ and the first consonant of the second root is either a dental or a palatal plain obstruent /t, c, s/, these consonants are tensed (42). This tensification does not occur in the case of other lax consonants /p, k/, which are just voiced following the regular phonological rule in such an environment (43):

- |      |    |                 |                         |                                |
|------|----|-----------------|-------------------------|--------------------------------|
| (42) | 發達 | <i>pal-tal</i>  | [palttal]               | 'development (charge-through)' |
|      | 實質 | <i>sil-cil</i>  | [silccil]               | 'substance, essence'           |
|      | 必須 | <i>phil-swu</i> | [p <sup>h</sup> ilssu]  | 'one's lifetime'               |
| (43) | 發見 | <i>pal-kyen</i> | [palg <sup>h</sup> jɛn] | 'discovery'                    |
|      | 筆記 | <i>phil-ki</i>  | [p <sup>h</sup> ilgi]   | 'writing down'                 |
|      | 說伏 | <i>sel-pok</i>  | [sɛlbok]                | 'persuasion'                   |

As in most foreign language adaptation in any language, the use of SK words is usually considered more complicated, less personal, more formal and distinguished, and more scholarly and scientific. When NK vocabulary coexists with SK words, it is the SK words that often sound more pompous and also outdated.

Sometimes both SK and NK lexical items are used together in a quasi-redundant way, but the effort is to mitigate the cold, formal impression of SK. Thus an NK word may

be attached to an SK word redundantly, e.g. 五月달 *owel-tal* ‘the month of May’, where SK *owel* already means ‘the fifth moon’ but NK *tal* ‘moon, month’ is attached.

### 5.6.2.2 Loanwords

Koreans have embraced literary Chinese as their own out of appreciation for the entire Sinitic civilization and tradition, but loanwords from other languages were usually mere adaptations of foreign words mainly for practical purposes. Koreans thus borrowed from the languages of their occupying forces, including the Mongols, Manchus, Jurchens, Japanese, and finally Russians in North Korea and Americans in South Korea. Loanwords from the languages of Korea’s northern neighbors are minimal; however, Japanese and English loanwords are noteworthy.

Many loanwords have seeped into Korean as a result of the South’s particular relationships with Japan and the United States over recent years. Their treatment by Korean speakers is quite different. The so-called ‘language purification movement’ in both Koreas – *Kwuke Swunhwa Wuntong* in South Korea and *Mal Tatumki Wuntong* in North Korea – seems to shun Japanese loanwords in favor of seeking truly native Korean-only vocabulary. Many Japanese loanwords still remain, but it is only because their Japanese origin is not obvious. Any Japanese-sounding vocabulary has been almost completely eradicated.

#### *Japanese loanwords*

As Koreans grew up bilingual during the Japanese colonial era (1910–1945), Japanese influence on Korean language and culture has become significant. Many Japanese loanwords became almost second nature to them. However, after independence, Koreans tried to extirpate anything that resembled Japanese from their language. Therefore, many formerly quotidian Japanese terms have been replaced by translated Korean names, although older people still may find these neologisms artificial, e.g. *tanmuci* ‘sweet pickled radish’ instead of *takwuang* (Japanese *takuan*), and *chopap* ‘vinegar rice’ instead of *susi* ~ *swusi* (J *sushi*).

Those Western loanwords that came to Korea via Japanese sound too Japanese and therefore have been replaced by Korean-style readings of foreign words, e.g. *thulek* instead of *tolakkwu* ‘truck’ (J *toraku*).

Many borrowings, including grammatical and idiomatic expressions, escaped the Korean campaign for eradication; these are mainly Sino-Japanese (SJ) forms based on Sino-Japanese roots that can easily be considered Sino-Korean. Most of these SJ words were coined by the Japanese for nontraditional objects and ideas and were adopted by the Chinese also. In most cases Koreans were not aware that these were coined in Japan; as long as they could be considered Sino-Korean, Koreans seem to have embraced them as their own. Such vocabulary includes 入口 *ipkwu* (< NJ *iriguchi*) ‘entrance’, 雜誌 *capci* (< SJ *zaqshi*) ‘magazine’, 新聞 *sinmun* (< SJ *shinbun*) ‘newspaper’.

#### *Loanwords from Western languages*

Koreans’ first contact with the West through the Chinese, especially in the eighteenth century, was when Koreans took an interest in the Roman Catholic Church and religion. Loanwords from that period are somewhat removed from the original pronunciation, as we can see in such examples as *mathay* ‘Matthew’, *nwuka* ‘Luke’, *paolo* ‘Paul’, *pangcike*

'Francisco', and *yohan* 'John'. These are now used mainly as Christian names by Korean Catholics. These and other words whose pronunciation is remote from the original are replaced by something deemed more closely approximating it to the Korean ear, e.g. *weshingthen* instead of *hwasesngton* 'Washington' and *pullanso* or *phulangsu* instead of *pepkwuk* 'France'. These changes had nothing to do with rejecting loanwords, but in fact reflected efforts to bring them closer to their original pronunciation.

Korea's recent tumultuous history has seen the unprecedented involvement of the United States in the life of Koreans politically and socially. First, during the 1880s American missionaries were at the helm of social reform by educating women and training doctors in modern medicine. Then the United States occupied Korea briefly after its liberation from the Japanese, and soon after the Korean War (1950–3) broke out, in which the U.S. was a major defender of South Korea. Subsequently there were the American Peace Corps volunteers sent to help in postwar reconstruction during the 1960s.

In spite of strong protests from ardent language-purification activists, the tide of globalization has arrived. One only has to pick up a newspaper or magazine to realize that now the great majority of trademarks and everyday expressions are loanwords, most of them from English. It is actually difficult to find a sentence without any English or other Western loanwords. A few strategies employed by Koreans in loanword adaptation include: (i) *u*-insertion to break up impermissible clusters, as in *sutha* 'star'; (ii) neutralization of vowel length, as in *mithing* 'meeting', *weylping* 'well-being'; (iii) Law of Initials is not observed in loanwords, so initial *l*'s are not nasalized and *ni*-/ny-sequences are not altered, e.g. *lipsuthik* 'lipstick', *nithu* 'knit', *nyusu* 'news'; (iv) some foreign sounds that have no close Korean equivalent are pronounced like – or at least with the goal of pronouncing like – in the original language, e.g. [f, v, z, ð, θ] are better pronounced now than before as English is studied more seriously and efficiently than before; (v) long words are truncated, e.g. *aphathu* 'apartment building', *eyekhon* 'air conditioner', *khompi* 'combination', *limokhon* 'remote control', *phesukhom* 'personal computer', *syuphe* 'supermarket', *theyleypi* 'television', and, less recognizably, *khātu* 'credit card'; (vi) many acronyms and initialisms are used as in English, e.g. *ci ai* (GI), *si ti* (CD), *phi si* (PC), *ti ceyi* (DJ), *yu si eyl eyi* (UCLA), and *yu eyn* (UN).

Today it seems that there is no dam against the English flood to protect any language anywhere – not even in traditionally language-chauvinistic France. But in Korea the impact of English is of such magnitude that language purifiers feel they are again being colonized. Young people, however, do not connote the use of English with anything like foreign domination, but they feel it is modern, playful, and even sophisticated to use foreign terms and that they are members of today's great civilization, just as their medieval ancestors regarded using literary Chinese.

Koreans today exercise an interesting style of neologism from native words. Names for products, stores, and programs are coined in such a way as to sound like loanwords from a foreign country, usually a Western nation, but also Japan. The most popular, and invariably humorous, form seems to be one ending in a vowel, so that the coinages sound Spanish, Italian, or even Japanese. For example, small shops or trademarks carry unfamiliar newly coined, foreign-sounding expressions as their names such as *mukka-makka* ('Shall I bite/eat it or not?' imitating the Kyengsang provincial dialectal form), *mekullay sakallay* ('Wanna eat [here] or take out?'), *nwuneyttiney* ('It catches the eye!').

Some loans probably have started by a mislearning of the original words or expanding of their semantic field, e.g. *khening* (< cunning) [via Japanese loan *khanningkwu* < *J kanningu*], 'cheating on a test', *maynsyon* (< mansion) 'condominium', *meymo* 'message', *mithing* (< meeting) 'blind date', *pilla* (< villa) 'town house'. Sometimes, two loanwords

are combined to make a new word, e.g. *eyi eysu* (A. S. = ‘after service’) ‘product warranty’, *opisutheyl* (< office + [ho]tel) ‘office-apartment’; and sometimes native and loaned words are combined, e.g. *maikha sitay* (< my-car + ‘era’), *sayngsen kkassu* (< ‘fish’ + J *katsu* ‘cutlet’), *kosok pesu* (< ‘express’ + bus).

Lexical borrowing is not the only way of adapting foreign expressions. Borrowing extends to syntax and discourse-pragmatics. When only a sentence structure is taken but disguised in translation – sounding like original Korean – such as *cohun nal toyseyyo!* (‘Have a good day!’), the result is adopted even by purification activists, whose main concern is lexical borrowing.

Neologism is quite actively practiced by purification advocates, who perceive borrowing as a ‘foreign intrusion’ into Korean. However, for some, the purification effort is mainly directed against Sino-Korean and Japanese loanwords.

As Korea has recovered from its low point of history and is now poised to be part of the world community, even primary-school students are learning English as a compulsory subject. Many families have relatives living abroad whom they go visit regularly or who come to see them in Korea. Koreans in the diaspora, as well as in Korea, are gaining international prominence in almost every field. Once again Koreans are borrowing foreign words as individual agents, in a spirit of being modern and civilized, not as something imposed by an unwelcome occupying force.

## NOTES

- 1 This chapter draws heavily from the author’s book, *Korean: An Essential Grammar* (Kim-Renaud 2009), which offers a more comprehensive coverage of Modern Korean grammar. The following abbreviations are used: ADV = adverbializer, ADN = adnominal, ATTR = attributive, CL = classifier, COM = comitative, COP = copula, DAT = dative, DECL = declarative, EXH = exhortative, FUT = future, GEN = genitive, HON = honorific, HUM = humble, IMP = imperative, INST = instrumental, NEG = negative, NMR = nominalizer, Nr = Number, ACC = object, PASS = passive, PL = plural, PRES = present, PROM = promissive, Q = interrogative, RET = retrospective, NOM = subject, SUP = suppositional, TOP = topic, VOC = vocative. Speech styles: PLN = plain, EQ = deferential equal, PAN = panmal, POL = polite, DEF = deferential, SUPERDEF = superdeferential.
- 2 Some linguists (e.g. K.-M. Lee 1961/1972; I. Lee and Ramsey 2000; K.-M. Lee and Ramsey 2011) use ‘Contemporary Korean’ to refer to the language of today and reserve ‘Modern Korean’ to refer to ‘Pre-Contemporary Korean’ or ‘Recent Korean’ (called *Kuntay kwuke* in Korean) from the seventeenth century until the end of the nineteenth.
- 3 In 1966 in an official language regulations book (*Cosenmal kyupemcip*) North Korea defined their standard, ‘Culture Language’ (*munhwa-e*), as ‘the language spoken by the workers living in the Pyongyang area’.
- 4 In the Korean lexicon, there is no item ending in a tense affricate.
- 5 There is some dialectal variation in applying palatalization. Some items where the /i/ in question originates from *-uy*, such as /muni/ (< /munuy/) ‘pattern’, do not undergo palatalization. Also many young speakers today do not palatalize /s/ and /ss/.
- 6 /ney-ka/ and /cey-ka/ show casual speech variants [niga] and [ciga]. All variants demonstrate historical change in progress. Cf. also *mues ~ mue* ‘what’, *kes ~ ke* ‘thing’, *uy ~ e* genitive. Another pronominal form, originating from demonstrative + *ai* ‘child’, commonly used in casual speech, shows another type of contraction, i.e. *i+ay* [jɛ:] ‘this (kid)’, *ku+ay* [kjɛ:] ‘that kid, s/he’, *ce+ay* [cɛ:] ‘s/he over there,’ now used by young speakers for inanimate as well as human nouns.

- 7 *haye* is allowed only as a conjunctive, and not as a sentence ender in Standard Korean.
- 8 Yeon and Brown (2011: 402) use the term ‘adnouns’ rather than ‘adnominals’ because of the special characteristics of this grammatical category, compared to, for example, English adnominals.
- 9 Existential *iss-*, *eps-*, and increasingly *kyesi-* and *an kyesi-* take the *-nun* ending, even though they are stative.

## REFERENCES

- Bak, Sung-Yun (1994) ‘The semantics of conditionals’, in Kim-Renaud, Young-Key (ed.) *Theoretical Issues in Korean Linguistics*, Stanford, CA: Center for the Study of Language and Information, Stanford Linguistics Society, 217–31.
- Brown, Roger and Gilman, Albert (1960) ‘The pronouns of power and solidarity’, in Sebeok, Thomas A. (ed.) *Style in Language*, Cambridge, MA: MIT Press.
- Chang, Suk-Jin (1973) *A Generative Study of Discourse: Pragmatic Aspects of Korean with Reference to English*, *Language Research* 9: 2 (Supplement), Seoul: Seoul National University Language Research Institute.
- (1996) *Korean*, Amsterdam and Philadelphia: John Benjamins.
- Cho, Young-Mee Yu and Sells, Peter (1995) ‘A lexical account of inflectional suffixes in Korean’, *Journal of East Asian Linguistics*, 4: 119–74.
- Choe, Hyon Sook (1995) ‘Focus and topic movement in Korean and licensing’, in Kiss, Katalin É. (ed.) *Discourse Configurational Languages*, Oxford: Oxford University Press, 269–334.
- Hong, Cong-sen; Kim, Yu-pem; Ko, Kyeng-thae; Kim, Swu-ceng; Hwang, Kwuk-ceng; Yi, Tong-hyep; Pak, Cin-wan; Yi, Tong-sek; Cha, Cwun-keng; Pay, Seng-wu; and An, Ceng-a (2000) *Hyentay Kwuke uy Hyengsengkwa Pyenchen I* [The formation and evolution of Modern Korean, vol. I], Seoul: Pagijōng.
- Kim-Renaud, Young-Key (ed.) (1997) *The Korean Alphabet: Its History and Structure*, Honolulu: University of Hawai‘i Press.
- (1999) ‘Pragmatic embedding for honorification and politeness’, in Kuno, Susumu; Lee, Ik-Hwan; Whitman, John; Maling, Joan; Kang, Young-Se; and Kim, Young-joo (eds) *Harvard Studies in Korean Linguistics VIII*, Seoul: Hanshin, 419–35.
- (2009) *Korean: An Essential Grammar*, London: Routledge.
- Kuno, Susumu and Kim-Renaud, Young-Key (1987; reprinted 2004) ‘The position of quantifier-like particles in Korean’, *Harvard Studies in Korean Linguistics II*, Seoul: Hanshin, 251–76; reprinted in Kim-Renaud, Young-Key and Whitman, John B. (eds) *Studies in Korean Syntax and Semantics by Kuno, Susumu with Kim, Young-Joo; Kim, Soo-Yeon; Kim-Renaud, Young-Key; Lee, Ik-Hwan; Takami, Ken-Ichi; and Whitman, John B.*, Seoul: Pagijōng Press, 17–43.
- Lee, Chungmin (1996) ‘Generic sentences are topic constructions’, in Fretheim, Thorstein and Gundel, Jeanette K. (eds) *Reference and Referent Accessibility*, Amsterdam/Philadelphia: John Benjamins, 214–22.
- Lee, Hyo Sang (1994) ‘Discourse-pragmatic functions of sentence-type suffixes in informal discourse in Korean’, in Kim-Renaud, Young-Key (ed.) *Theoretical Issues in Korean Linguistics*, Stanford, CA: Center for the Study of Language and Information, Stanford Linguistics Society, 517–39.
- Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, Albany: State University of New York Press.
- Lee, Keedong (1993) *A Korean Grammar on Semantic-Pragmatic Principles*, Seoul: Hankwuk Munhwasa.
- Lee, Ki-Moon (1961/1972) *Kwuke sa Kaysel* [Introduction to the history of the Korean language], Seoul: Minjoong Publishing Co.

- (1997) ‘The inventor of the Korean alphabet’, in Kim-Renaud, Young-Key (ed.) *The Korean Alphabet: Its History and Structure*, Honolulu: University of Hawai‘i Press, 11–30.
- Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge and New York: Cambridge University Press.
- Lee, Yongsung (2008) ‘Allomorphy in Korean particles: multiple inputs and default allomorph’, *Korean Linguistics*, 14: 65–89.
- Lim, Hong-Bin (1996) ‘Kyochakso-wa kyochakpep-uy cenglip-ul wihaye’ [Toward a theory of agglutinating elements and agglutination], Paper presented at the 24th Presentation of Joint Research Projects, Korean Linguistic Society, Seoul.
- Lukoff, Fred (1982) *An Introductory Course in Korean*, Seoul: Yonsei University Press.
- Martin, Samuel E. (1992) *A Reference Grammar of Korean*, Rutland, VT and Tokyo: Tuttle.
- Pak, Miok (2008) ‘Types of clauses and sentence end particles in Korean’, *Korean Linguistics*, 14: 113–55.
- Pak, Miok, Portner, Paul, and Zanuttini, Raffaella (2008) ‘Agreement and the subjects of jussive clauses in Korean’, in Elfner, Emily and Walkow, Martin (eds) *Proceedings of NELS*, 37(2): 127–38, University of Massachusetts, Amherst, Graduate Linguistic Students’ Association.
- Sohn, Ho-Min (1999) *The Korean Language*, Cambridge: Cambridge University Press.
- Song, Jae Jung (1997) ‘The so-called plural copy in Korean as a marker of distribution and focus’, *Journal of Pragmatics*, 27: 203–24.
- (2005) *The Korean Language: Structure, Use and Context*, London and New York: Routledge.
- Song, Seok Choong (1975) ‘Rare plural marking and ubiquitous plural marker in Korean’, *Papers for the Eleventh Regional Meeting of the Chicago Linguistic Society*, 536–46.
- Sung, Kychul (1970/1984) *Kwuke Taywuep Yenkwu* [Korean honorifics], Chwungpuk University Papers, IV; reprinted in Conghwun Kim (ed.) (1984) *Kwukekyengep Yenkwu* [Korean honorifics], Seoul: Cipmuntang, 405–49.
- Yeon, Jaehoon and Brown, Lucien (2011) *Korean: A Comprehensive Grammar*, Abingdon, UK and New York: Routledge.
- Yoon, James H.-S. (1995) ‘Nominal, verbal and cross-categorial affixation in Korean’, *Journal of East Asian Linguistics*, 4: 325–56.
- (2005) ‘Non-morphological determination of nominal particle ordering in Korean’, in Heggie, Lorie, and Ordóñez, Francisco (eds) *Clitic and Affix Combinations: Theoretical Perspectives*, Amsterdam: Benjamins, 239–82.

# KOREAN DIALECTS: A GENERAL SURVEY

*Jaehoon Yeon*

## 6.1 INTRODUCTION

The Korean language is relatively homogeneous and the dialects from different areas can be mutually intelligible to a great extent. Nevertheless, the dialects of Korean exhibit considerable variety in phonology, morphology, and vocabulary. They are finely differentiated into a number of areas based on regional differences. There is no obvious correlation between the modern dialects and the ancient historical divisions of Korea, i.e. the Three Kingdom period. Silla and Paekche roughly coincide with the current southeastern dialect and southwestern dialect respectively, but northeastern, northwestern, central, and Cheju dialects cannot be correlated with any one ancient historical kingdom in Korea. Since Korea is mountainous, the language is quite naturally divided finely into different dialects according to topography.

Most scholars seem to agree on six major dialectal zones based roughly on different geographical regions:

- (1) The northwestern dialects (P'yŏngan province)
- (2) The northeastern dialects (Hamgyŏng province)
- (3) The central dialects (including Kyŏnggi, Hwanghae, Kangwon, and Ch'ungch'ŏng provinces)
- (4) The southwestern dialects (Ch'ŏlla province)
- (5) The southeastern dialects (Kyŏngsang province)
- (6) Cheju dialect (Cheju island)

The dialect used by the Korean community in the Yanbian autonomous prefecture of China in Manchuria can be included in the Hamgyŏng dialects because their mutual similarity is due to the early immigration of Hamgyŏng people to that area and their subsequent linguistic contact. The language spoken by Koreans in Central Asia, i.e. Kazakhstan and Uzbekistan, has evolved from the Yukchin dialect that is part of the Hamgyŏng dialects, but it has many archaic forms as well as innovations compared with the original Yukchin dialects (J. R. P. King 1992). In addition to the regional dialects, North Korean and South Korean reveal a considerable linguistic divergence resulting from North and South division in 1945 and the subsequent Korean War in 1950. In this section we will examine the dialectal differences in Korean, salient phonological isoglosses, typical features of each dialect, and the linguistic divergence between North and South Korean. In the body of the text, we transcribe all Korean words in Yale Romanization, but Korean proper names are transcribed in the McCune-Reischauer system according to the common practice. In addition, we also adopt a slightly modified version of phonetic symbols for some linguistic/phonetic examples when necessary.

Dialect subzones can be classified by identifying the isoglosses of certain phonological features and morphological/grammatical features as well as lexical features. From the distribution of isoglosses, one can establish many dialectal subareas, but two or more dialectal areas may overlap in certain particular features. Earlier studies on Korean dialects include Ogura (1944); S.-N. Lee *et al.* (1971); Ramsey (1978); Y.-H. Kim (1982); T. K. Kim (1986); Choy (1987); J. R. P. King (1987); Ik-seop Lee *et al.* (1997); K.-G. Lee *et al.* (1998); Sohn (1999); Iksop Lee and Ramsey (2000); Pangen Yenkwuhoy (2001), among many others. The description below is largely based on Sohn (1999) and Iksop Lee and Ramsey (2000), and additional information and data are taken from J. R. P. King (2006), and Pangen Yenkwuhoy (2001). Linguistic divergences between North Korean and South Korean are largely based on Yeon (2006).

## 6.2 REPRESENTATIVE ISOGLOSSES

### 6.2.1 Tones

Tone functioned as a distinctive feature in Middle Korean (MK). In Modern Standard Korean, these tonal distinctions have been lost, and vowel length differences remain as their trace (see 5.2.3.1). However, tone is still distinctive in certain dialects. If tone is taken as a criterion to set up larger dialect divisions, Korea can be divided into two parts: an eastern half and a western half. Tone functions as a distinctive feature in the eastern half of the country consisting of Kyōngsang, Hamkyōng, and the eastern part of Kangwon (Yōngdong), while tonal distinction is not a distinctive feature in the western half. However, vowel length instead functions as a distinctive feature for most parts of the western half.

Just as in MK, tone in the modern dialects consists of an opposition between high and low pitch. However, the distinction mechanism in the modern dialects is not exactly the same as the earlier system. In a Modern Korean dialect, a ‘tone’ in isolation is not absolutely distinctive unlike in Chinese. For example, when pronounced in isolation, the Hamgyōng syllable *pay* can mean either ‘pear’ or ‘belly’. However, if a particle or the copula follows, it reveals the underlying tone of the noun: *pay-NUN* (low-high) means ‘pear-TOP’, while *PAY-nun* (high-low) means ‘belly-TOP’. In other words, the perception of the tone of a syllable depends not on its absolute pitch, but rather its pitch relative to that of a neighboring syllable.

Moreover, there is a clear difference between the tonal system of Hamgyōng and that of Kyōngsang. For example, in the Hamgyōng dialects, ‘head, hair’ is pronounced *meLI* (low-high), while in Kyōngsang the same word is pronounced *MELi* (high-low); Hamgyōng *ciLUM* (LH) ‘oil’ corresponds to Kyōngsang *Cilum* (HL); Hamgyōng *AYki* (HL) ‘baby’ corresponds to Kyōngsang *AYKI* (HH).

Some examples of minimal pairs depending on tone difference are given below (Lee and Ramsey 2000: 317, transcription modified here to Yale):

- (1) Hamgyōng
  - a. *mal(-I)* (LH) ‘horse’/*MAL(-i)* (HL) ‘language; quart’
  - b. *pay(-KA)* (LH) ‘pear’/*PAY(-ka)* (HL) ‘belly’
- (2) (North) Kyōngsang
  - a. *MAL(-i)* (HL) ‘horse’/*MAL(-I)* (HH) ‘quart’ / *MA:L(-i)* (H-LongL) ‘language’
  - b. *PAY(-ka)* (HL) ‘pear’/*PAY(-KA)* (HH) ‘belly’ / *PAY:(-ka)* (H-LongL) ‘double’

Although it is a tendency that tone and vowel length stand in complementary distribution with each other, vowel length can also be found in some of the dialect areas with tones. Most of the dialects of North Kyōngsang have both tone and vowel length. The dialect of the Yōngdong area of Kangwon has both tone and vowel length functioning to distinguish the meanings of words. There are some dialects that have neither tones nor vowel length. Cheju dialect is a representative one that has neither tones nor vowel length, and there are also other such dialects scattered around North Korea.

### 6.2.2 The MK vowel $\text{ㅓ}$ (ㅓ)

The vowel letter  $\text{ㅓ}$  representing the sound [ʌ] is no longer used in modern han'gŭl writing. This vowel has been largely lost in Contemporary Korean. In most of the dialects, it has developed into [a], [u], or [o]. However, Cheju dialect has preserved this vowel as [ʌ], a direct descendant of the MK 'arae a' (ㅓ), which is one unique feature of Cheju dialect. The following are words in Cheju dialect where this vowel can be seen: *tal* (**tal**) 'moon', *tali* (**tali**) 'bridge', *sal* (**sal**) 'flesh', *hata* (**hata**) 'to do', *masal* (**maul**) 'village', *hak* (**hulk**) 'earth', *mancita* (**mancita**) 'to feel', *phali* (**phali**) 'fly', *nongsal* (**nongsa**) 'farming'. MK [ʌ] tends to be preserved mainly in first syllables although a few examples are found in other than the first syllable. The exact phonetic quality of the vowel transcribed here as [ʌ] has been a matter of dispute. Kang (1988) described it as "neither [a] nor [o] nor [ə] but something in between." The presence of the diphthong [jʌ] is an even more unusual and unique feature in Cheju dialect. Examples of [jʌ] are *yatup* (**yetelp**) 'eight' and *yalum* (**yelmay**) 'fruit'.

### 6.2.3 The MK consonant $\text{ㅈ}$ (ㅈ)

The MK consonant  $\text{ㅈ}$  [z] has completely disappeared in Contemporary Korean, but in some dialects the consonant *s* is sometimes found in its place. Some examples are *kasay/kasikay* (**kawi**) 'scissors', *yesei/yasi/yeswu* (**yewu**) 'fox', *kasil/kasul* (**kaul**) 'autumn', and *masil/masul/mosil* (**maul**) 'village'. The following are some examples of words with this consonant in Kyōngsang dialect compared with MK:

(3)	N. Kyōngsang	MK	Seoul
'scissors'	<i>kasay/kasigay</i>	kózáj	<b>kawi</b>
'autumn'	<i>kasil</i>	kózól	<b>kaul</b>
'village'	<i>masul</i>	mózól	<b>maul</b>
'fox'	<i>yasi</i>	jezó	<b>yewu</b>
'turnip'	<i>musu/musi</i>	mwuzwu	<b>muu</b>

There are two broad areas where *s* (corresponding to MK *z*) has been preserved. In the south, it is found in Kyōngsang, Chōlla, and Ch'ungch'ōng, and in the north, in Hamgyōng. Standard *s*-irregular verbs (see 5.4.2.1) such as *is-ta* 'to connect', *ces-ta* 'to stir', *cis-ta* 'to build, make', and *nas-ta* 'to recover' are regular in Kyōngsang, Chōlla, Ch'ungch'ōng, and Hamgyōng provinces as in *is-umyen* (**i-umyen**) 'if [he] connects', *ces-ela* (**ce-ela**) 'stir!', and *nas-ase* (**na-ase**) 'as [he] has recovered'.

### 6.2.4 The MK consonant $\text{ㅂ}$ (ㅂ)

The MK voiced bilabial fricative consonant  $\text{ㅂ}$  [β] has also completely disappeared,<sup>1</sup> but in some dialects [b] (phonemically /p/) is found corresponding to MK  $\text{ㅂ}$ . For example,

MK *saβi* ‘shrimp’ is reflected variously as *saywu* (Central and P’yŏngan dialects), *sayo* and *saypi/saypi* (Hamkyŏng and Chŏlla, Kyŏngsang dialects), *saypayngi* (Ch’ungch’ŏng and southern Kyŏnggi province), etc. In general, these areas overlap with those where *s* is preserved in place of MK *z*. The following are some words where *p* has been preserved in Kyŏngsang dialects:

(4)	N. Kyongsang	MK	Seoul
‘shrimp’	<i>saypi</i>	saβi	<b>saywu</b>
‘silkworm’	<i>nwupey</i>	nwuβej	<b>nwuey</b>
‘cold’	<i>chwupun</i>	chiβŏn	<b>chwu:un</b>
‘pretty’	<i>kopun</i>	koβŏn	<b>kowun</b>

Standard *p*-irregular verbs such as **chwup-** ‘to be cold’, **tep-** ‘to be hot’, **musep-** ‘to be scary’, **kop-** ‘to be pretty’, and **mip-** ‘to be hateful’ are regular verbs in the Kyŏngsang dialects, as in *chwup-ela* (**chwuw-ela**) ‘it’s cold’ and *kop-ase* (**kow-ase**) ‘as [it] is pretty’.

### 6.2.5 Word-medial *k*

Quite a few words manifest alternation between the presence and absence of *k* in word-medial position depending on dialect, as in *pakwu/pangkwi* (**pawi**) ‘rock’, *kaykol/kaykwul* (**kaywul**) ‘brook’, *olkay* (**olhay**) ‘this year’, *tolkaci/tolkay* (**tolaci**) ‘Chinese bellflower’, and *silkeng/sikeng* (**sileng**) ‘wall shelf, rack’. This historical word-medial *k* appears in Kyŏngsang, Hamgyŏng, and part of Chŏlla, while it has been mostly dropped in the central dialects and P’yŏngan province.

### 6.2.6 Word-initial *l* and *n*

Northern dialects including P’yŏngan and Hamkyŏng retain the pronunciation of initial *l* in Sino-Korean words, whereas the rest of the dialects have either lost it (before *i* and *y*) or replaced it with *n* otherwise. The P’yŏngan dialect and Yukchin dialect of North Hamgyŏng province also retain the pronunciation of initial *n* before *i* or *y*, whereas the other dialects have lost it, as in *nima* (**ima**) ‘forehead’, *niphakwu* (**ipsakwi**) ‘leaf’, *ni* (**i**) ‘tooth’, *nilkwup* (**ilkop**) ‘seven’, *nyeca/neca* (**yeca**) ‘woman’, *nelum* (**yelum**) ‘summer’, and *nwuwel* (**yuwel**) ‘June’.

### 6.2.7 Palatalization

Using palatalization as the criterion for classification, the dialects of Korea are divided into three areas. In one area, the dialects undergo no palatalization whatsoever. In the second area, dental consonants (*t*, *th*, and *tt*) are palatalized, but velar consonants (*k*, *kh*, *kk*, and *h*) are not. In the third area, both of these two consonant types are palatalized.

The palatalization of *t*, *th*, and *tt* refers to the change of these dental stops to *c*, *ch*, and *cc* when the consonants occurred before *i* or *y*. Most of the dialects of Korea, including standard Seoul speech, underwent this change quite early after the seventeenth century. However, the P’yŏngan dialects in the northwest have not undergone this change. As a result, the unpalatalized forms are the most noticeable and typical characteristic of these North Korean dialects. *t*-palatalization began in the southern dialects and gradually spread to the north, affecting nearly all dialects except P’yŏngan and the far northeast Yukchin dialect. Here are some examples that did not undergo palatalization in the P’yŏngan dialect: *thita* (**chita**) ‘to hit’, *ttiluta* (**cciluta**) ‘to thrust’, *hetiman* (**haciman**)

‘but’, *kathi* (**kathi** [katʰi]) ‘together’, *kwuti* (**kwuti** [kudʒi]) ‘firmly’, and *kwuthita* (**kwuthita** [kutʰida]) ‘to harden’.

The term “*k*-palatalization” refers to the change of the velars *k*, *kh*, *kk* to *c*, *ch*, *cc* when the consonants occur before *i* or *y*. Some representative examples are as follows: *cil* (**kil**) ‘road’, *cilum* (**kilum**) ‘oil’, *citong* (**kitong**) ‘pillar’, *cwul* (**kyul**) ‘an orange’, *cilta* (**kilta**) ‘to be long’, *cyenwuta* (**kyenwuta**) ‘to take aim’, and *cciwuta* (**kkiwuta**) ‘to insert’. The area where *k*-palatalization occurred includes Kyōngsang, Chōlla, Ch’ungch’ōng, Cheju, the eastern part of Kangwon (Yōngdong) in the south, and Hamgyōng in the north. In this area, the velar fricative *h* can also be palatalized (*h*-palatalization), as in *sim* (**him**) ‘strength’, *seng* (**hyeng**) ‘older brother’, and *sungnyen* (**hyungnyen**) ‘bad crop year’.

### 6.2.8 Umlaut

In the South Hamgyōng dialect there is a productive umlaut system in which *a*, *e*, *u*, *o*, *wu* are fronted to *ay*, *ey*, *i*, *oy*, *wi* ([ɛ], [e], [i], [ø], [y]), respectively when they are immediately followed by a non-coronal consonant plus a high, front, non-consonantal segment. Here are some examples where fronting due to umlaut has become lexicalized: *eymi* (**emi**) ‘mother’, *koyki* (**koki**) ‘meat’, *payppwiki* (**paykkop**) ‘navel’, *nwipi* (**nwui**) ‘sister’, and *acwimi* (**acwumi**) ‘aunt’. Umlaut does not occur if the interceding consonant is a coronal: *kaci* ‘eggplant’, *poli* ‘barley’, *keli* ‘street’, *wuli* ‘cage’. Umlaut also applies quite frequently to Kyōngsang dialects as well.

### 6.2.9 Inflection of *t*-irregular verbs

In standard Korean, certain verb stems end in *-t* before a consonant but *-l* [ɾ] before a vowel; for example, **tut-ko** ‘hear and . . .’, **tul-umyen** ‘if [you] hear’. However, in certain dialects, a *t*-irregular verb like *tut-* ‘to hear’ can occasionally have the regular inflection of verbs. In P’yōngan dialects, for example, although most *t*-irregular verbs have the same irregularities as the standard language, the one verb *tut-* ‘to hear’ is regularly inflected as following: *tut-ko*, *tut-uni*, *tut-eto*, . . .

### 6.2.10 Phoneme inventory

As far as consonants are concerned, almost every dialect has an identical set of phonemes. The only exception to this generalization is that in Kyōngsang there is no contrast between *s* and *ss*. In some parts of Kyōngsang dialect, the words **sal** ‘flesh, meat’ and **ssal** ‘rice’ cannot be distinguished in pronunciation, nor can **sata** ‘to buy’ and **ssata** ‘to be cheap’. However, considerable differences are found in vowel inventory.

In the Kyōngsang dialects there is no contrast between *e* [ə] and *u* [i]. Thus, the words **thul** ‘frame’ and **thel** ‘hair’ are both pronounced the same in these dialects. The Kyōngsang pronunciation of **sungca** ‘winner’ and **sengca** ‘saint’ is heard as identical by Seoul standard speakers. As speakers of Kyōngsang dialects have merged *e* and *u* into one phoneme, the phonemic inventory has one fewer vowel than other dialects.

Many dialects do not distinguish *ay* [ɛ] from *ey* [e]. The area where this merger occurs is also Kyōngsang. In these dialects, the vowel that has resulted from the merger is usually neither [ɛ] nor [e], but rather somewhere between the two sounds. Speakers of Kyōngsang dialects do not distinguish between **kay** ‘dog’ and **key** ‘crab’ or between **nay** ‘my’ and **ney** ‘your’.

In many dialects, the vowels *oy* [ø] and *wi* [y] are not pronounced as monophthongs. The standard pronunciation of *oy* may be stipulated as [ø], but the actual pronunciation of the vowel in most dialects tends to be the diphthong [we] rather than a monophthong. Nevertheless, there are areas such as Ch'ungch'ōng, Chōlla, Hwanghae, and eastern Kangwon (Yōngdong) where the vowels *oy* and *wi* are pronounced as the monophthongs.

### 6.3 CHARACTERISTICS OF SIX DIFFERENT DIALECTS

#### 6.3.1 Northwest (P'yōngan) dialects

P'yōngan dialects have eight simple vowels *i* [i], *ey* [e], *ay* [ɛ], *u* [i], *e* [ə], *a* [a], *wu* [u], *o* [o]. P'yōngan dialects are virtually unique in that they never underwent the change of palatalization which took place in the south and swept over most of Korea in the eighteenth century. The Yukchin dialects in northernmost Hamkyōng are also free from palatalization. Thus, P'yōngan dialects show neither the *t*-palatalization nor the *k*- and *h*-palatalizations prevalent in southern and northeastern dialects.

In these dialects, unlike standard Korean and other dialects, *n* occurs before *i* and *y* in the word-initial position, as in *ni* (**i**) 'tooth', *niwus* (**iwus**) 'neighbor', and *nima* (**ima**) 'forehead', etc. P'yōngan dialects are also peculiar in that they retain the historical intervocalic *k*, as in *silkeng* (**sileng**) 'wall shelf', *naykwuli* (**nay**) 'smoke', and *molka* (**molay**) 'sand'.

In addition to the subject marker *-i*, P'yōngan dialects have another subject particle *-lay/ley*. *-i* is used after a consonant and *-lay/ley* is used after a vowel. For example, *nay-ley kulaysiyo* (the closest semantic equivalent in standard Korean being **nay-ka kulaysseyo**) 'I did it', *nay-ley katuleytiyo* (**nay-ka kassessciyo**) 'I went', *ku nengkam-i na-lkwa tongkap iti* (**ku yengkam-i na-wa tongkap ici**) 'He is the same age as me'. *-lay/ley* is a form peculiar to P'yōngan. The verb ending *-tuleyt-* marks the "remote past tense" and corresponds to **-ess.ess-** in standard Korean. The final ending *-tiyo* beginning with a *-t-* illustrates that these dialects did not undergo *t*-palatalization.

Many idiosyncratic sentence enders that are partly shared by the Hamgyōng dialects are:

- (5) Declarative: *-(u)m-ney/mey*, *-(u)wey* (functionally equivalent to standard **-e**, **-ney**, **-uo/so**)  
*-swu-ta*, *-(su)p-ney-ta*, *-(su)p-mey-ta*, *-(u)wa-yo*, *-(u)p-ti*, *-(u)wey-ta*  
**(-eyo, -(su)pnita)**
- Interrogative: *-(u)m-mey?*, *-wu?*, *-kan?* (**-e?**, **-na?**, **-uo/so?**)  
*-(su)p-ney-kka?*, *-(su)p-mey-kka?* (**-(su)pnikka?**)
- Imperative: *-(u)si-tana*, *-(u)m-mey* (**-key**, **-sey**, **-uo/so**)  
*-(u)si-la-yo*, *-(u)si-kyo*, *-(u)si-p-syo* (**-(u)seyyo**, **-(u)sipsio**)

The following are representative sentence examples (the sentence examples throughout this section are predominantly from Sohn (1999) and Lee and Ramsey (2000), whose examples were originally taken from Ik-seop Lee *et al.* (1997).)

- (6) a. *etumey ka-si-p-ney-kka*  
 where.to go-HON-POL-IND-Q  
 'Where are you going?'  
**(eti-ey ka-si-pnikka)**

- b. *na ampulla an ka-mun ekha-kan*  
 I even not go-if how.do-Q  
 'If even I don't go, what could we do?'  
 (**na-cocha an ka-myen etheh-key ha-keyss-na**)
- c. *onel ka-ss-ta o-kas-swu-ta*  
 today go-PAST-and come-will-POL-DEC  
 'I will go there and come back today.'  
 (**onul ka-ss-ta o-keyss-supnita**)

Like some other Korean dialects, P'yŏngan dialects have just three speech levels: low, equal, and respect. As in many dialects, P'yŏngan dialects have many unique lexemes. For example: *chanpap* (vs standard **cemsim**) 'lunch', *nephchakay* (**hocwumeni**) 'pocket', *peylkwuti* (**pyelwuk**) 'flea', *wuthi* (**os**) 'clothes', *elkheni* (**menilka**) 'distant relative', *mulwu* (**wupak**) 'hail', *omani* (**emeni**) 'mother', *eyminey* (**anay, yeca**) 'wife, woman', *punthu* (**aytul sin**) 'kid's shoes', *ssata* (**pissata**) 'to be expensive', *nwukta* (**ssata**) 'to be cheap', *kaykata* (**kacyekata**) 'to take along', *hanang* (**hamkkey**) 'together', *wuteng* (**ilpule**) 'on purpose', *sepsepi* (**hepha**) 'lung', *yel* (**ssulkay**) 'gallbladder'.

### 6.3.2 Northeast (Hamgyŏng) dialects

A major feature of Hamgyŏng dialects, which they share with Kyŏngsang and Yŏngdong (east of Kangwon) dialects, is distinctive pitch accent (tone). Hamgyŏng dialects are usually described as lacking distinctive vowel length.

One phonological characteristic of Hamgyŏng dialects is the weakening of [n] and [ŋ] when these nasal consonants follow a vowel and precede *i* or *y*. When this weakening occurs, the preceding vowel is nasalized, and in place of [n] there is a constriction at the glottis. A stereotypical feature of Hamgyŏng dialects is the local pronunciation of the negative morpheme **ani**, which becomes [ã'i] as a result of this (Lee and Ramsey 2000: 332).

Hamgyŏng dialects also retain the historical word-medial or intervocalic  $\beta z k$  as *p s k*, respectively, as in *nwupey* (**nwuey**) 'silkworm', *hapulaypi* (**holapi**) 'widower', *kasay* (**kawi**) 'scissors', *mosi* (**moi**) 'feed (for chickens)', *mol kay* (**molay**) 'sand', and *nolki* (**noIwu**) 'roe deer'.

Hamgyŏng dialects preserve a number of alternations in both nouns and verbs which can be traced back to MK. For example, **namu** (MK *namk-*) 'tree': *nayngki* (subject), *nangku* (object; possessive), *nangkey* (dative), *nangkullu* (instrumental), *nangkumu* (topic), *namuka* (comitative). A characteristic of Hamgyŏng dialects is the use of *-u/-lu* as the accusative marker instead of **-ul/-lul** in standard Korean. In this case the final consonant of the particle has been weakened.

The representative sentence enders characteristic to Hamgyŏng dialects are:

- (7) Declarative: *-(u)m-mey, -(u)m-ney, -(u)cipi, -(u)way (-e, -ney, -uo/so)*  
*-ota, -wuta, -weta, -op-cipi, -(u)m-mey-ta (-eyo, -(su)pnita)*
- Interrogative: *-wu?, -(u)m-mey?, -cipi?, -m-nungka?, -(u)m-twu(ng)? (-e?, -na?, -uo/so?)*  
*-sswu-ta?, -(u)m-mengi?, -(u)p-syo?, -(u)p-mi-kka? (-eyo?, -(su)pnikka?)*
- Imperative: *-(u)p-sey, -(u)p-so, -cipi (-e, -key, -sey, -uo/so)*  
*-sosey, -wuta, -si-p-so (-eyo, -(u)sipsio)*
- Propositive: *-op-sey, -op-ci(pi) (-(u)sipsita)*

Representative sentence examples are as follows:

- (8) a. *polipap-pokwu-sa nas-cipi*  
 barley.rice-than-EMPH better-DEC  
 ‘It is certainly better than barley rice.’  
 (**polipap-pota-ya nas-ci**)
- b. *tep-untey kule-ng ke ip-hiwu-ci ma-wuta*  
 hot-as that-MOD thing wear-CAUS-CON stop-IMP  
 ‘Don’t make him wear such a thing as it is hot.’  
 (**tew-untey kule-n kes ip-hi-ci ma-seyyo**)

Hamgyōng dialect vocabulary examples are as follows: *meykwuli* (vs standard **kykwuli**), ‘frog’, *yekki/eykki/yengkki* (**yewu**) ‘fox’, *phusungkay* (**hepha**) ‘lung’, *kwusay* (**kwulttuk**) ‘chimney’, *esi* (**pumo**) ‘parents’, *kasieypi* (**cangin**) ‘(male’s) father-in-law’, *tongsami* (**kyewul**) ‘winter’, *solay* (**tayya**) ‘wash basin’, *anakkan* (**anay**) ‘wife’, *haym* (**panchan**) ‘side dishes’, *kayngkay* (**kamca**) ‘potato’, *nwuli* (**wupak**) ‘hail’, *twuley* (**tul**) ‘field’, *misikkan* (**oyyangkan**) ‘stable’, *naco* (**cenyek**) ‘evening’, *polthi* (**acwu**) ‘very’.

### 6.3.3 Central dialects

The central dialects contain the standard speech of Korean as a subset. Although Seoul speech is considered the standard speech, the dialects outside of Seoul tend to show a few salient phonological characteristics. First, in colloquial speech, the vowel *o* is frequently raised to *wu* in final syllables of certain native morphemes, especially when it occurs as part of a suffix, as in **kuliko** > *kulkwu* ‘and’, **mek-eto** > *mek-etwu* ‘eat but’, **na-to** > *na-twu* ‘I also’, **sikol** > *sikwul* ‘countryside’, and **ka-ss-o?** > *ka-ss-wu?* ‘Has (he) gone?’. Second, palatalization is not as prevalent as in the southern dialects, but not as weak as in the P’yōngan dialects. The central dialects have undergone *t*-palatalization, but neither *k*- nor *h*-palatalization. Palatalization is more widespread in Ch’ungch’ōng province and grows stronger as one moves further south as it is closer to Kyōngsang dialects. Third, the historical word-medial (or intervocalic)  $\beta$   $\approx$  *k* disappeared completely or weakened while some other dialects have retained them as *p* *s* *k*, respectively.

Technically speaking, the central dialects include at least three sub-dialects, corresponding to Kyōnggi, Kangwon, and Ch’ungch’ōng provinces.<sup>2</sup> Some linguists include parts of Hwanghae province, too.

Kyōnggi and Ch’ungch’ōng dialects have distinctive vowel length, as do most areas of Kangwon, but the Yōngdong dialects (east of Kangwon) also have pitch accent (High and Low tones). Some relatively idiosyncratic aspects of the Ch’ungch’ōng dialects are slow tempo of speech<sup>3</sup> and general retention of the historical intervocalic *z* as *s*, as in *masil* (**maul**) ‘village’, *asi* (**awu**) ‘younger brother’, *is-e* (**i-e**) ‘connect and’, and *muswu* (**muu**) ‘turnip’. The most stereotyped grammatical feature of Ch’ungch’ōng dialects is the use of the sentence ending *-yu*, corresponding to standard **-yo**. Some idiosyncratic sentence enders in Ch’ungch’ōng dialects are:

- (9) Declarative: *-ta-ya* (**-ta**), *-swu* (**-uo/so**), *-eyu* (**-eyo**)  
 Interrogative: *-wu?* (**-uo/so?**), *-nam?* (**-na?**), *-eyu?* (**-eyo?**), *-sup-ni-kkya?*  
 (**-supnikka?**)  
 Imperative: *-keyna* (**-key**), *-wu* (**-uo/so**), *-eyu* (**-eyo**)  
 Propositive: *-ciyu* (**-ciyo**)

Ch'ungch'ŏng dialect vocabulary examples are: *tochi* (vs standard **tokki**) 'axe', *momci/mondayki* (**menci**) 'dust', *kengkeni* (**panchan**) 'side dishes', *pokumchi/pokwuni/pokwuli* (**pakwuni**) 'basket', *elleng* (**ellun**) 'quickly', *sipang* (**cikum**) 'now', *ei* (**ese**) 'right away', *naytwung* (**yethay-kkaci**) 'until now', *kecin/kecipan* (**keuy**) 'almost', *talputa* (**taluta**) 'to be different', *mikkalmacta* (**mipta**) 'to be hateful', *swukkumhata* (**coyonghata**) 'to be quiet'.

### 6.3.4 Southwest (Chŏlla) dialects

Chŏlla dialects have nine vowels as a result of merging *ey* [e] and *ay* [ɛ]. Some salient phonological features characteristic to Chŏlla dialects are as follows: first, palatalization is widespread including *k*- and *h*-palatalization, as in *cimchi* (**kimchi**) 'kimchee', *cil* (**kil**) 'road', *cim* (**kim**) 'steam, seaweed', *sey* (**hye**) 'tongue', and *seng* (**hyeng**) 'older brother'. Second, vowel fronting or raising frequently occurs, as in *kwusil* (**kwusul**) 'bead', *kasim* (**kasum**) 'chest', *kali* (**kalwu**) 'powder', *kochi* (**kochwu**) 'red pepper', *chimey* (**chima**) 'skirt', *ki* (**key**) 'crab', *pita* (**peyta**) 'to cut', *pikey-lul pita* (**peykey-lul peyta**) 'to rest one's head on a pillow'. Umlaut also occurs frequently, for example: *seymil* (**samil**) 'three days', *keylita* (**kalita**) 'to hide', *soyk-i* (**sok-i**) 'inside-NOM', *peyp-i* (**pap-i**) 'rice-NOM'. Third, Chŏlla dialects tend to exhibit extensive monophthongization of standard diphthongs, as in *ppey* (**ppye**) 'bone', *ppam* (**ppyam**) 'cheek', *sengnang* (**sengnyang**) 'matches', *kwusin* (**kwisin**) 'ghost', *peyl* (**pyel**) 'star', *pey* (**pye**) 'unhusked rice'. Fourth, these dialects retain the historical word-medial *z* and *k* as *s* and *k*, respectively, in many words, as in *kasey/kasikey* (**kawi**) 'scissors', *kasim* (**kam**) 'material', *masil* (**maul**) 'village', *tolkaci* (**tolaci**) 'Chinese bellflower', *pakwu* (**pawi**) 'rock', *nangkwu* (**namu**) 'tree', *pelkeci* (**peleci, pelley**) 'worm'.

Chŏlla dialects have the following interesting pronouns: *nuku* (vs **nehuy**) 'you-plural', *cuku* (**ku-tul, caki-tul**) 'they', *caku* (**caki**) 'self'. A productive adverbial suffix in Chŏlla dialects is *-heni*, which corresponds to **-key** in standard Korean, as in *kkeykkasheni* (**kkaykkushakey**) 'cleanly', *mianheni* (**mianhakey**) 'regretfully', and *nwule(he)ni* (**nwulehkey**) 'in yellow'. Other typical Chŏlla shapes are *-meyngilo* (vs **-chelem**) 'like', *-kanti* (**-killay**) 'as, since', and *-ullako* (**-ulyeko**) 'intending'.

A typical form characteristic to Chŏlla dialects is *-ttamsi* or *-ttamse* meaning 'because of'. Hearing this form, most Koreans would associate it with Chŏlla dialect. Another characteristic is that *-lau* is used as a polite-style ending instead of **-eyo**. Other conjunctive suffixes somewhat unique to Chŏlla dialects include *-(u)ngkkey/ningkkey* (vs **-(u)nikka**) 'because' as in *ka-ngkkey* (**ka-nikka**) 'because [he] goes' and *kuleng-kkey/kung-kkey* (**kule-nikka**) 'thus, therefore'; *-(u)msilong* (vs **-(u)myense**) 'while doing/being' as in *coh-umsilong* (**coh-umyense**) 'while [he] feels happy'. Also, the pre-final ending *-ke/-kye* serves as the honorific marker instead of **-si**, as in *ka-kye-lau* (**ka-seyyo**) 'goes', *ka-kye-ssnya?* (**ka-sy-essnunya?**) 'Did (someone) go?', *kumseyyo o-kye-lau?* (**pelsse o-sey-yo?**) 'Is [someone] coming already?'.

Sentence enders unique to Chŏlla dialects are as follows:

- (10) Declarative: *-eya* (**-ta**), *-elawu* (**-eyo**), *-lsey*, *-si* (**-ney**), *-ye* (**-ya**)  
 Interrogative: *-eya?*, *-nya?* (**-ni?**, **-(n)unya?**), *-elawu?* (**-eyo?**), *-(su)p-ni-kkye?*  
 (**-(su)pnikka?**)  
 Imperative: *-so* (**-key**)  
 Propositive: *-tulako* (**-sey**)

In addition, the form *-ing* [-in], with question intonation, is often used at the end of a sentence as a confirmation seeker ('... isn't that right?') which is more or less comparable

to standard **ung**, but [-ij] can be used with polite or formal speech styles. Since standard **ung** can only be used in plain or panmal conversation style, the Chölla form [-ij] in more polite contexts is often misunderstood by Seoulites as rude and impolite. The following are representative sentence examples:

- (11) a. *me ttamsi os-ul ip-ess-elau*  
 what because.of clothes-ACC wear-PAST-Q  
 ‘Why have you put on clothing?’  
 (**mo ttaymun-ey os-ul ip-ess-eyo**)
- b. *kumsey-po o-kye-lau-ij*  
 already come-HON-Q-TAG  
 ‘You are here already, aren’t you?’  
 (**pelsse o-sey-yo**)
- c. *inca ka po-tulako-ij* (K.-G. Lee *et al.* (1998))  
 now go see-let’s-TAG  
 ‘Let’s go now.’  
 (**icey ka po-sey/po-siciyo**)

The following are some words characteristic to Chölla dialects: *neyngkal* (vs **yenki**) ‘smoke’, *koypi* (**hocwumeni**) ‘pocket’, *cichen* (**kkwucilam**) ‘scolding’, *nucakwu* (**canglaysiaeng**) ‘a future’, *icengsulepta* (**kkomkkomhata**) ‘to be meticulous’, *pungkepta* (**twukkepta**) ‘to be thick’, *photasi* (**kyewu**) ‘barely’, *teypte-ylo* (**tolie**) ‘on the contrary’, *molleyngi* (**kkoktayki**) ‘tip’, *phaykkakcil* (**ttalkkwukcil**) ‘hiccup’.

### 6.3.5 Southeast (Kyöngsang) dialects

Kyöngsang dialects have just six vowels: *i* [i], *e* [ə], *ey* [e], *a* [a], *wu* [u], *o* [o]. In these dialects, standard Korean [e] and [ɛ] have merged as [e], and standard *u* [i] and *e* [ə] have merged as *e* [ə]. Kyöngsang dialects do not distinguish between lax *s* and tense *ss*. These dialects have undergone many simplifications in the course of their evolution. Thus, semivowels tend to disappear after a consonant, as in *pho* (**phyo**) ‘ticket’, *saka* (**sakwa**) ‘apple’, *ki* (**kwi**) ‘ear’, *haksilhi* (**hwaksilhi**) ‘surely’, *munha* (**munhwa**) ‘culture’, *keni* (**kwenwi**) ‘authority’, and *kakang* (**kwankwang**) ‘sightseeing’.

The Kyöngsang dialects form the other principal area together with Hamgyöng where the distinctions of MK tones have been preserved. These dialects resemble the Hamgyöng dialects in that (most) tonal patterns depend on an accent locus characterized by a high pitch.

Another salient feature is umlaut. Umlaut applies quite freely in Kyöngsang dialects. *i* or *y* fronts preceding *wu*, *e*, *o*, *a* unless blocked by an intervening coronal consonant, as in *pangmeyngi* (**pangmangi**) ‘mallet’, *eymi* (**emi**) ‘mother’, *keyki* (**koki**) ‘meat’, *kameyni* (**kamani**) ‘bag’.

Kyöngsang dialects preserve MK *z* as *s*, and MK *β* as *p*. For example: *chwupi/chipi* (**chwuwi**) ‘cold’, *tepi* (**tewi**) ‘heat’, *kapuntey* (**kawuntey**) ‘middle’.

Palatalization is widespread: indeed this region is known as the epicenter for *k*-palatalization. Examples are *citwung* (**kitwung**) ‘pillar’, *cil* (**kil**) ‘road’. A typical example of *h*-palatalization is *swungnyen* (**hyungnyen**) ‘bad harvest’.

In this dialect, we can witness extensive contractions occurring in colloquial speech. Some sentence examples are *eps-sim-te* (**eps-supnita**) ‘(I) don’t have (it)’, *me-la kha-no?* (**mwes-i-la ko ha-ni?**) ‘What are you talking about?’, and *wa ikha-no?* (**way ileh-key ha-ni?**) ‘Why do you do (it) this way?’.

Representative sentence enders characteristic of Kyōngsang dialects are as follows:

- (12) Declarative: *-e/a yey (-eyo)*, *-(si)m-te*, *-si-te*, *-ni-te (-su)pnita*  
 Interrogative: *-no (-ni)*, *-neng-kyo?* (*-(n)unkayo?*), *-((si)p)-ni-kkye?*  
*(-su)pnikka?*  
 Imperative: *-si-i-so (-u)sipsio*  
 Propositive: *-ip-si-te (-u)sipsita*

The following are representative sentence examples of Kyōngsang dialects:

- (13) a. *cip-ey iss-nen twung eps-nen twung mol-si-te*  
 home-LOC be-MOD if be.not-MOD if not.know-HON-DEC  
 ‘I don’t know whether (he) is at home or not.’  
**(cip-ey iss-nun-ci eps-nun-ci molu-keyss-supnita)**
- b. *phettek o-si-i-so*  
 quickly come-HON-POL-DEC  
 ‘Please come quickly.’  
**(ppali o-si-psi)**
- c. *ni-khang ney-khang talm-ess-cey*  
 you-and I-and resemble-PAST-TAG  
 ‘You and I are alike, aren’t we?’  
**(ne-lang na-lang talm-ass-ci)**
- d. *atel-manchilo wa kela-no*  
 children-like why do-Q  
 ‘Why are [you] acting like children?’  
**(aytul-chelem way kule-ni)**

Some of the vocabulary unique to Kyōngsang dialects is as follows: *kapuntali (cintuki)* ‘quietly, in earnest’, *sikkepheysssta (nollassta)* ‘was surprised’, *melakhwunta (kkwucicta)* ‘to scold’, *hamo (am)* ‘of course’, *ppucileyki (pusuleki)* ‘fragments’, *tongkeyta (phokayta)* ‘to overlap’, *hwupita/totikkhita (hwumchita)* ‘to steal’, *hama (pelsse)* ‘already’, *kkapchita (caychokhata)* ‘to urge, hurry’.

### 6.3.6 Cheju dialects

Cheju dialects have nine simple vowels [i, e, ε, i, ə, a, u, o, ʌ] and 13 diphthongs [je, jε, jə, ja, ju, jo, jʌ, wi, we, wε, wə, wa, ii]. The most salient feature of the vowel system is the existence of the vowel [ʌ], a direct descendant of the MK *ó*.

Umlaut is not a characteristic of Cheju dialects, but many nouns ending in a vowel show traces of the addition of an *-i*, which has fronted the original vowel, as in *kotungey (kotunge)* ‘mackerel’, *kamcwi (kamcwu)* ‘sweet rice wine’, *nwuey (no)* ‘oar’. The palatalization behavior for Cheju dialects is similar to those of other southern dialects. Cheju dialects have experienced *t*-palatalization and have also undergone *k*- and *h*-palatalization as in *cilta (kilta)* ‘to be long’, *sey (hye)* ‘tongue’.

Some fossilized suffixes peculiar to Cheju dialects are *-ang*, *-(ay)ngi*, and *-ayki*, as observed in *patang (vs pata)* ‘sea’, *halwupang (halapeci)* ‘grandfather, old man’, *apang (apeci)* ‘father’, *sayngi (say)* ‘bird’, *keyngi (key)* ‘crab’, *kangsayngi (kangaci)* ‘puppy’, *cokayngi (cokay)* ‘shell’, *talksayngi (talkyal)* ‘egg’, *songayngi (songaci)* ‘calf’, and *paksayngi (pakaci)* ‘gourd dipper’.

There are numerous sentence enders unique to Cheju dialects:

- (14) Declarative    *-(u)khiye, -em-ce, -em-se, -em-chwu, -khwu-ta (-ta, -e, -ney, -uo/so)*  
                           *-em-swu-ta, -(s)wu-ta (-eyo, -(su)pnita)*
- Interrogative:    *-em-ti(ya)?, -em-sini?, -esinya? (-ni?, -(n)unya?)*  
                           *-m-kka?, -m-kko?, -em-se?, -em-singa? (-e?, -na?, -uo/so?)*  
                           *-em-swu-kkwa?, -(wu)kkwa? (-eyo?, -(su)pnikka?)*
- Imperative:        *-(u)p-se, -(u)p-ce, -(u)sim, -cwu (-ela, -key, -e, -uo/so, -eyo, -(u)sipsio)*
- Propositive:      *-(u)p-se, -(u)p-ce, -(u)sim, -cwu (-ca, -sey, -e, -eyo, -(u)sipsita)*

The following sentences are typical examples from Cheju dialects:

- (15) a. *etu-ley ka-m-swu-kkwa*  
           where-to go-IND-HON-Q  
           ‘Where are you going?’  
           **(eti-lo ka-si-pnikka)**
- b. *ka-tang mul-eng ka-khwu-ta*  
           go-while ask-and go-will-DEC  
           ‘I will go asking around.’  
           **(ka-taka mul-ese-ka-keyss-o)**
- c. *na-yeng hanti sala-m-ce*  
           I-with together live-IND-DEC  
           ‘(He) is living with me.’  
           **(na-lang hamkkey sal-ko iss-ta)**

There are also many words unique to Cheju dialects that contain many old forms. Examples are *puay* (vs **hepha**) ‘lung’, *kwulmay* (**kulimca**) ‘shadow’, *keyyemci* (**kaymi**) ‘ant’, *taysani* (**manul**) ‘garlic’, *cisay* (**kiwa**) ‘roof tile’, *wuley* (**wulthali**) ‘fence’, *cengcey* (**puekh**) ‘kitchen’, *tochi* (**tokki**) ‘ax’, *pipali* (**kyeycip.ay**) ‘girl’, *simpang* (**mutang**) ‘shaman’, *namphi* (**muu**) ‘turnip’, *kkwang* (**ppye**) ‘bone’, *pappuli/papchwuli/pammeli* (**camcali**) ‘dragonfly’, *sayngi/chopsayngi* (**chamsay**) ‘sparrow’, *hwangkoci* (**mucikay**) ‘rainbow’, *tokwayngi* (**hoiyoli palam**) ‘whirlwind’, *montok* (**menci**) ‘dust’, etc.

#### 6.4 SOVIET KOREAN YUKCHIN DIALECT

In addition to six different dialectal groups, it is worth mentioning that there is another peculiar dialect outside of the Korean peninsula. As briefly mentioned at the beginning, the language spoken by Koreans in the former Soviet Union, i.e. in Kazakhstan and Uzbekistan, has evolved from the Yukchin dialect that is part of the Hamgyŏng dialects, but it has many characteristic features distinct from current Hamgyŏng dialects. The Koreans in Central Asia are those who originally lived in the Russian Far East and were forcefully deported in 1937 by Stalin. The first Korean emigrants to the Russian Far East were mainly from the Yukchin area of North Hamgyŏng province, and they have been isolated from Korea for approximately a hundred years, preserving many archaic as well as peculiar forms. Koreans in Central Asia call their language ‘Kolyemal’ [Koryŏmar], thus we adopt this term to refer to Soviet Korean Yukchin dialect. The following are representative characteristics of this dialect (R. King and Yeon 1992):

Standard Korean *l* is pronounced as a tap, or as a rolled [r] in all positions of Kolyemal, except before another *l*, as in *ppalli* (**ppalli**) ‘quickly’, *targari* (**talkyal**) ‘egg’, *harmi* (**halmi**) ‘grandmother’, *murkkoki* (**mulkoki**) ‘fish’, *cerpani* (**celpan**) ‘half’. Kolyemal

has a pitch accent system, composed of High and Low pitches, as in *SURi* ‘spoon’ vs *suRI* ‘alcohol’, *maRI* ‘horse’ vs *MAri* ‘words/speech’.

In Kolyemal, *p*-irregular verbs and *s*-irregular verbs conjugate like regular verbs. In other words, intervocalic *-p-* and *-s-* are retained as the result of this dialect never undergoing lenition, as in *tep-ese* (**tew-ese**) ‘is hot and’, *kop-un* (**kow-un**) ‘pretty’, *cis-umu* (**ci-umyen**) ‘if [X] builds’, *pus-ese* (**pu-ese**) ‘pour and so’.

This dialect retains MK pronunciations /su, cu, chu/, whereas they are pronounced /si, ci, chi/ in standard Korean, as in *surehata* (**silhehata**, cf. MK *sulhóta*) ‘not to like’, *sucip kata* (**sicip kata**) ‘to get married’, *surita* (**silita**) ‘to be painfully cold’, *sukkemaycita* (**sikkemaycita**) ‘to become black’, *curta* (**cilta**) ‘to be muddy’, *achum* (**achim**) ‘morning’.

Kolyemal preserves MK *n* as [n] before *i* or *y*, as in *ni* (**i**) ‘tooth; louse’, *nemmari* (**veys-mal**) ‘old story; folk tale’, *nyayki* (**iyaki**) ‘story’.

Kolyemal frequently deletes *l* before coronal sounds, as in *kita* (**kilta**) ‘to be long’, *twungguta* (**twungkulta**) ‘to be round’, *tumuta* (**tumulta**) ‘to be rare’, *saci ana* (**salci anha**) ‘don’t live’.

The most salient feature in case-marking in Kolyemal is the absence of nominative marker *-ka*. There is only one nominative marker *-i*, which is also added to nouns in citation form, as in *kwurumi* (**kwulum**) ‘cloud’, *moki* (**mok**) ‘neck’, *napi* (**nep**) ‘lead’. Other case-markers used in this dialect are as follows: accusative *-u/-ru* (**-ul/lul**), instrumental *-ri/-illi* (**-(u)lo**), dative-locative *-ey* (**-ey**) for inanimates, and *-(u)key* (**-eykey**) for animates, ablative *-eyse* (**-eyse**) for inanimates, and *-(u)keyse* (**-eykeyse**) for animates.

Kolyemal preserves MK *k*-irregular nouns, as in *namwu* ~ *nangk-i* (< MK *namo* ~ *namk-i*) ‘tree(-NOM)’, *kwuna* ~ *kwungk-i* (< MK *kwumwu* ~ *kwumk-i*) ‘hole(-NOM)’, *karki* ~ *karkwu* (< MK *kóló* ~ *kólh-i*) ‘powder’.

The following are representative sentence enders unique to Kolyemal:

- |                              |  |
|------------------------------|--|
| (16) Declarative:            | <i>-kkuma</i> ( <b>-(su)pnita</b> ), <i>-(u)o</i> ( <b>-eyo</b> )                  |
| Declarative retrospective:   | <i>-(u)pti-kkuma</i> ( <b>-(su)ptita</b> ), <i>-(u)ptey</i> ( <b>-teyyo</b> )      |
| Interrogative:               | <i>-(u)mtu?</i> ( <b>-(su)pnikka?</b> ), <i>-(u)o?</i> ( <b>-eyo?</b> )            |
| Interrogative retrospective: | <i>-(u)pti-mtu?</i> ( <b>-(su)ptikka?</b> ), <i>-(u)ptey?</i> ( <b>-tenkayo?</b> ) |
| Imperative:                  | <i>-(u)pso</i> ( <b>-(u)sipsio</b> ), <i>-(u)o</i> ( <b>-eyo</b> )                 |
| Propositive:                 | <i>-keypso</i> ( <b>-(u)sipsita</b> ), <i>-kio</i> ( <b>-eyo</b> )                 |

There are many words peculiar to Kolyemal, e.g.: *papputa* (vs **elyepta**) ‘to be difficult’, *thikhata* (**cinhata**) ‘[tea] to be strong’, *say natta* (**cichita**) ‘to be exhausted’, *khun payri* (**khun changca**) ‘large intestine’, *hapulluse* (**honcase**) ‘alone’, *meyceyki* (**engmang**) ‘mess’, *picekhata* (**pisushata**) ‘to be similar’, *pencinta* (**palumhata**) ‘to pronounce’, *phang* (**acwu**) ‘very much’, *nupi* (**nwui**) ‘sister’, *epang* (**taypupun**) ‘most’, *phayki* (**ttalkkwukcil**) ‘hiccup’, *karu pota* (**hulkye pota**) ‘to look hatefully’, *heri* (**swipkey**) ‘easily’.

## 6.5 LINGUISTIC DIVERGENCES BETWEEN NORTH AND SOUTH KOREAN

There is a large linguistic gap between South and North Korea due to a series of systematic innovations introduced into the North Korean language, motivated to a great extent by political as well as ideological considerations. There are many noticeable differences between North and South Korean at present, but for a matter of space, only the most outstanding ones will be mentioned in this section.

1. *Alphabetical order*: First of all, most people would be surprised to find that dictionary alphabetical order is different. For example, in a North Korean dictionary, words beginning with ㅇ come at the end of consonants, namely after ㅎ, because ㅇ is considered not a consonant but a symbol of silence.

2. *Vocabulary*: The field of vocabulary shows probably the most serious divergence. Abolition of many Sino-Korean words, standardization of words originating from northern dialects and archaic words, or coining of new words during the so-called *Mal tatumki wuntong* (Language purification movement) have resulted in vocabulary divergence between the North and the South. Some examples are shown below.

a. Words with different form but for the same concept: *tapseykita* (**sechakey ttaylita**) ‘to strike hard’, *palkacita* (**phoklo toyta**) ‘to be exposed’, *wisayngsil* (**hwacangsil**) ‘toilet’, *hosang* (**sangho**) ‘mutual’, *yangsayngwen* (**yanglowen**) ‘old people’s home’, *kwun-cwungkayo* (**taycwungkayo**) ‘popular song’, *lyohay* (**yanghay**) ‘approval’, *pap-kwak* (**tosilak**) ‘lunch box’.

b. Words that originated from northern dialects and became the standard in North Korean: *kesani* (**kewi**) ‘duck’, *pulwu* (**sangchwu/sangchi**) ‘lettuce’, *namsay* (**chayso**) ‘vegetable’, *kangnayngi* (**okswuswu**) ‘corn’, *talkwuci* (**swuley**) ‘cart’, *mangtol* (**maystol**) ‘grinding stone’, *masuta* (**puswuta**) ‘to break’, *incha* (**kot**) ‘soon, immediately’.

There are other differences, such as words with the same form but whose concept changed due to the influence of sociopolitical factors in North Korea. Examples are: *epei* ‘parents’ (S-K), ‘symbolic title of Kim Il-sung’ (N-K); *akassi* ‘girl, miss’ (S-K), ‘slave of feudalism’ (N-K); *kwungcen* ‘palace’ (S-K), ‘big, palace-like building used for social activities’ (N-K).

There are many loanwords from Russian: for example, *kkommuna* (**kongtong ciptan**) ‘community’, *kkamppania* (**cipcwung saep**) ‘concentration work’, *tulakttolu* (**thuleykthe**) ‘tractor’, *ppioneylu* (**sonyentan**) ‘boy scout’. Some loanwords borrowed from foreign languages other than Russian are often spelled with a Japanese-style pronunciation, e.g. *koppu* (**khep**, J *kooppu*) ‘cup’, *lacio* (**latio**, J *rajio*) ‘radio’, *ppomato* (**phomatu**, J *pomādo*) ‘pomade, hair grease’.

North Koreans tend to prefer native words to Sino-Korean words, and Sino-Korean words to loanwords (17):

(17)	S-K	N-K	
	Sino-Korean	Native	
	<b>hwamulsen</b>	<i>cimpay</i>	‘cargo ship’
	<b>taynoy</b>	<i>khunkol</i>	‘the cerebrum’
	<b>hongswu</b>	<i>khunmul</i>	‘flood’
	<b>phaakhata</b>	<i>thulecwita</i>	‘grasp’
	Sino-Korean	Native+Sino-Korean	
	<b>swu-cawen</b>	<i>pata-cawen</i>	‘marine resource’
	<b>cenkwu</b>	<i>cenki-al</i>	‘electric bulb’
	<b>yangkok</b>	<i>al-kok</i>	‘grains’
	<b>hoyngtanmyen</b>	<i>kalo-calun-myen</i>	‘cross section’

Loanword	Native or Sino-Korean	
<b>heyl-ki</b>	<i>cik-sung-ki</i>	'helicopter'
<b>nokhu</b>	<i>son-kichek</i>	'knock'
<b>phama</b>	<i>pokkum-meli</i>	'perm'
<b>sulliphe</b>	<i>kkul-sin</i>	'slippers'
<b>ssekhesu</b>	<i>kyoyeytan</i>	'circus'

3. *Phonetics and Phonology*: Phonetic and phonological differences are apparent but do not impede North and South Korean speakers in understanding each other. There are some well-known features in which North Korean differs from the Seoul standard.

In Pyongyang speech, the alveolar flap *l* [ɾ] occurs freely in word-initial position, whereas it is either dropped or replaced by *n* in Seoul speech, as in *lilon* (**ilon**) 'theory', *lyokum* (**yokum**) 'fare', *lotong* (**notong**) 'labor'.

Similarly, the alveolar nasal *n* occurs freely word-initially in N-K even when followed by the semivowel *y*, whereas it is dropped in S-K, as in *nyeca* (**yeca**) 'woman', *nyelum* (**yelum**) 'summer'. The word-initial occurrence of *l* and *n* in N-K is well known as one of the typical dialectal features of Pyongyang speech, and now has been officially recognized in both speech and writing.

The back vowels *a* and *e*, when followed by *i* within the word, often change to the front vowels *ay* and *ey* respectively through umlaut. This is widespread in spoken language in both N-K and S-K, but in N-K such assimilated forms are accepted as standard in both speech and writing. Some examples are: *engteyngi* (**engtengi**) 'hip, bottom', *muteyki* (**muteki**) 'pile', *kkwuleymi* (**kkwulemi**) 'bundle', *thopayki* (**thopaki**) 'native', *acilayngi* (**acilangi**) 'haze'.

In N-K, an additional *y* is inserted before *e* in verbal inflection, as in *kayyessta* (**kayessta**) 'got cleared', *peyyessta* (**peyessta**) 'have cut', *toyyessta* (**toyessta**) 'became', *heyyecita* (**heyecita**) 'be separated'. In spoken form, semivowel insertion takes place both in S-K and N-K, but only the forms without insertion are taken as standard in S-K writing.

Tensified pronunciation is more prevalent in N-K than S-K, as in *wensswu* (**wenswu**) 'enemy', *poksswu* (**pokswu**) 'revenge', *soktto* (**sokto**) 'speed', *haykssim* (**hayksim**) 'core', *ttwuk* (**twuk**) 'dam'.

One of the most noticeable features in N-K phonetics and phonology is in stress and intonation. Stronger stress and higher pitch are used in N-K compared to S-K. To S-K speakers, N-K sounds provocative and militant, and such an appalling feeling is no doubt conveyed by N-K's unique stress and intonation system. Interestingly, such intonation is interpreted in the North as "high-spirited, courageous and revolutionarily refined," as borne out by a speech made by Kim Il-sung at a meeting of North Korean linguists: "The pronunciation of Korean is high-spirited, courageous, and revolutionarily refined . . . , thus suited to express ideas and feelings of the people who are engaged in revolutionary activities . . ." (H. B. Lee 1990).

4. *Syntax and Morphology*: Although few differences are observed between N-K and S-K, N-K shows some interesting differences in a few cases, particularly morphology. The plural, expressed formally by *-tul*, is used much more frequently in the North, which can sound odd to South Koreans, but without affecting meaning. Consider the following examples: *ketayhan sengkwa-tul-ul ihwukhayssko* . . . 'achieved great results . . .', *motun hakkyo-tul-eyse* 'at all schools'. In S-K, plural *-tul* in these examples would often be omitted.

Another characteristic of North Korean syntax is that verbal clause structure is preferred to nominal phrases, where they are interchangeable, especially as newspaper headlines, etc. In contrast, in S-K a nominal phrase is more often used. This is presumably due to the fact that North Koreans are reluctant to use Sino-Korean words. Whereas Sino-Korean words allow concise nominal expression, North Korean depends on verbal phrases to express the same content, using native Korean verbs.

- (18) **mokcek talseng-ulo (sunglica-ka toyca)** (S-K)  
 ‘With goal attainment, (let’s be a victor)’  
**mokcek-ul ilwukhaye (sunglica-ka toyca)** (N-K)  
 ‘Attaining the goal, (let’s be a victor)’

Another interesting point was made by Shin (1995: 332f) regarding syntactic differences. After comparing newspaper headlines in S-K *Cosen Ilpo* and N-K *Lotong Sinmun*, she concludes that the former preferred sentence structures in which a patient was in subject position and an agent was sometimes deleted or left in oblique position, whereas the latter preferred sentence structures in which an agent was normally positioned as subject.

In N-K, the prospective modifier may occur where the present modifier would occur in S-K. Thus, the construction *-(ul) tey tayhaye* ‘regarding’ is frequently used, which could be a counterpart of S-K **-nun tey tayhaye**, **-nun kes-ey tayhaye**. Also *-(ul) taysin-ey* ‘instead of’ is used for S-K **-nun taysin-ey**. Consider the following examples: *cosene-uy thukseng-ul sallye naka-l tey tayhaye* ‘on developing the characteristics of Korean language’, *hyepcoha-l tey tayhaye* ‘concerning cooperation’, *cip-ey tolaka-l taysin-ey . . .* ‘instead of going back to the house . . .’.

Some other different phrasal forms are also observed, as in *yey-ha-myen (yey-lul tul-myen)* ‘for instance’, *il-tte-seta (him-cha-key ile-seta)* ‘rise up firmly’, *tul-ssuy-wuta (twicip-e ssuy-wuta)* ‘put (blame) on’, *cwuk-ul nayki lo (cwuk-tolok)* ‘to the end’, *yekhal-ul nolta (yekhal-ul hata)* ‘play a role’, *cakyong-ul nolta (cakyong-ul hata)* ‘do a function’.

In terms of morphology, a large number of words have been created using some derivational suffixes, such as *-cek*, *-hi-*, *-lop-*, *-wu-*, *-ci-*, *-mac-*, and *-ca*, beyond the function to which they are usually limited in Seoul Korean.

- (19) *-hi-*: *kip-hi-* ‘deepen’  
*-na-*: *coha-na-* ‘to like very much’  
*-cek*: *hakkyo-cek* ‘of the school’

Another noteworthy different trend is simplification of speech levels in North Korean. Traditionally, six different speech levels have been recognized in Korean depending on the relationship between speaker and hearer, although there are some discrepancies between scholars (see Chapter 5: Table 5.10; section 5.5.6.3). However, in North Korean, these six levels of speech style seem to have been reorganized and reduced to three levels, namely, (i) polite style, (ii) equal style, and (iii) low style. The sentence-ender forms of these three levels are as follows: (i) polite level: *-supnita/pnita* (sometimes *-yo* may be allowed), (ii) equal level: *-o/so*, *-yo*, *-suptey/ptey*, *-psey*, (iii) low level: *-ta*, *-nunya*, *-kela*, etc. (Nam and Ceng 1990: 36–40). Some reasons for this simplification may be found in the following statement: “We have three levels of speech style . . . suitable for the people engaged in the construction of a socialist country . . . whereby we can respect seniors, express affection to juniors, and be friendly with colleagues and friends” (H.-B. Lee 1990: 79).

Perhaps complete abolition of speech levels is much more desirable, considering speech levels are undesirable and incompatible in a socialist ideology. On the other hand, it is noted that the honorification suffix *-si-* is used more extensively in North Korean. This might have been encouraged by the excessive cult for Kim Il-sung, but contrasts with the simplification of speech levels.

5. *Stylistics*: In N-K, style, as a powerful weapon for revolution, is regarded as one of the most important means for carrying out the social function of language. Kim Il-sung claimed that the stylistic characteristics of 'Cultured language' must reflect the working class's needs. The characteristics of N-K stylistics are as follows:

- (20) a. They prefer short sentences to express militant emotion.
- b. Commands and exclamation styles are preferred.
- c. Emphatic style via repetition is preferred.
- d. Titles tend to be verbal clauses rather than nominals (see above).
- e. Spoken style is preferred to written style.

I have examined some noticeable differences in linguistic use between North and South Korea. Depending on viewpoint, such differences could be regarded as serious divergence or minor divergence. Some scholars minimize the extent of differentiation between North and South Korean, whereas others claim that the linguistic differentiation is not only a real phenomenon but it has historical (pre-division of the peninsula) and social reasons. Chinese had been the official written language of the elite until the late nineteenth century and at the end of the colonial period Korean was banned by the Japanese. Consequently, there was no standardized language in the Korean peninsula at the time when it was divided. Both states have implemented separate policies since division. In North Korea, language has been shaped by social requirements dictated by Communist ideology. Therefore, most North Korean language characteristics today are related to that ideological thought. Nevertheless, both languages have common phonological and grammatical structure and both peoples have no real difficulties in communicating except a few different vocabulary usages. In short, although some differences are found in morphology, phonetics, phonology, and stylistics we conclude that the extent of differentiation cannot reverse the belief in the homogeneity of the North and South Korean languages.

## NOTES

- 1 This consonant has changed to *w* in the central dialects.
- 2 Some linguists including Sohn (1999) separated Ch'ungch'öng off as an independent dialect, and classified Korean dialects into seven large groups.
- 3 Ch'ungch'öng speakers are normally stereotyped as speaking really slowly.

## REFERENCES AND FURTHER READING

- Choy, Hakkun (1987) *Hankwuk Pangen Sacen* [Dictionary of Korean dialects], Seoul: Myengmun-tang.
- Hankwuk Cengsin Munhwa Yenkwuwen (1986–1994) *Hankwuk Pangen Calyocip* [A collection of Korean dialect data], Seoul: Hankwuk Cengsin Munhwa Yenkwuwen.
- Kang, Ceng-Huy (1988) *Cecwu Pangen Yenkwu* [A study of Cheju dialect], Daejeon: Hannamday Chwulphan-pu.

- Kim, Thay-Kyun (1986) *Hampuk Pangen Sacen* [Dictionary of North Hamkyeng dialects], Suwon: Kyengki Tayhakkyo Chwulphan-pu.
- Kim, Yong-Hwang (1982) *Cosene Pangenhak* [Korean dialectology], Pyongyang: Kim Ilsung University Press.
- King, J. R. P. (1987) 'An introduction to Soviet Korean', *Language Research*, 23: 233–77.
- (1992) 'Archaisms and innovations in Soviet Korean dialects', *Language Research*, 28: 201–23.
- (2006) 'Dialectal variations in Korean', in Ho-min Sohn (ed.) *Korean Language in Culture and Society*, Honolulu: University of Hawai'i Press, 264–80.
- King, R. and Yeon, J. (1992) 'Cwungang Asia Hanintul-uy ene: Kolyemal' [The language of Koreans in Central Asia: Kolyemal], *Hankul*, 217: 83–134.
- Lee, Hyun-bok (1990) 'Differences in language use between North and South Korea', *International Journal of the Sociology of Language*, 82: 71–86.
- Lee, Ik-seop, Lee, Sang-oak, and Chae, Wan (1997) *Hankwuk-uy Ene* [The language of Korea], Seoul: Sinkwu Munhwasa.
- Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, New York: State University of New York Press.
- Lee, Ki-gap, Ko, Kwangmo, Ki, Sekwan, Chung, Jemun, and Song, Hajin (1998) *Cennam Pangen Sacen* [Dictionary of Southwestern dialects], Seoul: Tower Press.
- Lee, Sung-Nyong, et al. (1971) *Kwuke Pangenhak* [Korean dialectology], Seoul: Hyengsel Chwulphansa.
- Nam, Seng-wu and Ceng, Cay-yeng (1990) *Pukhan-uy Ene Saynghwal* [Linguistic life in North Korea], Seoul: Kolyewen.
- Ogura, Shinpei (1940) *The Outline of Korean Dialects*, Tokyo: Tōyō Bunko.
- (1944) *Chōsen Hōgen No Kenkyū* [A study of the Korean dialects], Tokyo: Iwanami Bunko.
- Pangen Yenkwuhoy (2001) *Pangenhak Sacen* [Dictionary of dialectology], Seoul: Thahaksa.
- Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects*, Seoul: Thap Chwulphansa.
- Shin, Eun-Kyung (1995) 'A lexical analysis of linguistic divergence in the North and South Korean presses', unpublished thesis, University of Wales.
- Sohn, Ho-min (1999) *The Korean Language*, Cambridge: Cambridge University Press.
- Yeon, Jaehoon (2006) "'Standard language" and "cultured language"', in Ho-min Sohn (ed.) *Korean Language in Culture and Society*, Honolulu: University of Hawai'i Press, 31–43.



PART III

**JAPANESE**



# OLD JAPANESE

*John R. Bentley*

## 7.1 INTRODUCTION

‘Old Japanese’, hereafter OJ, generally refers to the written language of Central Japan during the Nara period (710–794 CE). This language is richly preserved in a variety of works compiled during this period: *Kojiki* (712 CE), *Nihon shoki* (720 CE), and *Man’yōshū* [MYS] (ca. 759 CE). Because sections of these three works predate the Nara era, OJ can also refer to the language of the Asuka period (592–710 CE), which preceded the Nara. Based on the archaic nature of the spellings and lexicon of a number of *norito* ‘liturgies’, I label the language of the Asuka period Early Old Japanese.

A careful examination of the poetry in books 14 and 20 of *Man’yōshū* makes it clear that at least two dialects existed in Japan during the OJ period: the central or capital dialect, and an eastern dialect, hereafter referred to as Eastern Old Japanese (EOJ). Because of this split, some prefer to call OJ ‘Western Old Japanese’, but I will continue to call the language spoken in the center of the Yamato Basin simply OJ.

First, there are several distinctive features that separate OJ from Classical Japanese (ca. 794–1192 CE). The most obvious, and most often cited, difference is the number of vocalic distinctions. OJ had an eight-way distinction, while Classical Japanese had only five. The eight vocalic distinctions of OJ are preserved in the phonetic script, called *man’yōgana*. Over time this somewhat complex vocalic system underwent a partial collapse, where a number of vowels merged with others (cf. 7.2). Another OJ distinction is found in the vocalic difference of the evidential and imperative forms of quadrigrade verbs: *kaku* ‘write’ becomes *kake* EVID but *kakye* IMP. With the loss of these eight vocalic distinctions, the evidential and imperative forms of quadrigrade verbs become identical, leading to the misguided conclusion of some that the verbal suffix *-ri* ‘progressive’ attaches to the imperative *base*, when it actually attaches to the evidential. Finally, stative verbs (adjectives) in OJ retain a rather primitive morphology (acting more like nouns), where the stem can attach directly to nouns and verbs: *wakakusa* ‘young herbs’ (from *waka-si* ‘young’), *wakayaru* ‘become younger’. This primitive feature disappears in the classical era, aside from fossilized examples such as *wakayama* ‘young mountain’.

It is helpful here to briefly discuss the other dialect, Eastern Old Japanese, and the main differences between it and OJ. The most striking difference is that EOJ only preserves five vocalic distinctions, a fact that is somewhat obscured because the EOJ poets employed the capital dialect’s orthography. Thus, the rather wild divergence in spellings leads us to conclude that EOJ did not make certain vocalic distinctions that OJ did. EOJ also preserves vestiges of an older difference between the conclusive and attributive forms in quadrigrade verbs, while OJ does not: OJ *pur-u* ‘fall, precipitate’ and *pur-u yuki* ‘the snow that falls’. Compare this with EOJ *pur-u* ‘fall’, but *pur-o yuki* ‘the snow that falls’ (MYS 3424). Also OJ *pik-u* ‘pull’ and *pik-u pune* ‘the boat that is pulled in’, compared with EOJ *pik-u* ‘pull’ and *pik-o pune* ‘the boat that is pulled in’ (MYS 3431). It should be noted that Ryukyuan also preserves evidence that the form of quadrigrade verbs was Stem-*o*.

**TABLE 7.1 OLD JAPANESE VERSUS EASTERN OLD JAPANESE**

Gloss	Old Japanese	Eastern Old Japanese
‘as I gaze’	<i>mitutu</i>	<i>mitoto</i> (MYS 4421)
‘good’	<i>yoki</i>	<i>yoke</i> (MYS 4419)
‘snow’	<i>yuki</i>	<i>yoki</i> (MYS 3423)
‘though it includes’	<i>pupumeredo</i>	<i>popomaredo</i> (MYS 4387)
‘words of parting’	<i>sakiku</i>	<i>sakeku</i> (MYS 4368)

As *yoki* in the first example above makes clear, a number of EOJ nouns also have a vowel lower than that found in the same word in OJ. If we may be permitted to assume that EOJ is a more conservative dialect than OJ, then perhaps the logical conclusion is that OJ has undergone vowel raising, while EOJ has not (cf. Frellesvig and Whitman 2004). Consider Table 7.1. Another interesting difference is the change in the verbal ending when a *kakari-musubi* particle is inserted into a sentence. In OJ the insertion of the particle *kamo* results in an attributive ending for a verb: *yuki kamo pur-u* (fall-ATTR) *to* ‘Is it snowing [at my beloved’s house]?’ (MYS 844). In EOJ, however, when a *kakari-musubi* particle is inserted into a sentence, the quadrigrade verb undergoes the change of adding *-ar-* to the verb stem: *yuki kamo pur-ar-u* (fall-ATTR) ‘I wonder if it is snowing [on Tukupu Peak]?’ (MYS 3351), where *puru* ‘fall’ becomes *puraru*. These are just a few differences one notices between the central and eastern dialects. It is difficult to be more precise about how different EOJ was compared to OJ because of the dearth of evidence regarding this important dialect.

In the description of OJ that follows below, examples are cited from the following texts:

*Kojiki* (KJK, Onoda 1977)                      *Nihon shoki* (NS, Ienaga *et al.* 1986)  
*Man’yōshū* (MYS, Satake *et al.* 1999)      *Norito* (NT, Bentley 2001)  
*Senmyō* (SM, Kitagawa 1982)

## 7.2 PHONOLOGY

One of the main phonemic characteristics of Old Japanese compared to later stages of the language is the presence of eight distinct vowels, or ‘rhymes’. Aside from the vowels *a* and *u*, traditional Japanese scholarship has labeled the remaining six vowels *kō* ‘one’ and *otsu* ‘two’, because of the superficial relationship between the two types that later merge into *e*, *i*, and *o*; e.g.  $e_1$  and  $e_2 > e$ . Later work by Lange (1972) and Martin (1987) led to the romanization scheme called Yale Romanization, where a glide is posited either before or after the nuclear vowel: *ye* versus *ey*, *yi* versus *iy*, and *wo* versus *o*. As quickly becomes apparent, this romanization is actually more of a handy transcription, rather than an attempt to record the phonology of the language.

Until the work of Miyake (2003), the actual phonetics of these distinctions was heavily debated, and there were a number of competing theories. Miyake (2003: 262–4) convincingly shows through a large database of comparative evidence that OJ actually had seven vowels and one diphthong. Regarding the collapse of this system, Miyake (2003: 264) has argued that the following three mergers took place in the latter stages of OJ: *i* merged with *i* (*otsu i > kō i*),  $\partial$  merged with *o* (*otsu o > kō o*), and  $\partial y$  [əi] merged with *e* (*otsu e > kō e*). Compared to Classical Japanese, OJ was phonologically very simple, as Tables 7.2 and 7.3 illustrate.

**TABLE 7.2 OLD JAPANESE CONSONANTS**

	Labial	Dental	Palatal	Velar
Plosive	p <sup>m</sup> b	t <sup>n</sup> d		k <sup>ŋ</sup> g
Fricative		s <sup>z</sup>		
Nasal	m	n		
Flap		r		
Approximant	w		y	

**TABLE 7.3 OLD JAPANESE VOWELS**

i			i	u
		(əy)		
	e		ə	o
			a	

When citing examples, I cite the original, then provide a transcription and an English tag translation. When longer, complex examples are cited, I add a line of grammatical analysis. My transcription is as follows, comparing the traditional system and the phonemic system adopted in this chapter:<sup>1</sup>

Traditional	Phonemic
Ci <sub>1</sub>	Ci
Ci	Ci
Ci <sub>2</sub>	Cwi
Ce <sub>1</sub>	Cye
Ce	Ce
Ce <sub>2</sub>	Ce
Co <sub>1</sub>	Cwo
Co	Co
Co <sub>2</sub>	Co

When transcribing OJ represented in phonograms, the romanized form appears in lower case, but when transcribing semantograms (rebus script), the romanization appears in capital letters: 紀 *kwi* ‘tree’ versus 木 *KWI* ‘tree’.

### 7.2.1 Consonants

OJ had 13 consonantal phonemes. As Table 7.2 illustrates there was not a traditional voiceless/voiced contrast, but one of ± prenasalization: p ~ <sup>m</sup>b, t ~ <sup>n</sup>d, s ~ <sup>z</sup>, and k ~ <sup>ŋ</sup>g. Furthermore, historically there are two consonants that have also been subjected to much debate regarding their phonological status: (i) did Old Japanese actually have *p* or was this a fricative ([ϕ] or [f]) that originated from lenition of a much earlier, pre-OJ *p*; and (ii) did OJ have an affricate or not? The first question can be answered based on data from *Nihon shoki*. The phonogrammatic system used in *Nihon shoki* is based on Late Middle Chinese (LMC) *sans* a peninsular filter. The phonological inventory of LMC allowed the *Nihon shoki* scribes to transcribe so-called *ha-gyō* syllables with either \*p-graphs or \*f-graphs. An overview of the transcription in *Shoki* shows, however, that both were used, but this is easily explained due to the limitations of the underlying LMC

**TABLE 7.4 S-SERIES IN *KOJIKI* AND *NIHON SHOKI***

<i>Kojiki</i>	vowel	<i>Nihon shoki</i>
ts-	<i>a</i>	s-, ts-, ɕ-
c- or s-	<i>i</i>	s-, c-, ts-, sʃi-
s-	<i>u</i>	s-, c-, ɕ-
ɕ-	<i>e</i>	s-, c-, ɕ-
ts-	<i>o</i>	c-
s-	<i>wo</i>	s-

phonology. In other words, OJ *pa*, *pi*, *pe*, and *po* were transcribed with LMC graphs with \*p-, \*ph-, \*pʰ-, or \*b- initials, but OJ *pu* had to be transcribed with \*f-, \*fʰ-, and \*fʰ- initial LMC graphs. This skewed distribution demonstrates that OJ *p* was not a fricative.

The second issue is whether an affricate was present in OJ. Some have argued that the s-series in OJ actually consisted of *tsa*, *si*, *su*, *se*, *so*. Miyake's analysis of the phonogrammatic use of the LMC s-series in *Kojiki* and *Nihon shoki* (Table 7.4) points to an interesting overlap (2003: 178–81).

This chaotic distribution strongly argues more for the limitations of the Chinese-based orthography rather than the distribution of the actual phonology in OJ. I thus conclude that OJ had no affricate at this time; it only had an alveolar fricative.

### 7.2.2 Vowels

As Table 7.3 illustrates, OJ had seven vocalic phonemes and one diphthong. In the early stages of Old Japanese each vowel could follow every consonant, theoretically. This was not true, however, of the approximants, because there appears to have been no *yi* or *wu* in OJ. If these two phonemic distinctions actually did exist in OJ, then they fell into a gap in the Chinese syllabic inventory (at least for *yī*), because Chinese did not make a distinction between *i* and *yi*. Thus, I do not reconstruct *yi* or *wu*. As far as we can tell, a gradual collapse of the *kō-otsu* distinction began in the early days of the Nara era, with the merger of *mwo* and *mo*. Then roughly over a 150-year span of time all the other distinctions were neutralized, creating the five-vowel system Japanese has today.

### 7.2.3 Suprasegmentals

The pitch accent system of OJ is poorly understood. Scholars generally take the Heian pitch accent system, as preserved in texts such as *Wamyō ruijushō* or *Ruiju myōgishō* which contain accent marks, and project this system back into the Nara era or earlier. Some studies of a pitch accent type of orthography, as preserved in the poetry and annotational notes in *Nihon shoki* (cf. Takayama 1981, 1984), have shown that the pitch accent in the OJ period *appears* to be similar to that preserved in the Heian era, but we cannot say that with any certainty. One recent study (Matsumori 2008) cautions scholars not to project the Heian system into the proto-system. There are several issues here that hinder this study of pitch accent orthography in OJ texts: Chinese phonology did not preserve all the possible phonemic distinctions found in OJ; also a perfect one-to-one match of *both* Chinese phonology and tone is nearly impossible to achieve. Because of this, little can be definitely claimed about pitch accent in OJ, other than to remark that it *appears* to be very similar to that preserved in Heian texts.

### 7.2.4 Morphophonology

OJ had a phonotactic constraint where consonant or vowel clusters were not permitted. The rare cases of OJ words such as *kai* ‘rudder’ or *kui* ‘regret’ likely actually contain the segment *-yi-* (*\*kayi* and *\*kuyi*), as the OJ orthography had no way of making a distinction between *-i-* and *-yi-*. The phonotactic constraint mentioned above results in a number of morphophonological changes, namely: monophthongization, contraction, and nasalization.

Monophthongization usually occurs when a morpheme boundary is lost in a compound, and the resulting series of vowels monophthongizes. This process is very apparent in verbal morphology. When the vowels *-i-* and *-a-* occur together, monophthongization results in the vowel *-ye-*: *-iki-* retrospective + *-am-* tentative > *-ikyem-* ‘retrospective-tentative’; *-iki-* retrospective + *ar-* ‘exist’ > *-ikyer-* ‘retrospective’; *tono* ‘(palace) building’ + *iri* ‘enter’ > *toneri* ‘chamberlain’. When this combination is reversed, with the vowels *-a-* and *-i-* occurring together, the resulting vowel is *-e-*: *uka* ‘food?’ + *ip-* ‘say’ > *ukep-* ‘make a promise, contract’; *taka* ‘high’ + *iti* ‘market’ > *taketi* ‘Takechi, toponym’.

Contraction results when two vowels occur together, and then one elides. This is quite apparent in many court titles and toponyms, because the sinographs provide a crucial clue to the meaning: 淡海 *apa-umi* ‘bland-sea’ > *apumi* ‘province of Afumi’. The name of this province later undergoes another step of monophthongization in the Heian era when *-p-* lenites and then drops: *apumi* > *ahumi* > *aumi* > *oumi*. 朝臣 *asa-omi* ‘morning-minister’ > *asomi* ‘title of Asomi’. 河内 *kapa-uti* ‘inside the river’ > *kaputi* ‘Kawachi province’.

Nasalization occurs in the following environment: NYC > NC, where N is a nasal, and C is a voiceless obstruent: 粟津 *apadu* ‘Awazu, toponym’ = [apa<sup>n</sup>du] < *apa no tu* ‘port of Apa’; 美夜麻賀久理弓 *mi-yamagakurite* ‘hiding within the mountains’ = [mi-yama<sup>n</sup>gakurite] < *mi-yama ni kakurite* ‘hiding in the august mountains’; 多奈婆多 *tanabata* ‘weaver maid’ = [tana<sup>n</sup>bata] < *tana no pata* ‘loom [that is like] a shelf’; 畝火 *unebwi* ‘Mount Unebi’ = [une<sup>n</sup>bwi] < *une no pwi* ‘fire of Une’; 山邊 *yamabye* ‘Yamabe, toponym’ = [yama<sup>n</sup>bbye] < *yama no pye* ‘to the side of the mountain’.

## 7.3 SCRIPT

Inhabitants of the Japanese islands were originally illiterate, and appear to have been introduced to the complexities of the Chinese script (sinographs) by scholars from the Korean kingdom of Paekche. With the passage of time the Japanese altered the syntax of Classical Chinese (SVO) to fit the syntax of their vernacular (SOV). They also adopted sinographs, as other East Asian nations had, and used certain graphs for the phonetic value to create a kind of syllabary. This process of amalgamation resulted in OJ texts preserving three types of writing: (i) pure Chinese, (ii) a mixture of Chinese (semantograms) and phonograms, and (iii) phonetic script. The system of writing in phonograms is called *man’yōgana*, because of its copious use in the poetic anthology *Man’yōshū*. The following examples are illustrative of these three systems. The verb *noboru* ‘climb, go upstream’ and its transitive counterpart *nobosu* could be written in Chinese with a number of graphs that have the semantic value of the OJ verbs: 登・上・昇・泝. The verb could also be written with the same Chinese graph but with phonetic script attached to show Japanese verbal declension. In most cases the graphs used to represent verbal morphology were written in a smaller size: 泝須良牟 *NOBOS-uram-u* ‘likely being caused to go upstream’ (MYS 50). The word could also appear completely in phonetic script: 能煩理 *nobor-i* ‘climbing’ (KJK).

TABLE 7.5 DIFFERENT STRATA OF CHINESE IN JAPANESE RECORDS

Sinograph	Old Chinese	EMC	LMC	Pre-Shoki	Nihon shoki
支	*ke	*tei	*tʃi	ki	avoided
都	*təi	*tə	*tuə	tu	two
婆	*buai	*ba	*pʃia	ba	pa
摩	*məi	*ma	*mbuə	ma	ba
母	*mōʔ	*məw	*məw	mo	mo

It is primarily through the phonetic script that we are able to analyze the grammar of OJ more precisely, as well as determine the actual phonology of OJ consonants and vowels. One important point to make the reader aware of is that the inventory of OJ phonograms actually represents a mixture of different Chinese phonological strata, depending on the source of the system. The phonograms preserved in *Kojiki* and *Man'yōshū* belong to a rather old system apparently influenced by Sino-Paekche, the system known to the Paekche scholars. This has relied on a strain of Chinese variously labeled as Early Middle Chinese or Later Han. The orthographic system found in *Nihon shoki*, however, is a newer, *avant-garde* system taught directly to the Japanese court by students who had studied in the capital of China, Chang'an. This system is generally called Late Middle Chinese. Table 7.5 offers a concise comparison of these systems and the differences.

## 7.4 MORPHOLOGY

### 7.4.1 Nouns and case markers

OJ noun morphology can roughly be divided into the following three groups: prefixes, suffixes (including case markers), and pronouns.

There are a number of prefixes in OJ: honorific prefixes *mi-* and *opo-*, and the reinforcement prefix *ma-*. *mi-* is one of the most common prefixes found in almost any stage of the language, and it attached to both animate and inanimate nouns. The following example from *Norito*, the liturgies of Shintō, illustrates this aptly.

- (1) 皇 御 孫 命 能 宇 豆 乃  
 SUMYE MI-MA [N-O] MI-KOTO n-o udu n-o  
 imperial HON-grandchild DV-ATTR HON-deity DV-ATTR noble DV-ATTR  
 幣 帛 乎  
 MI-TEGURA wo  
 HON-offering ACC  
 '[present] the noble offerings of the honorable imperial grandchild . . .' (NT 7)

The prefix *opo-* meant 'greatness' and its prefixation elevated both animate and inanimate nouns. In the example below, *opo-* is prefixed before the prefixation of *mi-*.

- (2) 皇 大 御 神 能 大 前 爾  
 SUMYE OPO-MI-KAMWI n-o OPO-MAPYE ni  
 imperial HON-HON-deity DV-ATTR HON-before LOC  
 '(place the offerings) before the great, honorable imperial deity . . .' (NT 1)

The prefix *ma-* is often transcribed in the OJ corpus with the sinograph 真 'truth', and so I label it a reinforcing prefix, as the idea of 'truth' reinforces the nature of the noun that it is prefixed to. The following song illustrates this:

- (3) 摩蘇餓豫      蘇餓能古羅破      宇摩奈羅麼  
*ma-swoga yo      swoga n-o kwo-ra pa      uma nar-aba*  
 true-Soga-EMPH    Soga DV-ATTR child-PL TOP    horse be-if  
 譬武伽能古摩      多智奈羅麼    勾禮能摩    差比  
*pimuka n-o kwoma      tati nar-aba    kure n-o ma-sapi*  
 Himuka DV-ATTR steed    sword be-if    Koryō DV-ATTR true-blade  
 ‘The true Soga the posterity of the Soga. If you were horses, you would be  
 Himuka steeds. If you were swords, you would be true blades of Koryō (NS 103).

There are a number of other suffixes, but these can be classified as either plural markers case markers, or particles. There are three productive plural markers in OJ: *-tati*, *-domo* and *-ra*. The plural marker *-domo* is traditionally believed to be pejorative in nature but because of its limited distribution, and the fact that it only attaches to animate objects I would rather interpret this plural marker as a group classifier, much like English ‘bunches’. The following examples are illustrative: 祖止母 *OYA-domo* ‘ancestors’ (SM 13), 人止母 *PITO-domo* ‘people’ (SM 16), 古騰母 *kwo-domo* ‘children’ (MYS 4408), 惡奴止母 *ASI-KI YATUKWO-domo* ‘evil knaves’ (SM 43), 雷等 *KAMWI-DOMO* ‘the (gods of) thunder’ (NS Book 1).

*-tati*, on the other hand, functions as a simple plural for animate objects. Because *-domo* is often interpreted as a pejorative, *-tati* is often construed to be more appropriate for more elevated nouns. This, however, is not the case. Consider that edict 17 (of *Senmyō*) contains strongly worded chastisement to the empress’s servants, where she calls them 汝多知 *IMASI-tati* ‘you [who are like nephews]’. The following examples are illustrative: 皇神等 *SUMYE KAMWI-TATI* ‘the imperial deities’ (NT 1), 人等 *PITO-TATI* ‘people’ (MYS 340), and 臣多知 *OMI-tati* ‘ministers’ (SM 13).

The plural marker *-ra* is important, as it attaches to both animate and inanimate nouns: 俱梅能固邏 *kume no kwo-ra* ‘[you] children of the Kume’ (NS 9), 絢綿良 *KINU WATA-ra* ‘silk and cotton (garments)’ (MYS 900).

OJ had a number of case markers, listed alphabetically: *-ga* genitive<sup>1</sup>, *-i* nominative, *-made* terminative, *-ni* locative, *-no* genitive<sup>2</sup>, *-pye* direction, *-to* comitative, *-tu* locative-genitive, *-wo* accusative, *-yworī* (with *-yu*, *-ywo*, and *-yuri* as allomorphs) ablative. An example of each case marker follows: 伊豆毛多祁流賀 *idumwo takyeru ga* (GEN) ‘the brave Izumo’ (KJK 23); 志斐伊波奏 *sipwi i* (NOM) *pa MAWOSE* ‘[the maiden of] Shihī speaks [thus]’, (MYS 237); 比得能美流麻提 *pito no miru made* (TERM) ‘until a person sees [mistakes the mist for snow]’ (MYS 839); 于旒能和多利珥 *udi no* (GEN) *watari ni* (LOC) ‘at the ferry of Uji’ (NS 42); 倭辺遣登 *YAMATO pye YARU to* ‘I will send him back to Yamato’ (MYS 105); 岐備比登登等母 *kibwi pito to* (COM) *tomo ni* ‘with the person from Kibi’ (KK 54); 天津御量 *AMA TU* (LOC-GEN) *MI-PAKARI* ‘the august measure of heaven’ (NT 8); 種々色物乎備 *KUSAGUSA [NO] IRO MONO wo SWONAPE* ‘prepare the various colored objects’ (NT 1); 虚能莽由毛 *ko no ma yu* (ABL) *mo* ‘even from among the trees’ (NS 12); 許能麻用 *ko no ma ywo* (ABL) ‘from among the trees’ (KK 14); 本由利 *MOTO yuri* (ABL) ‘from the origin (beginning)’ (SM 7); 於朋耆妬庸利 *opoki two yworī* (ABL) ‘from the great door’ (NS 18).

A number of these case markers can also be combined with other case markers or particles: 嶋乎母家跡 *SIMA wo mo* (ACC INCL) *IPYE to* ‘[the bird] also [would make] the island his home’ (MYS 180); 年渡麻呂爾毛人者 *TOSI WATARU made ni mo* (TERM LOC INCL) *PITO PA* ‘Even though we pass through a whole year, a person . . .’ (MYS 3264); 枳彌波夜那祇 *kimi pa ya* (FOC PT) *naki* ‘Do (you) have no lord?’ (NS 104).

OJ also had zero-case marking, for both the subject and object. Below is an example of each:

- (4) a. 是 木 靈 也  
 KORE Ø KWI [N-O] TAMA NAR-I  
 this Ø tree DV-ATTR spirit is-CONC  
 ‘This is the spirit of trees’ (NT 8)
- b. 千木 高 知 弓  
 TIGWI Ø TAKA SIR-Ite  
 thousand.trees Ø high raise-GER  
 ‘raise high the roof rafters . . .’ (NT 2)

### 7.4.2 Verbs

OJ essentially has two classes of verbs, consonant-stem and vowel-stem verbs, and then a separate class of stative verbs, often labeled adjectives. There are also two irregular verbs, *s-* ‘do’ and *k-* ‘come’. Table 7.6 illustrates the traditional classification of these verbs, with a structural description of each verb. As this table makes clear, the traditional classification of verbs overcomplicates the analysis, and in reality there are only two classes. Also, what traditional grammarians consider to be particles (*ni* and *to*) I analyze as defective verbs,<sup>2</sup> so labeled because they do not have a complete verbal paradigm. Thus the copular *nar-* ‘to be’ is analyzed as a contraction of *n-i* (the INF of *n-*) plus *ar-* ‘to exist’. This is different from *nar-* ‘to become’. The texts only preserve two forms of the defective verb *n-*: *n-i* ‘INF’ and *n-o* ‘ATTR’. Note the vowel of attributive *n-o*; EOJ quadrigrade verbs sometimes also show *-o*, e.g. *ap-o sida* ‘when we meet’ (MYS 3478). Below are examples of *n-* ‘to be’, *to* ‘to be’, and *to* ‘to say’.

- (5) a. 歳 真尼久 傷 故 爾  
 TOSI MANeku YABUR-URU [GA] YUWE *n-i*  
 year repeatedly break-ATTR GEN reason DV-INF  
 ‘Because [crops] were damaged for so many years . . .’ (NT 4).
- b. 美之 比 乃 其等久 安里 等 都気 己曾  
*m-isi* *pi n-o* *gotoku ar-i to tuge koso*  
 see-RETRO.ATTR day DV-ATTR like be-CONC DV.INF tell.INF FOC  
 ‘Please tell [them] that I [am the same] as the day they saw me.’ (MYS 4473)
- c. 仲 麻呂 伊 忠 臣 止 之天  
 NAKamaro *i* TADASI-KI OMI *to s-ite*  
 Nakamaro NOM loyal-ATTR servant DV.INF do-GER  
 ‘Nakamaro has served as a faithful servant.’ (SM 34)
- 侍 都  
 PABER-It-u  
 be.HUM-PFV-CONC
- d. 諸 聞 食 止 宣  
 MOROMORO KIKOSIMYES-E *to* NOR-U  
 everyone give.ear-IMP DV.INF proclaim-CONC  
 ‘I declare, saying, “Give ear, all of you.”’ (NT 10)

**TABLE 7.6 OLD JAPANESE VERB CLASSES**

Quadrigrade ‘write’		Upper Bigrade ‘get up’	
IMPF	kak-a	IMPF	ok-wi
INF	kak-i	INF	ok-wi
CONC	kak-u	CONC	ok-u
ATTR	kak-u	ATTR	ok-uru
EVID	kak-e	EVID	ok-ure
IMP	kak-ye	IMP	ok-wi yo
Lower Bigrade ‘approach’		Upper Monograde ‘wear’	
IMPF	yos-e	IMPF	ki
INF	yos-e	INF	ki
CONC	yos-u	CONC	ki-ru
ATTR	yos-uru	ATTR	ki-ru
EVID	yos-ure	EVID	ki-re
IMP	yos-e yo	IMP	ki yo
<i>n</i> -irregular ‘die’		<i>r</i> -irregular ‘exist’	
IMPF	sin-a	IMPF	ar-a
INF	sin-i	INF	ar-i
CONC	sin-u	CONC	ar-i
ATTR	sin-uru	ATTR	ar-u
EVID	sin-ure	EVID	ar-e
IMP	sin-e	IMP	ar-e
<i>s</i> -irregular ‘do’		<i>k</i> -irregular ‘come’	
IMPF	s-e	IMPF	k-o
INF	s-i	INF	k-i
CONC	s-u	CONC	k-u
ATTR	s-uru	ATTR	k-uru
EVID	s-ure	EVID	k-ure
IMP	s-e yo	IMP	k-o (yo)

Stative verbs are also usually classified in the traditional grammars into two classes: *ku-katsuyō* and *shiku-katsuyō*. What is often overlooked is the fact that the stem of the stative verb is not necessarily a bound form in OJ, which fact is clear because of the insertion of other words between the stem and the noun: 布刀詔戸言 *putwo NORITOGOTO* ‘thick words [of a] liturgy’ (*KJK*); 高津鳥 *TAKA TU TORI* ‘birds high [in the sky]’ (*NT 10*); 比佐箇多能阿梅箇儺麼多 *pisa kata n-o ame kanabata* ‘long and sturdy metal loom (steady like the sky)’ (*NS 59*).

Stative verbs also inflect, but with a limited number of paradigms: [Table 7.7](#) illustrates the basic inflections of two stative verbs, one from each class.

With this limited range of inflection, stative verbs could not express a broad variety of meanings. In order to expand the semantic range of stative verbs, a verbal element was added to the infinitive and to this was attached other suffixes: *taka-ku ar-az-u* high-INF be-NEG-CONC, and because of the morphophonological constraint noted above, this coalesced to *takakarazu* ‘it is not high’. Consider a number of examples:

TABLE 7.7 OLD JAPANESE STATIVE VERBS

<i>Ku-katsuyō</i> ‘high’		<i>Shiku-katsuyō</i> ‘correct’	
INF	taka-ku	INF	tadasi-ku
CONC	taka-si	CONC	tadasi-Ø
ATTR	taka-ki	ATTR	tadasi-ki
EVID <sub>1</sub>	taka-kyere	EVID <sub>1</sub>	tadasi-kyere
EVID <sub>2</sub>	taka-kye	EVID <sub>2</sub>	tadasi-kye
NML <sub>1</sub>	taka-sa	NML <sub>1</sub>	tadasi-sa
NML <sub>2</sub>	taka-kyeku	NML <sub>2</sub>	tadasi-kyeku

- (6) 一 豆 乃 善 有良努 行 爾 在  
*PITOtU no YO-KU Ar-an-u WAZA ni AR-I*  
 one GEN good-INF be-NEG-ATTR behavior LOC be-CONC  
 ‘Here is one behavior that is not good.’ (NT 7)

This example is important, because while the orthography is a mixture of semantograms and phonograms, the sinograph 有 ‘exist’ demonstrates that at least in the Asuka era contraction had not yet occurred, strongly suggesting that the morpheme boundary had as yet to be lost. The accuracy of this analysis is underscored by the next poem, a phonetic rendition from *Kojiki* and *Nihon shoki*.

- (7) a. 斯良 多麻 能 岐美 何 余曾比 斯 多布斗久  
*sira tama n-o kimi ga yosopi si taputwo-ku*  
 white jewel DV-ATTR lord GEN form PT noble-INF  
 阿理祁理  
*ar-ikyer-i*  
 be-RETRO-CONC  
 ‘How noble is your form, like white jewels . . .’ (KJK 7)
- b. 未 迺 於朋鷄句 塢  
*mwi n-o opokye-ku wo*  
 fruit DV-ATTR many-INF ACC  
 ‘(one) with lots of fruit . . .’ (NS 7)

### 7.4.3 Numerals

Native numerals in OJ are primarily cardinal, but can also function as ordinals. OJ also had a classifier system that was still being developed. Table 7.8 outlines the primary numbers. Asterisks mean the number is not historically attested, but theoretically should be of this shape.

There is also *yorodu* ‘ten thousand’, but when combining with this number, OJ added a classifier plus *amari* ‘extra’ to alert the reader that a certain digit would be repeated within a larger number. An important example appears in *Nihon shoki*, in the Jinmu record, where the number ‘1,792,470 years’ is glossed as:

- (8) *mwomwo yorodu-tose amari nanaswo yorodu-tose amari*  
 one.hundred ten.thousand-years and seventy ten.thousand-years and  
*kokono yorodu-tose amari putati-tose amari*  
 nine ten.thousand-years and two.thousand-years and  
*yopo-tose amari nanaswo-tose amari*  
 four.hundred-years and seventy-years and

**TABLE 7.8 OLD JAPANESE NUMERALS**

	Ones	Tens	Hundreds	Thousands
1	<i>pito</i>	<i>towo/swo</i>	<i>mwomwo/po</i>	<i>ti</i>
2	<i>puta</i>	<i>pata</i>	<i>*putapo</i>	<i>putati</i>
3	<i>mi</i>	<i>miswo</i>	<i>*mipo</i>	<i>*miti</i>
4	<i>yo</i>	<i>yoswo</i>	<i>yopo</i>	<i>*yoti</i>
5	<i>itu</i>	<i>iswo</i>	<i>ipo</i>	<i>*iti</i>
6	<i>mu</i>	<i>muswo</i>	<i>*mupo</i>	<i>*muti</i>
7	<i>nana</i>	<i>nanaswo</i>	<i>*nanapo</i>	<i>*nanati</i>
8	<i>ya</i>	<i>yaswo</i>	<i>yapo</i>	<i>yati</i>
9	<i>kokono</i>	<i>*kokonoswo</i>	<i>*kokonopo</i>	<i>*kokonoti</i>

The following classifiers are found in the OJ corpus: *-ka* ‘days’, *-moto* ‘plants’, *-pasira* ‘deities’, *-pye* ‘things that can be folded or layered’, *-ri* ‘people’, *-tabi* ‘times’, *-tose* ‘years’, *-tu* ‘object’. Some of these are fossilized forms of the object being counted, and as phonological changes took place, the fossilized form lost its original meaning and became a classifier. Others refer to a part of the object being counted. Table 7.9 illustrates this.

Examples include: 比苔菟麻菟 *pito-tu matu* ‘one pine tree’ (NS 27), 比邇波登袁加袁 *pi nipa towo-ka wo* ‘as for days, ten days [have passed]’ (KJK), 八神 *YASIRA [N-O] KAMWI* ‘eight deities’ (NS), 介瀾羅毘苔茂苔 *kamira pito-moto* ‘one leek’ (NS 13), 和賀布多理泥斯 *wa ga puta-ri nesi* ‘how the two of us slept together’ (KJK), 夜幣賀岐都久流 *ya-pye-gaki tukuru* ‘building the eight-fold fence’ (KJK 1), 一年二年爾不在 *PITO-TOSE PUTA-TOSE ni AR-AZ-U* ‘[It] is not one or two years’ (NT 4).

These numbers could also be attached to a few parts of speech without a classifier, apparently calling a classifier to the mind of the listener. This is perhaps the most primitive form of the Japanese number system, as the following illustrate: 那那由久 *nana yuku* ‘going seven times’ (KJK), 毛毛那比苔 *mwomwo napi to* ‘(have seen) one hundred battles’ < *mwomwo no api to* (NS 11).

## 7.5 SYNTAX

### 7.5.1 Noun phrase structure

The most productive form of noun phrase construction in OJ utilized the defective verbal *n-*, in the attributive, where nouns are simply tacked together in a string, with the head noun at the end of the string:

**TABLE 7.9 NOUN VERSUS CLASSIFIER IN OLD JAPANESE**

Noun		Classifier	
Form	Meaning	Form	Counts
<i>ke</i> < <i>*ka-i</i>	day	<i>-ka</i>	days
<i>moto</i>	trunk, base	<i>-moto</i>	plants
<i>pasira</i>	pillar	<i>-pasira</i>	deities
<i>pye</i>	layer	<i>-pye</i>	clothes, fabric
<i>(a)re</i>	thing, person	<i>-ri</i>	people
<i>tabi</i>	time	<i>-tabi</i>	number of times
<i>tosi</i>	year	<i>-tose</i>	years
<i>tu</i>	thing?	<i>-tu</i>	objects in general

- (9) 聞 食 々 國 乃 東  
*KIKOSIMES-U WOS-U KUNI n-o PIMUKASI [N-O]*  
 rule.HON-ATTR govern.HON-ATTR country DV-ATTR east DV-ATTR
- 方 陸 奧 國 乃 小田  
*KATA MITI [N-O] OKU [N-O] KUNI n-o WODA*  
 direction path DV-ATTR interior DV-ATTR land DV-ATTR Oda
- 郡 爾  
*[N-O] KOPORI ni*  
 DV-ATTR district LOC  
 ‘... in the Oda district of the province of Michinoku’ (lit. ‘land that is in the interior down the path) of the direction of the east of the land ruled over [by the emperor]’ (*SM* 13)

Adjectival modification of nouns was achieved with the addition of a suffix *-ki*. Consider the following example from the liturgies:

- (10) 狭 國 者 広 久 峻 國 者 平 久  
*SA-KI KUNI PA PIRO-ku SAGASI-KI KUNI PA TAPIRAKE-ku*  
 narrow-ATTR land TOP wide-INF steep-ATTR land TOP flat-INF  
 ‘[May] the narrow land be broad, and the steep land be flat...’ (*NT* 1)

Relative clauses in OJ, much like later stages of the language, are complex and constructed essentially as adnominalized segments. Consider the following example from the ostentatious language of imperial edicts:

- (11) 是 者 聞 母 威 岐 近江  
*KO PA KAKE-M-AKU mo KASIKO-ki APUMI [N-O]*  
 this TOP consider-TENT-NML INCL awe-ATTR Afumi DV-ATTR
- 大 津 宮 御宇  
*OPOTU [N-O] MIYA [NI] AME [N-O] SITA*  
 Ōtsu DV-ATTR palace LOC heaven DV-ATTR below
- 大倭 根子 天皇  
*SIR-AS-I-MES-ISI YAMATO nekwo SUMYERA MIKOTO*  
 rule-HON-INF-HON-RETRO.ATTR Yamato Neko ruler august.thing
- 乃 与 天地 共 長 与 日月  
*no AMETUTI TO TOMO [N-I] NAGA-KU PITUKWI*  
 GEN heaven.earth COM with DV-INF long-INF day.month
- 共 遠 不改  
*TO TOMO [N-I] TOPO-KU KAPAR-UMASIZI-KI*  
 COM with DV-INF far-INF change-NEG.TENT-ATTR
- 常 典 止 立 賜 比  
*TUNE [N-O] NORI to TAT-I TAMAp-i*  
 always DV-ATTR law DV.ATTR stand-INF- HON-INF
- 敷 賜 霸 留 法 乎 受 被賜  
*SIK-I-TAMAp-yer-u NORI wo UKE-TAMAPAR-I*  
 lay.out-INF-HON-PROG-ATTR law ACC receive-INF-HUM-INF

‘We (have) humbly received this, the law which was established by the divine ruler, Yamato Neko, who ruled everything under heaven from the Afumi Ōtsu Palace, whose (name or rule) we consider from our hearts with awe, who set up the law that will last as long as heaven and earth, and be enduring, even as the sun and moon are distant, a law that is unchanging and eternal . . .’ (SM 3)

### 7.5.2 Pronouns and anaphora

OJ has a number of pronouns, which can be divided into two categories, personal and demonstrative. Table 7.10 illustrates the pronouns in these two categories. Parentheses mark suffixes.

The meaning of the suffix *-re* that appears on most of the pronouns is unclear. It is clear, however, that the form with the suffix has a broader usage, while the form *sans* suffix is limited to a nominative/genitive or accusative role. Compare the following two sets of examples, with and without the suffix.

(12) Without a suffix

- a. 和 賀意富岐美      和 賀禰勢 流  
*wa ga opo-kimi      wa ga kyes-er-u*  
 I GEN great-lord      I GEN wear.HON-PROG-ATTR  
 意須比能 須蘇 爾  
*osupi no suswo ni*  
 outer.garment GEN hem LOC  
 ‘my great lord . . . on the hem of my outer garment that I wear . . .’ (KJK 28)
- b. 夜麻登幣 迹      由玖 波 多賀 都麻  
*yamato pye ni      yuk-u pa ta ga tuma*  
 Yamato ALL LOC go-ATTR TOP who GEN spouse  
 ‘Whose spouse is it who heads off toward Yamato?’ (KJK 56)
- c. 阿加等伎乃 加波 多例等枳爾  
*akatoki no ka pa tare doki ni*  
 dawn GEN over.there TOP who time LOC  
 ‘At dawn who is it over there . . .?’ (MYS 4384)

(13) With a suffix

- a. 久知比比久      和禮 波 和須禮 士  
*kuti pipik-u      ware pa wasure-zi*  
 mouth sting-ATTR I TOP forget-NEG.TENT  
 ‘I will not forget the stinging of my mouth.’ (KJK 12)
- b. 於彌能 烏苔咩 烏      多例挪始儺播務  
*omi no wotomye wo      tare yasinap-am-u*  
 palace GEN maiden ACC who take.care.of-TENT-CONC  
 ‘Who will take care of the maiden of the palace?’ (NS 44)
- c. 安我 毛布 支見我      弥不根 可母 加礼  
*a ga mop-u kimi ga      mi-pune kamo kare*  
 I GEN think-ATTR lord NOM HON-boat PT that  
 ‘Is that august boat my lord’s whom I long for? That out there?’ (MYS 4045)

TABLE 7.10 OLD JAPANESE PRONOUNS

Personal Pronouns		Demonstrative Pronouns		
Gloss	Form	Gloss	Form	Place
1 <sup>st</sup> person	<i>wa(re), a(re)</i>	proximal	<i>ko(re)</i>	<i>ko(ko)</i>
2 <sup>nd</sup> person	<i>na(re)</i>	mesial	<i>so(re)</i>	<i>so(ko)</i>
3 <sup>rd</sup> person	<i>si</i>	distal	<i>ka(re)</i>	∅
interrogative	<i>ta(re)</i>	interrogative <sub>1</sub>	<i>idu(re)</i>	<i>idu(ku)</i>
reflexive	<i>ono(re)</i>	interrogative <sub>2</sub>	<i>iku</i>	∅
		interrogative <sub>3</sub>	<i>itu</i>	∅
		interrogative <sub>4</sub>	<i>nani</i>	∅

### 7.5.3 The basic sentence

#### 7.5.3.1 Declarative

The declarative sentence in OJ consisted of the essential structure SOV, but in the majority of cases a subject in OJ is implied, and a nominative marking system appears to have weakened. Below is an illustrative example where the nominative is marked:

- (14) 仲末呂 伊 許 紆 流 心 乎  
*nakamaro i ITUPAR-I KADAM-Er-u KOKORO wo*  
 Nakamaro NOM lie-INF deceive-PROG-ATTR heart ACC  
 以 天 兵 乎 發  
*MOT-Ite IKUSA wo OKOS-I*  
 have-GER soldiers ACC raise-INF  
 ‘Nakamaro, having a lying and deceiving heart, raised troops . . .’ (SM 28)

There are a few rare instances, mainly in poetry, where the subject of a main clause is marked with *ga*:

- (15) 多伽機 珥 辭藝和奈 破蘆 和 餓末菟  
*taka kwi ni sigi wana par-u wa ga mat-u*  
 high fort LOC snipe trap lay-ATTR I NOM wait-CONC  
 ‘In the high fort setting a trap for snipe I wait . . .’ (NS 7)

In declarative sentences of the type X *nar-* ‘It is X’, the subject is either not marked or a topic marker *pa* is inserted, leaving the subject implied. Of the ten *nar-* sentences I have found in the liturgies and edicts, seven take *pa*, and three have the subject unmarked. Below is an example of each marking.

#### (16) Topicalized

汝 多知 諸 者 吾 近  
*IMASI-tati MOROMORO PA WA [GA] TIKA-KI*  
 you-PL various TOP I GEN close-ATTR

姪 奈利  
*WOPI nar-i*  
 nephew be-CONC

‘As for all you various servants, you are [like] my close nephews.’ (SM 17)

(17) Unmarked

位 名 乎 繼 止 念 天 在  
*KURAWI NA wo TUG-AM-U to OMOP-Ite AR-U*  
 throne name ACC inherit-TENT-CONC DV.INF think-GER exist-ATTR

人 奈利  
*PITO nar-i*  
 person be-CONC

‘(He) is a person who thought to inherit the throne and the name [of the imperial family].’ (SM 28)

7.5.3.2 *Tense/Aspect*

I follow Martin (1991: 272) in analyzing Japanese as having an aspectual system, within which certain features that others would interpret as ‘tense’ are also incorporated. The two most obvious features are the perfective and imperfective. What I label as the conclusive form of the verb, ending in *-u* (or *-i* for *ar-* ‘exist’, *nar-* ‘be’, *-itar-* and *-iky-*), represents the imperfective for our purposes here. There are also a number of suffixes which represent imperfection, but with an added nuance. Table 7.11 illustrates these.

The feature of perfective, however, needs to be expanded to take care of a number of nuances in OJ (Table 7.12). It should be noted that the difference between a number of these perfectives is not readily clear in the present.

It should be noted that the retrospective *-iki* has a number of suppletive forms: *-isi* (ATTR), *-ise* (IMPF), and *-isika* (EVID) that due to limits on space I have elected to ignore. There are a number of other suffixes that have features other than perfection or imperfection, but these are found under 7.5.3.4 ‘Modality’ or 7.5.5 ‘Passive and causative’.

One other feature of OJ verbs related to aspect is that of telicity. While telicity often deals with whether an action is completed or not, there is some evidence that a number of verbs made a distinction about the realization of a goal. Two such OJ verbs are *twor-* ‘hold, support’ and *tor-* ‘take, pick up’. The following example from a song in *Kojiki* illustrates this:

- (18) 美那 曾曾久 淤美能 袁登賣  
*mina sosok-u omi no wotomye*  
 water pour-ATTR servant GEN maiden  
 ‘The pouring water - the maiden of the palace
- 本陀理 登良須 母 本陀理 斗理  
*pwodari tor-as-u mo pwodari twor-i*  
 goblet pick.up-HON-ATTR INCL goblet hold-INF  
 picks up the goblet. With the goblet in hand

TABLE 7.11 OLD JAPANESE IMPERFECT SUFFIXES

Feature	Suffix	Examples
Conclusive	<i>-u</i>	<i>yuk-u</i> ‘to go’; <i>s-u</i> ‘to do’
Iterative	<i>-ap-</i>	<i>kakus-ap-u</i> ‘conceal continuously’ <i>mat-ap-u</i> ‘continue waiting’
Negative Tentative	<i>-azi</i>	<i>ar-azi</i> ‘there probably is no . . .’ <i>mot-azi</i> ‘probably will not hold . . .’
Subjunctive	<i>-amas-</i>	<i>ar-amas-eba</i> ‘if that were the case . . .’ <i>mi-se-mas-i</i> ‘would have liked to show . . .’

TABLE 7.12 OLD JAPANESE PERFECT SUFFIXES

Feature	Suffix	Examples
Perfective <sub>1</sub>	<i>-in-</i>	<i>kanap-in-u</i> ‘it is now suitable’ <i>tir-in-am-u</i> ‘will have scattered’
Perfective <sub>2</sub>	<i>-ite-</i>	<i>mi-te-ba</i> ‘when I have seen . . .’ <i>ar-it-u</i> ‘it has been’
Perfective Progressive	<i>-itar-</i>	<i>sak-itar-aba</i> ‘if it bloomed’ <i>uwe-tar-aba</i> ‘if you had planted’
Retrospective <sub>1</sub>	<i>-iki</i>	<i>ar-iki</i> ‘there was . . .’ <i>matur-iki</i> ‘came down [from heaven]’
Retrospective <sub>2</sub>	<i>-ikyer-</i>	<i>ar-ikyer-i</i> ‘there [once] was . . .’ <i>tir-ikyer-i</i> ‘blossoms scattered’
Tentative Retrospective	<i>-ikyem-</i>	<i>pur-as-ikyem-u</i> ‘she waved [the scarf], right?’ <i>pur-ikyem-u</i> ‘surely it had snowed . . .’

加多久斗 良勢	斯多賀多久
<i>kata-ku twor-as-e</i>	<i>sita gata-ku</i>
firm-INF hold-HON-IMP	base firm-INF
hold it firmly!	By the base
夜賀多久斗 良勢	本陀理斗 良須古
<i>yagata-ku twor-as-e</i>	<i>pwodari twor-as-u kwo</i>
firm-INF hold-HON-IMP	goblet hold-HON-ATTR child
hold it even tighter;	the child holding the goblet’ ( <i>KJK</i> 103).

The first example of *tor-* appears at the beginning of the poem, where we imagine that the maiden has picked up the goblet, with the apparent goal of drinking the contents of the goblet. The remaining four examples of *twor-* mean ‘to hold’ or ‘support’ the goblet, which is reinforced by the adverb ‘firmly’.

### 7.5.3.3 Negation

OJ has a number of suffixes that attach to verbs for simple negation, or a number of negative moods. Attaching the suffix *-az-* or *-an-*: 一年二年爾不在 *PITO-TOSE PUTA-TOSE ni AR-AZ-U* ‘[It] is not one or two years’ (*NT* 4); 多陀爾阿波須 *tada ni ap-az-u* ‘(on years) when we do not meet directly’ (*MYS* 809); 恩乎不報奴乎波 *MEGUMI wo MUKI-N-U woba* ‘. . . and do not reward (with mercy)’ (*SM* 41); 雖見飽奴 *MIR-EDO AK-An-u* ‘I do not tire (of seeing Yoshino River) no matter how often I see it’ (*MYS* 27).

When these attach to a vowel-stem verb, the initial vowel *-a-* is dropped: 登賀米受 *togame-z-u* ‘not censuring’ (*KJK*); 美受弓夜和礼波 *mi-zu-te ya ware pa* ‘without looking, I am . . .’ (*MYS* 862); 二遍美延農知知波波袁 *PUTATABI miye-n-u titi papa wo* ‘not being able to see my father and mother again’ (*MYS* 891).

There also is a stative verb of negation: *na-*: 安布余志勿奈子 *ap-u yosi mo na-si* ‘There is no way (for us) to meet’ (*MYS* 807). There also is a rather rare negative, *-anap-* in the Eastern Songs of *Man’yōshū*: 阿抱思太毛 安波乃散思太毛 奈尔己曾与佐礼 *ap-o sida mo ap-anop-ye sida mo na ni koso yos-are* ‘even when we meet, and when we do not meet (my thoughts) will be drawn to you’ (*MYS* 3478).

### 7.5.3.4 Modality

There are a number of suffixes in OJ which have a feature of plausibility or possibility. When the verbal feature is that of perfection or imperfection, these suffixes appear in 7.5.3.2 ‘Tense/Aspect’. The relevant suffixes for modality are listed in Table 7.13.

**TABLE 7.13 OLD JAPANESE SUFFIXES OF POSSIBILITY**

Feature	Suffix	Examples
Debitive	-ube-	<i>imasap-ube-si</i> ‘should remain [safe]’ <i>mi-t-ube-ku</i> ‘I can see [the jewels]’
Hypothetical Conditional	-aba	<i>nor-aba</i> ‘if he proclaims . . .’ <i>yuk-aba</i> ‘if it goes . . .’
Negative Debitive	-umasi-zi-	<i>yor-umasi-zi</i> ‘I do not want you to come’ <i>wasur-ay-umasi-zi</i> ‘I cannot forget’
Tentative <sub>1</sub>	-am-	<i>nak-am-u</i> ‘probably will cry out’ <i>ar-am-u</i> ‘perhaps it is . . .’
Tentative <sub>2</sub>	-uram-	<i>nor-uram-u</i> ‘maybe riding [in the boat]’ <i>mit-uram-u</i> ‘maybe the tide will be high’

7.5.3.5 *Non-declarative sentence types*

There are a number of non-declarative sentence types in OJ, including questions, exclamations, prohibitives, and imperatives. Questions, exclamations, and prohibitives are formed by adding a particle either in the middle of the sentence, or at the end of the sentence, after the verb: 四宝三都良武香 *sipo Mitur-am-u ka* [PT] ‘Is the tide coming in?’ (MYS 40); 御念八君 *OMOPOS-U YA* [PT] *KIMI* ‘Will you think (about the capital), my beloved?’ (MYS 330); 聞可毛 *KIK-ASE kamo* [PT] ‘Is it because you have heard . . .?’ (MYS 680); 欲寸君可毛 *POSI-KI KIMI kamo* [PT] ‘How I want (to meet) you’ (MYS 2992); 吾者余利爾思乎 *A PA yor-in-isi wo* [PT] ‘Yet I have followed after you!’ (MYS 3377); 吾乎奈見給比曾 *WA wo na* [PT] *MI-TAMAp-i so* [PT] ‘(She said) “Do not look at me!”’ (NT 12).

Imperatives, on the other hand, conjugate the verb in such a way that it is non-declarative, by adding either an imperative suffix to consonant-stem verbs, or a particle *yo* at the end of the stem of vowel-stem verbs: *yuk-ye* ‘Go!’ *yom-ye* ‘Read!’ *ki yo* ‘Wear!’ *se yo* ‘Do it!’

7.5.4 **Topic, focus, and emphasis**

OJ uses a number of particles, *mo*, *pa*, and *so* (later *zo*), to place focus on a unit in a sentence. The focus particle *pa* is often classified as a topicalizing particle. I consider all three particles here to be particles of focus, where the noun, or noun phrase preceded by the particle, is highlighted.

In the liturgies, the officiator of the liturgy asks a question, and the focus particle highlights the ‘august will’ of the deities:

- (19) 百 能 物 知 人 等 乃  
*MWOMWO n-o MONO SIR-I PITO-TATI n-o*  
 hundred DV-ATTR thing know-INF person-PL DV-ATTR
- 卜 事 爾 出 牟 神 乃  
*URAGOTO ni IDE-m-u KAMWI no*  
 divination.thing DAT appear-TENT-ATTR deity GEN
- 御 心 者 此 神 止 白 止  
*MI-KOKORO PA KONO KAMWI to MAWOS-E to*  
 HON-mind TOP this deity DV-INF say-IMP DV-INF

負 賜 支

*OPOSE-TAMAP-Iki*

command.INF-HON-RETRO.CONC

‘The ruler commanded that the various geomancers divine [and determine] the august will of the deity and find out who this deity is [lit. say it is this deity].’ (NT 4)

The focus particle *pa* is also used to contrast two different things, by marking both with *pa*:

- (20) 麻都良我波 奈奈 勢能 與騰 波

*matura gapa nana-se n-o yodo pa*

Matsura River seven-sandbars DV-ATTR pool TOP

与等武 等毛 和礼 波 与騰麻 受

*yodom-u tomo ware pa yodom-az-u*

stagnate-CONC although I TOP stagnate-NEG-INF

吉美遠 志麻多 武

*kimi wo si mat-am-u*

beloved ACC PT wait-TENT-CONC

‘Though the pool of the seven bars of the Matsura River should stagnate, I will not stagnate, but will continue to wait for you.’ (MYS 860)

The particle *mo* shifts the focus to the noun it follows. *so* (later prenasalized to *zo*) functions much like *mo*, but in sentences with *so*, the verb ends in the attributive, instead of the conclusive. Below is an example of each:

- (21) 今 母 去 前 母 天皇 我 朝廷  
*IMA mo YUK-U SAKI mo SUMYERA ga MI-KADWO*  
 now INCL go-ATTR before INCL ruler NOM HON-gate

乎 平 久 安 久

*wo TAPIRAKE-ku YASURAKE-ku*

ACC safe-INF peaceful-INF

‘[We petition] that the ruler’s court may be safe and peaceful, both now and in the future . . .’ (NT 2)

- (22) 持 五 兵 弓 追 走  
*ITU-KUSA [N-O] TUPAMONO MOT-Ite OP-I PASIR-I*  
 five-kinds DV-ATTR warriors use-GER chase-INF run-INF

刑殺 物 曾 登 詔

*KOROS-AM-U MONO so to NOR-U*

kill-TENT-ATTR thing FOC DV-INF declare-ATTR

‘Use the five kinds of warriors, run and chase after [that person] and kill him [lit. so that he will be a thing that is killed].’ (NT 16)

### 7.5.5 Passive and causative

OJ had a rather simple system for passives and causatives. In the natural word order of OJ, the causative or passive suffix always followed the verb stem, and then all other verbal suffixes are attached after.

There were two causative suffixes in OJ: *-asime-* and *-ase-*. Consider the following examples: 佐加叡志米 *sakaye-sime* ‘cause to prosper’ (NT 2), 兵癸之武 *IKUSA OKOS-Asim-u* ‘caused soldiers to be gathered’ (SM 28), 倭例儻魔柯斯每 *ware ni*

*mak-asim-e* ‘caused me to wrap (my hand) around hers’ (NS 96); 荒風浪爾安波世受 *ARA-KI KAZE NAMI ni ap-ase-z-u* ‘I will not let you run into violent winds or waves . . .’ (MYS 4245).

There are two passives in OJ: *-aye-*, and *-rare-*, but this last one has a number of alloforms: *-rare-*, *-rar-*, *-are-*, and *-ar-*. The passive *-aye-* appears to be the older passive, with *-rare-* a newer innovation, as *-aye-* falls out of use by the end of the OJ period. A number of representative examples include: 伊喻之之乎 *i-y-u sisi wo* ‘[Chasing] the game that had been shot’ (NS 117); 人爾不知所知 *PITO ni SIR-AYE-Z-U* ‘not known to people’ (lit. ‘not having been made known’) (MYS 1018); 火結神生給弓 美保止被燒 弓 *POMUSUBI [NO] KAMWI UM-I-TAMAP-Ite mipoto YAK-AYE-te* ‘[She] gave birth to the deity Pomusubi [the fire deity] and her vulva was burned’ (NT 12).

**7.5.6 Speech levels and respect**

As far as we can tell, all stages of recorded Japanese preserve a complex system of speech levels. This is true of the OJ period, where speech directed to the deities or a member of the royal family required elevated forms of speech. When the speaker addressed someone higher in status but referred to himself, this required a level of humble speech. Addressee or subject honorification could be achieved in a number of ways: by attaching an honorific auxiliary to the verb, using an honorific suppletive verb, or attaching a prefix to a noun that elevated it.

The following honorific prefixes are found in OJ: *mi-*, *o-*, *opo-*. Consider the following example with combined honorific prefixes and an honorific verb:

- (23) 天 御 蔭 日 御 蔭 登 隱 坐 弓  
*AME MI-KAGE PI MI-KAGE to KAKUR-I-MAS-Ite*  
 heaven HON-light sun HON-light DV.ATTR ide-INF-be.HON-GER  
 ‘(He = emperor) will hide (reside) in the heavenly light, in the solar light . . .’ (NT 1)

Table 7.14 illustrates verbs and verbal suffixes used in the OJ period to show higher or lower status of the addressee or hearer. Based primarily on prose texts preserved in *norito* and *senmyō*, we can divide the system of elevating the subject or object into two levels: ultra-elevated (for deities and royalty) and elevated (for other people).

The following two examples contain both elevated and humble verbs:

- (24) 大 御 巫 能 辞 竟 奉  
*OPO-MI-KAMUNAGI no KOTO WOPE-MATUR-U*  
 HON-HON-shaman GEN words end-HON-ATTR  
 皇 神 等 能 前 爾 白 久  
*SUMYE KAMWI-TATI n-o MAPYE ni MAWOS-Aku*  
 imperial deity-PL DV-ATTR before LOC say.HUM-NML  
 ‘Finishing the words of the great august shaman, we humbly speak before the imperial deities . . .’ (NT 1)

- (25) 此 乃 食 國 天 下 乎  
*KOno WOS-U KUNI AME [NO] SITA wo*  
 this.GEN rule.HON-ATTR land heaven GEN below ACC  
 調 賜 比 平 賜 比 天 下  
*TOTONOPE-TAMAp-i TAPIRAGE-TAMAp-i AME [N-O] SITA*  
 put.in.order.INF-HON-INF pacify.INF-HON-INF heaven DV-ATTR below

TABLE 7.14 OLD JAPANESE HONORIFIC VERBS

Elevated Status		
Gloss	Ultra-elevated verbs	Elevated verbs
be	<i>opomasimas-</i>	<i>masimas-</i>
think	<i>omoposimyēs-</i>	<i>omopos-</i>
hear, rule	<i>kikosimyēs-</i>	<i>kikos-</i> ‘say, partake’
rule over	<i>sirosimyēs-</i>	<i>siros-</i>
see, view	∅	<i>misonapas-</i>
be, go	∅	<i>imas-</i>
give (down)	∅	<i>tamap-</i>
say	∅	<i>notamap-</i>
eat, wear, rule	∅	<i>wos-</i>
Humble Status		
Gloss	Humble verbs	
give (up), offer	<i>tatematur-</i>	
go, attend	<i>maude-</i>	
receive	<i>tamapar-</i>	
say	<i>mawos-</i>	

乃 公 民 乎 惠 賜 比  
*n-o OPO-MI-TAKARA wo UTUKUSIBWI-TAMAp-i*  
 DV-ATTR great-HON-treasure ACC bless.INF-HON-INF

撫 賜 牟 止 奈母 隨神 所  
*NAGE-TAMAP-Am-u to namo KAMUNAGARA*  
 bring.peace.INF-HON-TENT-CONC COM PT follow.divine

思 行 佐久 止 詔 天皇 大 命  
*OMOP-OSIME-s-aku to NORITAMAP-U SUMYERA OPO-MI-KOTO*  
 think-HON-CAUS-NML DV.INF declare-ATTR ruler HON-HON-thing

乎 諸 聞 食 止 詔  
*wo MOROMORO KIKI-TAMAPE-YO to NORITAMAP-U*  
 ACC various listen.INF-HON-IMP DV.INF declare.HON-CONC

‘We declare that all you people will give ear to the Great Existence, the ruler, who will declare the divine thoughts and actions about how this land that is ruled under the heavens will be put into order, and ruled in peace, and how he will bless the myriad people under the heavens and bring peace to their hearts.’ (SM 1)

By adding a verbal auxiliary, any verb could theoretically be used to elevate the subject or the object. The most productive verbal auxiliary was *-as-*, as found in examples such as 行波須 *OKONAp-as-u* ‘acted (HON)’ (NT 4), or 婆婆止在須 *papa to AR-Asu* ‘who is (HON) your mother’ (SM 7). A number of the verbs in Table 7.14 could also be used as auxiliaries.

### 7.5.7 Adverbials

The most common way to form an adverbial in OJ was to attach *-ku* to the stem of a stative verb:

- (26) 狭 國 者 広 久 峻 國 者 平 久  
*SA-KI KUNI PA PIRO-ku SAGASI-KI KUNI PA TAPIRAKE-ku*  
 narrow-ATTR land TOP wide-INF steep-ATTR land TOP flat-INF  
 ‘[May] the narrow land be broad, and the steep land be flat . . .’ (NT 1)

7.5.8 Coordination

Coordination in OJ is constructed through a number of gerunds: *-ite* subjunctive, *-itutu* coordinative, and *-do(mo)* concessive. The subjunctive functions much like ‘and’, where actions can be strung together in a set order:

- (27) 大 海 爾 舟 滿 都都氣弓 自陸 往  
*OPO UMI ni PUNE MIT-I tutuke-te KUGA YORI YUK-U*  
 great sea LOC ship fill-INF continue-GER land ABL go-ATTR  
 道 者 荷 緒 縛 堅 弓 磐 根 木 根  
*MITI PA NI [N-O] WO YUP-I KATAME-te IPA-NE KWI-NE*  
 path TOP burden DV-ATTR string tie-INF tighten-GER rock-root tree-root  
 履 佐久彌弓 馬 爪 至 留  
*PUM-I sakum-ite UMA [NO] TUME [NO] ITAR-I TODOMAR-U*  
 step.on-INF divide-GER horse GEN hoof NOM reach-INF stop-ATTR  
 限 長 道 無 間 久 立 都都氣弓  
*KAGIRI NAGA MITI PIMA NA-ku TAT-I tutuke-te*  
 limit long path break not.be-INF stand-INF continue-GER  
 狭 國 者 広 久 峻 國 者  
*SA-KI KUNI PA PIRO-ku SAGASI-KI KUNI PA*  
 narrow-ATTR land TOP broad-INF steep-ATTR land TOP  
 平 久 遠 國 者 八十 綱 打  
*TAPIRAKE-ku TOPO-KI KUNI PA YASO TUNA UT-I*  
 flat-INF far-ATTR land TOP 80 ropes hit-INF  
 掛 弓 引 寄 如 事  
*KAKE-te PIK-I YOSUR-U KOTO [N-O] GOTO-KU*  
 attach-GER pull-INF close-ATTR thing DV-ATTR be.like-INF  
 ‘. . . with ships continuing to fill the great sea, and regarding the path that leads into the interior of the land, tying tight the ropes that hold the load [of tribute], and until the hooves of the horses step on and divide the rocks and trees, the horses shall continue to stand and continue down that long path constantly, [and we shall follow] as if the narrow land was broad, and the steep land was flat, as if 80 ropes had pulled together the far land . . .’ (NT 1)

The coordinative particle *-itutu* means that while A is going on, B is also happening, as the following example illustrates:

- (28) 烏梅乎 乎岐都都 多努 之岐乎倍 米  
*ume wo wok-itutu tanwosi-ki wope-me*  
 plum ACC invite-COOR merry-ATTR end-TENT.IMPV  
 ‘While I invite the plum we will make merry.’ (MYS 815)

The concessive gerund *-do* and its sibling, *-domo*, attach to the evidential form of the verb, and show that although A is happening, B is the reality:

- (29) 伊能知遠志 家 騰                      世 武 周弊 母 奈新  
*inoti wosi-kye-do                      se-m-u subye mo na-si*  
 life regrettable-EVID-although do-TENT-ATTR method INCL not.be-CONC  
 ‘Although my life is regrettable there is nothing I can do.’ (MYS 804)
- (30) 等伎波奈周                      迦久斯 母 何 母 等                      意母悶 騰母  
*tokipa nas-u                      kaku s-i mo ga mo to                      omop-e-domo*  
 eternal.rock do-ATTR thus do-INF INCLPT DV-INF think-EVID-although  
 ‘Although I think that I would like to remain as these eternal rocks . . .’ (MYS 805)

### 7.5.9 Subordinate clauses

The most productive way to nominalize in OJ was to attach the suffix *-aku* to verbs. This nominalizer was used after the attributive form of action verbs (as opposed to stative verbs). When affixed to the attributive form of the verb, the final *-u* of the attributive marker elided (*-uru* > *-ur-*). Consider the following examples, with the final example of a nominalized stative verb:

- (31) 是 東 人 波 常 爾 云 久  
*KORE ADUMABITO pa TUNE n-i IP-Aku*  
 this eastern.people TOP always DV-INF say-NML  
 ‘And this [is what] the people in the east are always saying . . .’ (SM 45)
- (32) 夜 渡 月 之 隱 良 久 惜 毛  
*YWO WATAR-U TUKWI NO KAKUr-aku WOSI MO*  
 night cross-ATTR moon GEN hide-NML regret.CONC PT  
 ‘It is even regrettable, the moon hiding that traverses the night [sky].’ (MYS 169)
- (33) 固 奈瀾 餓 那 居波佐麼 多智曾麼能  
*kwonami ga na kopas-aba tatisoba no*  
 main.wife NOM side.dish beg-if hawthorn GEN
- 未迺那鷄句塢                      居氣襖 被惠 禰 宇破奈利餓  
*mwi no na-kyeku wo kokisi piw-e ne upanari ga*  
 fruit GEN not-NML ACC much give-IMP PT consort NOM
- 那 居波佐麼 伊智佐介幾 未迺 於朋鷄句 塢 居氣襖被惠禰  
*na kopas-aba itisakaki mwi no opo-kyeku wo kokida piw-e ne*  
 side.dish beg-if Hisakaki fruit GEN many-NML ACC many give-IMPPT  
 ‘If your main wife asks for a side dish, give her but a small slice of meat, barren like the hawthorn. If your favorite wife asks for a side dish, give her a generous cut of meat like the abundant Hisakaki.’ (NS 7)

## 7.6 LEXICON

The OJ lexicon, as far as the extant texts illustrate, was fairly free from Chinese loans, a situation that would change as the Nara era came to an end, as the effects of the wholesale importation of Chinese government and laws resulted in a wave of Chinese loans. This is not to say there are **no** Chinese loans in OJ: *uma* ‘horse’ from 馬 LH \*ma, or *ume* ‘plum’ from 梅 LH \*me (compare Old Northwest Chinese \*mai). There were, however, a great number of peninsular loans, mainly from Paekche, that had entered OJ with the importation of Paekche technology, and Buddhism. This is quite

TABLE 7.15 PÆKCHE LOANS IN OLD JAPANESE

English gloss	OJ	From Paekche
father	<i>titi</i>	<i>kaso</i>
mother	<i>papa</i>	<i>omo</i>
mountain	<i>yama</i>	<i>mure</i>
village	<i>mura</i>	<i>suki</i>
river	<i>kapa</i>	<i>nari / nare</i>

clear because of a number of doublets in the language. Table 7.15 lists a number of the most important examples.

A number of important loans also came via Paekche when Buddhism was imported: *potoke* ‘Buddha’, *tera* ‘temple’, *kapara* ‘cement roof tile (for temples)’, and *pati* ‘bowl for alms’.

## NOTES

- 1 My transcription follows Frellesvig and Whitman (2004). An example like Ce<sub>1</sub> Ce<sub>2</sub> means that Ce without a subscript number is a syllable either lacking the distinction, or a merger has taken place before recorded records, and scholars are not completely sure of the underlying phonemic distinction.
- 2 The term ‘defective verb’ is from Alexander Vovin (2005).

## REFERENCES

- Bentley, John R. (2001) *A Descriptive Grammar of Early Old Japanese*, Leiden: E. J. Brill.
- Frellesvig, Bjarke and Whitman, John (2004) ‘The vowels of proto-Japanese’, *Japanese Language and Literature*, 38: 281–99.
- Ienaga, Saburō, Inoue, Mitsusada, Ōno, Susumu, and Sakamoto, Tarō (eds) (1986) *Nihon shoki*, 2 vols, Tokyo: Iwanami Shoten.
- Kitagawa, Kazuhide (1982) *Shoku Nihongi Senmyō: Kōhon, Sōsakuin*, Tokyo: Yoshikawa Kōbunkan.
- Lange, Roland (1972) *The Phonology of Eighth-Century Japanese: A Reconstruction Based upon Written Records*, Tokyo: Sophia University Press.
- Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven: Yale University Press.
- (1991) *A Reference Grammar of Japanese*, Tokyo: Tuttle.
- Matsumori, Akiko (2008) ‘On the reconstruction of the proto-accentual system of Japanese’, in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese: Issues and Prospects*, Amsterdam: John Benjamins.
- Miyake, Marc Hideo (2003) *Old Japanese: A Phonetic Reconstruction*, London: RoutledgeCurzon.
- Onoda, Mitsuo (1977) *Kojiki*, vol. 1 of *Koten-hen*, Tokyo: ST, Shintō Taikē Hensankai.
- Satake, Akihiro, Kudō, Rikio, Ōtani, Masao, Yamada, Hideo, and Yamazaki, Yoshiyuki (eds) (1999–2003) *Man’yōshū*, 5 vols, Tokyo: Iwanami Shoten.
- Takayama, Michiaki (1981) ‘Gen’on seichō kara mita Nihon shoki ongana hyōki shiron’, *Gobun Kenkyū*, 51: 13–121.
- (1984) ‘Nihon shoki no ongana to sono gen’on seichō ni tsuite: jōdai akusento to no sōkansei o kangaueru’ in *Kindaichi Haruhiko Hakase Koki Kinen Ronbunshū*, vol. 1: 45–69.
- Vovin, Alexander (2005) *A Descriptive and Comparative Grammar of Western Old Japanese*, London: Global Oriental.

# CLASSICAL JAPANESE\*

*Nicolas Tranter*

## 8.1 INTRODUCTION

The year 794 saw the move of the Japanese capital westwards from Nara to a new site, modern Kyoto. The new capital was named *Feian-kyau* 平安京 ‘capital of eternal peace’, though it was more commonly known simply as *kyau* 京 ‘capital’ or *miyako* 都/京. Despite a brief relocation southwards in 1180, a rival emperor and imperial capital between 1336 and 1392, and the location of actual government in different cities further east during the Kamakura (1185– or rather 1192–1333) and Edo (1603–1868) periods of shogunal rule, and the lack of any centralized control during most of the period in-between, the city remained the official imperial capital of Japan till the late nineteenth century. The 200 years after the establishment of the capital saw the golden age of Japanese literature, the Heian period (794–1185), and the literary language that developed during this period was the basis of most Japanese-language writing till the start of the twentieth century.

This chapter has been deliberately loosely named ‘Classical Japanese’ (CJ), thus covering both the language of the Heian period, strictly known as Early Middle Japanese (EMJ), and the use of this as a fossilized literary form, at least in much of its grammar, in later periods including the modern period. In Japan the latter is known as **bungo** 文語 ‘literary language’ – in contrast to **kōgo** 口語 ‘vernacular language’ – whereas the language of the earlier period tends instead to be known as **kogo** 古語 ‘old language’. Despite its later association as the language of the golden age of Japanese literature, the tendency in the Heian period was for men to write prose in Classical Chinese. As a result, much of the classic literature of the earlier and mid-Heian period was composed by women, such as *Kagerō nikki* 蜻蛉日記 ‘The Dragonfly Diary’ (c. 975, a daughter of Fujiwara no Tomoyasu), *Makura no sōshi* 枕草子 ‘The Pillow Book’ (1002, Sei Shōnagon), *Genji monogatari* 源氏物語 ‘The Tale of Genji’ (c. 1010, Murasaki Shikibu), *Sarashina nikki* 更科日記 ‘The Sarashina Diary’ (early 11C, a daughter of Sugawara no Takasue), etc. A common genre was the composition made up of short entries, i.e. travelogues, diaries or notebooks, as exemplified by three of the above. There is, of course, male prose writing too. The first attributable prose composition of the period – and the earliest extant example of the short-entry genre – was *Tosa nikki* 土佐日記 ‘The Tosa Diary’ (935, Ki no Tsurayuki), written by a major male poet of the period. It concerns the slow boat-trip of the retiring governor of Tosa province – Tsurayuki himself – and his retinue back to the capital, but the author uses the conceit of writing in the first person as a female member of his own retinue, showing that both the genre and Japanese itself as a written form had already become associated with female writing.

By the late Heian and early Kamakura periods, however, there is a noticeable increase of highly regarded male writing in Japanese, such as *Hōjōki* 方丈記 ‘Description of a Monk’s Hut’ (1211, Kamo no Chōmei), and late-Kamakura *Tsurezuregusa* 徒然草 ‘Jottings in Idleness’ (1332, Yoshida Kenkō), a male equivalent of Sei Shōnagon’s notebook

style. We also encounter the development of literature aimed not just at the aristocratic elite but also at the military classes and the common people, such as the military chronicle *Heike monogatari* 平家物語 ‘Tale of the Taira’ (early 13C, anon.), the Buddhist tales and parables of *Konjaku monogatari-shū* 今昔物語集 ‘Collection of Tales of Times Now Past’ (early 12C, anon.) and *Uji shūi monogatari* 宇治周囲物語 ‘Tales Picked from Uji’ (mid-12C, anon., though the Uji is meant to refer to Minamoto no Takakuni).

*Tsurezuregusa* and *Heike monogatari* are clearly Late Middle Japanese (Irwin and Narrog, this volume), but the authors of *Hōjōki*, *Konjaku monogatari-shū* and *Uji shūi monogatari* were alive at the end of the Heian period, and it is useful to include them as EMJ, not least because their deviations from the CJ ‘standard’ give insights into how the Heian language was evolving. The language of texts of most of the Heian period is thought to approximate to the actual spoken language, whereas later Heian texts onwards are arguably more artificial and some grammatical forms more fossilized, e.g. the *-er<sup>R</sup>*- and *-n-* verb-forms probably vanished in later Heian spoken language because they are largely restricted to set forms (e.g. *-tamaf-er-i* honorific, *nar-i-n-u* ‘became’). Therefore, though their continued use in later texts does not represent productive grammar, it is *how* they are used in these texts that shows us how the spoken language has changed. I consider even the classics to have a degree of artificiality just as any written language is not the same as speech, as evidenced by the rarity of *t-‘onbin’* sandhi in writing even though sporadic examples occur at the very start of the period; however, despite their use of fossilized verb-forms, later texts in their sentence syntax still display writing-as-you-speak, such as mid-sentence afterthoughts.

## 8.2 PHONOLOGY

Phonologically, EMJ represents a development from Western Old Japanese. Eastern Old Japanese vanishes totally from the written record, except for *kekere* = WOJ/EMJ *kokoro* ‘heart’ in one poem collated in *Kokin waka-shū* (920).

### 8.2.1 Consonants

EMJ likely distinguished 14 consonant phonemes, one of which was peripheral (Table 8.1). The prenasalized consonants were phonetically [<sup>h</sup>b], [<sup>h</sup>d], [<sup>h</sup>g] and [<sup>h</sup>(d)z]; *kana* orthography largely did not distinguish them from *f*, *t*, *k* and *s*. The distinction of *f* and *w* word-medially was lost, or at least confused, quite early on. Word-medial *f* became *w*, and the change was probably complete in the first half of the Heian period. As a result, there

TABLE 8.1 EARLY MIDDLE JAPANESE CONSONANTS

	Bilabial	Dental/ Alveolar	Palatal	Velar
Plosive:	{p}	t		k
Prenasalized:	b	d		g
Nasal:	m	n		
Fricative:	f	s		
Prenasalized Fricative:		z		
Flap:		r		
Approximant:	w		y	

is an orthographic confusion of <fV> and <wV> word-medially in manuscripts. In the later Heian period, *w* was lost before front vowels, so *we* > *ye* and *wi* > *i*.

*s* and *z*, and perhaps *t*, *d* and *n*, are likely to have been distinctly palatalized before a front vowel, e.g. *si* [sʲi] or [ɕi] or [ʃi]. No consonant clusters were allowed in native words, but *kw*, *gw* and sequences of consonant + *y* occur in Chinese loans. *C* + *y* will have displayed the same palatalization as *C* + front vowel. Thus, *sy* would have been [sʲ] or [ɕ] or [ʃ], as it corresponds to palatal or retroflex consonants in Chinese and is a palatal consonant in later Japanese. /p/ vs /ɸ/ is of uncertain status: we know a single OJ phoneme *p* in later Japanese results word-initially in LMJ *f*- > **h**-/f- in native content and grammatical words but /p/- in ideophones, but when this split happened, i.e. when OJ *p*- fricated, is unclear.

Gemination applied in bimorphemic Chinese loans where it is the result of assimilation of Middle Chinese -*t*/-*k*/-*p* + J *p*-/*t*/-*k*/-*s*-, e.g. EMC 日記 *nyit.ki* ‘diary’.<sup>1</sup> However, where it occurs in Modern Japanese, it is normally not orthographically represented in EMJ, e.g. Modern Japanese **niō.ki** ‘diary’ corresponds to older <ni.ki>. The rare example of native word gemination, e.g. *uotafu* < OJ *urutapu*, shows that it must have existed in the spoken language and not just have been a feature of learned pronunciation of Chinese.

EMJ also had a nasal segment *N* that did not exist in Old Japanese, and which was most likely a moraic consonant like its modern descendent. *N* has four sources: (1.) Certain sequences of *nV* and *mV*, especially *mu* and especially word-finally, reduce to *N*, e.g. *namu* (emphatic particle) > *nan*; (2.) *-ru*, almost always a (compound with a) form of *ari* ‘be (located)’, often reduces to *N* before *n* or *m* in certain grammatical forms, e.g. *ar-umeri* ‘he seems to be (located)’ ~ *anmeri*; (3.) Word-medial onbin sandhi, prenasalizing the following consonant where possible (see below); (4.) Middle Chinese word-final *-n* and *-m* > *-N* in loanwords: 山 *sren* ‘mountain’ and 三 *sam* ‘three’ both give *san*. Middle Chinese *-ŋ*, on the other hand, reduces to *-u/-i*: 樣 *yang* ‘manner’ > *yau*. Since Middle Chinese word-final *-t* and *-k* result in *-tu/-ti* and *-ku/-ki*, we may hypothesize an earlier stage of \**yanju* ‘manner’, but there is no manuscript evidence to support the existence of a /ŋ/ phoneme in EMJ; it would have been peripheral and short-lived at best. As *N* derives often from earlier *mu*, the same set of symbols serve to represent both *N* and *mu* in early hentaigana. For example, it is the prenasalization of *g* in <ya.mu.go.to.na.si> ‘be impressive’ that shows that it is actually *yangotonasi*. Before *mV* and *nV*, moreover, *N* is often not even written, so *anmeri* above is written <a.mu.me.ri> or <a.me.ri>.

In Chinese loanwords, we occasionally have examples of what we may call ‘nasal drift’ in later forms of words, whereby *N* + (*w/y*)*V* results in *N* + *n*(*y*)*V*, e.g. EMC 天皇 *then-hwang* > EMJ *ten.wau* ‘emperor’ > later **ten.nō**. There are a few cases where nasal drift gives *N* + **m**(*y*)*V* rather than *N* + **n**(*y*)*V*, in all cases the *N* corresponding to Middle Chinese *-m*, e.g. EMC 陰陽 *pin.yang* > EMJ *on.yau* ‘yin-yang’ > *on.myau* > **on.myō**. This may point to a distinction of /*N*/ vs /*M*/, i.e. the EMJ form above is actually *om.yau*, but there is no other evidence of a distinct /*M*/. We may, in fact, be looking at distinctions in Middle Chinese preserved originally only in very learned idiolects of Japanese, or even at nasal drift in the source form.

### 8.2.2 Vowels

EMJ had five vowels /i e o u/. OJ *i* [i] and *wi* [i] had merged as /i/, *ye* [e] and *e* [əy] as /e/, and *wo* [o] and *o* [ə] as /o/. The only restrictions that apply to consonant + vowel sequences involve approximants: there is no orthographic and probably no phonemic distinction between /u/ and /wu/ (= conventionally transcribed *u*), and /i/ and /ji/ (= *i*).

During the Heian period, orthographic confusion points to the developing merger of /o/ and /e/ with /wo/ and /je/ (= *ye*). (See Martin 1987: 79 for a timescale summary of changes.)

Native words do not have sequences of vowels; rare examples, such as <siuto> ‘father-in-law’, probably represent \**siwuto*. Sequences of vowels *iu*, *ei*, *eu*, *ai*, *au*, *ou*, *ui*, *uu* are frequent in Chinese loans.

Another source of the sequence *au* is a morpheme-final contraction of *awV* (or of *afV* > *awV*), e.g. *mawir-* ‘go; come (HUM)’ > *maur-* via \*/*maw*r-/. Cases tend to be lexicalized even in the earliest Heian period, e.g. *tukafe-matur-* ‘serve; do (HUM)’ > *tukau-matur-*, \**mawi(ri)-ide-* ‘go [out]; come [out] (HUM)’ > *maude-*, suggesting the contraction, as well as *f/w* merger, dates back to the mid-ninth century. As this involves vowel loss and the syllabification of the preceding *w*, the change is probably part of the onbin phenomenon (see below).

### 8.2.3 Suprasegmentals

EMJ was a pitch-accent language, in which each mora was either high-pitch (H) or low-pitch (L). Words may either have continuous pitch throughout or a change of pitch between morae; some words have a second change of pitch, always H to L, which means that we cannot talk in terms of an ‘accented’ mora as we do with regard to the modern Tokyo Japanese. The basic patterns (without indication of number of morae in each pitch) are therefore H, L, HL, LH and LHL. A limited number of words have a change of pitch *within* a mora, often due to the imposition of a two-mora pitch pattern on a monomoraic word, e.g. LH on *te* ‘hand’ gave a rising pitch. Pitch is not represented normally in writing, but there are certain texts, most notably the lexicographic *Ruiju myōgishō* of 1081, which include a Chinese-derived tone-dot system (Martin 1987: 167). The following account is based on Martin’s (1987: 138ff) extensive discussion of accent in Japanese.

As in the modern standard language, verbs and adjectives have just two accent-classes. The two verb accent-classes, for example, respectively have fundamental H pitch and fundamental L pitch, and basic inflections – by which we largely mean the different ‘stems’ – are divided between those that take the opposite pitch to that of the preceding mora, and those that do not. The conclusive *-u*, for example, takes the opposite pitch, whereas the attributive *-u/-uru* involves no pitch change. The infinitive *-i*, moreover, appears to be the opposite pitch as a verbal form, but not as a deverbative noun, suggesting the conventional ‘infinitive’ actually represents two different morphemes. An additional rule that affects the accent pattern in verbs appears to be that the first H mora can be no later than the third mora of the word, so in fundamentally L pitch verbs, where a change of pitch of LLLH might be expected in a four-mora form it is realized as LLHL.

Nouns, in contrast, display more variety of accent patterns. Whereas the presence or absence of pitch change in verbs is determined by the basic inflection (‘stem’), the presence or absence of pitch change in nouns is lexically determined; if a noun has pitch change, the mora on which it occurs is also lexically determined, multiplying the number of classes. Following Martin (1987), Japanese historical noun classes are named by a formula *x.y*, where *x* refers to the number of morae in the noun, and the formula covers all possible contrasts in not only EMJ but also all Modern Japanese varieties. Table 8.2 includes the five EMJ classes for two-mora nouns, as well as the modern Kyoto (four classes) and modern Tokyo (three classes) equivalent patterns. There is much research dealing with whether the EMJ/Kyoto or the Tokyo patterns are closest to proto-(Mainland) Japanese (pJ). Shimabukuro (2007) reconstructs the latter as essentially those of EMJ.

TABLE 8.2 TWO-MORA NOUN ACCENT IN JAPANESE-RYUKYUAN

	2.1	2.2	2.3	2.4	2.5
Tokyo:	LH(-H)		LH(-L)		HL(-L)
Kyoto:	HH		HL	LH	LF
	HH-H		HL-L	LL-H	LH-L
EMJ:	HH(-H)	HL(-L)	LL(-L)	LH(-H)	LH*(-L)
	<i>kaze</i>	<i>fito</i>	<i>yama</i>	<i>kata</i>	<i>ame</i>
	‘wind’	‘person’	‘mountain’	‘shoulder’	‘rain’
pJ:	HH(-H)	HL(-L)	LL(-L)	LH(-H)	LH(-L)
pJR:	2.1	2.2	2.3 B and C	2.4 B and C	2.5 B and C
pR:	A		B	C	
	HH(-L)		LH(-L)	LHH(-L)	
Shuri:	HL(-L)			LL(-L)	

\*: In isolation, H(igh) or F(all) is marked in *Ruiju myōgishō*.

However, proto-Ryukyuan (pR) is reconstructable as having three patterns, of which what we may term ‘A’ corresponds to 2.1 and 2.2, while both ‘B’ and ‘C’ include nouns in each of classes 2.3, 2.4 and 2.5, which Shimabukuro (2007) explains as due to the existence of vowel-length in some proto-Japanese-Ryukyuan (pJR) nouns (corresponding to Ryukyuan C) and different patterns of mergers in Japanese and Ryukyuan from the eight pJR class types. Modern dialects of both Japanese and Ryukyuan display the results of further mergers (e.g. 2.2 and 2.3 in Kyoto and Tokyo; 2.4 and 2.5 in Tokyo; B and C in Shuri) and splits (see ‘Tsuruoka’, Matsumori and Onishi, this volume). A simplified presentation of Japanese-Ryukyuan accent correspondences based on Shimabukuro (2007) is given in Table 8.2, where -H and -L indicate the boundary and pattern when a particle without its own accent is attached.

Noun-following particles, as well as verb/adjective-following tense/aspect/negation markers and auxiliaries, largely have their own independent accent in the *Ruiju myōgishō*, e.g. topic *-fa*, comitative *-to* and gerund *-te* are H; emphatic *zo*, quotative *to* and adjective attributive *-ki* are L; and focal *-koso* is HL (Wenck 1959: 403). Only a couple appear to be accentually enclitic (though onbin shows *-te* to be morphologically enclitic), having pitch that is determined by the accent class of the preceding word, most notably genitive *-no*.

### 8.2.4 Morphophonology

A factor affecting morphology is a form of euphony applying to monomoraic grammatical morphemes attached to another monomoraic morpheme. If the two are identical, the sequence is avoided. This applies to the marked past morpheme *-ki* when attached to *su* ‘do’ and *ku* ‘come’. Its expected conclusive and attributive forms are *si-ki* and *\*si-si*, and *\*ki-ki* and *ki-si*. *\*ki-ki* is avoided entirely, though its humble equivalent *mawiri-ki* occurs; while *\*si-si* is replaced by the ‘irregular’ *se-si*. The same process applies to the conclusive of *si/zi*-root adjectives, itself typically a morpheme added to an older root. For example, the conclusive of *fosi*<sup>A</sup>- ‘want’ and *tanomo-si*<sup>A</sup>- ‘be reliable’ (related to *tanom-* ‘rely on’) are not *\*fosi-si* and *\*tanomo-si-si*, but *fosi-Ø* and *tanomo-si-Ø*.

Prenasalization and concomitant voicing (corresponding to Modern Japanese sequential voicing, or 連濁 *rendaku*) is frequent in noun-noun or noun-verb compounds, and compounds sometimes co-exist with their sources with the noun particles *no* or *ni*, e.g. *ko-no suwe* [konosuwe > konosuje] ~ *kozuwe* [ko<sup>n</sup>(d)zuwe > ko<sup>n</sup>(d)zue] ‘treetop’. Prenasalization in both *rendaku* and in onbin (below) is blocked if the second element of a compound already contains a prenasalized segment, e.g. *kokage* ‘tree’s shade’ but not *\*kogage*.

Other sandhi (known as onbin, 音便) constitute morphophonemic changes whereby a mora is changed to *N*, *Q*, *U* or *I*, or *Ū* or *Ī*. *U* and *Ū* are phonemically both /u/, and *I* and *Ī* both /i/, but *N/Ū/Ī* impose prenasalization (resulting from the nasality of the affected mora) on the following consonant where this is possible, while *Q/U/I* do not. This occurs especially: (1.) in certain compound words, commonly with *-miti* ‘path’ and *-fito/-bito* ‘person’, which result in *-udi* ~ *-Ndi* /u<sup>n</sup>di ~ N<sup>n</sup>di/ and *-udo* ~ *-Ndo* /u<sup>n</sup>do ~ N<sup>n</sup>do/ respectively; (2.) between a quadrigrade (consonant-root) verb and a following *t*-initial morpheme, namely *-te*, *-tari* and *-tamafu/-tabu*, but not *-tu*. Examples of this *t*-sandhi occur in early Heian texts, e.g. *yondaru* (actually written <yo.mu.ta.ru>) and *yontabi* <yo.mu.ta.fi> for *yom-i-tar-u* (recite-INF-PERF-ATTR) and *yom-i-tab-i* (recite-INF-HON-INF) occur in *Tosa nikki* (Month 1 Days 7 and 20); cf. modern **yon-da** (read-PAST). This and the fact that it applies to quadrigrade verbs but not bigrade verbs, e.g. quadrigrade *uti-te* > *uQ-te*, but not bigrade *oti-te* > *\*oQ-te*, suggest that it originated when the stem-final *i* of the two conjugations were phonetically different, i.e. OJ *i* vs *wi*. (This in turn raises questions in the case of *ti*, which was already neutralized in OJ.) Despite the early, probably pre-Heian, emergence of *t*-sandhi, it is only sporadic in texts for over 600 years, as manuscripts tended to be conservative in spelling and grammar. Nevertheless, a few onbin-derived forms are standard, e.g. *ikaga/ikade* ‘how’ < *ika-ni* + *-ka/-te*. Rhythmic considerations may have been a factor, as shown by *ob-oye-* ‘think/feel (involuntarily)’ and *ob-os-* ‘think/feel’ (honorific), but *omof-*, never *\*ob-* ‘think’.

Frellesvig (1995) discusses in detail the phonetics of the change as a reduced style of diction. Tables 8.3 and 8.4 are adapted from Frellesvig (1995: 108–19); I have

TABLE 8.3 CLASSICAL JAPANESE ONBIN

OJ	EMJ		
	Full	Reduced	Resulting onbin
pi	/phi/	[ɸ <sup>j</sup> ]	u/Q
pu	/phu/	[ɸ <sup>w</sup> ]	u/Q
ki	/ki/	[ç <sup>]</sup>	i
ku	/ku/	[x <sup>w</sup> ]	u
bi	/ <sup>m</sup> bi/	[ <sup>m</sup> β <sup>j</sup> ]	ū/N
bu	/ <sup>m</sup> bu/	[ <sup>m</sup> β <sup>w</sup> ]	ū
gi	/ <sup>ŋ</sup> gi/	[ <sup>ŋ</sup> j <sup>]</sup>	ī
gu	/ <sup>ŋ</sup> gu/	[ <sup>ŋ</sup> ɣ <sup>w</sup> ]	ū/N
mi	/mi/	[m <sup>i</sup> ]	ū/N
mu	/mu/	[m <sup>w</sup> ]	ū/N
ni	/ni/	[n <sup>j</sup> ]	N
nu	/nu/	[n <sup>w</sup> ]	N
ri	/ru/		Q
ru	/ri/		Q

TABLE 8.4 EXAMPLES OF T-SANDHI IN CLASSICAL JAPANESE QUADRIGRADE VERBS

kogi-te	/ko <sup>o</sup> gi-te/	>	[ko <sup>o</sup> j <sup>i</sup> te]	koi-de	/koi- <sup>o</sup> de/
yaki-te	/jaki-te/	>	[jaç <sup>i</sup> te]	yai-te	/jai-te/
yomi-te	/jomi-te/	>	[jom <sup>i</sup> te]	yON-de	/jon- <sup>o</sup> de/
yupi-te	/juφi-te/	>	[juφ <sup>i</sup> te]	yuQ-te	/juq-te/

included *ru/ri* onbin, as it also enters in *t*-sandhi, though Frellesvig (1995: 38–9) excludes it on a number of grounds, treating it as a separate phenomenon.

Where there are two possible outputs, e.g. *mi* > *ū/N*, dialect differences appear to be involved, e.g. *ū* is typical of the capital and *N* of western areas. (See Frellesvig 1995: 136–49 for the identification of three theoretical ‘varieties’.) Whether onbin occurs or not may be determined by morphological or syntactic factors. For example, the adjectival infinitive *-ku* frequently alternates with *-u*. However, this onbin is much more frequent with *si/zi*-root adjectives than with other adjectives. Moreover, *-ku + ari-* > *-kari-* is immune to onbin, even in uncontracted form: *wokasi-kar-azu* and *wokasi-ku-fa ar-azu* ‘isn’t nice’, vs \**wokasi-u-fa ar-azu*, etc.

### 8.3 SCRIPT

Old Japanese was written entirely in Chinese characters, whether entirely phonographically (poetry of *Kojiki* and *Nihon shoki*, and much of the poetry of *Man’yōshū*), in a mixture of phonograms and semantograms (the rest of the poetry of *Man’yōshū*), or in pseudo-Classical Chinese (*hentai kanbun*, the prose of *Kojiki*). In contrast, the early part of the Heian period saw the development of two syllabaries that were visually distinct from Chinese characters. These came to be known as kana 仮名 ‘temporary names’, connected with the idea of the ‘temporary borrowing’ (**kashaku** 仮借, = Chinese *jiǎjiè*) of characters to write homophones that is one the six categories of Chinese characters established by the Chinese *Shuōwén jiězì* 說文解字 (c. AD 100). The one system was hiragana, but in more recent times has been termed ‘hentaigana’ to distinguish it from the narrow use of ‘hiragana’ for its modern descendent characterized by a one-to-one kana-to-syllable relationship. Although modern published editions always transcribe the hentaigana of manuscripts into modern hiragana, hentaigana retained the principle of man’yōgana of typically having more than one kana available per phonemic syllable; a one-to-one relationship was only established in the second half of the nineteenth century. Hentaigana was derived from the fast-written grass-script style of characters. Unlike in China, however, a functional difference between grass script = phonogram vs square or cursive script = semantogram made hentaigana a new script, separate from Chinese characters.

The other kana system to develop was katakana. This evolved from the phonographic use of characters in the margin of texts composed as Chinese entirely in character script to indicate how the characters should be read out in Japanese. However, each such character was written only with no more than four of its distinctive strokes. During the Heian period, katakana migrated from the margins of texts into the texts themselves. Unlike hentaigana, which maintained a many-to-one letter-to-syllable relationship, katakana virtually had a one-to-one relationship. It did include the occasional polysyllabic device, such as 𛀀 *site*. Because it was used to gloss Chinese texts, katakana had a male rather than a female association.

The EMJ period saw therefore the establishment of three scripts – character script, hentaigana and katakana – and different texts chose different combinations of these in a startling variety of orthographies, ranging from writing clearly in Japanese (*wabun*) to writing in Chinese (*kanbun*) or encoded in Chinese but intended to be decoded into Japanese (*kanbun kundoku*). The female(-style) writing of the Heian period was a mixed script of hentaigana and character script, though hentaigana predominated, but in the early Heian *Tosa nikki* (Table 8.5 (a.)) only the diary dates are in characters – everything else is entirely hentaigana. Other texts, such as the early Kamakura *Hōjōki* (Table 8.5 (c.)) were written in a mixed script of katakana and character script, again with katakana predominating. Another style of writing in the period is exemplified by the late Heian *Konjaku monogatari-shū* (Table 8.5 (b.)): absolutely anything that can be written in character script is, and a few grammatical endings are written in Chinese prepositional order, especially causative 令  $V = V-[a]sime-$ , passive/honorific 被  $V = V-[r]are-$  (at the end of (b.)), debitive 可  $V = V-ube^A-$ , negative 不  $V = V-[a]zu/V-[a]zar^R-$ ); whatever remains is written in small katakana. Other texts, such as a later version of *Hōjōki* (Table 8.5 (d.)) were written entirely as Classical Chinese, using only right-marginal katakana to indicate the pronunciation of occasional words and to indicate endings, and left-marginal 二 ‘two’ and 一 ‘one’ to indicate which order the Chinese should be read out to be Japanese word order, and the symbol ㄥ between two adjacent characters to indicate they should be read in reverse order. Because the match between Chinese and Japanese is difficult, certain endings are doubly written. For example, the negative ending  $-[a]zu$  of *tayezu site* ‘doesn’t cease’ is written 不 Chinese-style before the verb *taye-* 絶 whereas the gerund ending *site* is written 而 after the 絶; however, because the negative gerund form is a sequence  $-[a]zu site$  in CJ, the *site* is also written ず in katakana beside the 不. *Heike monogatari* (Table 8.5 (e.)) illustrates a three-script mix: the main body uses characters and hentaigana mixed script, with katakana resuming its historical role of indicating the pronunciation in the margin; the katakana – but not the hentaigana – in (e.) is the only script in the five illustrations to distinguish *f/b*, *t/d*, *s/z*, *k/g*, using the same diacritics as Modern Japanese kana. This lack of distinction in most texts allowed graphic puns, e.g. *fati na-si* ‘hasn’t got the bowl’ ~ *fadi na-si* ‘has no shame’ in *Taketori monogatari*, summarizing a suitor throwing away the bowl he has tried to pass off as Buddha’s and still trying to woo the girl who has asked for it.

Hentaigana-character mixed script ultimately became the norm in later periods and is the basis of modern Japanese writing. However, some other combinations still occurred in relatively recent times. For example, letter- and postcard-writing in the first half of the twentieth century used a style known as *sōrōbun* that is characterized linguistically by the use of classical grammar with the polite ending *-safurafu -sōrō* and orthographically by being written predominantly in characters, with minimal hiragana and with an even greater use of Chinese-derived sequences than in *Konjaku monogatari-shū*.

## 8.4 MORPHOLOGY

### 8.4.1 Nouns and particles

Nouns and pronouns are ambiguous as to grammatical number, i.e. singular vs plural, e.g. *fito* = ‘person’ or ‘people’. There are explicit plural endings used with sentient referents and, exceptionally, inanimates: *-domo*, honorific *-tati* or humble *-ra*: *fito-domo* ‘people’, *koto-domo* ‘matters’, *mi-ko-tati* ‘princes’, *takumi-ra* ‘artisans’ (all *TM*). The few reduplicated nouns have meanings that are more than just unspecific plurals,

**TABLE 8.5 VARIETY OF MIDDLE JAPANESE ORTHOGRAPHIES**

(a) *Tosa nikki*:

十九日 ひあし<sup>シ</sup> 逢<sup>フ</sup>はふねいささす  
 廿日 き<sup>シ</sup>はふねいささすみ<sup>ミ</sup>ひと、う<sup>レ</sup>れへな<sup>ナ</sup>き<sup>キ</sup>く、る<sup>ル</sup>しくこ、僕<sup>ボク</sup>もと<sup>ト</sup>み  
 け<sup>ケ</sup>逢<sup>フ</sup>はさ、ひのへぬ<sup>ヘ</sup>る<sup>ル</sup>す<sup>ス</sup> 嶺<sup>ミネ</sup> 今<sup>イマ</sup>ふ<sup>フ</sup>い<sup>イ</sup>く<sup>ク</sup> む<sup>ム</sup>つ<sup>ツ</sup> み<sup>ミ</sup>そ<sup>ソ</sup> とか<sup>ト</sup>さ<sup>サ</sup>ふ<sup>フ</sup>逢<sup>フ</sup>は<sup>ハ</sup>お<sup>オ</sup>よ<sup>ヨ</sup>ひ<sup>ヒ</sup>も<sup>モ</sup>う<sup>ウ</sup>こ<sup>コ</sup>み<sup>ミ</sup>ハ  
 逢<sup>フ</sup>ぬ<sup>ヌ</sup>へ<sup>ヘ</sup>し

TOWOKA [amari] KOKON UKA *fi a si ke re fa fu ne i ta sa su*  
 FAT UKA *ki no fu no ya u na re fa fu ne i ta sa su mi na fi to ” u re fe na ke ku ” ru si ku*  
*ko ” ro mo to na ke re fa ta ” fi no fe nu ru ka su wo ke fu i ku ka fa tu ka mi so ka to ka so*  
*fu re fa o yo fi mo so ko na fa re nu fe si*

‘19th day: The day was terrible, so we didn’t push the boat out.

20th day: It was like yesterday, so we didn’t push the boat out. Everyone was complaining and moaning. They were terribly anxious, so their fingers could have even been hurt just by their [trying to] count [on them] the number of days that had past – 20? 30?’

(b) *Konjaku monogatari-shū*, 27/10:

(<http://edb.kulib.kyoto-u.ac.jp/exhibit/konjaku/frame/kj27/kj27fr07.htm>)

今昔延喜, 御代仁壽殿臺代, 御燈油夜半許, 物来取南殿様去ル事毎夜有ル比有テリ  
 天皇此レ目キナマシ事思食シテ何テ此レヲ見顯ト被仰ニ

IMA [fa] MUKASI YEN GI no MI YO ni ZIN ZYU DEN no TAI SIRO no ON TON ABURA wo  
 YO NAKA BAKARI ni MONO KI te TORI te NA DEN ZAMA ni SARU ni KOTO MAI YA ni  
 ARU ni KORO ARI ke ri TEN WAU KORE re wo ME sa ma si ki KOTO ni OBOSI MESI si te IKA de  
 KORE re wo MI ARAFASU sa mu to RARE OFOSE ke ru ni

‘Long ago, during the Engi era, there was a time around midnight every night that something came and took the lamps on the walkway to the [abandoned] Jinjuden Palace and left [with them] in the direction of the Naden Palace. The Emperor thought this annoying, and said “How shall find out what this is?” ...’

(c) *Hōjōki* (Daifukubon MS):

ユク河ノナカレハタエスシテシカモ、トノ水ニアラスヨトミニウカフウタカタハカツ  
 キエカツムスヒテヒサシク、マリタルタメシナシ

yu ku KAFA no na ka re fa ta ye su si te si ka mo ” to no MIDU ni a ra su yo to mi ni u ka  
 fu u ta ka ta fa ka tu ki ye ka tu mu su fi te fi sa si ku to ” ma ri ta ru ta me si na si

‘The flow of the rolling river is ceaseless, yet it is not the original water. The bubbles that float up in a whirlpool first vanish then reform but never stay around for long.’

(d) *Hōjōki* (Shinjibon MS):

行<sup>ユク</sup>川<sup>カハ</sup>之水<sup>ノミヅ</sup>不<sup>ズ</sup>絶<sup>ズ</sup>而<sup>シテ</sup>然<sup>ル</sup>非<sup>ズ</sup>本<sup>ノ</sup>水<sup>ニ</sup>濺<sup>ル</sup>浮<sup>ル</sup>転<sup>ル</sup>浮<sup>ル</sup>且<sup>シテ</sup>消<sup>ス</sup>且<sup>シテ</sup>結<sup>ス</sup>久<sup>ク</sup>無<sup>ク</sup>留<sup>ル</sup>事<sup>ナシ</sup>

YUKU KAFA no NAGARE ZU TAYE TE SIKI NIARAZU 2 MOTO MIDU 1  
 YODOMI UKABU UTA KATA KATU KIYE KATU MUSUBI FISASI NASI 2  
 TODOMARU KOTO 1

(e) *Heike monogatari*:

東山<sup>フモトシノ</sup>乃<sup>ニ</sup>麓<sup>タニ</sup>鹿<sup>カ</sup>乃<sup>ニ</sup>谷<sup>コ</sup>と云<sup>フ</sup>所<sup>ト</sup>云<sup>フ</sup>所<sup>ト</sup>ハう<sup>シ</sup>ろ<sup>シ</sup>ろ<sup>シ</sup>ハ三<sup>ミ</sup>井<sup>イ</sup>寺<sup>ジ</sup>、<sup>シヤククワフ</sup>ほ<sup>ハ</sup>い<sup>イ</sup>て<sup>テ</sup>ゆ<sup>ユ</sup>、し<sup>シ</sup>城<sup>シヤククワフ</sup>郭<sup>クワフ</sup>み<sup>ミ</sup>て<sup>テ</sup>そ<sup>ソ</sup>あり<sup>リ</sup>け<sup>ケ</sup>  
<sup>シユンクワンソウツ</sup>る<sup>ル</sup>俊<sup>シユン</sup>寛<sup>クワン</sup>僧<sup>ソウツ</sup>都<sup>ツ</sup>乃<sup>ニ</sup>山<sup>サン</sup>莊<sup>ザウ</sup>あり<sup>リ</sup>かれ<sup>レ</sup>、<sup>サンザウ</sup>ほ<sup>ハ</sup>ね<sup>ネ</sup>き<sup>キ</sup>よ<sup>ヨ</sup>り<sup>リ</sup>あ<sup>ア</sup>ひ<sup>ヒ</sup>、<sup>メダ</sup>平<sup>ヘイ</sup>家<sup>カ</sup>乃<sup>ニ</sup>ろ<sup>ロ</sup>ほ<sup>ホ</sup>さ<sup>サ</sup>む<sup>ム</sup>す<sup>ス</sup>は<sup>ハ</sup>ら<sup>ラ</sup>り<sup>リ</sup>ま<sup>マ</sup>  
 と<sup>ト</sup>嶺<sup>ミネ</sup>そ<sup>ソ</sup>廻<sup>マエ</sup>ら<sup>ラ</sup>し<sup>シ</sup>ける<sup>ケル</sup>

FINGASI YAMA no FUMOTO SISI no TANI to IFU TOKORO *fa u si ro fa mi WI*  
 DERA *ni tu ” i te yu ” si ki ZYAU KWAKU ni te so a ri ke ru SYUN KWAN SOU*  
 DU no SAN ZAU *ar i ka re ni tu ne fa yo ri a fi ” ” ” ” FEI KE fo ro fo sa mu su ru fa*

*ka ri ko to wo so MEGURASU ra si ke ru*

‘Beyond the place called Shishi-no-tani at the base of the Higashiyama mountains, it is an impressive walled area adjacent to the Miidera temple. Bishop Shunkan’s mountain villa was there. They would often meet and meet again there and devise plots to destroy the Taira Clan.’

NB: The original manuscripts all used vertical script, so the direction of writing here has been rotated for typesetting purposes. This is followed by a grapheme-by-grapheme analytical transcription: Chinese characters are transcribed in UPPER CASE, katakana in lower case, hentaigana in *italic*, unwritten morphemes in [ ] and repetition marks as ” or ” ” ” ”.

e.g. *fitobito* ‘(many) people’, *fibi* ‘every day’; similarly, a few time nouns + *-goro* < *-no koro* ‘-GEN time’, e.g. *fi-goro* ‘(several/many) days’, *yo-goro* ‘(several/many) nights’. Nouns may also take prefixes, especially honorific *mi-* (20c) and *ofoN-* (later *oN-*) (9a, 23b).

There are two formal sub-classes. (1.) Some pronouns have a short form and a long form in *-re*. *-re* is lost before a genitive particle (*-no*, *-ga*), and occasionally in subject or object function. Otherwise it is obligatory. The pronouns in this category are *ko* ‘this’, *so* ‘that’ (near the addressee [deictic], or previously mentioned in the conversation [anaphoric]), *ka* (or *a*) ‘that over there; (s)he’, *ta* ‘who?’, *wa* ‘I’, *na* ‘you’, and optionally *ono* ‘I; you; (s)he’, which have the genitive forms *ko-no*, *so-no*, *ka-no* (or *a-no*), *ta-ga*, *wa-ga*, *na-ga*, *ono(re)-ga*, but otherwise normally use the stems *kore*, *sore*, *kare* (or *are*), *tare*, *ware*, *nare*, *onore*. In explicit reference to a human, *sore* gives *sore-ga* or rarely *so-ga*. On the other hand, *idure* ‘which?’ in theory should give the genitive form \**idu-no*, but does not; it gives *idure-no* instead, and its short form *idu* only appears in a few frozen derivatives, e.g. *iduko* ‘which place? = where?’. Other simple or derived pronouns (e.g. *nani* ‘what?’, *nandi* ‘you’) have no *re*-form. (2.) A few nouns also have two stems. Normally these end in *-e* or *-i*, e.g. *fune* ‘boat’, *sake* ‘alcohol’, *ki* ‘tree’, *fi* ‘fire’. However, in established compound words derived from contractions of noun + *-no* + noun, they have alternative stems in *-a-* or *-o-* respectively, i.e. *funa-*, *saka-*, *ko-*, *fo-*, e.g. *funabito* ‘seaman; ship’s captain; passenger’ (cf. *fito* ‘person’), *sakaduki* ‘sake cup’ (cf. *-tuki* classifier for dishes or cups), *kozuwe* ‘treetop’ (cf. *suwe* ‘end’), *fomura* ‘flames’ (cf. *mura* ‘group’). In some cases, especially *ki/ko-* ‘tree’, we find the alternative stem used before *-no* in established phrases, e.g. *ko-no suwe* (~ *kozuwe*) ‘treetop’, *ko-no sita* ‘under trees’ (cf. *sita* ‘underneath’), *ko-no kage* (~ *kokage*) ‘the shade of a tree’ (cf. *kage* ‘shade/shadow; reflection’). See also *fonofo* ‘flame’ < *fo-no fo* ‘tail of fire’. An isolated combinatory form is *mina-* corresponding to *midu* ‘water’ (*minato* ‘harbour’; *minamoto* ‘source’; clan name). The OJ alternation *kamwi/kamu-* is archaic, OJ *kamuyo* ‘age of the gods’ usually appearing as *kamiyo*.

Nouns may be followed by one or more of various types of suffix or particle, in the order (a.1) coordinating, (a.2) *-bakari* ‘the extent of; approximately’ (16a), (b.) case, (c.) delimitating, (d.) question (*-ka*, *-ya*) and emphatic/focal (*-zo*, *-namu/-nan*, *-koso*), and (e.) pragmatic (*-fa*, *-mo*) particles. Strictly, (a.–c.) are noun-following particles; (d.–e.) follow any appropriate phrase in the sentence and are dealt with in the syntax section.

Slot (a.1) particles consist of exemplifying *-nado* ‘and such like, etc.’ (2, 3, 6, 16b), and comitative *-to* ‘and’ when used after the last of a list of coordinated nouns (*musasi-to*, (2)). *-nado* and *-bakari* occasionally co-occur in either order (*yonaka-bakari-nado-ni* ‘around midnight or so’ (MS), *yufugure-nado-bakari-zo* ‘around dusk or so’ (GM 5)).

Slot (b.) particles mark case. There is no nominative particle, and the subject of a main clause is always unmarked. The object of a verb is optionally marked by accusative *-wo*. Other case particles are dative *-ni* (also location, direction of movement, agent of a passive, causee of a causative), ablative *-yori* (also standard of comparison ‘than’, (18a)), allative *-fe*, terminative *-made*, genitive *-no* and *-ga*, comitative *-to* (‘with’), essive *-to* (mostly used with verbs such as *nar-* ‘become’ or *se<sup>S</sup>-* ‘make’ to express ‘become X’ or ‘make Object into X’: (16b)), causal *-kara*, and copular *-ni* (used like essive *-to*, (5)). The copular (and rarely the essive) particle also combines with *ar<sup>R</sup>-* to create the copular verb. Comitative *-to* also functions as an ‘and’ between nouns (*kuni-to*, (2)), though CJ frequently listed coordinated nouns with no explicit link, some frequent sequences forming dvandva compounds: *no-yama-ni* ‘on the plains and in the mountains’, *kogane sirogane ruri-iro-no midu* ‘water the colour of gold, silver and lapis lazuli’ (TM).

Old Japanese genitive *-tu* survives frozen in just a few set expressions, e.g. *simo-tu* ‘lower’, *kami-tu* ‘upper’, *kata-tu* ‘one (of a pair)’. Not a particle is *-ti*, attached only to roots of correlatives to express direction: *koti* ‘hither’, *soti* ‘thither’, *iduti* ‘whither’, later *ati* ‘over there’ (*a-no wonoko koti yor-e* ‘That boy! Come here!’ (SN)). *konata / sonata / kanata* (usually + *-ni*) have a similar function.

Rarely, *-wo* does not mark the object, but seems to have some kind of unclear pragmatic function:

- (1) *fune-no uti-wo-namu semete mi-ru*  
 boat-GEN inside-ACC-EMPH against.their.will look-ATTR  
 ‘[Everyone] inside the boat couldn’t help but stare.’ (TM)

The functional difference between the two genitive forms is complicated. Names take *-ga*, but clan names linked to a personal name take *-no*, while pronouns can only take one or the other (*ta-ga* ‘whose?’ not *\*ta-no*, but *ka-no* ‘his/her’ not *\*ka-ga*). They are used to mark: possessor (*taka-ki iyasi-ki fito-no sumafi* ‘homes of high-born [and] lowly people’, *wa-ga kimo* ‘my liver’); the subject in any subordinate clause whose verb is in the attributive form, including those whose verb is in the realis form; the complement of *-goto<sup>A</sup>*- ‘be like’ (*ko-no ife-wo dairi-no goto-ku tukur-i-te* ‘he built this house like the imperial palace’ (SN), *afugi-wo firoge-tar-u-ga goto-ku* ‘like opening out a folding fan’ (H)); part-whole (*sAN-bUN-ga iti* ‘1/3’, *ni-sANzifu-nIN-ga uti-ni* ‘amongst the 20 or 30 people’); (*-no* only) a link between an adverbial phrase such as of location (minus such case particles as *-ni* or *-yori*) and the head noun, in which case *-no* is interchangeable with *-ni ar<sup>R</sup>-u* [-DAT ‘exist’-ATTR]; (*-no* only) a general link between nouns where either the first is just some kind of modifier almost functioning like an adjective, or the second is some sort of title (*X-no tono* = ‘Lord X’, *X-no mikoto* = ‘The God X’). In frequent one noun + *-no* + one noun sequences *-no* often reduces to mere prenasalization of the second component (rendaku), e.g. *kamo-no kafa* [kamonokaɸa] ‘the Kamo river’ > *kamogafa* [kamo<sup>ŋ</sup>gaɸa]. Where the second element already contains a prenasalized consonant, it cannot take a second one, and so even the prenasalization is lost as an indicator of the loss of a *-no*, e.g. *ko-no kage* [konoka<sup>ŋ</sup>ge] ‘shade of a tree’ > *kokage* [koka<sup>ŋ</sup>ge], not *\*kogage*.

Slot (c.) consists of *-nomi* ‘only’, *-safe* ‘even’, *-dani* ‘even (as a minimum)’, *-sura/-sora* ‘even (as a maximum)’: *fotaru-bakari-no fikari-dani na-si* ‘there wasn’t even a firefly’s amount of light’, *yuk-u kata-sora-mo oboye-zu* ‘I had no idea even where I was going’ (TM).

Other forms that are often treated as case particles but which are actually complex forms include the instrumental *-ni-te* and *-ni s-i-te* (< copular *-ni*) and the explicit locative *-ni-te* (< dative *-ni*, showing an actual or perceived etymology of dative *-ni* from copular *-ni*). (See also (15).)

- (2) *simotusa-no kuni-to musasi-to-no sakafi-ni-te*  
 Shimōsa-GEN province-COM Musashi-COM-GEN border-DAT-GER  
*ar-u futowigafa to if-u-ga kami-no se*  
 exist-ATTR Futoigawa QUOT call-ATTR-GEN top-GEN shallows  
*matusato-no watari-no tu-ni tomar-i-te yoru fito-yo*  
 Matsusato-GEN crossing-GEN port-DAT stay-INF-GER night one-night  
*fune-ni-te katugatu mono-nado watas-u*  
 boat-DAT-GER in.stages thing-EX take.across-CONC

‘We stopped at the port at the Matsusato crossing, the upper shallows of [the river] called Futoigawa which is on the border between Shimōsa province and Musashi, and took our things and such like over by boat throughout the night.’ (SN)

There are a number of other particle-like forms, e.g. *-goto-ni* ‘at/on every . . .’ (*ma-goto-ni* ‘in every room’, *asa-goto yufu-goto-ni* ‘every morning and evening’, (TM)), *-goro-ni* ‘on . . . or thereabouts’ (*sawototosi-no kisaragi-no towo-ka-goro-ni* ‘on around the tenth of the second month, three years ago’ (TM)), and *-zama-ni/fe* ‘towards’ (*soba-zama-ni fus-i-te* ‘he would lay down onto his side’ (USM 2/7)). These are contractions of *-no* + noun (*koto* ‘fact’, *koro* ‘time’, *sama* ‘direction’), and are perhaps best treated as phrasal postpositions or compounds. This is also the origin of *-bakari* (< *fakari* ‘measurement; size’), but this is more grammaticalized as a particle.

#### 8.4.2 Verbs and adjectives

CJ verbs are inflected for a range of tense, aspect, modal, passive, causative, coordinating and subordinating forms. Inflections are suffixal, though there are a number of prefixes of unclear function, e.g. emphatic *uti-* in (16c). The inflection of verbs is conventionally presented in modern Japan in terms of six ‘stems’: *shūshikei* 終止形 (‘conclusive’, ‘final’, ‘finite’, ‘sentence-final’), *rentaikei* 連体形 (‘attributive’, ‘adnominal’), *ren’yōkei* 連用形 (‘infinitive’, ‘conjunctive’), *mizenkei* 未然形 (‘irrealis’, ‘imperfective’), *izenkei* 已然形 (‘realis’, ‘perfective’, ‘evidential’, ‘conditional’), *meireikei* 命令形 (‘imperative’). Western attempts to translate these terms are varied. There are problems with this conventional presentation. Firstly, not only does the *ren’yōkei* have both a verbal function conjoining verbs or clauses and a nominal function as a verb-noun, these two functions are known to have had different tone patterns, suggesting different origins. Secondly, there is evidence that certain endings are conventionally misparsed, such as the debitive. The conventional explanation that it is the *shūshikei* of most verbs (but *rentaikei* of *r*-irregular verbs) + *be<sup>A</sup>*, e.g. *kuf-u be-si* ‘he should/could eat it’, ignores the Old Japanese (and occasional EMJ) evidence of *ube* as an independent word. This makes the analysis root + *ube<sup>A</sup>*, e.g. *kuf-ube-si*, both more likely etymologically and neater because it does not have to include a sub-rule for *r*-irregular verbs. Thirdly, the *mizenkei* does not have an independent existence, and in some cases, such as the addition of passive endings, the *a* that characterizes it in quadrigrade verbs appears to be part of the ending *-[r]are-* when the ending-initial *r* is dropped. This raises the suspicion that the stem-final *a* of the *mizenkei* before the few other endings that are added to it actually belong to the ending (Vovin 2003: 168), and I cite such forms as *-[a]zu* negative, etc., the brackets indicating a sound that appears in some conjugations and not in others as the result of the simplification of the prohibition of CC and VV sequences. I do not go as far as Vovin in presenting the large number of endings that follow the *ren’yōkei*, such as the *te*-form, as though they began with *i*, i.e. *-ite*; in these cases the *i* is part of the verb. However, in the case of the continuous ending *-er<sup>R</sup>*, where the *e* is a fusion of *ren’yōkei i* with ending-initial *a* in Old Japanese, the treatment as root + ending-initial is just convenient.

Verb inflection is presented in Table 8.6. The conjugation names approximate the conventional Japanese terms: quadrigrade (四段), *n*-irregular (ナ変) and *r*-irregular (ラ変) verbs have roots ending in a consonant; bigrade (二段) and monograde (一段) have roots ending in a vowel; and *s*-irregular (サ変) and *k*-irregular (カ変) verbs – just one verb each – show a mix of quadrigrade and bigrade features. I use superscript <sup>M</sup> (monograde),

TABLE 8.6 CLASSICAL JAPANESE VERB MORPHOLOGY

	<i>s</i> -Irreg. 'do'	<i>k</i> -Irreg. 'come'	<i>n</i> -Irreg.	Quadri- grade	<i>r</i> -Irreg.	Bigrade 'Lower'	'Upper'	Mono- grade
Root:	√ <i>se</i> <sup>S</sup>	√ <i>ko</i>	√ <i>n</i>	√ <i>-C</i>	√ <i>-r</i> <sup>R</sup>	√ <i>-e</i>	√ <i>-i</i>	√ <i>-i</i> <sup>M</sup>
Independent Forms:								
Conclusive:	<i>s-u</i>	<i>k-u</i>	<i>-n-u</i>	<i>-C-u</i>	<i>-r-i</i>	<i>-Ø-u</i>	<i>-Ø-u</i>	<i>-i-ru</i>
Attributive:	<i>s-uru</i>	<i>k-uru</i>	<i>-n-uru</i>	<i>-C-u</i>	<i>-r-u</i>	<i>-Ø-uru</i>	<i>-Ø-uru</i>	<i>-i-ru</i>
Realis:	<i>s-ure</i>	<i>k-ure</i>	<i>-n-ure</i>	<i>-C-e</i>	<i>-r-e</i>	<i>-Ø-ure</i>	<i>-Ø-ure</i>	<i>-i-re</i>
Imperative:	<i>se-yo</i>	<i>ko-yo</i>	<i>-n-e</i>	<i>-C-e</i>	...	<i>-e-yo</i>	<i>-i-yo</i>	<i>-i-yo</i>
Infinitive:	<i>s-i</i>	<i>k-i</i>	<i>-n-i</i>	<i>-C-i</i>	<i>-r-i</i>	<i>-e-Ø</i>	<i>-i-Ø</i>	<i>-i-Ø</i>
Verb-noun:	<i>s-i</i>	<i>k-i</i>	<i>-n-i</i>	<i>-C-i</i>	<i>-r-i</i>	<i>-e-Ø</i>	<i>-i-Ø</i>	<i>-i-Ø</i>
Suffixes:								
<i>u</i> -Initial:								
e.g. <i>-ube</i> <sup>A</sup> :	<i>s-ube</i> <sup>A</sup>	<i>k-ube</i> <sup>A</sup>	<i>-n-ube</i> <sup>A</sup>	<i>-C-ube</i> <sup>A</sup>	<i>-r-ube</i> <sup>A</sup>	<i>-Ø-ube</i> <sup>A</sup>	<i>-Ø-ube</i> <sup>A</sup>	<i>-i-rube</i> <sup>A</sup>
[ <i>a</i> ]-Initial:								
e.g. <i>-[a]m</i> :	<i>se-m</i>	<i>ko-m</i>	<i>-n-am</i>	<i>-C-am</i>	<i>-r-am</i>	<i>-e-m</i>	<i>-i-m</i>	<i>-i-m</i>
<i>-er</i> <sup>R</sup> :	<i>s-er</i> <sup>R</sup>	...	...	<i>-C-er</i> <sup>R</sup>	...	...	...	...
<i>-[r]are</i> :	<i>se-rare</i>	<i>ko-rare</i>	<i>-n-are</i>	<i>-C-are</i>	...	<i>-e-rare</i>	<i>-i-rare</i>	<i>-i-rare</i>
<i>-[s]ase</i> :	<i>se-sase</i>	<i>ko-sase</i>	<i>-n-ase</i>	<i>-C-ase</i>	...	<i>-e-sase</i>	<i>-i-sase</i>	<i>-i-sase</i>

<sup>R</sup> (*r*-irregular) and <sup>S</sup> (*s*-irregular) to identify roots belonging to three of the conjugations. (<sup>A</sup> identifies a form inflected as an adjective.) Roots not so marked are unambiguous. The *n*-irregular conjugation is not strictly irregular, since all three *n*-root verbs in the language follow it rather than the quadrigrade pattern: *in-* 'go', *-n-* Perfective, *sin-* 'die'.

The number of monograde verbs is small and mostly monosyllabic, the commonest being *mi*<sup>M</sup> 'see', *ki*<sup>M</sup> 'put on, wear', *wi*<sup>M</sup> 'sit; dwell', *ni*<sup>M</sup> 'resemble'. A process of monogradization of bigrade verbs must have been happening in at least some sectors of the spoken language, and in the Kamakura period we start to encounter isolated examples of bigrade attributives in *-eru* rather than *-uru*.

Verbal endings are normally inflected as verbs in their own right and in theory have the full range of stems, though, mostly because of a lack of need of certain stems, several are defective. Irregularly inflected endings are the past tense (conclusive *-ki*, attributive *-si*, realis *-sika*), negative (conclusive and infinitive *-[a]zu*, attributive *-[a]nu*, realis *-[a]ne*), negative-tentative (conclusive, attributive, and 'realis' triggered by *-koso* all *-[a]zi*), and counterfactual (conclusive and attributive *-[a]masi*, realis *-[a]masika*).

Gerund *-te* 'and' is a particularly frequent ending, normally attached to the infinitive, which by itself has a similar function. Only the latter, however, is also used to create compound verbs (*kafer-i-ko* 'come back', *omof-i-wabi* 'feel sad', *sinob-i-ir* 'sneak in', etc.). There are a few cases in which *-te* is attached not to an infinitive (*sa* 'thus' > *sa-te*, *\*-[a]n(u)* negative > *-[a]de*, *-ni* dative > *-ni-te*), or attached to forms that have been reinterpreted as infinitives (*-[a]zu* negative > *-[a]zu-te*, *to* quotative > *to-te*, *-ni* copula *-ni-te*). A frequent form <mo.te> 'carrying' usually in *mote ko-* 'bring' appears to derive from *mot-* 'carry' + *-te*. This may represent onbin (*mot-i-te* > *moq-te*), though onbin is rarely written for other verbs, or it may be another case of *-te* attached directly to a root rather than an infinitive. (A perfect, usually attributive, <mo.ta.ru> also occurs.)

Adjectives (roots tagged here with <sup>^</sup>) have a simpler basic morphology, consisting of a root, e.g. *faya*<sup>^</sup>- ‘be quick’, which normally only occurs in certain compound words or in exclamations (*ana kasiko*-Ø ‘How lovely!’ (*TM*)), a conclusive form in *-si*, an attributive form in *-ki*, a realis form in *-kere*, and an infinitive form in *-ku* or *-u*. *si*- and *zi*-root adjectives are an exception, in that their conclusive form is zero (*kanasi*<sup>^</sup>- > *kanasi*-Ø ‘he is sad’, *imizi*<sup>^</sup>- > *imizi*-Ø ‘it is terrible’); Japanese grammarians distinguish *si*-/*zi*-root adjectives as ‘*siku*-conjugation adjectives’ (シク活用形容詞), as opposed to other ‘*ku*-conjugation adjectives’ (ク活用形容詞). The *zi*-root adjective *onazi*<sup>^</sup>- ‘be the same’ also tends to lose attributive *-ki*. The infinitive of adjectives is used either as an adverb (*faya*-*ku* ‘quickly’) or – optionally reinforced by *-te* or *si*-*te* – as a coordinating form similar to the infinitive of verbs (*faya*-*ku* ‘he was quick, and’). One difference between *-ku* and *-u* is given in the next paragraph; otherwise, there is probably no real difference, and it is noticeable that the *si*-root adjectives prefer *-u* whereas other adjectives prefer *-ku*. However, unlike verbs, the infinitive of adjectives takes no verbal endings other than *-te* to reinforce coordination. Deadjectival nouns derive from root + *-sa* (*uresi*<sup>^</sup>- ‘be happy’ > *uresi*-*sa* ‘happiness’); those from colour adjectives, though, are identical to the root (*siro*<sup>^</sup>- ‘be white’, *siro* ‘white(ness)’).

Similarly, the infinitive of *-[a]zu*, the negative of verbs, can take no verbal ending other than *-te*. Both adjectives and the negative morpheme take endings only by means of an auxiliary verb *-ar*<sup>R</sup>-, which occurs in a contracted form, i.e. adjective-*ku* (never adjective-*u*) + *ar*<sup>R</sup>- > adjective-*kar*<sup>R</sup>-, negative *-[a]zu* + *ar*<sup>R</sup>- > *-[a]zar*<sup>R</sup>-. These forms occur in all possible verb stems, except the conclusive. (The conclusive *ofo*-*kar*-*i* ‘are many’ corresponds to *ofo*<sup>^</sup>-, but the *ar*-*i* here represents *ar*<sup>R</sup>- not in an auxiliary use but as a full verb ‘there are’.)

The copula *-ni* is heavily defective, and, apart from a very occasional alternative conjunctive form *-no* (which never takes *-te*) and the possible identification of the genitive particle *-no* as an attributive form (but which never enters into *kakari*-*musubi* or functions as a nominalization), it only has the one stem *-ni* (COP) which, because of its coordinating function and the fact that it can be extended with *-te*, is classified as an infinitive. All other stems and inflections of the copula are achieved by the addition and frequent fusion of the auxiliary *ar*<sup>R</sup>-, i.e. *-ni ar*<sup>R</sup>- ~ *-nar*<sup>R</sup>-. The shorter version is normal in affirmative sentences, but the longer version is common in negative sentences, or where a topic, focal, inclusive or emphatic particle is attracted to it, e.g. *-nar*-*azu* ~ *-ni ar*-*azu* ~ *-ni*-*fa ar*-*azu* ‘isn’t’. Unlike the forms with the auxiliary in the previous paragraph, the copula + auxiliary does occur in the conclusive, as there is no other conclusive copula.

In addition to true adjectives, there is a range of adjectives that are formed from a nominal – some independent nouns, others not – combined with the copula, e.g. *siduka*-*nar*<sup>R</sup>- ‘be quiet’, *afare*-*nar*<sup>R</sup>- ‘be a shame’ (hereafter ‘*nari*-adjectives’). A few adjectives have a related *nari*-adjective, e.g. *ofoki*<sup>^</sup>- ~ *ofoki*-*nar*<sup>R</sup>- ‘be big’, *tifisa*<sup>^</sup>- ~ *tifisa*-*yaka*-*nar*<sup>R</sup>- ‘be small’.

### Splitting

Non-case particles (8.4.1) whose scope is or includes the verb can be inserted into the verb. This involves splitting the verb. Copular *-nar*<sup>R</sup>- and adjectives, which involve fusions with auxiliary *ar*<sup>R</sup>-, are split to accommodate this particle. Note that negative *-zar*<sup>R</sup>- is not split. Verbs can also be ‘split’ by transforming to verb-noun + *se*<sup>S</sup>- ‘do’ (3, 20c):

- (3) *yuk-u-mo*            *tomar-u-mo*            *mina nak-i-nado*    *s-u*  
 go-ATTR-INCL    stay-ATTR-INCL    all    cry-VN-EX    do-CONC  
 ‘Both those going and those staying all cried and such like.’ (SN)

### 8.4.3 Numerals and classifiers

CJ had a full range of native numerals as far as *yorodu* 10,000, though this really has the function of indicating just a very large number or ‘very many’ in the same way as the Chinese 万 ‘10,000’ does. The other numbers are: *fito-/fitotu* ‘1’, *futa-/futatu* ‘2’, *mi-/mitu* ‘3’, *yo-/yotu* ‘4’, *itu-/itutu* ‘5’, *mu-/mutu* ‘6’, *nana-/nanatu* ‘7’, *ya-/yatu* ‘8’, *kokono-/kokonotu* ‘9’, *towo(-)* ‘10’, *fata-/fatati* ‘20’, *miso-/misoti* ‘30’, *yoso-/yosoti* ‘40’, etc., *momo(-)* ‘100’, *ti(-)* ‘1,000’; *iku-/ikutu* ‘how many?’. Multiples of hundred and thousand are probably already extinct. The numbers are used either independently (in which case the units take a suffix *-tu* and the tens, other than *towo*, take *-ti*), or with a few native classifiers, e.g. *-uka* ‘days’, *-yo* ‘nights’, *-tokoro* ‘places’, *-sudi* ‘long thin objects’, *-mura* ‘rolls’. The classifier *-(ta)ri* ‘people’ appears only to occur with 1 to 4 (*fito-ri*, *futa-ri*, *mi-tari*, *yo-tari*) and in *iku-tari* ‘how many?’, higher numbers taking Sino-Japanese number + *-nin*. The sequences with *-uka* involve the elision of one of the two vowels at the morpheme boundary: *fut-uka* ‘2 days; 2nd day’, *mi-ka* ‘3 days; 3rd day’, *kokon-uka* ‘9 days; 9th day’. Compound numbers are formed using a linking *amari* ‘additionally’ and where there is a classifier it is attached to each component. This necessarily results in long phrases, e.g. *fat-uka amari kokon-uka* ‘29 days; 29th day’.

Consequently, CJ made increasing use of Chinese numerals which allowed for shorter phrases, since classifiers were only attached to the final element and no link word like *amari* was used: units *iti/itu*, *ni/zi*, *san*, *si*, *go*, *riki/roku*, *siti*, *fati*, *kiu/ku*; tens *zifu*, *nizifu*, *sanzifu*, etc.; hundreds *fyaku*, *nifyaku*, etc.; thousands *sen*, *nisen*, etc. Chinese numbers generally take Chinese-derived classifiers, e.g. *-niti* (*-zitu*?) ‘day(s)’, e.g. *nizifu kiu-niti* ‘29 days; 29th day’. However, they could also be used without a classifier, especially in certain texts, e.g.:

- (4) *iti-no*    *sima-ni*    *kome-ok-e-ba*  
 1-GEN    island-DAT    seclude.INF-put-REAL-DISC  
 ‘They secluded them on **one/an** island, and [. . .]’ (KMS 15/44)

Most of the time, numbers are written in Chinese characters, so it is difficult to tell in most cases whether the author intended them to be read as native or as Chinese numerals, e.g. 十九日 ‘19th day’ could represent either *towo-ka amari kokon-uka* or *zifu kiu-niti*. Similarly, month-names tended to be written in characters, disguising whether native name or Chinese enumerated name was intended, e.g. 六月 *minaduki* (also explicitly 水無月 ‘waterless month’) ~ *roku-gwatu* (‘6th month’).

## 8.5 SYNTAX

### 8.5.1 Noun phrase structure

CJ modifiers largely precede their head, and are marked either by a genitive particle or by an attributive form of the verb. Quantity expressions usually precede their head with a linking *-no*, though they may also follow their head as adverbs (e.g. *amata-no* Noun vs Noun *amata* in (8) and (9, 13b) below).

EMJ used two relative-clause structures. The first precedes its head noun, and the verb or adjective at the end of the relative clause is in an attributive form. The structure is essentially a nominalization, and so the subject within a relative clause tends to be marked as a genitive.

The head of a relative clause may be ellipted. The attributive form is a nominalization, and so can be followed directly by noun-following inflections (2, 26a).

Internally headed relative clauses were also common. The head retains its place within the relative clause, and the clause-final attributive form can be followed directly by noun-following inflections. In the following examples, the head noun is in bold, the relative clause is indicated by [ ], and (b.) examples provide what a head-final structure would have produced.

- (5) a. [*kinofu fito-no medeta-ki firatake-wo tab-i-*  
yesterday person-GEN lovely-ATTR oyster.mushroom-ACC give.me-INF-  
*tar-i-si]-wo imono-ni s-i-te*  
PERF-INF-PAST.ATTR-ACC stew-COP make-INF-GER  
*mes-ase-mu to-te mausisaburaf-i-t-uru-nar-i*  
eat.HON-CAUS-TENT QUOT-GER call.HUM-INF-REC-ATTR-COP.AUX-CONC  
'I called you here in order to feed you as a stew some lovely oyster mushrooms  
which someone gave me yesterday.' (KMS 28/18)
- b. ?[*kinofu fito-no tab-i-tar-i-si*] *medeta-ki firatake-wo nimono-ni s-i-te tabe-sase-mu*  
*to-te mausisaburaf-i-t-uru-nar-i*

It is possible, where the head is modified by another modifier, for the head to be ellipted as in head-final relative clauses, and it is the other modifier that as a nominalization occupies the head's clause-internal position.

- (6) a. *nma-fa [ito kuro-ki-ga tada isasaka siro-ki tokoro-nado*  
horse-TOP very be.black-ATTR-GEN just slightly be.white-ATTR place-EX  
*ar-u]; murasaki-no mon tuk-i-tar-u; asige;*  
have-ATTR, purple-GEN patch be.attached-INF-PERF-ATTR, dapple.grey  
'Horses: very black ones which just have slightly white patches; ones with  
grey patches on them; dapple-greys; [ . . . ]' (a list of the author's favourite types  
of horse, MS)
- b. ?*muma-fa [tada isasaka siro-ki tokoro-nado ar-u] ito kuro-ki . . .*

As illustrated above, the head-internal structure is preferred if the relative clause is not the only modifier of the head noun, as *medetaki* 'lovely' and *ito siroki* 'very white' above. Sequences of relative clause + another modifier + head are exceptional.

However, where the head noun itself would be followed by *-no* + noun if it were not clause-internal, the *-no* does not appear, because both the attributive and *-no* are simply perceived to be markers of noun-modification. It may be that *-no* is in origin the attributive of the copula, so attributive + *-no* would be a double attributive.

- (7) a. [*vingasi-no kata-ni take-no sukosi of-i-*  
east-GEN direction-DAT bamboo-GEN a.little grow-INF-  
*tar-i-ker-u] naka-yori ko-no aka-ki*  
PERF-INF-UNW-ATTR middle-ABL this-GEN be.red-ATTR

*fitofeginu rei-no yau-ni fafe-tob-i-te*  
 summer.kimono usual-GEN manner-DAT flap-INF-fly-INF-GER

*watar-i-ker-u-wo*  
 go.across-INF-UNW-ATTR-ACC

‘This red summer kimono flew out as usual from the middle of [a patch of] bamboo growing towards the east and flew across [the garden], and [. . .]’ (KMS 27/4)

- b. [*fiŋgasi-no kata-ni sukosi of-i-tar-i-ker-u*] *take-no naka-yori ko-no aka-ki fitofeginu rei-no yau-ni fafe-tob-i-te watar-i-ker-u-wo*

- (8) a. *kudara-no kuni-no yabure-ker-u toki*  
 Paekche-GEN country-GEN be.destroyed-INF-UNW-ATTR time

*ka-no kuni-wo tasuke-m-u-ga tame-ni [ofoyake*  
 that-GEN country-ACC help-TENT-ATTR-GEN purpose-DAT court

*amata-no ikusa-wo tukafas-u] naka-ni ko-no atafi-wo*  
 lots-GEN troops-ACC dispatch-ATTR middle-DAT this-GEN atai-ACC

*tukafas-i-ker-i*  
 dispatch-INF-UNW-CONC

‘Amongst the many troops that the court dispatched to Paekche to support it at the time that country was destroyed, it dispatched this *atai*.’ (KMS 15/44)

- b. . . [*ofoyake tukafas-u*] *amata-no ikusa-no naka-ni ko-no atafi-wo tukafas-i-ker-i*

*Genji monogatari* opens with a sentence which, though not long, contains many modifiers, and therefore has two head-internal relative clauses. The equivalent head-final structure (b.) has so many modifiers that it would at the very least have been stylistically ugly.

- (9) a. *idure-no ofon-toki-ni ka [nyougo kauli amata*  
 which-GEN HON-time-DAT Q first.rank.lady second.rank.lady lots

*saburaf-i-tamaf-i-ker-u] naka-ni [ito yangotona-ki*  
 serve-INF-HON-INF-UNW-ATTR midst-DAT very important-ATTR

*kifa-ni-fa ar-anu-ga sugurete tokimek-i-*  
 rank-COP-TOP AUX-NEG.ATTR-GEN extremely be.favoured-INF-

*tamaf-u] ar-i-ker-i*  
 HON-ATTR exist-INF-UNW-CONC

‘In a particular reign, there was amongst the many first- and second-rank ladies who served [the Emperor] one who was not [of] very important rank who was extremely favoured [by the Emperor].’ (GM 1)

- b. ?*idure-no ofon-toki-ni ka amata saburaf-i-tamaf-i-ker-u nyougo kauli-no naka-ni sugurete tokimek-i-tamaf-u ito yangotona-ki kifa-ni-fa ar-anu-ga ar-i-ker-i*

### 8.5.2 Pronouns and anaphora

CJ has the personal pronouns *wa(re)* ‘I’, rare *na(re)* ‘you’, and *ono(re)* ‘I; you; (s)he’. From *na(re)* appears the derivative *nandi*. However, the hierarchical society of the Heian period gave rise to a range of nouns used as status-dependent pronouns.

The demonstrative pronouns and modifiers display a three-way system of proximal *ko(re)*, mesial *so(re)*, distal *ka(re)* or *a(re)*, with mostly *so(re)* used in anaphoric function. Derived from these are the place pronouns *koko/soko/kasiko* and direction *koti/soti/ati*, *konata/sonata/kanata*. Demonstratives of manner or comparison are from different roots: *sa/sika* ‘like that’ vs *kaku* ‘like this’, *sa-te* vs *kaku-te*, *sa-yau-ni* vs *ka-yau-ni*; also to ‘like that’, only used together with *kaku* in set phrases, e.g. *to(-mo/-ni) kaku* ‘this way and that; all over; all sorts’.

The interrogative pronouns are: *ta(re)* ‘who’, *nani(goto)* ‘what’, *itu* ‘when’, *iduku/iduko* ‘where’, *iduti* ‘whither’, *idukata* ‘which direction’, *idure* ‘which one’, *iku-* ‘how many’, *nazo/nado* ‘why’, *nande* or *ikani* ‘how’, *ika-yau-ni* ‘like what’. The sequence *ikani* + question particle *-ka* already fuses as *ikaga* at the start of the period, and later in the period a second *-ka* can be added to the sentence. The stative verbs *sar<sup>R</sup>-/sika<sup>R</sup>-* ‘be thus’, *kakar<sup>R</sup>-* ‘be like this’, *ikanar<sup>R</sup>-* ‘be like what?’, or *sa-/ka-/ika-yau-nar<sup>R</sup>-*, are fusions of the adverbs of manner + auxiliary *ar<sup>R</sup>-*.

There are no indefinite pronouns. Instead, *fito* ‘person > someone’, *mono* ‘thing > something’, *toki* ‘time > sometime’, *tokoro* ‘place > somewhere’ are used. However, to express the same concept the language also resorted to embedded question-word questions. For example, *Genji monogatari* begins with the embedded question *idure-no ofon-toki-ni-ka* (9), literally ‘in which reign was it?’, but which corresponds to indefinite forms such as ‘in a certain reign’ in English. Later in the history of Japanese these embedded questions developed into the modern indefinite pronouns, *dareka* ‘someone’, *nanika* ‘something’, etc.

### 8.5.3 Basic sentence structure

#### 8.5.3.1 Declarative

CJ is a topic-oriented SOV language. The verbal predicate – verb or adjective – is strictly clause-final, but topicalization and information-packaging override the neutral Subject—Adjunct—Object order of the rest: *kore-wo kaguyafime kik-i-te* = OSV ‘Kaguyahime listened to this, and’ (TM).

The copula *-nar<sup>R</sup>-*, or *-ni ar<sup>R</sup>-* when incorporating a pragmatic or question particle or often in the negative, is used with a preceding noun complement: Noun<sup>1</sup> Noun<sup>2</sup>-*nar<sup>R</sup>-* = ‘Noun<sup>1</sup> is a Noun<sup>2</sup>’. Unlike the OJ period, the single unit version *-nar<sup>R</sup>-* is more common, and in the conclusive the split version *\*-ni ar-i* does not occur. A sentence-final copula is commonly ellipted. (10a.–b.) are both essentially copula-less noun predicates ‘[It was] the 28th of the 4th month of the Angen era’ and ‘The source of the fire [was] Higuchitomi Alley’, but each is followed by a sequence of three particles that express opinion, questioning and assertion.

- (10) a. *inzi angen san-nen si-gwatu nizifu-fati-niti-ka to yo*  
 now.passed Angen.era 3-year 4-month 20-8-day-Q QUOT ASSERT  
 ‘I think it might have been the 28th of the 4th month of the Angen era!’ (H)
- b. *fomoto-fa figutitomi-no koudi to-ka ya*  
 source.of.fire-TOP Higuchitomi-GEN Alley QUOT-Q Q/ASSERT  
 ‘They say the source of the fire was Higuchitomi Alley, don’t they?’ (H)

When it is used in coordination, the *-ni* underlying the copula *-n(i)ar<sup>R</sup>-* is used as an infinitive, with or without *-te*, alongside the Nara-inherited *-n(i)ar-i-te*. As a result, the usual copula *-n(i)ar<sup>R</sup>-* developed an alternative *-ni-te ar<sup>R</sup>-* (24), which in the Kamakura period competes with *-ni ar<sup>R</sup>-*. At the same time, the contraction *-ni-te > -de* results in

-*de ar<sup>R</sup>*-. A rare alternative copula -*tar<sup>R</sup>*- appears, mostly in *kanbun kundoku* and other male-style writing, although some forms do appear in other literature.

Copular -*nar<sup>R</sup>*- is distinct from *nar*- ‘become’, which creates inchoative structures: it follows the simple infinitive of adjectives and negative -[*a*]zu, and Noun-*nar<sup>R</sup>*- ‘be Noun’ > Noun-*ni*/-*to nar*- ‘become Noun’ (20c, 23c).

Existence, location and possession are all expressed with *ar<sup>R</sup>*-; possession has the structure Possessor-*ni* Possessed + *ar<sup>R</sup>*-, literally ‘there is a Noun<sup>2</sup> at Noun<sup>1</sup>’ = ‘Noun<sup>1</sup> has a Noun<sup>2</sup>’. Unlike in the OJ period, -*ni ar<sup>R</sup>*- does not normally contract, except in poetry (*kasuga-n’ ar-u/mikasa-no yama-ni/ide-si tuki kamo* ‘I wonder if it’s the moon that [I remember] emerged from the Mikasa mountains which **are in** Kasuga’ (TN 1/20)). The negative of *ar<sup>R</sup>*- as an independent verb is either regular *ar-azu* or the adjectival *na<sup>A</sup>*-, but as an auxiliary only *ar-azu* is standard, i.e. -*nar<sup>R</sup>*- > -*nar-azu*-, -*ni ar-azu*-, but not \*-*ni na<sup>A</sup>*-. (By the twelfth century, the latter does occur, e.g. *tadabito-ni-fa na-ki mono* ‘someone who **wasn’t** a normal human being’ (KMS 31/33).)

The language tends not to express subjects, or even objects, unless they are not contextually clear. Because the language has both switch-reference coordination devices and a rich system of honorific and humble verb(-forms), this can result in sizeable stretches of narrative with no explicit identification of two or more previously established protagonists as subjects of verbs.

### 8.5.3.2 Tense and aspect

The basic tense distinction is between non-past (unmarked) and past (three forms). The three past markers are infinitive + -*te*- (REC) for recent past (typically an event occurring in or continuing to the last diurnal day), infinitive + -*ki* (PAST) for past witnessed by the speaker, and infinitive + -*ker<sup>R</sup>*- (UNW) for past unwitnessed by the speaker (or about which the speaker expresses emotion or other subjectivity, e.g. *tatu-wo koros-am-u to omof-i-ker-i* ‘I thought that I’d kill a dragon’ (TM)). Fictional narratives vary in their narrative tense: high Heian texts such as *Genji monogatari* use the unmarked form (i.e. a ‘historical present’), whereas late Heian and Kamakura tale collections such as *Konjaku monogatari-shū* and *Uji shūi monogatari* use -*ker<sup>R</sup>*-.

- (11) a. *wotoko* “*ko-no nak-i-t-uru sika-no ne-fa kik-i-*  
man this-GEN cry-INF-REC-ATTR deer-GEN sound-TOP hear-INF-  
*tamaf-i-t-u-ka*” *to if-i-ker-e-ba*  
HON-INF-REC-CONC-Q QUOT say-INF-UNW-REAL-DISC  
‘The man said, “Did you hear the sound of that deer that cried [just now]?”  
and [. . .]’ (KMS 30/12)
- b. *makoto-ni fourai-no ki-ka to-koso omof-i-t-ure*  
truth-DAT Penglai-GEN tree-Q QUOT-FOC think-INF-REC-REAL  
‘“Is it really a tree from Penglai?” is what I thought [just now].’ (TM)

There was originally no future tense, and in cases where the speaker is certain the future can be expressed by the conclusive form.

- (12) *tada katafarani zeokon-wo yatof-i-te fusyau-no amida-*  
just anyway tongue-ACC use-INF-GER unsolicited-GEN Amitābha-  
*butu ryau-san-ben maus-i-te yam-i-n-u*  
Buddha two-three-times say-INF-GER stop-INF-PFV-CONC

‘Anyway I’ll just use my tongue and say a half-hearted prayer to the bodhisattva Amitābha two or three times, and then stop.’ (*H*, end)

The ending *-[a]m-* (TENT) is used for a range of modal functions, including cohortative (‘let’s . . .’) and expression of uncertainty (‘shall I . . .?’, ‘maybe’, etc.). Its conclusive and attributive forms tend to reduce *-[a]m-u* to *-[a]N* (25), though which pronunciation is intended tends to be obscured by the use of the same hentaigana for both <mu> and <N>. By itself, it largely involves future events. An extended version *-[a]m-u* to *se<sup>S</sup>*- or its contraction *-[a]Nze<sup>S</sup>*- tends to express intention (‘intend to . . .’, ‘try to . . .’, (13, 17)).

- (13) a. *iti-ni kaf-am-u to s-ure-ba so-no*  
 market-DAT buy-TENT-CONC QUOT do-REAL-DISC it-GEN  
*atafi na-si*  
 price not.exist-CONC  
 ‘When he tried to buy [some meat] at the market he didn’t have the money for it.’ (*KMS* 19/6)
- b. *iduti-mo iduti-mo asi-no muk-i-tar-an*  
 whither-INCL whither-INCL leg/foot-GEN head-INF-PERF-TENT.ATTR  
*kata-fe in-anz-u*  
 direction-ALL go-CONC.QUOT.do-CONC  
 ‘We shall go in absolutely whichever direction our feet are facing.’ [*TM*]

*-[a]m-* also occurs in two extended forms, *-kem-* (20c) and *-uram-* (14), the first representing an earlier fusion of *-ki* (past) + *-[a]m-*, the second a root that may also appear in *makura* ‘pillow’ < *mak-* ‘sleep’, *sakura* ‘cherry blossom’ < *sak-* ‘bloom’. These two forms typically refer to the past and to a present/habitual respectively, though the distinction between *-uram-* and *-[a]m-* can be unclear.

- (14) *nukadukimusi mata afare-nar-i. sar-u*  
 kowtow.beetle also pitiful-COP.AUX-CONC be.thus-ATTR  
*kokoti-ni dausin okos-i-te tuk-i*  
 character-DAT faith make.arise-INF-GER attach-INF  
*arik-uram-u yo*  
 wander-HAB.TENT-CONC ASSERT  
 ‘The kowtowing beetle also moves me. It experienced a deep religious faith in a little heart like that, and so it wanders around with [its head] touching [the ground]!’ (*MS*)

Finally, there is a counterfactual ‘tense’ formed with the irregular ending *-[a]masi-*. It is particularly common in counterfactual conditions.

- (15) *nandi kasiko-ni-te if-amasika-ba wa-ga kimo-mo*  
 you there-DAT-GER say-CF.REAL-DISC I-GEN liver-INCL  
*mata foka-no saru-no kimo-mo tor-i-te*  
 also other-GEN monkey-GEN liver-INCL take-INF-GER  
*tatematur-i-te-masi*  
 give.HUM-INF-REC.INF-CF(.CONC)  
 ‘If you had said that back there, I would have got both my liver and also other monkeys’ livers and given them to you.’ (*KMS* 5/28)

Aspect is expressed independently of tense through two sets of affix. The first are perfect *-tar<sup>R</sup>-* and *-er<sup>R</sup>-*, which express continuation of a state resulting from an earlier action. (The few uses that appear to be progressive, e.g. *umi-no ufe-ni tadayof-er-u yama* ‘a mountain that **was floating** on top of the sea’ (*TM*), are probably still perfect, i.e. *tadayof* = ‘enter into a floating position’.) Functionally there seems no difference between the two, even though many have attempted to identify such a difference (e.g. Takeuchi 1987: 162–98). Formally, though, *-er<sup>R</sup>-* only occurs on *se-* ‘do’ and quadrigrade verbs; *-tar<sup>R</sup>-* is used on any verb that semantically allows it – although Kamo no Chōmei stands out in almost restricting *-tar<sup>R</sup>-* to those verbs that cannot take *-er<sup>R</sup>-*. Though *-er<sup>R</sup>-* outnumbers *-tar<sup>R</sup>-* at the very start of the period, *-tar<sup>R</sup>-* quickly supercedes *-er<sup>R</sup>-*, which seems to become increasingly common in just a few combinations, such as honorific *-tamaf-er-i* or reciprocal *-af-er-i*.

The second are perfective *-n-* and probably *-te-* which can be used, regardless of tense, to express completion. The two never co-occur on a verb, but their difference is problematic, and scholars have noted that the transitivity of the verb or the freedom of the subject to choose to act show a statistical *preference* for one over the other. Sandness (1999) argues well that treating *-n-* and *-te-* as two perfectives is wrong, showing that *-n-* is a marker of punctuality, expressing either ‘attainment’ or ‘inception’ (i.e. ‘started to . . .’), and arguing that *-te-* only expresses recent past. She does, however, accept that only a perfective explanation can explain sequences such as *-te-m-* or *-t-ube<sup>A</sup>-*, and the sequence *-te-ki*, though rare, does occur. And there *are* examples where a recent past function cannot apply. In any case, by the end of the Heian period *-n-* itself had probably largely fallen out of the spoken language, perhaps because of confusion with negative attributive *-[a]nu* as a result of the increasing use of the attributive instead of the conclusive, and the number of verbs that *-n-* appears attached to in texts becomes smaller.

- (16) a. *kaze fagesi-ku fuk-i-te*  
wind be.violent-INF blow-INF-GER
- siduka-nar-azari-si*  
calm-COP.AUX-NEG.AUX-PAST.ATTR
- yo inu-no toki-bakari miyako-no tounan-yori fi*  
night dog-GEN time-APPROX capital-GEN southeast-ABL fire
- ide- k-i-te seifoku-ni itar-u*  
go.out-INF- come-INF-GER northwest-DAT reach-CONC
- ‘At around the hour of the dog on a night which was not calm and on which a wind blew violently, a fire broke out from the south-east of the capital and reached the north-west.’ (*H*)
- b. *fate-ni-fa syusyakumon daikokuden daigakureu minbusyau-nado-made*  
end-DAT-TOP Shushakumon Daikokuden Daigakuryō Minbushō-EX-TERM
- utur-i-te fito-yo-no uti-ni dinkwai-to*  
move-INF-GER one-night-GEN within-DAT dust.and.ashes-ESS
- nar-i-n-i-ki*  
become-INF-PFV-INF-PAST.CONC
- ‘In the end it moved as far as the Shushakumon, Daikokuden, Daigakuryō, Minbushō and other buildings and in the space of one night they became dust and ashes.’ (*H*)

- c. *fitosire-zu* *uti-nak-are-n-u*  
 be.known.to.others-NEG.INF EMPH-cry-PASS-PFV-CONC  
 ‘I couldn’t help but burst into tears when no one else was around.’ (SN)

Where there is more than one tense/aspect suffix, the order is PERF-PFV-tense.

CJ also made use of aspectual auxiliaries such as infinitive + *ok-* (lit. ‘put’) to indicate an action that is done in readiness for some future occasion (17),<sup>2</sup> or + directional verbs (e.g. *fasir-* ‘run’ + *ir-* ‘enter’ > *fasir-i-ir-* ‘run in’), or occasionally + *wor<sup>R</sup>-* ‘be sitting; be’ as an explicit progressive.

- (17) *“ima. akikaze fuk-am-u wori-zo ko-m-u*  
 now autumn.wind blow-TENT-ATTR time-EMPH come-TENT-ATTR  
*to s-uru. mat-e yo” to if-i-ok-i-te*  
 QUOT do-ATTR wait-IMV EXCL QUOT say-INF-put-INF-GER  
*nige-te in-i-ker-u-mo sir-azu*  
 flee-INF-GER go-INF-UNW-ATTR-INCL know-NEG.INF  
 ‘[the demon parent,] said “Now – when the autumn wind blows is when I’ll come – wait!”’, and [the child] did not even know [the parent] had run away, and [. . .]’ (MS)

### 8.5.3.3 Negation

Verb and adjective negation is marked by the suffix *-[a]zu* (CONC; INF used in coordination only), which has the attributive *-[a]nu* and the realis *-[a]ne*. When it needs other stems, or the infinitive takes endings, the extended form *-[a]zar<sup>R</sup>-* is used. This also supplies alternative attributive and realis forms. The negative precedes tense and aspect endings. However, tentative *-[a]m-* is always replaced in the negative with the special form *-[a]zi*. *ar<sup>R</sup>-* ‘exist/have’ has an alternative suppletive negative *na<sup>A</sup>-*. Debitive *-ube<sup>A</sup>-* has a suppletive negative *-umazi<sup>A</sup>-* (18c) as well as the regular *-ube-kar-azu*. In the negative, copular *-nar<sup>R</sup>-* often splits to its etymological *-ni ar<sup>R</sup>-*.

- (18) a. *fi-ni yake-nu koto-yori-mo keura-nar-u koto*  
 fire-DAT burn-NEG.ATTR fact-ABL-INCL beautiful-COP.AUX-ATTR fact  
*narabi na-si*  
 comparison not.exist-CONC  
 ‘That it was beautiful was more incomparable even than the fact that it didn’t burn in fire.’ (TM)
- b. *yuk-u kafa-no nagare-fa taye-zu s-i-te sikamo*  
 go-ATTR river-GEN flow-TOP cease-NEG.INF do-INF-GER yet  
*moto-no midu-ni ar-azu*  
 origin-GEN water-COP AUX-NEG.CONC  
 ‘The flow of the rolling river is ceaseless, yet it is not the original water.’ (H, start)
- c. *yo-ni ar-umazi-ki kokoti-no s-i-ker-e-ba*  
 world-DAT exist-NEG.DEB-ATTR feeling-GEN do-INF-UNW-REAL-DISC  
 ‘He felt he couldn’t/shouldn’t go on living, and . . .’ (TM)

## 8.5.3.4 Modality

Modality is expressed by verb-following expressions. Epistemic forms include *-[a]m-* possibility/probability (neg. *-[a]zi*: e.g. *sarani i-zi* ‘You’ll definitely **not** hit it’ (KMS 27/4)), *-unar<sup>R</sup>-* hearsay (26a), attributive + copula *-nar<sup>R</sup>-* assertion/explanation (neg. *-[a]nu-nar<sup>R</sup>-* ~ *-[a]zan-nar<sup>R</sup>-*),<sup>3</sup> *-umer<sup>R</sup>-* (or *-ubera-nar<sup>R</sup>-* in male-style prose) deduction or assumption (19a). *-ar<sup>R</sup>-* and *-nar<sup>R</sup>-* tend to fuse with the last three: e.g. *-nar-umer<sup>R</sup>-* > *-na(N)mer<sup>R</sup>-*, *-ar-u-nar<sup>R</sup>-* > *-a(N)nar<sup>R</sup>-*. An evidential based on visual or aural evidence is derived from true adjectives and *nari*-adjectives by root + *-ge* + copula (19b, 24). Root + *-gar-* expresses an evidential of an adjective of emotion (19c); it can also be attached to a noun root (‘look like a’).

- (19) a. *fotoke afare to obosimes-i-tar-i-ker-unamer-i*  
 Buddha pity QUOT think(HON)-INF-PERF-INF-UNW-MOD-CONC  
 ‘It seems that Buddha actually took pity on him.’ (USM 6/4)
- b. *imizi-u nagekasi-ge-ni omof-i-tar-i*  
 terrible-INF be.upset-EVID-COP.INF feel-INF-PERF-CONC  
 ‘She felt terribly upset.’ (TM)
- c. *kaguyafime ayasi-gar-i-te mi-ru-ni*  
 Kaguyahime be.suspicious-EVID-INF-GER look-ATTR-DISC
- fati-no naka-ni fumi ar-i*  
 bowl-GEN inside-DAT letter exist-CONC  
 ‘When Kaguyahime looked at it suspiciously, there was a letter inside the bowl.’ (TM)

The main deontic device is debitive *-ube<sup>A</sup>-* ‘should; have to’ (neg. *-ube-kar-azu*, *-umazi<sup>A</sup>-*). Volition is expressed primarily by the desiderative construction *-[a]m-afosi<sup>A</sup>-* ‘I want to’ (neg. *-[a]m-afosi-kar-azu*), which appears in early texts as *-[a]m-aku fosi<sup>A</sup>-*, and ultimately derived from a nominalization + *fosi<sup>A</sup>-* ‘I want to have’ (20a). A wish is frequently expressed by *-[a]ba ya* ‘If only . . .’ (neg. *-[a]zar-aba ya*). The debitive is also used with potential meaning (20b). Unambiguous potential meaning can be expressed by infinitive + *-kate-* for ‘can’, and *ye(-fa)* + negative verb (20c) or infinitive + *-kane-* for ‘cannot’. The passive morpheme *-[r]are-* is also used in potential function (20d). The debitive can also express a deduction (20e).

- (20) a. *yuk-amafosi-ku omof-u-ni seuto-nar-u fito*  
 go-want-INF feel-ATTR-DISC elder.brother-COP.AUX-ATTR person  
*idak-i-te wi-te yuk-i-tar-i*  
 hold.in.arms-INF-GER take-INF-GER go-INF-PERF-CONC  
 ‘I wanted to go to her, so the one who was my elder brother had taken me there in his arms.’ (SN)
- b. *ofo-ku make-te watas-ube-ki mono*  
 be.much-INF lose-INF-GER hand.over-DEB-ATTR thing  
*na-kar-i-ker-u-ni ita-ku*  
 not.have-INF.AUX-INF-UNW-ATTR-DISC painful-INF  
*seme-ker-e-ba omof-i wab-i-te*  
 attack-INF-UNW-REAL-DISC think-INF- apologize-INF-GER

‘He lost a lot [at backgammon] and did not have anything he could/should hand over [to pay his debts], and as his opponent abused him a lot [for this], he apologized and [. . .]’ (USM 6/4)

- c. *motome-tatematur-e-do-mo mi-sin-i-mo-ya s-i-tamaf-i-*  
 search.INF-HUM-REAL-but-INCL HON-die-VN-INCL-Q do-INF-HON-INF-  
*-kem-u ye mituke-tatematur-azu nar-i-n-u*  
 PAST.TENT-ATTR POT find.INF-HUM-NEG-INF become-INF-PFV-CONC  
 ‘Even though they searched for him, it seemed he might even have died [lit. an aside: – might he have even died? – ] and they ended up unable to find him.’ (TM)
- d. *i-mo ne-rare-zu*  
 sleep-INCL sleep-PASS-NEG.CONC  
 ‘I couldn’t even sleep a wink.’ (TM)
- e. *ware-fa ko-no wauzi-ni make-n-ube-si*  
 I-TOP this-GEN prince-DAT lose-PFV-DEB-CONC  
 ‘It must be that I’ve lost to this prince.’ (TM)

#### 8.5.3.5 Non-declarative sentence types

Questions have the same word order as statements. There are two question particles, *-ka* and *-ya*. They follow the verb at the end of the sentence only when the verb itself constitutes their scope, in which case *-ka* is preceded by the attributive and *-ya* by the conclusive. Usually, however, they follow the phrase within their sentence that constitutes their scope. In this case they always trigger a change of the verb at the end of the sentence from its expected conclusive form to the attributive. The difference between the two particles is that question-word questions take *-ka* (21), whereas yes/no questions normally take *-ya* (20c, 27a), though *-ka* is used for particular effects, such as a rhetorical question. The sequence *ikani* ‘how’ + *-ka* normally reduces to *ikaga*, which in later texts becomes reinterpreted as a single morpheme and attracts a second *-ka*. Where the complement contains a question-word, the copular verb and question particle are commonly expressed with an otherwise unused question particle *-zo* as *X(-fa) tare-zo* ‘Who’s X?’.

- (21) *um-are sin-uru fito idukata-yori k-i-tar-i-te*  
 bear-PASS-INF die-ATTR people where-ABL come-INF-PERF-INF-GER  
*idukata-fe-ka sar-u*  
 where-ALL-Q go.away-ATTR  
 ‘Where have the people who are born and who die come from and where do they go to?’ (H)

Commands are expressed in the affirmative by the imperative (17, 26a), and in the negative by root + *-una* (or conclusive + *-na?*), or *na* + infinitive + *-so* (22a, b). Tentative *-[a]m-u* is used to express cohortative function (‘let’s; I shall’).

- (22) a. *ko-no tama tor-i-e-de-fa ife-ni*  
 this-GEN jewel acquire-INF-get-NEG.GER-TOP house-DAT  
*kafer-i-k-una*  
 return-INF-come-PROH  
 ‘Don’t come back home unless you’ve got hold of that jewel.’ (TM)

- b. *fito na ita-ku wabi-sase-tatematur-ase-tamaf-i-so*  
 people PROH painful-INF be.upset-CAUS-HUM-HON-HON-INF-PROH  
 ‘Don’t upset people too much!’ (TM)

There are also a few sentence-final particles that add emotive effects to a preceding sentence: assertive *yo*, exclamatory *ya*, tag or rhetorical ATTR + *kana*. The *yo* can be attached to quadrigrade imperatives for greater impact (17), but not to bigrade/monograde imperatives as these already contain a *-yo*.

#### 8.5.4 Topic, focus and emphasis

The topic marker is *-ba* after the particles *-wo* [ACC] and *-koso* [FOC], and also in frozen verbal forms, and *-fa* elsewhere. It marks: the topic (old information) at the start of a sentence; contrast, typically marking the phrases in two parts of a sentence that are in contrast (a function which underlies frozen expressions such as *aruifa . . . aruifa . . .* ‘on the one hand . . . on the other . . .’); and a ‘selective’ function that Modern Japanese does not share, where the item marked with *-fa/-ba* represents a choice from several. The latter can be seen especially in the sequence *-koso-ba* [FOC-TOP] or the occasional combination of question-word + *-fa*, e.g. *ikaga-fa se-m-u* ‘What shall we do?/How shall we do it?’ (TM). (Question-word + *-wa* is ungrammatical in Modern Japanese, but see the Yonaguni ‘selective’ particle *-ba*, Izuyama, this volume). *-fa* in all functions is optional (cf. *kore-wo kaguyafime kik-i-te*, 8.5.3.1).

*-fa* is also sometimes attracted to a negative verb, coming as close to the verb at the end of the sentence as possible. Where the verb is, for example, copular *-ni* + the auxiliary *ar<sup>R</sup>-*, *-fa* appears in the middle of the combination, splitting the two elements, e.g. *tori-nar-azu ~ tori-ni-fa ar-azu* ‘it isn’t a bird’. Although this *-fa* is again optional, even when it is not explicit the splitting very often still occurs: *tori-ni ar-azu*.

Inclusive *-mo* ‘also; even’ occupies the same syntactic slot as, and cannot co-occur with, *-fa*. It is also used after two or more phrases as ‘both X and Y’ (3, 15). The adverb *mata* ‘also’ can reinforce a preceding *-mo*, or even replace it (14).

There are three main emphatic particles: focal *-koso* (11b, 25), emphatic *-zo* (17) and the mild *-namu/-nan* (1). Quite what the function of the latter was is unclear, and its high frequency in certain later texts suggests it had become clichéd.

These three particles follow the phrase within a clause that constitutes their scope. They consistently trigger the replacement of the expected conclusive form of the sentence-final verb with the realis form in the case of *-koso* and the attributive with *-zo* and *-namu/-nan*. This phenomenon, also found with the question particles *-ka* and *-ya*, is known by Japanese linguists as *kakari-musubi*, and the five triggering particles are known as *kakari* particles.

A fourth emphatic device *-si-mo* (OJ emphatic *si* + inclusive *-mo*) is occasionally encountered, and very rarely even *-si*.

#### 8.5.5 Passive and causative

The passive is formed with the verb suffix *-[r]are-*. The agent, when expressed, is marked with *-ni*. The passive also is used as an honorific marker, a potential marker (20d), and to express an action that is spontaneous and beyond the control of the agent (16c). There are two causative morphemes, *-[s]ase-* (5, 22b) and *-[a]sime-*; the former also functions as an honorific (22b).

The language has many pairs of verbs formed from a common root, of which one is normally intransitive (or a verb of motion whose path is marked with *-wo*) and the other transitive, though in a few cases both verbs are transitive, especially where the action involves two agents and the verbs take the two agents' different perspectives. The pairs make use of four morphemes: transitive *-[a/o/u]s-* (which may be related to causative *-[s]ase-*), intransitive *-[a]r-* (which may be related to passive *-[r]are-*), intransitive *-i-*, and transitive or intransitive *-[y]e-*. Which is used is lexically determined, e.g.: *ki*<sup>M</sup>- 'put on O' ↔ *ki-s-* 'dress O', *kafe-* 'change' ↔ *kafe-s-* 'change O', *sug-i-* 'pass' ↔ *sug-us-* 'pass O', *forob-i-* 'be destroyed' ↔ *forob-os-* 'destroy O', *f-i*<sup>M</sup>- 'dry' ↔ *f-os-* 'dry O', *yak-e-* 'burn' ↔ *yak-* 'burn O' (contrasted in (25)), *tamaf-e-* > *tab-e-* 'be given O' ↔ *tamaf-* > *tab-* 'give O', *fuy-e-* 'increase' ↔ *fuy-as-* 'increase O', *id-e-* 'go out' ↔ *id-as-* 'take O out', *ko-ye-* 'go over' ↔ *ko-s-* 'take O over', *tuk-* 'be attached' ↔ *tuk-e-* 'attach O', *ir-* 'go in' ↔ *ir-e-* 'put O in', *tat-* 'stand up; be built' ↔ *tat-e-* 'stand O up; build O', *ag-ar-* 'rise' ↔ *ag-e-* 'raise O', *kak-ar-* 'hang' ↔ *kak-e-* 'hang O', *k-ar-i-* 'borrow O' ↔ *k-as-* 'lend O', *tafu-r-* 'fall over' ↔ *tafu-s-* 'knock O over', *kafe-r-* 'go back' ↔ *kafe-s-* 'give O back'.

Related to these are three pairs of verbs of perception, of which the basic member expresses the experiencer's deliberate action and the other, through the suffix *-[o]ye-*, indicates the experiencer's automatic or non-volitional involvement in the event: *kik-* 'hear' > *kik-oye-*, *mi*<sup>M</sup>- 'see' > *mi-ye-*, and *omof-* 'think' > *omof-oye-* > *ob-oye-*.

### 8.5.6 Speech levels and respect

CJ reflects the highly stratified society of the time, and so the actions of anyone of higher rank or of lower rank to the speaker/author are marked to reflect their status, using 'honorific' and 'humble' forms respectively. The most common devices are to mark the verb with infinitive + *-tamaf-* > *-tab-* (lit. 'give (HON)') and infinitive + *-tatematur-* (lit. 'give (HUM)') respectively. However, certain verbs have suppletive forms, e.g. *if-* 'say' > *notamaf-* (HON), *mawos-/maus-* (HUM); *yuk-* 'go' / *ko-* 'come', *ofase*<sup>S</sup>- / *ofasimas*<sup>(S)</sup>- / *imas*<sup>(S)</sup>- (HON), *maude-* ~ *mavir-* (HUM); *ar*<sup>R</sup>- 'exist; have' > *ofase*<sup>S</sup>- / *ofasimas*<sup>(S)</sup>- / *imas*<sup>(S)</sup>- (HON), *saburaf-/safuraf-* (HUM, lit. 'serve'). Copular *-n(i)ar*<sup>R</sup>- 'be' is always 'split' when auxiliary *ar*<sup>R</sup>- is replaced with an honorific/humble equivalent: *sama-mo yo-ki fito-ni ofas-u* 'the Prince is a good-looking man' (TM). Some compounds involve a (sometimes redundant) semantic matching of the honorific/humble element, e.g. *inab-* 'refuse' > *inab-i-maus-* 'refuse (HUM)', *ko-* 'come' > *maude-* ~ *maude-ko-* 'come (HUM)', *ide-* 'leave' > *makar-* ~ *makar-i-ide-* 'leave (HUM)'.

The language possessed a range of honorifics that reflected different levels of high rank. These devices include the causative morpheme *-[s]ase-*, lexicalized survivals of OJ *-as-/os-* (e.g. *obos-* 'think; feel'), the passive morpheme *-[r]are-* and multiple marking of respect. In actions that involved two people of different ranks other than the speaker, verbs could be marked simultaneously as honorific and as humble. In (22b), the speaker shows respect to the addressee, hence honorific *-[s]ase-* and *-tamaf-*, but her status is lower than that of the 'people' who are courting her and whom she constantly refuses to marry, hence humble *-tatematur-*.

Politeness (addressee honorification) was not widely expressed in the types of material from the Heian period that we have, but the language had a polite equivalent of *ar*<sup>R</sup>- 'exist; have', *faber*<sup>R</sup>- = OJ *pabyer*<sup>R</sup>-. Infinitive + *-faber*<sup>R</sup>- created a polite verb, and even copular *-nar*<sup>R</sup>- could be replaced by *-ni faber*<sup>R</sup>-. Later, humble *(-)safuraf-* also became debased into a polite form, and this form developed into the polite auxiliary that characterized the epistolary style *sōrōbun* of more recent times.

### 8.5.7 Adverbials

CJ has a range of adverbials, from basic (e.g. *ito* ‘very’, cf. *ita-ku* ‘painfully; very’; *faya* ‘early; soon/already; really’, cf. *faya-ku* ‘early’; *tada* ‘just’; *mata* ‘also; again’; *nafo* ‘still’), nominal + copular (or dative?) *-ni* (e.g. *tatimati-ni* ‘suddenly’), and verbal + gerund *-te* (e.g. *sugurete* ‘extremely’ < *sugure-* ‘excel’). Some of the latter have modal function (e.g. *karauzite* + Verb = ‘manage to Verb’). The simple infinitive of adjectives and *nari*-adjectives functions as an adverb, e.g. *imizi-u* ‘terribly’ (cf. conclusive *imizi-Ø* ‘it’s terrible’), *yasu-ku* ‘easily’ (cf. *yasu-si* ‘it’s easy’), *afare-ni* ‘pitifully’ (cf. *afare-nar-i* ‘it’s a shame’). Unlike when they are coordinated as predicates, adjective infinitives used as adverbs do not take an additional *-te*.

Case particles are used to produce adverbial phrases. To indicate location in time, *-ni* is optional, and expressions of relative time such as *kefu* ‘today’, *kinofu* ‘yesterday’, *kozo* ‘last year’, *ima* ‘now’, are used without any case particle.

The language has a range of grammatical nouns that occur as the second noun in sequences of Noun<sup>1</sup> + *-no/-ga* + Noun<sup>2</sup> + case particle, and which function as complex postpositions, e.g. *naka* ‘middle; midst’, *soto* ‘outside’, *ufe* ‘on; above’, *sita* ‘under’, *tame* ‘purpose; sake’, *mafe* ‘front; before; presence’, *moto* ‘chez’, *fotori* ‘vicinity’, *kata* ‘direction’, *uti* ‘inside’, *yau* ‘appearance; manner’, *fodo* ‘period’.

### 8.5.8 Coordination

‘Coordination’ can be interpreted in two ways. Functionally, it can be taken to be the simplest and common equivalent(s) to English ‘and’; syntactically, it refers to clause linkage that has none of the features that characterize syntactic subordination. In the syntactic sense, Japanese only has (a.) infinitive, infinitive + *-te* ‘and’ (adjectives and copula, also simple infinitive + *s-i-te*), and their negative equivalents *-[a]zu* (16c), *-[a]zu(si)-te* (18b), and *-[a]de*. In the functional sense we can include (b.) realis + *-ba*, attributive + *-ni/-wo*, and their negative equivalents *-[a]n-e-ba* and *-[a]zar-u-ni/-wo*.

I take the functional approach here, because INF and INF-*te* in most cases indicate that the two linked verbs share a common subject, which in the syntactic sense means that there are no devices for coordinating verbs with different subjects.

The vast majority of verbs and clauses in CJ are linked by the devices listed above as (a.) or (b.), which allows all of them to be translated as ‘and’, ‘and then’, ‘and so’, ‘when’, ‘because’. There are functional and syntactic differences between them. The functional difference between them is that (b.) express some sort of discontinuity. In the vast majority of cases, this discontinuity manifests itself as switch-subject. Overwhelmingly, then, (a.) indicate a retention of a subject, whereas (b.) indicate a change of subject, which allows CJ texts to express explicit subjects so sparingly. However, there are significant numbers of exceptions. Quite a number of exceptions can be characterized as cases of Clause<sup>1</sup> + Clause(s)<sup>2</sup> + Clause<sup>3</sup>, where the coordinating marker is attached at the end of Clause<sup>1</sup>, where Clause<sup>1</sup> grammatically coordinates with Clause<sup>3</sup>, and Clause(s)<sup>2</sup> and Clause<sup>3</sup> have different subjects. In these cases, there appears to be some hesitation whether the clause to which Clause<sup>1</sup> grammatically links or the immediately following clause should determine the choice of coordinator ((23a–b), the latter reinterpreted from McAuley 2002). Most other exceptions are probably because there is perceived to be a logical closeness or a logical distance between two coordinated situations, regardless of the subject (23c).

- (23) a. *oni-no*                      *um-i-tar-i-ker-e-ba*<sup>1</sup>                      *oya-ni*  
          demon-GEN      bear-INF-PERF-INF-UNW-REAL-DISC      parent-DAT

*ni-te*<sup>2</sup>                      *kore-mo osorosi-ki*                      *kokoro ar-am-u*<sup>2</sup>  
 resemble.INF-GER it-INCL be.frightening-ATTR heart have-TENT-CONC

*to-te*<sup>3</sup>                      *oya-no*                      *ayasi-ki*                      *kinu fikikise-te*  
 QUOT-GER parent-GEN be.crude-ATTR cloth swaddle.INF-GER  
 ‘A demon had given birth to it, and thinking that it too might have a terrible  
 nature resembling its parent, its parent swaddled it in rough cloth and [. . .]’  
 (MS)

- b. *furu-ki*                      *miya-fa*                      *kaferite*                      *tabikokoti*  
 be.old-ATTR palace-TOP rather journey.feeling

*s-i-tamaf-u-ni-mo*<sup>1</sup>                      *ofon-satozumi*                      *taye-tar-u*<sup>2</sup>  
 do-INF-HON-ATTR-DISC-INCL HON-living.at.home cease.INF-PERF-ATTR

*tosituki-no*                      *fodo*                      *obosimegurasar-ube-si*<sup>3</sup>  
 years.months-GEN period think.about.HON-DEB-CONC

‘With the old palace, [Fujitsubo] rather felt she was just passing through, and  
 she must have been thinking of the months and years since her living there  
 had come to an end.’ (GM 10)

- c. *aruifa*                      *kozo*                      *yake-te*                      *kotosi tukur-er-i*.  
 on.one.hand last.year burn (intr.).INF-GER this.year build-PERF-CONC

*aruifa*                      *ofo-ife*                      *forob-i-te*                      *ko-ife-to*  
 on.one.hand big-house be.destroyed-INF-GER small-house-ESS

*nar-u*

become-CONC

‘On the one hand, [the houses] burnt down last year and we’ve built [new  
 ones] this year. On the other hand, big houses are destroyed and become  
 small houses.’ (H)

The syntactic differences are threefold. Firstly, INF and INF-*te* do not allow the verb to which they are attached to be marked for tense; the tense of the verb that they link with grammatically – if it is marked – is taken to apply also to the verb marked by INF or INF-*te*. However, unlike formal Modern Japanese, it is not necessarily the case that the scope of every verb-following construction at the end of a sequence of INF(-*te*)-linked verbs includes all such linked verbs. On the other hand, REAL-*ba* and ATTR-*ni* normally require the same tense marking as would occur sentence-finally, and so if a narrative is presented sentence-finally as unwitnessed with *-ker*<sup>R</sup> we normally encounter *-ker-eba* or *-ker-u-ni* clause-internally too.

Secondly, an explicit subject of a verb that ends in INF(-*te*) is zero-marked as a nominative, but because REAL-*ba* and ATTR-*ni* both derive from relative clause structures an explicit subject of a verb that ends in either of these is typically marked with a possessive *-no* or *-ga*.

REAL-*ba* and ATTR-*ni* have a number of other functions beyond their coordinative role. For example, prose sometimes makes use of afterthoughts, whereby the writer remembers mid-sentence some important fact that has not been mentioned, and so introduces this fact mid-sentence in the form of an aside. REAL-*ba* and ATTR-*ni* are used to rejoin the narrative of the sentence.

- (24) *aki*                      *kita-no*                      *kata-ni*                      *yamazato-ni-te*  
 autumn north-GEN direction-DAT mountain.village-COP-GER



to the introduction of quotations (see below). Other clause nominalizations involve attributive + a formal noun, such as *koto* ‘fact’, *yau* ‘manner’, *sama* ‘appearance’, etc.

Quoted speech and thoughts appear always to be direct quotation, with no indication of subordination within the quotation relative to the verb of quotation, i.e. an unmarked quoted statement ends in the conclusive, the scope of *kakari* particles to trigger *kakari-musubi* does not extend beyond the end of the quotation, and the inside of the quotation does not affect the choice of switch- and non-switch-reference coordinators outside it. Quotations typically are followed by either (a.) *to* by itself, or (b.) *to* + an explicit verb of quoting, e.g. *if-* ‘say’, *kotafe-* ‘reply’, *tof-* ‘ask’, *omof-* ‘think’. *to* appears to be a frozen verb ‘say’, and so *to* in construction (a.) may be regarded as a conclusive form, and *to* in construction (b.) may be regarded as a coordinating infinitive form. The prime evidence supporting this analysis of a verbal origin, other than parallel verbs in both the above constructions in other northern Asian languages (cf. Mongolian *ge-* ‘say’, *gež* QUOT; Manchu *se-* ‘say’, *seme* QUOT), is the fact that *to* itself can carry gerund *-te*, so that ‘and’ (same subject) is expressed as quotation *to*( *if/omof-i*) or quotation + *to*( *if/omof-i*)-*te* (5, 26a). A similar development is seen in the Yonaguni quotative particle *ndi*, derived from *ndun* ‘say’ (Izuyama, this volume). Evidence against *to* being synchronically verbal is the fact that it can readily be followed by emphatic/focal/topic particles (8.4.1, slots (d.–e.)) or even *-bakari*. A verb of quoting is obligatory after the *to*, also, if endings other than *-te* are required, e.g. switch-subject *-ba*. Also, *to* has no attributive form, and the fuller form *to if-u* fulfils this function (*mi-ka to if-u fi* ‘on the day they’d designated the third day’).

Because quotations can be long, and involve a significant break in the normal flow of a CJ sentence, there is a tendency to introduce a quotation with a verb of quoting. This uses *-[a]ku*, ((26b), an otherwise unused survival of an OJ nominalization) or some other nominalization, such as attributive + *yau*. When a quotation is introduced in this way, the quoting structure after it becomes optional.

- (26) a. *fumi-wo kak-i-te* “*finezumi-no kafa to if-unar-u*  
 letter-ACC write-INF-GER fire.rat-GEN hide QUOT call-HEARSAY-ATTR  
*mono kaf-i-te okose-yo*” *to-te tukaumatur-u fito-no*  
 thing buy-INF-GER send-IMP QUOT-GER serve-ATTR person-GEN  
*naka-ni kokoro tasika-nar-u-wo erab-i-te*  
 midst-DAT heart certain-COP.AUX-ATTR-ACC choose-INF-GER  
 ‘He wrote a letter, saying, “Please buy and send me what’s apparently known  
 as a fire-rat’s hide,” and from amongst those who served him he chose one  
 who was reliable, and [. . .]’ (TM)
- b. *wonna kotafe-te if-aku* “*kore-fa fourai-no*  
 woman reply-INF-GER say-NML this-TOP Penglai-GEN  
*yama-nar-i*” *to kotaf-u*  
 mountain-COP.AUX-CONC QUOT reply-CONC  
 ‘The woman replied, saying: “This is Mount Penglai.”’ (TM)

Other examples of quotation are (11, 13b, 17). The exemplifying particle *-nado* (8.4.1) has been extended as a substitute to *to*, combining the exemplifying function of the former and the quoting function of the latter (27a). In this use it is preceded by the same forms that precede quoting *to*, i.e. it does require a nominal(ization). Tempting though it may be to derive this *nado* from some compound with quoting *to*, it appears to have

none of the verbal use of the latter: it cannot act like an infinitive without a following quotative verb, and \**nado-te* does not occur.

Noun + *to* (or *nado*) + *if-* means ‘they call X Noun’, i.e. ‘X is called Noun’. It is very common in linking two nouns (2, 26a, 27b).

- (27) a. *kaguyafime-ni sum-i-tamaf-u to na. koko-ni-ya*  
 Kaguyahime-DAT live-INF-HON-CONC QUOT ! here-DAT-Q  
*imas-u” nado tof-u*  
 be.HON-ATTR EX ask-CONC  
 ‘They asked things like: “They say he’s moved in [here] with Kaguyahime!  
 Is he here?”’ (TM)
- b. *mukasi simotusa-no kuni-ni mano.no.teu to*  
 long.ago Shimōsa-GEN province-DAT Mano.no.Chō QUOT  
*if-u fito sum-i-ker-i*  
 call-ATTR person live-INF-UNW-CONC  
 ‘Long ago, a man called Mano no Chō lived in the province of Shimōsa.’  
 (SN)

The language also made frequent use of embedding questions in the middle of statements with no overt grammatical link with the rest of the sentence. See (9) and (20c).

## 8.6 LEXICON

The EMJ period is characterized by an influx of Chinese vocabulary (‘Sino-Japanese’). At the start, most of this was learned and predominantly belonged to the spheres of the court, administration and religion, and the proportion of Chinese words is relatively small in the early texts. However, as the period progressed, the proportion rises, though Sino-Japanese was always just a fraction of the lexicon. Sino-Japanese entered in two waves, both probably pre-EMJ, but just not represented in OJ because OJ is mostly preserved only in poetry. The first wave, later known as go’on ‘Wu sounds’, is related to Early Middle Chinese and entered in the late Asuka/early Nara periods. The second wave, later known as kan’on ‘Han sounds’, is derived from Late Middle Chinese of the Tang dynasty. The ‘Wu’ and ‘Han’ are names of much older dynasties, but appear to have been used more as ethnic tags. ‘Han’ – as still used in Chinese – was used to represent ethnic Chinese, i.e. ‘proper Chinese’, and appears to have contrasted with ‘Wu’ as non-ethnic Chinese. It may be that go’on entered Japanese via a Korean intermediary (though there is no clear proof of this), or because it derived from an earlier variety before the promotion of the Tang spoken standard, but did not supplant the go’on stratum. Modern Japanese has inherited both sets of pronunciations for the majority of Chinese characters. The go’on :: kan’on strata reflect EMC voiced consonants :: LMC voiceless murmureds (CJ *b* :: *f*, *d* :: *t*, *g* :: *k*, *z* :: *s*), EMC *mj-* and *bj-* :: LMC fricatives (CJ *m* :: *b*, *b* :: *f*), EMC *ny* :: LMC (fricative) *r* (CJ *n(y)-* :: *z(y)-*). An unusual distinction is go’on *m-* and *n-* :: kan’on *b-* [ʰb] and *d-* [ʰd], where both EMC and LMC have *m-* and *n-*, implying the source variety of kan’on differed somewhat from the Tang standard.

Other than place names within the capital in (16b) and numbers, the following Sino-Japanese forms occur in the examples used in the Morphology and Syntax sections of this chapter: *amida* 阿彌陀 ‘Amitābha bodhisattva’, *-ben* 遍 ‘... times’, *butu* 佛 ‘Buddha’, *dairi* 内裏 ‘imperial palace compound’, *dausin* 道心 ‘deep faith’, *dinkwai* 塵灰 ‘dust and

ash', *fourai* 蓬莱 'Penglai, mythical island of immortality', *fusyau* 不請 'unsolicited', *-gwatu* 月 '... th month', *kau* 更衣 'second-rank court ladies', *mon* 紋 'crest; patch', *-nen* 年 '... years; ... th year', *-niti* 日 '... days; ... th day', *nyougo* 女御 'first-rank court ladies', *rei* 例 'custom' (*rei-no* 'usual'), *ruri* 琉璃 'lapis lazuli', *seifoku* 西北 'north-west', *tounan* 東南 'south-east', *wauzi* 皇子 'prince', *yau* 樣 'manner; appearance', *zeokon* 舌根 'tongue' (Buddhist term). Despite the 'correctness' of the kan'on stratum, reflected in *dairi* 内裏, most of the words above have EMC-derived, i.e. go'on, readings, e.g. *mon* 紋 < EMC *mjun* vs LMC *vjyn/vun*, *dausin* 道心 < EMC *daw.sim* vs LMC *thaw.sim*, *butu* 佛 'Buddha' < EMC *bjut* vs LMC *ffijyt/ffhut*, *zeokon* 舌根 < EMC *zyet.kon* vs LMC *ffhiat.kən*. Most of these are in the areas of the court, administration or Buddhism, or are specifically China-related concepts (Penglai is Taoist). Some have no such connections, such as *yau*, which is fully grammaticalized by the start of the period, and probably is the morpheme in *sa-* / *ka-* / *ika-yau-ni* (8.5.2). In many cases, the use of Chinese characters in a text disguises whether a Chinese or a native pronunciation was intended, so 塵灰 in *Hōjōki* could represent either *dinkwai* or *tiri fafi* ('dust [and] ash'), or 皇子 in *Taketori monogatari* could represent either *wauzi* or *mi-ko* (HON-child). Similarly, directions such as 西北 kan'on *seifoku* could also be go'on *saifoku*, native *nisi kita* ('west [and] north'), or zodiacal *usi tora* ('ox [and] tiger'). To a great extent, we have to rely on later readings to predict what pronunciation is hidden by Chinese characters, and either competing doublets or 'reading pronunciations' mean that pronunciations do vary over time, e.g. 愛敬 CJ *aigyau* vs LMJ <*aikyau*> vs modern *aiki* 'love, respect'.

Most Chinese loans were nouns, but verbs and adjectives were also borrowed, incorporated into CJ morphology through the addition of *-se<sup>s</sup>-*/*-ze<sup>s</sup>-* 'do' and the copula respectively: *nenze<sup>s</sup>-* 'think; pray' < 念, *anze<sup>s</sup>-* 'consider' < 案, *yeuze<sup>s</sup>-* 'want' < 要, *goranze<sup>s</sup>-* 'see (HON)' < 御覽, *beti-nar<sup>R</sup>-*/*-ni* 'be separate/otherwise' < 別. Other means of incorporation are reflected by *aigyauduk-* 'be attractive' < 愛敬 *aigyau* 'love' + *-ni tuk-* 'be attached to', and *sauzok-* 'get dressed' < 装束 *sauzoku* 'clothes', the latter being the only example of a verb borrowed without any additional morpheme.

## NOTES

- \* My thanks are especially to the years of students who have taken the 'Evolution of the Japanese Language' module (EAS342 and EAS232) at the University of Sheffield, and have proved by their performance that, contrary to the statistics, the *-te* vs *-ba* same/switch-subject 'rule' works in virtually every case where they actually needed such a rule.
- 1 Middle Chinese is given in Early Middle Chinese, corresponding strictly to the earlier so-called go'on stratum of Chinese loans in Japanese, following the Baxter (1992) system. Late Middle Chinese in section 6 follows Pulleyblank (1991). Indications of tone have been removed.
- 2 It is tempting to add *-te* + *mi<sup>M</sup>-* 'see' to mean 'try [do]ing', because this construction exists in the later language. However, virtually all examples that I have encountered can be interpreted literally as '[do] and look', and there is an absence of forms such as *kuf-i-te mi<sup>M</sup>-* 'try eating' in tales where the modern language would have the equivalent construction.
- 3 Being the copula, this *-nar<sup>R</sup>-* is sometimes ellipted. This leads to a distinction between a neutral sentence ending in the conclusive and an assertive/explanatory sentence ending in the attributive, e.g. *muma-ni nor-i-te, tukusi-yori tada nan-uka-ni nobor-i-maude-k-i-tar-u* 'he mounted the horse and came up [to the capital] from Tsukushi in just seven days' (*TM*).

## ABBREVIATIONS

As this chapter deals with the whole system of CJ, I have taken a maximal morphemic analysis, e.g. *nak-i-ker-e-ba* ‘cry-INF-UNW-REAL-DISC’. In a paper on a narrow linguistic topic, of course, this would be over-analysis, where *naki-ker-eba* would be more than enough.

APPROX = approximative, ASSERT = assertive, ATTR = attributive, AUX = auxiliary *ar<sup>R</sup>*-, CF = counterfactual, CONC = conclusive, COP = copular *-ni*, DEB = debitive, DIREC = directive, DISC = discontinuity, ESS = essive, EVID = evidential *-gar/-ge*, EX = exemplary, GER = gerund, HAB = habitual, INCL = inclusive, MOD = modal, PERF = perfect, PFV = perfective, REAL = realis, REC = recent past, TENT = tentative, TERM = terminative, UNW = unwitnessed, VN = verb-noun.

## TEXTS

Volume and tale numbers are added to *KMS* and *USM*, chapter numbers to *GM*, and month and day numbers to *TN* citations. *MS* citations are from the ‘Mushi-wa’ section, *SN* citations from the start (up to the ‘Takeshibadera’ tale), and *TM* citations are from the suitors’ tales. All examples have been checked against *Nihon Koten Bungaku Taikai* editions, and where readily available also against manuscripts. The only script example in [Table 8.5](#) that is not adapted directly from a manuscript is (d.), which is adapted from Yanase (1967).

<i>GM</i>	<i>Genji monogatari</i>	<i>SN</i>	<i>Sarashina nikki</i>
<i>H</i>	<i>Hōjōki</i>	<i>TM</i>	<i>Taketori monogatari</i>
<i>KMS</i>	<i>Konjaku monogatari-shū</i>	<i>TN</i>	<i>Tosa nikki</i>
<i>MS</i>	<i>Makura no sōshi</i>	<i>USM</i>	<i>Uji shūi monogatari</i>

## REFERENCES

- Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.
- Frellesvig, Bjarke (1995) *A Case Study in Diachronic Phonology: The Japanese Onbin Sound Changes*, Aarhus: Aarhus University Press.
- Ikeda Kikan, Kishigami, Shinji and Akiyama, Ken (eds) (1958) *Makura no Sōshi – Murasaki Shikibu Nikki*, *Nihon Koten Bungaku Taikai* vol. 19, Tokyo: Iwanami Shoten.
- Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.
- McAuley, T. E. (2002) ‘Switch-reference and semantic discontinuity in Late Old Japanese’, *Journal of Japanese Linguistics*, 18: 25–49.
- Mochizuki, Ikuko (1974) *Ruiju Myōgishō Yonshu Shōten-tsuki Wakun Shūsei* [Compilation of Japanese readings with four types of tone marks in *Ruiju Myōgishō*], Tokyo: Kasama Shoin.
- Nishio, Minoru (ed.) (1957) *Hōjōki – Tsurezuregusa*, *Nihon Koten Bungaku Taikai* vol. 30, Tokyo: Iwanami Shoten.
- Ōtsu, Yuichi and Tsukishima, Hiroshi (eds) (1957) *Taketori Monogatari – Ise Monogatari*, *Nihon Koten Bungaku Taikai* vol. 9, Tokyo: Iwanami Shoten.
- Pulleyblank, Edwin G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver: UBC Press.
- Sandness, Karen (1999) *The Evolution of the Japanese Past and Perfective Suffixes*, Ann Arbor: Center for Japanese Studies, The University of Michigan.

- Shimabukuro, Moriyo (2007) *The Accentual History of the Japanese and Ryukyuan Languages: A Reconstruction*, Folkestone, UK: Global Oriental.
- Suzuki, Tomotarō, Kawaguchi, Hisao, Endō, Yoshimoto and Nishishita Kyōichi (eds) (1957) *Tosa Nikki – Kagerō Nikki – Izumi Shikibu Nikki – Sarashina Nikki*, Nihon Koten Bungaku Taikei vol. 20, Tokyo: Iwanami Shoten.
- Takeuchi, Lone (1987) *A Study of Classical Japanese Tense and Aspect*, Copenhagen: Akademisk Forlag.
- Vovin, Alexander (2003) *A Reference Grammar of Classical Japanese Prose*, London and New York: RoutledgeCurzon.
- Watanabe, Tsunaya and Nishio, Kōichi (eds) (1960) *Uji Shūi Monogatari*, Nihon Koten Bungaku Taikei vol. 27, Tokyo: Iwanami Shoten.
- Wenck, Günther (1959) *Japanische Phonetik 4: Erscheinungen und Probleme des Japanischen Lautwandels*, Wiesbaden: Otto Harrassowitz.
- Yamada, Yoshio (ed.) (1959) *Konjaku Monogatari-Shū*, part 1, Nihon Koten Bungaku Taikei vol. 22, Tokyo: Iwanami Shoten.
- (ed.) (1959) *Konjaku Monogatari-Shū*, part 2, Nihon Koten Bungaku Taikei vol. 23, Tokyo: Iwanami Shoten.
- Yamagishi, Tokuhei (ed.) (1958) *Genji Monogatari*, part 1, Nihon Koten Bungaku Taikei vol. 14, Tokyo: Iwanami Shoten.
- (ed.) (1958) *Genji Monogatari*, part 2, Nihon Koten Bungaku Taikei vol. 15, Tokyo: Iwanami Shoten.
- Yanase, Kazuo (ed.) (1967) *Hōjōki*, Tokyo: Kadokawa Shoten.

# LATE MIDDLE JAPANESE

*Mark Irwin and Heiko Narrog*

## 9.1 INTRODUCTION

Of all the periods of Japanese-language history it is perhaps not unreasonable to say that Late Middle Japanese (LMJ) is the poor cousin. Sandwiched between the ‘classical’ Early Middle period and the Modern period, it is unsurprising that research has tended to focus on its chronologically contiguous brethren.

The slice of time taken to constitute LMJ has always been somewhat fluid. In the Japanese tradition, the period is generally referred to as *chūseigo*, ‘the medieval language’, but perhaps more common are the ‘language of Kamakura’ (1192–1333) and the ‘language of Muromachi’ (1392–1573), the two major historical periods during which LMJ was spoken. There is, of course, no necessity for linguistic periods to mirror historical ones and LMJ is often extended outside these dates, especially forward in time. Japan’s first regular contact with the European world began in 1543, initially with the Portuguese, first shipwrecked sailors on Tanegashima, then the Jesuit missionary Francis de Xavier in 1549, then further Iberian Catholic missionaries thereafter. This contact continued until 1641, by which time the Iberians had been expelled, the Dutch confined to an island in Nagasaki Bay, Christian proselytizing proscribed, and Japan had embarked on its long experiment with isolationism. Since this first period of foreign contact is extremely important both for the materials it left behind, the so-called Christian materials, and the wave of mostly Portuguese loanwords introduced into the language, we feel justified in including it in LMJ and so, in keeping with the tradition of fluid periodization, take it to run from the late twelfth to the early seventeenth century.

The primary sources from which our knowledge of LMJ is drawn can be divided into native and foreign (the abbreviated title used throughout the remainder of the chapter follows in parentheses where applicable). As well as dictionaries and encyclopaedias, the former comprise Buddhist materials, including *shōmono* treatises such as Shinran’s *Shittan Kuden* (1181: *Shittan*) on Sanskrit orthography; theatrical materials (especially *kyōgen*); and works of literature, such as the *Heike Monogatari* (early thirteenth century) and the *Hōjōki* (1216). Foreign sources include dictionaries, grammars and textbooks compiled by Chinese and Korean scholars (e.g. the Korean textbooks *Ilopha* (1492: *Ilopha*) and *Shōkai Shingo* (manuscript 1618, published 1676: *Shōkai*)). Even more important is the large trove of Christian materials, which comprise not only dictionaries such as *Vocabulario da Lingoa de Japam* (1603: *Vocabulario*) and grammars such as João Rodriguez’s *Arte da Lingoa de Iapam* (1604–1608: *Arte*), but also translations of Christian literature and other moral writings, as well as transcriptions in Roman orthography of Japanese literature, such as Fabian’s *Feiqe Monogatari* (1592: *Feiqe*).

One outstanding characteristic of LMJ, in contrast to EMJ, is the great diversity of literary genres, styles, sociolects and dialects which have been transmitted in the written documents. As LMJ covers almost 500 years of linguistic history, an author’s choice of sub-period or written materials can lead to highly variant descriptions of the language.

Our focus is on the sixteenth century where materials are most abundant and information on phonology and grammar available in its most explicit form. The language described in these materials is essentially that of the Kinki area, whose most important city was Kyoto. Even the Jesuits, Dominicans and Franciscans, based largely in Nagasaki, recognized this as the ‘standard’ and tended overwhelmingly to a description of this variety. It is important, however, to bear in mind that the standard modern language, being based on a sociolect of Tokyo, one of the ‘eastern dialects’, is thus not the direct descendant of LMJ. Contrary to Frellesvig’s (2010: 383) claim that Modern Japanese (ModJ) ‘represents a fairly unbroken tradition of the Kyoto-based common language’, the consensus in Japanese linguistics is that, during the shift in the political and cultural centre of Japan that occurred in the eighteenth century, the documented Japanese language underwent substantial changes. As these changes cannot be explained diachronically but instead reflect regional differences (cf. *inter alia* Shibatani 1990: 185–7; Sakanashi and Tsukimoto 2001: 149), it is problematic to see ModJ as a direct descendant of (Western) Middle Japanese.

For subjects covered in this chapter that are uncontroversial and receive widespread acceptance in the secondary literature we refrain from supplying bibliographical references and instead refer the reader in the first instance to general Japanese works on LMJ, such as Yuzawa (1958), Iwai (1970–1974), Yanagida (1985) or Y. Yamauchi (1989), then to general works covering the history of the Japanese language. These include, in Japanese, Nakata (1972), Tsukishima (1977), Okimori (1989), Yamaguchi *et al.* (1997), Sugimoto (1998), Hanzawa *et al.* (2002); in Western languages, Wenck (1957), Miller (1967), Martin (1987), Lewin (1989), Shibatani (1990), Rothaug (1991), Frellesvig (2010).

## 9.2 PHONOLOGY

Throughout the description of LMJ phonology to follow, it may be assumed by the reader that the phonetic values of the various phonemes discussed is, unless specifically stated, the same, or approximately the same, as in ModJ. Furthermore, it is our belief that a chapter such as this, whose primary function is to offer the reader a solid overview of the language of the period, is not the place to put forward radical phonemicizations. The phonemic inventory proposed here thus represents, to a large extent, an orthodox viewpoint and also, as few phonemic inventories can remain constant throughout a period of four centuries, a panchronic ‘best fit’.

Before examining the consonantal (9.2.1) and vocalic (9.2.2) phonemes of LMJ, a brief look at the general phonological processes of liaison and sequential voicing (known as *renjō* and *rendaku* respectively in the Japanese tradition) is in order. The former, as the English translation suggests, was a phenomenon internal to the phonological word, whereby a closed syllable with an *N* coda transferred the nasal feature of its final consonant onto the beginning of a following vowel-initial syllable: e.g. *kan-on* > *kannon* ‘Kannon (Buddhist goddess of mercy)’. The phenomenon occasionally occurred with *t* codas also: *set-in* > *settin* ‘toilet’. While the first attestations of liaison date from the tenth century and the phenomenon is now unproductive in ModJ, it would appear to have hit the peak of its productivity in LMJ when it became no longer confined to the Sino-Japanese (SJ) vocabulary stratum, but spread into the native vocabulary stratum also. Unlike liaison, sequential voicing, an allomorphic phenomenon whereby the non-initial element in a compound may undergo voicing under certain conditions, is still highly productive in ModJ. Some of the numerous examples from the *Vocabulario* include *yamagawa* ‘mountain stream’ and *fanadi* ‘nosebleed’.

### 9.2.1 Consonants

Table 9.1 shows the LMJ consonantal phonemic inventory. Although the palatal glide *y* represents [j], it represents a fricative or affricate when preceded by *t d s z* (e.g. *sya* = [ja], *tya* = [tja]). No comment is required for either *k* or *r* and these phonemes will not be treated further.

In the well-documented series of sound changes known collectively as labial lenition (*hagyō tenko-on*), the behaviour of (pre-)Old Japanese (OJ) *p* differed depending on its position in the phonological word: by LMJ it had lenited to *f* word-initially, while word-internally it had lenited to *w* before *a â o* and elided completely elsewhere. Labial lenition did not effect all lexical strata identically, however. Although in LMJ the relative overall frequency of *p* is extremely low, it does occur in mimetic vocabulary (e.g. *ponpon*), as the quasi-mimetic geminate *pp* in native lexemes (e.g. *mappira*), as a geminate and after *n* in SJ lexemes, and in a few lexemes borrowed from Portuguese at the very end of the period (cf. 9.6).

Throughout the vast majority of the LMJ period there is general agreement that this lenited *f* was articulated [ɸ]: the description in *Shittan* is unambiguous; the famous imperial riddle (answer: ‘lips’) in the 1516 *Gonara-in Gyosen Nazo* (‘they meet twice for mother (*fafa*), but for father (*titi*) not even once’) indicates a bilabial phone; and in the *Dictionarium sive Thesauri Linguae Iaponicæ Compendium* of 1632 *f* is described as being halfway between [f] and [h]. Although the Christian literature transcribes *f* consistently as <f>, which in the Portuguese of the time was [f], this is not surprising as Portuguese had no [ɸ]. While the further lenition of LMJ *f* to ModJ *h* is widely held to have begun in the early ModJ period, there is a slim possibility that by very late LMJ *f* may have already been [h] in some dialects. Evidence for this includes confusion in the sinographic transcriptions of the Chinese *Nihon Fudoki* (1592) and the use of both <h> and <f> (the former seemingly reserved for interjections) in *Arte*.

LMJ *w* had a wider distribution than it does in ModJ, since, as just described, it was admissible before *o* when syllable-initial (cf. 9.2.2). In addition, there existed a sizeable number of SJ morphemes beginning in *kw* and *gw*, although before *a â* only. Evidence for these pronunciations appears in the 1474 Japanese dictionary *Setsubyōshū* as well as in the later Christian literature, with spellings in the *Vocabulario* including <guaijin> *gwaiziv* ‘foreigner, outsider’ and <quōsai> *kwâāsai* ‘lustre’. According to Toyama (1972: 205–6), however, it would appear that these pronunciations may well have been restricted to educated usage only, with remarks in the fourteenth-century *Santai Shishō* indicating that at that period the ‘uneducated lower classes’ had no labialization.

The question of what phonetic value to assign to what we here phonemicize as *s z* is perhaps one of the more problematic issues in LMJ phonology. Since the consensus becomes stronger the closer one gets to the ModJ period, we shall begin by considering the orthography employed in the late LMJ Christian materials. Here the graphs <x>

TABLE 9.1 LATE MIDDLE JAPANESE CONSONANT PHONEMES

	Labial	Dental/Alveolar	Palatal	Velar	Mora consonant
Plosive:	<i>p b</i>	<i>t d</i>		<i>k g</i>	
Nasal:	<i>m</i>	<i>n</i>			<i>ŋ</i>
Fricative:	<i>f</i>	<i>s z</i>			
Liquid:		<i>r</i>			
Approximant:	<i>w</i>		<i>y</i>		

(voiceless) and <j i> (both voiced, word-internally and word-initially respectively) are employed before the two front vowels *i e*, while <s z> are used elsewhere. Since the contemporary Portuguese pronunciation of <x> was approximately [ʃ] and that of its voiced counterpart <j> approximately [ʒ], and we also find LMJ *sy zy* rendered as <x j> in the same materials, there is broad agreement that at the end of the period *s z* were in all probability [ʃ ʒ] (or [e z]) before the two front vowels *i e*, and [s z] elsewhere. Nevertheless, comments in *Arte* indicate that *se* was [se] in eastern dialects, as it is in standard ModJ. There is also broad agreement that this allophonic variation was the case right through the LMJ period, with han'gūl transcriptions in the *Ilopha* providing further evidence. Where the real debate lies with respect to LMJ *s z* is their allophonic variation before the three non-front vowels *a o u* prior to the LMJ Christian evidence. Theories abound as to the articulation of *s z* before these vowels in the EMJ period, including [s ʃ ts tʃ] for the former and [z ʒ dz dʒ] for the latter. Which, if any, of these allophones persisted into LMJ, for how long, and in which dialects is a subject that allotted space will not permit us to tackle and the interested reader is referred to Tsukishima (1969: 385–6), Kamei (1970), Toyama (1972: 188–90), Kobayashi (1981) or Martin (1987: 33–4) *inter alia* for further detail.

The dental stops *t d* had, unlike in ModJ, no affricate allophones before high vowels until late LMJ. Since, however, the Christian materials consistently transcribe the syllables *ti tu* as <chi tçu>, *t* is taken to have had by the end of LMJ a [tʃ] or [çe]<sup>1</sup> affricate allophone before *i* and the allophone [ts] before *u*. Similarly, the same materials transcribe their voiced counterparts *di du* as <gi zzu/dzu>, and *d* is believed to have had the very late LMJ allophones [dʒ] or [jz] before *i* and [dz] before *u*. Although there is overwhelming evidence from both Chinese and native Zen Buddhist sources that the affricate allophones of *t d* were not in existence in the first half of the LMJ period, strong support for this lack of allophony extending into the period's second half is provided by the *Ilopha* with its phonemic orthography: here *ti tu* were transcribed ㄸ <ti> [ti] and ㄸ <tu> [tu] respectively. From the sixteenth century, however, we start to see evidence for affricatization in the transcriptive sinographs employed by Chinese sources such as the 1523 *Nihon Kigo* and c. 1566 *Nihon Ikkan*, and it is for this reason that the turn of the sixteenth century is the most probable date for the onset of the allophonies in question.

In what is known in the Japanese tradition as the *yotsugana* ('four letters') problem (see also Matsumori and Onishi, this volume, 11.3.1.2), confusion in the orthographies for the syllables *zi di zu du* at the very end of the LMJ period points to the onset of the merger of voiced dental obstruent phonemes before high vowels (*zi* with *di* and *zu* with *du*) that is now complete in standard ModJ. While occasional examples of this confusion may be found in Japanese materials, especially dictionaries, of the late LMJ period, perhaps the clearest inconsistencies are in the Christian literature. These include <nezzumi> (*nedumi*) for expected <nezumi> (*nezumi*) 'mouse' in *Isopo Monogatari* (1593), comments in *Arte* on how speakers in Kyoto confuse <gi> *di* with <ji> *zi* and <dzu> *du* with <zu> *zu*, and double entries in the *Vocabulario* such as the Buddhist term <giocuacu> *dyokuaku* and <iocuacu> *zyokuaku* 'sinfulness' (濁悪). While it is likely the merger began in the spoken language some generations prior to the above cited written evidence, perhaps the early sixteenth century, there is also a general consensus that it started earlier in eastern dialects, an often-cited piece of evidence being confusion in the late thirteenth-century works of the Buddhist monk Nichiren.

Unlike in ModJ, *t* was permissible syllable-finally in LMJ, although this was restricted to the SJ vocabulary layer. In the thirteenth century, we find that Shinran in his 1224

*Kyōgyō Shinshō* divides the Middle Chinese entering tone (i.e. stop-final) syllables into two groups, ‘slow’ and ‘quick’, the former consisting of \*-p k final Middle Chinese morphemes, the latter of \*-t final morphemes only. These quick entering tones were written with the same *katakana* syllable he employed to indicate Middle Chinese \*-n final morphemes and this ‘has been taken as evidence that the stop-final pronunciation goes back at least to the early days of the Kamakura period’ (Martin 1987: 73). Further evidence for SJ *t*-final morphemes appears in the *Vocabulario*, as illustrated by (1a–b); native Japanese morphemes (1c–d) that in ModJ contain **tsu** are, however, with a few minor exceptions, spelt <τϰu> (*tu*):

(1)	<i>Vocabulario</i>	LMJ	ModJ	
a.	<suisat>	<i>suisat</i>	<b>suisatsu</b>	‘conjecture’
b.	<xetna>	<i>setna</i>	<b>setsuna</b>	‘transience’
c.	<matϰu>	<i>matu</i>	<b>matsu</b>	‘pine’
d.	<itϰumo>	<i>itumo</i>	<b>itsumo</b>	‘always’

Yet more evidence appears in *Arte*, where the author states that ‘all Japanese words end in either a vowel, N or T’, as well as in *Shōkai* (e.g. 인빔 <it.pit> for ModJ **ip.pitsu** ‘one brush’) and in traditional drama. What the phonetic realization of syllable-final *t* may have been is hinted at in the 1727 *Ongyoku Gyokuenshū*, where we find that when a SJ morpheme-final \*-t (called *tsumeji*) precedes a voiced obstruent or a nasal, then this *tsumeji* is ‘swallowed’. Given that syllable-final *t* has assimilated to a following voiceless obstruent to give geminates in ModJ (viz. the *Shōkai* example above), one can speculate that the pronunciation tended towards [ʔ], especially in the late LMJ period.

From comments in *Arte* it is clear that the voiced obstruents *d g* were prenasalized when word-internal [ʰd ʰg] and that the latter was probably even realized as [ŋ]. That these prenasalized allophones existed throughout the LMJ period is corroborated by orthographies used in contemporaneous Korean and Chinese materials. There is less written evidence for prenasalization of the two remaining voiced obstruents *b z* (evidence which includes *Arte* noting that only *a* was nasalized before *b*, as well as prenasalization of word-internal *z* in *Ilopha*) and the existence of prenasalization here is regarded as problematic. An oversimplification might be that *b z* were probably prenasalized in early LMJ but that this prenasalization had weakened or disappeared by the period’s close.

In EMJ, *M* codas had been permissible, but by the end of the period these had become confused orthographically with *N*, and by early LMJ the two had merged phonemically in coda position as *N*, although differentiation of the two persisted longer in writings in western dialects, e.g. *Hōjōki*.

### 9.2.2 Vowels

The LMJ vocalic phonemic inventory consisted of the six vowels *i e a u o â*, although peripheral dialects may have had slightly differing inventories. The last of these vowels, *â*, could only exist long, and we phonemicize such long vowels as *ââ*, *oo*, etc. Little comment is required regarding LMJ *i a u* and discussion of LMJ vowels will therefore focus on the three remaining vowels, *e o â*, as well as the important series of monophthongizations that took place during the period.

Textual evidence from the late EMJ period indicates confusion in the orthography of the graphs used for the syllables <e> and <ye>, with descriptions in *Shittan* and *Arte* neatly spanning the entire LMJ period: both appear to indicate that syllables written <e> were articulated [je] throughout. Further evidence from other Christian materials, such as the romanizations in the *Vocabulario* (<yebi> *yebi* ‘prawn’) and *Feiqe* (<coye> *koye* ‘voice’), bolster further the orthodox view that *e* could not appear syllable-initially in LMJ. Sinographs and han’gūl used in the contemporaneous Chinese and Korean materials describing the Japanese language also seem to point to a similar conclusion, although here the evidence is less clear cut and the consensus more fragile (see Toyama (1972: 178–9) for a summary of the debate). More contentious, however, is the phonetic value of *e* when following a consonant, i.e. when not syllable-initial. Debate as to the true vowel articulation of what the han’gūl medial ㅚ (Modern Korean *yey* = [je]) in a syllable such as ㅚ ㅚ <kyey> was meant to represent in *Ilopha* has led to claims (e.g. Lange 1971) that *e* was [je] everywhere in early LMJ and that the romanizations showing <ye> only syllable-initially in the Christian materials indicate simply that by very late LMJ [je] for *e* survived only syllable-initially and after *s z* (explaining [je ze] for *se ze* discussed above). This theory also offers a more concise explanation for the monophthongization of *eu* to *yoo* (see below).

In parallel fashion to <e>, the LMJ syllable written <o> in kana is held to have been pronounced [wo], and evidence for this syllable-initial *wo* again spans the entire LMJ period, with descriptions in *Shittan* and *Arte* as well as other Christian materials, e.g. <voca> *woka* ‘hill’ and <xiwō> *siwo* ‘tide’ in *Feiqe*. When *wo* follows tautomorphemic *o* the Christian materials do, however, exhibit consistencies (e.g. <vouoi> *wowoi* ‘many’, but <vōcami> *wookami*, not \**wowokami*, ‘wolf’ in the *Vocabulario*), perhaps a precursor of the delabialization of *wo* generally believed not to have begun until the following century. Unlike *e*, there have been no suggestions that LMJ *o* was pronounced [wo] when it was not in syllable-initial position.

During the LMJ period, a series of monophthongizations affected *Vu* vowel sequences: see (2) below where late EMJ examples are from the *Iroha Jiruishō* (1144–1181), and late LMJ examples from the *Vocabulario*. Exactly when these changes, which together with others are known in the Japanese tradition as *chōon*, began is a subject on which there is little agreement. Perhaps, according to Martin (1987: 44), in Kyoto around the turn of the fourteenth century when vowel sequences were tautomorphemic, and therefore largely confined to the SJ vocabulary stratum, but at a later date when heteromorphemic (for useful summaries of other views see also Toyama (1972: 209–22)). Even as late as the *Vocabulario*, however, we find instances of <xīū> (*siuu*) where <xū> (*syuu*) should suffice, suggesting the palatalization in (2a) was not yet complete. It is clear that both the date of initiation and the speed with which these changes progressed varied from monophthongization to monophthongization and from dialect to dialect.<sup>2</sup>

(2)	late EMJ		late LMJ		ModJ	
a.	<i>iu</i> キウコン	<i>kiukon</i>	<qiūcon>	<i>kyuukon</i>	窮困	<b>kyūkon</b> ‘abject poverty’
b.	<i>uu</i> フツウ	<i>futuu</i>	<futçū>	<i>futuu</i>	普通	<b>futsū</b> ‘usual’
c.	<i>eu</i> ケフ	<i>keu</i>	<qiô>	<i>kyoo</i>	今日	<b>kyō</b> ‘today’
d.	<i>ou</i> トウカイ	<i>toukai</i>	<tōkai>	<i>tookai</i>	東海	<b>tōkai</b> ‘eastern sea’
e.	<i>au</i> カウサ	<i>kausa</i>	<qōza>	<i>kāza</i>	高座	<b>kōza</b> ‘dais’

It can be seen from (2c–e), that ModJ *ō* has at least three sources: EMJ *eu ou au*. Monophthongization of *au* produced a long vowel which we phonemicize here as *āā*,<sup>3</sup> no longer extant in standard ModJ, and which is generally considered to have been

articulated as [ɔ:]. While there is no indication of this intermediate stage between EMJ *au* and ModJ *oo* in the Japanese orthography of the period, the Christian literature is very careful in keeping the two apart, employing <ô> for LMJ *oo* (2d) and <ö> for LMJ *ââ* (2e). Although opinion as to when *ââ* and *oo* fell together is uncertain, what is clear is that their coalescence had the adverse effect of producing a large amount of homophony in the SJ vocabulary stratum which persists today. It is possible that the general raising of *ââ* to *oo* occurred not much after publication of the Christian works, but this could have happened even earlier in some eastern dialects – indeed some Christian materials, especially *Feige*, exhibit occasional orthographical confusion. It is also possible the raising occurred very much later than this, since the 1743 *Ongyoku Gyokuenshū* still makes a distinction in the orthography and in a very few ModJ dialects, such as areas of Niigata and Nagano, *ââ* and *oo* are still separate today. Whatever the case may be, such speculation takes us firmly beyond the LMJ period.<sup>4</sup> While we do find minimal pairs such as <côqi> *kooki* ‘postscript’ and <cöqi> *kââki* ‘noble’ in the *Vocabulario*, it should nevertheless be borne in mind that *â* is not a phoneme that can be reconstructed for the entirety of LMJ, with *ââ* having probably been *au* nearly everywhere for approximately the first century of the period and possibly having raised to *oo* in some eastern areas at the very end of the period.

As the monophthongization of *iu* to *yuu* (2a) requires no comment, we turn finally to that of *eu* to *yoo* (2c). Although here the orthographies in the Christian literature are sometimes confused and a detailed analysis difficult, the monophthongization probably proceeded *eu* > *eo* > *eoo* > *ioo* > *yoo* if we follow the orthodox line of LMJ *e* being [e] when not syllable-initial; if we adopt Lange’s (1971) position (see above) that all LMJ *e* was [je] then we can envisage a less complex and somewhat more natural change of *yeu* > *yeo* > *yeoo* > *yoo*.

### 9.2.3 Suprasegmentals

Perhaps the most heavily researched primary source on historical accent is the late EMJ *Myōgishō*, affording us a picture of eleventh-century Kyoto dialects. While it is standard theory that this late EMJ Kyoto accent pattern is the direct descendant of that reconstructed for proto-Japanese and that the Modern Japanese patchwork of accent patterns can be derived from various developments since, there is a division of academic opinion as to whether the modern Tokyo-type pattern is an innovation from that observable in the *Myōgishō*, a position originally advocated by Kindaichi (1975), or whether in fact the *Myōgishō* is closest to modern Tokyo-type accent patterns and modern Kyoto-type is the innovator, a position originally advocated by Ramsey (1979). Whichever of these two competing theories may be the more accurate, the most important LMJ accent development was a shift in the accent locus (one syllable right for Tokyo-type patterns according to the former Tokyo-innovative school; one syllable left for Kyoto-type patterns according to the latter Kyoto-innovative school). This shift brought about a merger of noun accent classes 2.2 and 2.3 in probably the second half of the thirteenth century. Another issue that has attracted considerable research is whether the reconstructed noun accent classes 2.4 and 2.5, which correspond to the modern Kyoto-type patterns *kazu-ga* (LLH) ‘number + NOM’ and *kumo-ga* (LHL) ‘spider + NOM’ (both are HLL in modern Tokyo), were differentiable when in isolation in late EMJ. If they were not, then the differentiation, which would have been confined to cases where certain enclitic particles were attached, may have spread to nouns in isolation during LMJ. From descriptions in *Arte* it is clear the present Kyoto-type accent pattern was complete by the end of the

LMJ period. For a thorough review in English of the historical accent problem see Martin (1987: 162–363).

### 9.3 ORTHOGRAPHY AND TRANSCRIPTION

The early LMJ period saw the continuation of the *wabun* style which had developed in late EMJ, although the increasing number of SJ lexemes appearing in the language of LMJ (cf. 9.6) meant the later establishment of a new style known as *wakan konkōbun*, a combination of *wabun* and *kanamajiribun*, whose epitome is perhaps the *Heike Monogatari*. Although very early LMJ histories and chronicles were still written in *kanbun*, as official documents came to be compiled by warrior classes whose educational level was generally lower than that of their aristocratic predecessors, *junkanbun* (pure *kanbun*) gradually slipped into *hentaikanbun* (modified *kanbun*), which had been used for private documents in EMJ. This style was eventually adopted by the shogunate and continued to be used into the early ModJ period for official documents, thus meaning an end to the previous distinction between private and public documents. Nevertheless, *kanbun* did live on in LMJ in the official writings of most Buddhist sects and monks. For a thoroughgoing overview of LMJ writing styles see Sato Habein (1984: 43–66).

Perhaps the most important LMJ works dealing with orthography were *Gekanshū*, compiled by the poet Fujiwara Teika in approximately the first decade of the thirteenth century, and *Kanamojizukai*, compiled some time after 1363 by Minamoto Tomoyuki, later to become the Buddhist monk Gyōa. Both deal with problems of *kana* usage, specifically those related to the consequences of labial lenition. While the earlier *Gekanshū* treats only the eight *kana* い ひ る え へ ゑ お を <i fi wi e fe we o wo> the later *Kanamojizukai* plugs the gaps and treats は わ う ふ ほ <fa wa u fu fo> and others as well. Both essentially recommend following ‘historical precedent’ in their spelling guidelines, with the important exception of お を, which were distinguished on an accentual basis, the former indicating a low and the latter a high pitch. More detailed analyses of these works and LMJ *kana* spelling problems in general may be found in I. Yamauchi (1972: 562–610) and Seeley (2000: 104–17) *inter alia*.

Other points of note in LMJ orthographical history are the large number of *hiragana* variants (*hentaigana*) and the increase in use of *katakana*; a growing predilection for *ateji*, the phenomenon of assigning native Japanese lexemes phonetic sinographs, e.g. 馬糞 for *baka* ‘stupid’ in the 1474 dictionary *Setsuyōshū*; the gradual (although not yet complete by the end of LMJ) standardization of *dakuten*, the diacritics used to indicate the voiced variant in *kana* orthography, e.g. だ *da* vs た *ta*; the first use in print of *handakuten* (a circular diacritic used to distinguish *p* from *f*) in the 1598 Jesuit-published sinograph dictionary *Rakuyōshū*, e.g. ぱ *pa* vs は *fa*; and the first use of a Roman orthography to transcribe Japanese, this in the Christian literature in general (see 9.2.1 for greater detail).

### 9.4 MORPHOLOGY

#### 9.4.1 Nouns

As in EMJ, LMJ pronouns are morphosyntactically indistinguishable from other nouns. Plurality can be marked on animate nouns with the following suffixes, in order of declining degree of honorification: *-tati* > *-syu* > *-domo* > *-ra*. In a number of lexicalized items, plurality is indicated by reduplication, e.g. *fito-bito* ‘people (person-person)’.

The following is the set of basic particles suffixable to LMJ nouns and marking their syntactic and pragmatic functions:

- (i) case particles: *-ga* nominative, *-no(/-ga)* genitive, *-ni* dative, *-wo* accusative, *-ye* allative, *-yori/-kara* ablative, *-de/-ni-te* essive
- (ii) pragmatic particles: *-wa* topic, *-mo* focus, *-saye* extreme degree
- (iii) noun coordinating particles: *-to* ‘and’, *-ya* ‘and (amongst others)’, *-ka* ‘or’, *-yara* ‘and/or’.

This set is remarkably similar to that of ModJ, hinting that the foundation of ModJ case marking lies in LMJ. *-ga* as a genitive marker is in decline (see below); allative *-ye* is expanding at the expense of *-ni*; *-kara* is in the process of taking over from *-yori* as the ablative; *-de* is gradually replacing *-ni-te*, of which it is a contraction; and *-saye* is assuming the functions of EMJ *-sura* and *-dani*. Although nominative, accusative and topic nouns can also be unmarked (zero particle), overt case marking in LMJ is expanding in comparison to EMJ.

The genitive *-no* can be appended to *-ye*, *-kara*, *-yori*, *-de/-ni-te* and *-to*, as can topical *-wa*, which may be added to *-ni* also. A very frequent particle combination in LMJ is *-wo-ba*, topicalizing an object. As in ModJ, however, LMJ *-wo-ba* could be replaced by simple *-wa*.

#### 9.4.2 Verbs

EMJ verb classes are continued into LMJ. Only the irregular athematic verbs based on *ari* ‘be’ (Vr, *ra-hen*) are absorbed into the regular athematic paradigm (Vc, *yodan*), due to the fact that in the Kamakura period the adnominal inflection takes over the function of the finite inflection, which was irregular in the case of the Vr verbs.<sup>5</sup> Thus, we find three thematic classes (Vv monograde (*kami-*, *shimo-ichidan*), Vi bigrade (*kami-nidan*), Ve bigrade (*shimo-nidan*)), two athematic classes (Vc quadrigrade (*yodan*), Vn quadrigrade (*na-hen*)), and two irregular classes (Vs (*sa-hen*), Vk (*ka-hen*), consisting of *suru* ‘do’ and *kuru* ‘come’ respectively). By the Muromachi period, the following set of inflectional endings has developed (brackets indicate that the presence of a phoneme depends on verb class, the symbol ~ separates allomorphs, and small capitals indicate allophonic variation):

- (i) inflections added to the verb stem: *-(ur)u* non-past [CJ attributive conclusive], *-(ur)e* conditional I [CJ realis], used as the base of the *-ba* and *-do* forms and in *kakari-musubi*
- (ii) inflections added to the root of athematic verbs and the stem of thematic verbs: *-(y)oo* ~ *-(y)ãã* future I, *-(a)mai* negative future, *-(a)ba* conditional II, *-(a)ide* adverbial negation I, *-(a)zu* adverbial negation II, *-e* ~ *-i* ~ *-yo* imperative I, *-(s)ai* imperative II
- (iii) inflections added to the verb base: *-Te* gerund, *-Ta* past, (*na*) . . . *-So* prohibitive.

In the Muromachi period, Japanese verb forms reach a degree of assimilation and fusion between verb stem (or base) and endings unmatched in Japanese language history in the standard dialects. The most important of these involve the non-past tense ending and the inflections added to the verb base. Examples (3) and (4) below are representative:

TABLE 9.2 LATE MIDDLE JAPANESE VERB INFLECTION

ending	yom- 'read' (Vc)	ag- 'give' (Ve)	s- 'do' (Vs)
-(ur)u	yom-u	ag-uru	s-uru
-(ur)e	yom-e	ag-ure	s-ure
-(y)oo ~ -(y)ââ	yom-ââ	age-oo ~ ag-yoo	s-yoo
-(a)mai	yom-amai	age-mai	se-mai
-(a)ba	yom-aba	age-ba	se-ba
-(a)ide	yom-aide	age-ide	se-ide
-(a)zu	yom-azu	age-zu	se-zu
-e ~ -i ~ -yo	yom-e	age-yo ~ age-i	se-yo ~ se-i
-(s)ai	yom-ai	age-sai	s-ai
-Te	yoo-de	age-te	si-te
-Ta	yoo-da	age-ta	si-ta
(na) ... -So	na yoo-zo	na age-so	na se-so ~ na si-so

## (3) Verb stem + -(ur)u present:

negaw-	'pray' (Vc)	+ -u	>	nega-u	>	negââ
omow-	'think' (Vc)	+ -u	>	omo-u	>	omoo
iw-	'say' (Vc)	+ -u	>	i-u	>	yuu*

## (4) Verb base + -Te gerund:

yobi-	'call' (Vc)	+ -Te	>	yON-de*	>	yoo-de
nirami-	'stare' (Vc)	+ -Te	>	niran-de*	>	nirââ-de
isogi-	'hurry' (Vc)	+ -Te	>	isoi-de*		
daki-	'embrace' (Vc)	+ -Te	>	dai-te*		
yori-	'approach' (Vc)	+ -Te	>	yot-te*		
moti-	'hold' (Vc)	+ -Te	>	mot-te*		
sasi-	'stick' (Vc)	+ -Te	>	sai-te		
usina(w)i	'lose' (Vc)	+ -Te	>	usinââ-te		

Only the asterisked forms have found their way into standard ModJ, although some of the others can still be found in modern dialects.

Table 9.2 demonstrates the inflection of the two verb classes with the highest type frequency, the Vc and the Ve, as well as the high-token-frequency Vs *s-* 'do' which has a third stem form, the vowel stem *se-*, and behaves irregularly.

In LMJ, the Ve and Vi show signs of absorption into the Vv, and the Vn of absorption into Vc. The process, however, is not completed yet in this period. Furthermore, verbs have a set of inflecting derivational suffixes which can occupy a slot between verb stem and inflectional ending:

- (i) derivational suffixes added to the base all verbs: *-ta-* volition (Ai), *-T-* perfective, *-Tar-* resultative.
- (ii) derivational suffixes added to the base of athematic verbs and the stem of thematic verbs: *-(r)ar-* passive, *-(s)as-* causative, *-(s)asim-* honorific (present: *-(s)asim-u* ~ *-(s)asimo*), *-(s)asimas-* ~ *-(s)asemas-* honorific, *-(a)n-* negation, *-(a)nan-da* past negation, *-(y)ooz-* ~ *(y)ââz-* future II.

Table 9.3 shows these suffixes added to Vc and Ve verbs, with a present tense ending, or a past ending if the non-past is unavailable. The suffix *-(y)ooz-* ~ *(y)ââz-* has two

TABLE 9.3 LATE MIDDLE JAPANESE INFLECTING DERIVATIONAL SUFFIXES

ending	yom- 'read' (Vc)	ag- 'give' (Ve)	s- 'do' (Vs)
- <i>ta-i</i>	<i>yomi-ta-i</i>	<i>age-ta-i</i>	<i>si-ta-i</i>
- <i>T-uru</i>	<i>yoo-d-uru</i>	<i>age-t-uru</i>	<i>si-t-uru</i>
- <i>Tar-u</i>	<i>yoo-dar-u</i>	<i>age-tar-u</i>	<i>si-tar-u</i>
-( <i>r</i> ) <i>ar-uru</i>	<i>yom-ar-uru</i>	<i>age-rar-uru</i>	<i>se-rar-uru</i> ~ <i>si-rar-uru</i> ~ <i>s-ar-uru</i>
-( <i>s</i> ) <i>as-uru</i>	<i>yom-as-uru</i>	<i>age-sas-uru</i>	<i>se-sas-uru</i> ~ <i>si-sas-uru</i> ~ <i>s-as-uru</i>
-( <i>s</i> ) <i>asimo</i> ~ -( <i>s</i> ) <i>asim-u</i>	<i>yom-asimo</i> ~ <i>yom-asim-u</i>	<i>age-sasimo</i> ~ <i>age-sasim-u</i>	<i>s-asimo</i> ~ <i>s-asim-u</i>
-( <i>s</i> ) <i>asimas-</i>	<i>yom-asimas-u</i>	<i>age-sasimas-u</i>	<i>s-asimas-u</i>
-( <i>a</i> ) <i>n-u</i>	<i>yom-an-u</i>	<i>age-n-u</i>	<i>se-n-u</i>
-( <i>a</i> ) <i>nan-da</i>	<i>yom-anan-da</i>	<i>age-nan-da</i>	<i>se-nan-da</i>
-( <i>y</i> ) <i>ooz-u(ru)</i> ~ -( <i>y</i> ) <i>ããz-u(ru)</i>	<i>yom-ããz-u(ru)</i>	<i>age-ooz-u(ru)</i> ~ <i>ag-yooz-u(ru)</i>	<i>s-yooz-u(ru)</i>

alternative present endings, *-uru* and *-u*. A third class of inflecting suffixes occurs in the same structural position as particles, appended to fully inflected verb forms: *nar-* (Vr) 'be', *mazi-* (Ai) negative future and *be-* (Ai) admonition/possibility, e.g. *ar-u-be.ki* 'should be'.

While EMJ still had two inflectional classes of adjectives (Aku and Asiku), the fact that the finite form disappears eliminates the differences between the two classes in LMJ (i.e. they merge as Ai). Inflectional endings are *-(k)i* present, *-(k)u* adverbial, *-(k)ute* gerund and *-kere* conditional. The suffixation of *-kar-* (Vr) allows the addition of verbal inflections, e.g. *taka-kat-ta* 'was high'. The *k* of EMJ *-ki* is often lost in LMJ. As with verbs, fusion between stem and inflection occurs, particularly with adverbial inflections:

- (5) *taka-* 'high' + *-(k)u* > *taka-u* > *takãã*  
*kurusi-* 'painful' + *-(k)u* > *kurusi-u* > *kurusyuu*

### 9.4.3 Numerals and classifiers

The numeral systems of LMJ are already essentially the same as those of ModJ: three numeral paradigms, SJ (*iti, ni, san* . . .), the native lexical numerical series (*fitotu, futatu, mittu* . . .) and this latter series' prefixal counterparts (*fito-, futa-, mi-* . . .). Indefinite expressions with two juxtaposed numbers, e.g. *ni-san* 'two or three', and with *su(u)*- 'some' are also found, although the latter is apparently not yet prefixed to numerals (e.g. *sui-nen* 'some years', but not \**sui-zuyuu* 'a few tens'). There are a few other minor differences from ModJ. Firstly, numbers of 100,000 or more were expressed differently, with *-woku*, for example, primarily having the meaning of 100,000 and *-tyoo* standing for 10,000,000. Secondly, in limited contexts, *zi* was an alternative for *ni* (2), *sãã* for *san* (3), and *riku* for *roku* (6). Thirdly, native numerals, especially prefixal ones, were used more frequently and in more contexts than in ModJ, where the SJ numeral system has replaced them. Thus, one still said *ya-tabi* 'eight times' rather than the SJ *fati-do*, or *mi-kasira* 'three head' instead of *san-too*. A detailed description of numeral system and classifier usage is provided in *Arte*.

As in ModJ, ordinals are formed either with the suffixes *-me* or *-ban-me*, or the SJ prefix *dai-*. Fractions are expressed with the suffix *-bu*, decimals (10 percent) as *-wari*

and percents (1 percent) again as *-bu*. Thus, ‘a third’, for example, is *san-bu iti* (without genitive *-no*, in contrast with ModJ), 0.5 (50 percent) is expressed as *go-wari*, and 11 percent would be *iti-wari iti-bu*.

The classifier system too is essentially the same as in ModJ, although individual classifiers may differ. Fish are counted as *-kake* instead of **-hiki**, and swords as *-furi* rather than **-hon**. Many LMJ classifiers count things that have vanished from modern life or have lost their cultural importance. The general tendency, however, is that LMJ still had many indigenous Japanese classifiers in places where ModJ uses SJ ones. Assimilations between numeral and classifier are also already in place by LMJ, e.g. *iti+fai* ‘cup’ > *ip-pai*. While native numbers ending in *-tu* (*fitotu*, etc.) are morphologically independent, the SJ paradigm and, particularly, the native prefixal paradigm both need a following noun or nominal suffix (e.g. classifier) indicating the unit to be counted.

## 9.5 SYNTAX

LMJ has SOV as its basic word order, is strictly head-final, dependent-marking and is an accusative language. With respect to topic-prominence and subject-prominence in the sense of Li and Thompson (1976), it combines features of both. In these basic respects, LMJ does not differ from ModJ.

### 9.5.1 Noun phrase structure

*Demonstratives.* *kono*, *sono* and *ano/kano* ‘this, that’ make up the set of adnominal demonstratives in LMJ. *kono* refers to things close to the speaker, *sono* to things close to the hearer and *ano/kano* to things close to neither. Together with *dono* ‘which?’ these belong to the so-called *ko/so/a/do* paradigm (see 9.5.2).

*Possession.* Possessor precedes possessed, and is marked by a genitive particle, e.g. *feike-no senzo* ‘the ancestors of the Heike’. Although LMJ still has two genitive particles, *-ga* and *-no*, the former has now become a general nominative case marker, no longer confined to adnominal or nominal clauses, and its genitive function has declined. In LMJ, it can still be found with names and in fixed expressions, such as *wa-ga mi* ‘my life, me’. It is considered to be less polite than *-no*. Furthermore, possession can be indirectly indicated by honorification, such as using the prefixes *go-/wo-*, e.g. *go-ikke* ‘your (honoured) family’.

*Adjectival modifiers.* As at any stage in Japanese linguistic history, adjectives precede the head noun. Besides the word class of adjectives, which are morphosyntactically similar to verbs, Japanese has a noun-like class of adjectives, the nominal adjectives (An). While initially their syntactical functions were indicated mainly by the copular verbs *-tari* and *-nari*, in LMJ *-tari* becomes obsolete and *-nari* remains mainly in adnominal position where it is reduced to *-na* (e.g. *râzeki-na yatu* ‘ruffian’), although it can still also be found in finite position, where it competes with *-dya*. Other syntactic functions are usually indicated with the same particles and copulas as with nouns (*-ni* adverbial, *-de* essive)

There is no major change in *relative clauses* in LMJ. They usually precede the head noun directly, and are marked with *-ga* or *-no* only occasionally, especially when forming an adjunct clause with a ‘formal noun’ such as *tame* or *tokoro* as the head noun. As in EMJ, internally-headed relative clauses can be found.

*Quantifiers* (or quantifiers + classifiers) either follow the nouns unmarked (e.g. *sisai yottu* ‘four reasons’) or precede them with genitive marking (*yottu-no sisai* (Doi 1955: 762)).

### 9.5.2 Pronouns and anaphora

*Personal pronouns or pronoun equivalents.* Even more perhaps than in other periods of the language’s history, a wealth of pronoun equivalents can be found, mainly distinguished by degree of politeness and the relative social status (including gender and age) of the interlocutors or third-person referents. *ware*, *watakusi*, *konata* and *soregasi* are polite first-person usages, *mi* and *midomo* are less polite and *ura* and *wora* are vulgar. The second person is referred to politely as *sonata*, *konata*, *won-mi* or *go-fen*, and less politely as *wo-nusi*, *wonore*, *soti* or *sore*. The third person can be referred to as *kore*, *are*, *kare* or *ano fito/kono fito*.

*Demonstratives and anaphora.* The deictics of the *ko/so/a/do* paradigm carry, along with the zero anaphora characteristic of Japanese, the main load of anaphoric reference. By LMJ these paradigms are complete in their modern form, with the exception of the *konna/sonna/an-na/donna* set. In the Kamakura period, *asoko* ‘there’ is added to the spatial demonstratives. In Muromachi, the forms *kotira/sotira/atira/dotira* ‘here/there/where’ with the suffix *-ra*, and *koitu/soitu/aitu/doitu* ‘this/that/which person (disparaging)’ (< *(ko)-yatu*, etc.) develop. In contrast to ModJ, however, the demonstrative pronouns of the *kore/sore/are/dore* ‘this/that/which’ set can still refer to people, and not only things. Besides the *ko/so/a/do* sets, anaphoric words with *ka-* (equivalent to *a-*) are also used, especially *kano* (adnominal) and *kare* (pronoun), referring to something separate from both speaker and hearer, or something in common memory.

*Interrogatives.* In the Kamakura period, the interrogative forms shift from *idu-* to *do-* (e.g. *iduko* ‘where’ > *doko*). In Muromachi, *dono* ‘which’ is added to the adnominal *kono/sono/ano* paradigm (e.g. *dono kuni* ‘which country’). Common interrogative nouns which do not belong to the *ko/so/a/do* paradigm include *tare* ‘who’, *nani* ‘what’, and *itu* ‘when’. *iku-tu*, *iku-ra*, *iku-fodo*, *iku-baku*, etc. are used for ‘how many (things)’ and *iku-tari* for ‘how many (persons)’. *nani-to yââ-na*, *nan-to si-ta*, *ika-yââ-na*, etc. ask ‘what kind of’, similar to ModJ **donna**, and expressions such as *kono yââ-na*, ‘this kind of’, serve as the answer.

### 9.5.3 The basic sentence

Arguably, the biggest syntactic change from EMJ to LMJ is the loss of the distinction between adnominal and finite inflections. This means the loss of overtly marked nominalized main clauses, as ModJ **no-da** had not developed yet, and leads to the loss of *kakari-musubi* agreement. This change is further accompanied by the development of *-ga* as a general nominative case marker and by the expansion of case marking in general.

#### 9.5.3.1 Declaratives

*Noun-predicate sentences.* If a main clause has a noun as the predicate, it is usually marked by the essive particle *-de* and followed by a verb of existence, typically *ar-u* (Vr) (as well as its various honorific equivalents). *-de* and *ar-u* contracted in various

ways resulting in copular verbs such as *-dyar-u*, *-dea* and *-dya* (Eastern Japanese *-daru*, *-da*). Nominal predicates in declarative sentences can also be marked only by a sentence-final particle, in particular *zo* or *yo*, or, rather rarely, remain completely unmarked. The older copular forms *-ni ar-u* and *-nar-u* ~ *-na* decline but still survive in specific contexts. Thus *-ni ar-u* can, for example, be found in the form *-ni ar-oo* ‘may be’ and *-nar-u* is frequently found in subordinate clauses and fixed expressions.

*Locational, existential and possessive sentences.* As in ModJ, these are expressed by the pattern ‘X-*ni/wa* Y-*ga* + verb of existence’, X being the location or possessor, and Y the thing existing or possessed. In LMJ, *ar-u* essentially still stands for animate as well as inanimate subjects of existential sentences. As Kinsui (2006, ch. 3) shows, while *i-ru* is increasingly used for the animate subjects of locative-existential predications in the Muromachi period, it is still competing with *ar-u* in this function.

*Comparison.* Comparison is marked by the ablative particle *-yori*, or, if negated, by the suffix *-fodo*. The object compared is marked by *-wa*. The optional adverb *nawo* ‘more’ indicates a higher degree.

- (6) a. X-*wa* Y-*yori* (*nawo*) + predicate    ‘X is . . . -*er* than Y’  
 b. X-*wa* Y-*fodo* + predicate            ‘X is as . . . as Y’  
 c. X-*wa* Y-*fodo* + neg. predicate        ‘X is not as . . . as Y’

The noun *yãã* and the particle adjective *goto-si* indicate similarity, ‘X is like Y’.

### 9.5.3.2 Tense and aspect

While in the Kamakura period the diverse tense and aspect markers of EMJ were still in use, although subject to meaning changes and to increasing concatenation of two or more markers as one unit, the EMJ tense and aspect system finally broke down in Muromachi, paving the way for the emergence of the modern system. *-keri* and *-ki* ~ *-si* essentially disappeared, and *-Ta* (< EMJ resultative *-tari*) emerged as the new past marker, leading to the *-Ta* past vs *-(ur)u* non-past system of ModJ. Likewise, the four EMJ aspect markers (*-eri*, *-tari*, *-t-* and *-n-*) gradually decline, while *-Te i-* ~ *-Te wor-* and *-Te ar-* slowly emerge but have not yet fully evolved (cf. Ahn and Fukushima (2005)). *T-* is the last of the old markers to hold its ground, but only as a perfect and recent past form, and not as a perfective. LMJ *-Ta* retains the resultative function of EMJ *-tari* and takes over the perfective functions of former *-t-* and *-n-*, thus bearing a heavy functional load. The old form *-tari* (now *-Tari*) is retained as a base for the suffixation of conjunctive inflections and particles (e.g. *-Tar-e-ba*). The following tense and aspect system results:

Tense: *-Ta* past, *-T-* recent past/perfect, *-(ur)u* non-past  
 Aspect: *-(ur)u* dynamic progressive; *-Ta* resultative, perfective; *-Te i-* ~ *-Te wor-*, *-Te ar-* stative progressive, passive resultative (*-Te i-* ~ *-Te wor-* with animates only)

### 9.5.3.3 Negation

Clausal negation is marked on the predicate. In Muromachi, *-(a)n-u* takes over from *-(a)zu* as the sentence-final negation on verbal predicates and *-(a)nan-da* is the new common past tense negation. *-(a)zu* is now specialized as adverbial negation, together

with *-(a)ide* (new in Muromachi). In eastern dialect material, uninflected *-(a)nai*, the predecessor of ModJ **-(a)na-i** is attested, which probably goes back to EOJ *-(a)napye* (see 7.5.3.3) and not the adjective *na-i* (EMJ *na-si*). Nominal predicates are negated by *-de-wa* plus *na-i* or a negated verb of existence. For adjectives as well, *na-i* takes over the function of negation from EMJ *ar-azu*.

#### 9.5.3.4 Modality

*Epistemic modality*: While the EMJ markers *-kem-* and *-bera-nari* fall out of use and the epistemic uses of *be-* also decline, EMJ *-(a)m-u* fuses to become the inflection *-(y)oo ~ -(y)ââ* and thrives in LMJ, particularly in the extended form *-(y)ooz- ~ -(y)ââz-*, derived from *-(a)m-u-to s-*. This extended suffix is generally viewed as synonymous with simple *-(y)oo ~ -(y)ââ* and becomes the dominant epistemic marker of LMJ, broadly covering epistemic possibility and necessity, or inference in general. *mai* (also *mazi-*), which morphologically behaves either as an inflection or as a particle, is the negative counterpart of *-(y)ooz- ~ -(y)ââz-*. *-rââ* (< EMJ *-ram-u*) survives, first as a present-oriented conjecture in contrast to future-oriented *-(y)ooz- ~ -(y)ââz-*, and later as a conjecture about a past event, suffixed to perfect/recent past *-T-* as *-T.u-rââ*. Sentence-final *yara<sup>N</sup> ~ yaraâ ~ yara* (< EMJ *-ni-ya ar-am-u*) denotes uncertainty on the part of the speaker.

*Evidentiality*: LMJ loses visual *-umeri* and hearsay *-unari*. The new suffixes *-ge* and *-sââ* denote a judgement based on (visual) appearance, e.g. *ari-sââ*.

*Deontic modality*: *be-* (LMJ also *besi-*) virtually vanished from the spoken language of western Japan as early as the Kamakura period. Thus, standard LMJ had no productive grammaticalized expression for ‘should’ or ‘must’. Instead, the future is used as an indirect means to indicate obligation, or nominal predicates such as *mono-dya* and *koto-dya* (‘it is a thing that . . .’) are used. Furthermore, the double-negated construction *-(a)ide kanaw-an-u* (‘it does not fit if . . . does not . . .’) is attested, but it seems to be less common and grammaticalized than its modern counterparts. In the domain of mood, LMJ has imperative and prohibitive inflections.

*Volition*: *-(y)ooz- ~ -(y)ââz-* and *-(y)oo ~ -(y)ââ* express intention with first-person subjects, hortative with plural first-person subjects, and admonition with second-person subjects. *-(a)mai* (~ *-(a)masi-*) is the negative counterpart. Wishes are now expressed by the suffix adjective *-ta-i*, which is derived from EMJ *ita-ki* ‘painful’.

*Dynamic modality*: The passive-spontaneous suffix *-(r)are-* can now also be used in non-negated contexts as a potential. Also, the first examples of a new potential suffix *-(y)e-* for situational possibility and ability can be found, added to the stems of athematic verbs. More commonly, these notions are expressed lexically or periphrastically. Thus, the adverb *ye* and a negated predicate is frequently used to denote impossibility (see Genenz 1979).

#### 9.5.3.5 Non-declarative sentence types

*Interrogatives*. Interrogatives are formed without change of word order, and are usually marked by particles. The interrogative particle *ya* declines, and so does the interrogative marking of sentence-internal constituents. Instead, the sentence-final particles *ka*

(for self-directed questions in general and for yes/no-questions) and *zo* (wh-questions) are used. According to Yanagida (1985: 128f), at the end of LMJ sentence-internal *ka* remains only in rhetorical wh-questions (rhetorical questions are otherwise formally indistinguishable from other questions). In alternative questions, each of the alternatives is marked with *ka* or, if the choice is between a positive and a negative proposition, the negative proposition can be replaced by *ina* ‘not’ plus *ya*, resulting in the constructions *ka ina ya* or *ya ina ya*. There is no class of tag questions.

*Answers.* For wh-questions, the constituent which was asked for alone is a complete answer. Yes/no-questions can be answered with *ââ*, *sââ*, *a*, *nakanaka*, etc. for ‘yes’, and *iya* or *iya-iya* for ‘no’. *kasikomata* is a common positive response to a request.

*Commands and requests.* Imperative sentences are formed with the imperative inflection *-(y)e ~ -yo ~ -i* on the predicate. The imperative in *-(s)ai* is more respectful. Requests can be made more polite by adding honorific markers, such as *-(s)asimas-* or *-(s)ase-rare-* to the verb stem. Indirect requests can be made by using the future suffixes *-(y)oo ~ -(y)ââ* or *-(y)ooz- ~ -(y)ââz-* (‘you would/might’) in place of the imperative. Another common strategy for rendering requests more indirectly is to add the particle *kasi*, indicating self-directedness or third-person-directedness, onto the imperative form. Furthermore, in the sixteenth century benefactive constructions with *-Te kur-* and *-Te kudasar-* arise and are used for requests. Adhortatives (‘let’s . . .’) are expressed with future suffixes on the predicate. Prohibitions are expressed either via *-(u)na* or the construction *na . . . -So*.

Exclamations are expressed mainly by sentence-initial interjections (*ara*, *aa*, *aware*, *sate*, etc.) on the one hand, and sentence-final particles (*kana*, *wa*, *ya*, etc.) on the other. These forms are usually not specialized as to the expression of a specific emotion, but their interpretation varies with context. Sentence-final *mono-wo* expresses regret (‘If only I had . . .!’).

#### 9.5.4 Topic, focus and emphasis

Topicalization, focus, etc. are usually segmentally marked. *wa* is the particle used for topicalizing constituents. *mo* has an inclusive (‘also’) and a focusing function. Other important scalar and focusing particles are *saye* ‘even’, which displaces EMJ *dani* and especially *sura* to become the dominating scalar particle; *made*, for maximal degree; and *nomi* and *bakari* for restriction (‘only’). *koso* marks contrastive emphasis (‘just’, ‘especially’) and is the only case where *kakari-musubi* persists through LMJ, although gradually declining. While the *kakari-musubi* agreement of all other EMJ focusing particles depended on a distinction between adnominal and finite inflection which was lost in LMJ, *koso* triggered the conditional form which remained in LMJ.

Clefts and pseudo-clefts in Japanese are not distinguished from each other. In LMJ, sentences are clefted by the preposing of the predicate (plus arguments), followed by information-structure particles, particularly *-wa*. The preposed nominalized predicate is not yet marked by **no** as in ModJ, however.

#### 9.5.5 Passive and causative

The **passive** demotes the main participant of an event, including the possibility of deleting it, and is marked by the suffix *-(r)are-* on the verb. The agent is demoted to

an optional peripheral participant marked by *-kara*, *-yori* or *-ni*, and the patient is either marked with the nominative or, in the case of adversative passives, remains an accusative argument. The **causative** is a device to increase the number of arguments. It is marked by the suffix *-(s)ase-* on the verb, and introduces a new participant, marked as nominative, while the agent (now the causee) is either marked with dative *-ni* or accusative *-wo*. The new potential suffix *-(y)e-* also displays an affinity for voice in that it optionally highlights the patient and demotes the agent, but examples are still infrequent. There is no significant change in the intransitive/transitive verb pairings, most of which were already lexicalized by OJ.

### 9.5.6 Speech levels and respect

The development of honorific speech in LMJ is generally characterized by two features. First, the differentiation of speech levels and, concomitantly, the diversification of lexical and grammatical forms of honorific speech. Second, there is an increasing tendency towards expressing the speaker–hearer relationship in the speech situation, and this is reflected in the progressive development of addressee honorifics. As always, there is considerable fluctuation and constant renewal in the use of honorific forms. Typical components of respectful verb forms of the period are the causative *-(s)as-* (also the older causative *-(a)sim-*), the passive *-(r)ar-*, *-(s)ase-rar-*, *-(s)asimas-* ~ *(s)asemas-* (< EMJ *-(s)ase-ofasi-mas-*), *-(s)asi-tamãã* ~ *-(s)ase-tamãã* or *-(s)asimo* (< EMJ *-(s)ase-tamaf-u*), *o-V-ar-* and *o-V-nasar-*. Typical humble verbs are *-tatematur-*, *-mããs-* and later *-ma(i)ras-* (< EMJ *mawir-as-uru*), which later became ModJ polite **-mas-**. Typical polite forms added onto the predicate are *-sããrãã* ~ *-soro* (< EMJ *saburaf-u*), *-odyar-u* (< *o-ide ar-u*), *-oryar-u* (< *o-iri ar-u*) and *-gozar-u* (< *go-za ar-u*), attached to the stem. Towards the end of the period, humble *-ma(i)ras-* is also used politely while *-sããrãã* retreats to the domain of the written language.

Honorification on nouns is expressed productively by the prefixes *go-* on SJ nouns and *wo-* on indigenous vocabulary. *gyo-* (SJ) and *mi-* (NJ) are used only in lexicalized collocations, while *on-* (NJ) is highly formal. For the prefix *wo-*, see 9.6 below. *-sama*, *-koo* and *-dono* are added to personal names and rank names in order to express respect, e.g. *arima-sama* ‘Lord Arima’. In contrast, the suffixes *-domo* and *-ra* added onto names are impolite, while *-me* is pejorative, e.g. *akuniN-me* ‘scoundrel’.

### 9.5.7 Adverbials

LMJ has a number of lexicalized adverbs. They may obligatorily or optionally be followed by the particles *-ni* or *-to* (also *-to si-te*). Among these specialized adverbs are numerous onomatopoeic and mimetic words, many of which are already identical to those in ModJ. Adverbs are derived regularly from adjectives by the adverbial inflection *-(k)u*, which leads to fusions between stem and ending, e.g. *taka-ku* ‘highly’ > *takãã*. Nominal adjectives are regularly used adverbially with the particle *-ni*, e.g. *kakubet-ni* ‘especially’. In some cases, adverbs are derived from nouns and verbs. In the case of nouns, this is done with *-to*, *-ni* or zero-derivation, e.g. *tabun* ‘the greater part’ > ‘probably’. *sugure-te* ‘especially’ (< *sugure-* ‘surpass’) and *tatoye-ba* ‘for example’ (< *tatoye-* ‘liken’) are examples of adverbs derived from verbs.

Spatial location is marked with the particles *-ni* and *-de* (< *-ni-te*), location in time with *-ni*, direction with *-ni* or *-ye*, cause, purpose and goal with *-ni*, materials and instruments with *-de*, accompaniment with *-to*, and source with *-ni*, *-kara* or *-yori*. Thus, as already

mentioned, particle use in LMJ, at least in the Muromachi period, is already extremely similar to ModJ.

### 9.5.8 Coordination

Morphosyntactically, Japanese, including LMJ, does not have true coordination since all clauses except the main clause are marked as either adverbially or adnominally subordinate and bear restrictions with respect to tense and/or modality marking. Semantically, however, there are clauses that correspond closely to ‘and’- and ‘but’-clauses in English. Firstly, there are clauses marked with the particle *ga*, which can be interpreted both as ‘and’ or ‘but’, depending on context. *ga* emerged as the most common conjunctive particle in this function in LMJ. Secondly, there are clauses with a predicate ending in *-Te* or the negative *-(a)zu ~ -(a)ide*. *-Te* can denote a neutral sequence of events (‘and’) but, depending on context, can also be interpreted causally or concessively. The particles *si* and *keredo*, used in a quasi-coordinative function in ModJ, had not yet established themselves in this period.

### 9.5.9 Subordinate clauses

Since the distinction of adnominal and finite inflection disappeared in early LMJ and the particle **no** had not yet developed its function of nominalizing clauses, LMJ is a period in which most adnominal and nominal clauses are still segmentally unmarked. As in EMJ, *koto* nominalizes only a limited number of clauses. Quotations are mostly marked by the particle *to*, less frequently by *tote*. Indirect and direct quotations are not distinguished.

With respect to adverbial clauses, LMJ developed a wealth of marking. The inflection *-(a)ba* can still mark a hypothetical conditional, but was apparently declining by the Muromachi period, yielding its function to the other conditional *-(ur)e-ba* (also still used as a temporal condition for a specific event (‘as’) and causally (‘because’)) and to the frozen sequences *-Tar-aba* and sentence-final form + *-nar-aba*. Sentence-final form + *tokoro-de* is also used as a temporal condition (‘when’) as well as causally (‘because’). Sentence-final form + *fodo-ni* and *-ni yot-te* are the typical causal connectors of the Muromachi period, the latter also being used to express purpose. Concessive clauses are typically marked with *-(ur)e-do-mo* and *to-mo*, but there are already examples of *-Te-mo*, which is later to take over concessive functions from these markers. The adverb *tatoi* (*tatoye*) at the beginning of a clause anticipates hypothetical concessive *to-mo* or a hypothetical conditional, while *mosi* anticipates a hypothetical conditional. The particle *nagara* marks a simultaneous or a concessive clause. *-(ur)e-ba* and *-(ur)e-do-mo* are frequently found added not directly to the verb stem, but to the suffix *-Tari* or *-nari*, which by the end of the period are rarely found in any function other than to carry these conjunctive inflections and certain particles. Clauses in *-(ur)e-ba* have a clear switch reference function.

## 9.6 LEXICON

A useful snapshot of the LMJ lexicon can be gleaned from the work of Miyajima (1967), who surveyed what proportion of one thousand common words in ModJ have been in use at various periods in the past. Approximately 65 percent of such high-frequency modern words appear in the *Vocabulario*, chosen as his LMJ sample text, as against 45 percent for the mid-EMJ *Genji Monogatari*. When his analysis is broken down by part

of speech, however, it is clear that by the time of the *Vocabulario* a full 93 percent of high-frequency modern verbs are listed (85 percent by earlier *Genji*), as opposed to only 52 percent of nouns (30 percent by *Genji*), suggesting that the greatest changes in the LMJ lexicon took place amongst nouns.

In terms of lexical content, two general trends can be witnessed in the literature of the LMJ period. The first is a marked increase in colloquial vocabulary, especially mimetic lexemes, apparent in the military and narrative literature. The second is the steady growth in SJ vocabulary found in all types of writing and in many social milieux, especially political, cultural and religious. Particularly apparent was the trend towards writing native vocabulary with sinographs, then giving these sinographs their Chinese readings: e.g. native *fi-no-koto* 'fire' moves from a native *kana* orthography to 火事 and then comes to be pronounced SJ *kwazi*. Statistical support is given for the increase in the SJ lexicon by Miyajima (*ibid.*) where, of the high-frequency ModJ words that appear in the *Vocabulario*, 21 percent are SJ. This is in contrast to a mere 5 percent in the *Genji Monogatari*, although it should be borne in mind that the two works have very different agendas and emanated from hugely differing social milieux.

Tabooed vocabulary (*imikotoba*) has always been a feature of Japanese linguistic history and LMJ was no exception. Noteworthy in a period whose latter half was scarred by civil war is the avoidance by the warrior classes of any term connected to 'defeat'. Such *bukekotoba* or *mushakotoba* included *firaku* 'open' for 'retreat' and *fayasu* 'grow' for 'cut'. Another example of such *imikotoba* were so-called *nyōbōkotoba* or *jochūkotoba*, used by women serving in the imperial palace, and which included such terms as *kukon* for *sake* 'saké' and *musi* for *miso* 'miso', neologisms such as *syamozi* 'rice scoop' and *wo-fiya* 'water' still in use today, and widespread employment of the honorific prefix *wo-*. These words spread rapidly, at first through the aristocracy, and later came to be seen as refined: by early ModJ they were being used by ordinary people of either sex.

Lexical borrowing in LMJ can be divided into two major categories: borrowings from Chinese and Portuguese. Although, as just discussed above, the period saw steady growth in the number of SJ lexemes, it also saw a new, albeit comparatively small, set of sinograph readings. Most authorities regard sinographs as utilized in Japanese orthography as having three major SJ readings (*on*), *go-on*, *kan-on* and *tō-on*, the first two of which date from borrowings in the OJ and pre-OJ periods. The latter *tō-on* were the product of the spelling of a relatively small number of borrowed SJ lexemes brought back from the Wu region of China (corresponding roughly to modern Jiangsu province just north of Shanghai) by Japanese monks who had gone to the region to study Zen Buddhism in the first half of the LMJ period. Not surprisingly, therefore, many of the SJ words read in the *tō-on* manner are of a religious nature: *angya* 'pilgrimage', for example, is cited in the *Vocabulario*. In addition, Chinese merchants were also responsible for transmitting some *tō-on* lexemes and these include words such as *isu* 'chair' and *futon* 'futon'.

The other major category of foreign borrowings in LMJ, those from Portuguese (and to a far lesser extent Latin and Spanish), also have their roots in religion, this time in Catholic missionaries, active right at the end of the LMJ period mainly in the Nagasaki area. In his monumental dictionary of Nagasaki *gairaigo* compiled during the first half of the twentieth century, Koga (2000) records some 3,245 words of Portuguese origin and 227 of Latin. Spanish borrowings are yet rarer and probably include *meriyasu* '(type of) knitwear' (< *medias* 'stockings'). How many Iberian borrowings are still alive in standard ModJ is difficult to estimate, but the most common include **pan** 'bread', **tabako** 'tobacco, cigarette', **botan** 'button' and **tempura** 'tempura' (< Port. *tempero* 'seasoning'). Others, such as *waka* 'beef' and *sorudado* 'soldier', have disappeared from

the language. Still others possess through later borrowing a doublet whose usage is more frequent. Most of these doublets have been borrowed from English (e.g. **rĕnkōto** ‘raincoat’ for **kaopa**, **besuto** ‘waistcoat, vest’ for **chooki**, **berubeqto** ‘velvet’ for **birōdo**) but occasionally from other languages (e.g. **garasu** from Dutch *glas* ‘glass’ for **bīdoro**). Usage of the older Iberian loanword may now be restricted to elderly speakers (e.g. **birōdo**), have a limited semantic range (e.g. **bīdoro** is generally restricted to the specialized terminology of glassware), or be restricted to dialect. Finally, since the foreign missionaries had been in contact with other areas of Asia before arriving in Japan, they were responsible for introducing a few Asian borrowings also, including *kiseru* ‘pipe’ from Khmer. For a more detailed discussion of Iberian borrowings see Irwin (2011: 29–35).

## NOTES

- 1 Although in Modern Portuguese <chi> is [ʃi], in the language of the late sixteenth and early seventeenth centuries it was [tʃi].
- 2 These monophthongizations can be extended to vowel sequences where the second vowel is *i*, but these are still incomplete, and often marked, in the modern language. Frellesvig (1995: 84–5) sees these monophthongizations within ‘complex nuclei with post-peak vowel’ as an example of ‘a long term drift . . . not yet complete . . . towards nuclei with no internal contrast’.
- 3 There is no widely accepted phonemic symbol for LMJ [ɔ:]. Although our choice may be viewed as unorthodox, we believe *â* is more useful than something like *ɔ* because (i) it is similar in its articulation to that of the Swedish graph <â>; and (ii) it serves as a useful mnemonic for the derivation of *ââ* from *au*.
- 4 One puzzling aspect of the raising of *ââ* to *oo* is the fact that we fail to find ModJ forms (except in a very few peripheral dialects) such as \***kō** ‘buy’, which are clearly indicated as *kââ*, etc. in the *Vocabulario*. Instead we find, seemingly preserved, what were the late EMJ forms *ka-u*, etc. (along with late EMJ forms such as *araso-u* for what should have become \***arasō**, ‘vie’: see (2d)). Although an answer may appear to lie in the fact that we are dealing with a heteromorphemic environment, such a violation is not apparent in western ModJ adjective forms such as **tako-o** ‘highly’ (eastern = standard ModJ **taka-ku**) which are also heteromorphemic, cf. EMJ *taka-ku*, LMJ *takââ*. See Martin (1987: 44) for an appraisal of the various theories that have been advanced.
- 5 Verb forms referred to as ‘finite’, ‘adnominal’, ‘conditional’ and ‘stem’ in this chapter correspond to ‘conclusive’, ‘attributive’, ‘realis’ and ‘infinitive’ in the previous chapter.

## REFERENCES AND FURTHER READING

- Ahn, Pyeong-Ho and Takenobu Fukushima (2005) ‘Chūsei makki nihongo to gendai kankokugo no tensu, asupekuto taikai’, *Nihongo no kenkyū*, 1(3): 139–154.
- Cho, Seung-bog (1970) *A phonological study of Early Modern Japanese on the basis of the Korean source materials, vol. I: material, historical background, methodology and sound correspondences*, Stockholm: Almqvist and Wiksell.
- Doi, Tadao (ed.) (1955) *Nihon dai-bunten*, translated from *Arte da Lingoa de Japam* by João Rodrigues, Tokyo: Sanseido.
- Frellesvig, Bjarke (1995) *A case study in diachronic phonology – the Japanese onbin sound changes*, Aarhus: Aarhus University Press.
- (2010) *A history of the Japanese language*, Oxford: Oxford University Press.
- Genetz, Kay J. (1979) ‘Notizen zum Ausdruck des Potentials im vormodernen Japanischen’, *Oriens Extremus*, 26(1–2): 291–9.

- Hanzawa, Kan'ichi, Seiya Abe, Masahiro Ono and Hiroshi Kaneko (2002) *Kēsu sutadi nihongo no rekishi*, Tokyo: Ōfū.
- Irwin, Mark (2011) *Loanwords in Japanese*, Amsterdam and Philadelphia: John Benjamins.
- Iwai, Yoshio (1970–1974) *Nihongo hōshi*, Tokyo: Kasama Shoin.
- Kamei, Takeshi (1970) 'Suzume shiushiu', *Seikei Kokubun*, 3: 13–24.
- Kindaichi, Haruhiko (1975) *Nihon no hōgen: akusento no hensen to sono jissō*, Tokyo: Shinkōsha.
- Kinsui, Satoshi (2006) *Nihongo sonzai hyōgen no rekishi*, Tokyo: Hitsuji Shobō.
- Kobayashi, Akemi (1981) 'Ennin no kijutsu suru sansukuritto onsetsu 'ca' no onka: kyūseiki nihongo suion no kokoromi', *Osaka Gaikokugo Daigaku Gakuhō*, 52: 63–80.
- Koga, Jūjirō (2000) *Gairaigo shūran*, Nagasaki: Nagasaki Bunkensha.
- Lange, Roland (1971) 'Bunkenshiryō ni han'ei shita chūseinihongo e-retsū onsetsu no kōgaisei', *Kokugogaku*, 85: 36–42.
- Lewin, Bruno (1989) *Sprache und Schrift Japans*, Leiden: Brill.
- Li, Charles N. and Sandra A. Thompson (1976) 'Subject and topic: a new typology of language', in Li, Charles N. (ed.) *Subject and topic*, New York and London: Academic Press, 457–85.
- Martin, Samuel E. (1987) *The Japanese language through time*, New Haven and London: Yale University Press.
- Miller, Roy Andrew (1967) *The Japanese language*, Chicago and London: University of Chicago Press.
- Miyajima, Tatsuo (1967) 'Gendai goi no keisei', in Kokuritsu Kokugo Kenkyūjo (eds) *Kotoba no kenkyū daisanshū*, Tokyo: Shūei.
- Morita, Takeshi (1977) 'On'in no hensen (3)', in Ōno, Susumu and Shibata, Takeshi (eds) *Iwanami kōza nihongo 5: on'in*, Tokyo: Iwanami Shoten, 253–80.
- (1980) 'Hosetsu', in Doi, Tadao, Morita, Takeshi and Chōnan, Minoru *Nippo jisho*, Tokyo: Iwanami Shoten, 847–62.
- Nakata, Norio (ed.) (1972) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 58–171.
- Narrog, Heiko (1999) *Japanische Verbflexive und flektierbare Verbalsuffixe*, Weisbaden: Harrassowitz.
- Okimori, Takuya (ed.) (1989) *Nihongoshi*, Tokyo: Ōfū.
- Okumura, Mitsuo (1972) 'Kodai no on'in', in Nakata, Norio (ed.) *Kōza nihongoshi 2: On'inshi, mojishi*, Tokyo: Taishūkan, 58–171.
- Peiseler, Sabine (1990) *Morphologie des Verbs im Mitteljapanischen*, MA thesis, Philipps-Universität Marburg.
- Ramsey, S. Robert (1979) 'The Old Kyoto dialect and the historical development of Japanese accent', *Harvard Journal of Asiatic Studies*, 39(1): 157–75.
- Rickmeyer, Jens (1995) *Japanische Morphosyntax*, Heidelberg: Julius Groos Verlag.
- Rothaug, Petra (1991) *Abriß der japanischen Lautgeschichte*, Hamburg: Helmut Buske.
- Sakanashi, Ryūzō and Masayuki Tsukimoto (2001) *Nihongo no rekishi*, Tokyo: Hōsō Daigaku Kyōiku Shinkōkai.
- Sato Habein, Yaeko (1984) *The history of the Japanese written language*, Tokyo: Tokyo University Press.
- Schneider, Roland (1968) 'Zu den Hilfsverben des Tempus und Aspekts des Spaetmitteljapanischen', *Nachrichten der Gesellschaft für Natur- und Völkerkunde Ostasiens/Hamburg*, 104: 21–55.
- (1989) 'Sprachgeschichte', in Lewin, Bruno (ed.) *Sprache und Schrift Japans*, Leiden: Brill, 119–61.
- Seeley, Christopher (2000) *A history of writing in Japan*, Honolulu: University of Hawai'i Press.
- Shibatani, Masayoshi (1990) *The languages of Japan*, Cambridge: Cambridge University Press.

- Sugimoto, Tsutomu (1998) *Nihongo no rekishi*, Tokyo: Yasaka Shobō.
- Toyama, Eiji (1972) 'Kindai no on'in', in Nakata, Norio (ed.) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 173–268.
- Tsukishima, Hiroshi (1969) *Heianjidaigo shinron*, Tokyo: Tokyo University Press.
- (1977) *Kokugo no rekishi*, Tokyo: Tokyo University Press.
- Wenck, Günther (1957) *Japanische Phonetik 3: die Phonetik des Sinojapanischen*, Wiesbaden: Otto Harrassowitz.
- Yamaguchi, Akiho, Hideo Suzuki, Ryūzō Sakanashi and Masayuki Tsukimoto (1997) *Nihongo no rekishi*, Tokyo: Tokyo University Press.
- Yamauchi, Ikuo (1972) 'Kanazukai no rekishi', in Nakata, Norio (ed.) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 531–620.
- Yamauchi, Yōichirō (1989) *Chūseigo ronkō*, Osaka: Seibundō.
- Yanagida, Seiji (1985) *Muromachi jidai no kokugo*, Tokyo: Tōkyōdō.
- Yuzawa, Kōkichirō (1958) *Muromachi jidai gengo no kenkyū. Shōmono no gohō*, Tokyo: Kazama Shobō.

# MODERN JAPANESE

*Nicolas Tranter and Mika Kizu*<sup>1</sup>

## 10.1 STANDARDIZATION AND MODERNIZATION

When in 1868 the feudal shogunal Edo period ended, formal written Japanese was diverse, ranging from *bungo* 文語 (grammatically Classical but lexically a mix of Classical Japanese (CJ) and contemporary; orthographically diverse), through *sōrōbun* 候文 (kanji-heavy CJ with intricate turns of phrase and a polite marker *-sōrō* 候 < LMJ *-sāārāā*), to *kōgo* 口語 (more vernacular but still diverse).<sup>2</sup> *Kōgo* was found mostly only in sermon transcriptions and in dialogue in fiction. The following Meiji period (1868–1912), characterized by rapid modernization and competition with the West, saw some attempts to overcome this diversity – Western-influenced punctuation appeared in some texts (Twine 1984), woodblock printing shifted to type, and vocabulary was modernized (see 10.6) – and intellectuals discussed the idea of *genbun itchi* 言文一致 ‘spoken–written unification’ (Twine 1991: 74ff).

Yet the Ministry of Education only set up the *Kokugo Chōsakai* (National Language Enquiry Board) to consider language issues in 1899, and the twentieth century began with CJ still the norm in school textbooks of all levels and government publications, and *sōrōbun* the norm in formal letter-writing. Newspapers still exhibited a wide variety of styles, but predominantly CJ (see Appendix). The main sphere in which *kōgo* spread in the Meiji period was in fiction, as authors experimented with different Tokyo-based styles that often included different degrees of *bungo* retention. One frequently cited ‘measurement’ of *kōgo* diversity was the copula: CJ *-nari*, high-status *-de gozaru*, highly polite *-de goza(r)imasu*, academic formal *-de aru*, Tokyo colloquial *-da*, Kansai *-ja*, Tokyo polite *-desu* and sophisticated female *-de arimasu* (Twine 1991: 70–3); different works displayed different choices (Twine 1991: 75ff, 138ff). Even in the early twentieth century, some novelists still wrote in *bungo*, even retaining it for writing dialogue between characters. Market forces – people preferred to read a language they spoke – meant that the most popular fiction was in *kōgo*. A decade into the twentieth century saw more conformity in the form of *kōgo* written, and school textbooks and then newspapers were increasingly written in *kōgo*. However, the language of government documents except the Ministry of Education remained *bungo* till 1945, and wartime military communications were grammatically *sōrōbun*.

The post-war government rapidly involved itself in reform: it adopted modern kana spellings (*Gendai Kanazukai* 現代仮名遣い) and a maximal list of 1,850 kanji (*Tōyō Kanji-hyō* 当用漢字表) (Seeley 1991: 153–4) and issued punctuation guidelines (Twine 1984; Seeley 1991: 183–5) in 1946, marked the abolition of *bungo* in government with the new *kōgo*-written Constitution in 1947 (still CJ in a 1946 draft – Carroll 2001: 69), and adopted a subdivision of 881 of the kanji classified by school grade for use in schooling (*Tōyō Kanji Beppyō* 当用漢字別表) and an official streamlined list of acceptable readings of kanji (*Tōyō Kanji Onkun-hyō* 当用漢字音訓表) in 1948 (Seeley 1991: 155–7), and standardized of kanji forms (*Tōyō Kanji Jitai-hyō* 当用漢字字体表) in 1949 (Seeley

1991: 156). Post-1949 reforms of official kanji lists are dealt with in 10.3, but there were also later rules on writing endings in kana (*‘okurigana’*) after kanji and reforms of official kanji readings. For detailed discussion of twentieth-century writing reforms, see Seeley (1991) and Gottlieb (1995).

**10.2 PHONOLOGY**

Japanese is a moraic language. Morae have the limited structure of /V/, /CV/, /CjV/ or ‘moraic’ /Q/ or /N/. Edo-period Japanese also allowed sequences /kwa/ and /gwa/, and late nineteenth-century Western linguists and dictionary-compilers, such as Hepburn (1887), represented these in their romanization of Japanese. However, these appear to have been obsolescent at the time, and such sequences no longer exist. After 1946, spellings such as <kwaigi> くわいき (Hepburn’s (1887: 355) *kwaigi*) were modernized, i.e. <kaigi> かいき.

**10.2.1 Consonants**

*10.2.1.1 Consonant segments*

It is conventional to refer to conservative and innovating varieties of Japanese (Bloch 1950: 88; Vance 1987: 17). The most conservative variety supposedly has 14 consonants /b p m w t d s dz n r j k g h/, which combine with /a e o i u/ or /ja jo ju/ to create simple /C(j)V/ morae. /w/ is restricted to \_/a/, and /j/ to \_/a/o/u/, but otherwise all /CV/ /CjV/ sequences are possible. Various allophones occur: (1.) /g/ = [g] word-initially, [ŋ] ~ [g] word-medially according to speaker, register, type of morphemic boundary or lexical stratum (see the discussion in Vance (1987: 108–32)); (2.) /h/ (< earlier *f*) = [ϕ], /t/ = [ts] and /d/ = [dz] before /u/; (3.) /t/ = [tɕ], /s/ = [ɕ], /d/ = [dʒ], /h/ = [ç] before /i/ or /j/, the [j] phonetically absorbed; (4.) all other consonants reflect a degree of palatalization before /i/j/. This conservative variety reflects an older stage of the language, and linguists find it useful for describing morphological phenomena such as quinquagrade verbs, and most Western linguistic works use the *kunreishiki* transcription that reflects this variety. However, the most conservative variety arguably no longer exists; the impact of Western loanwords has led to the allophonic variations (1.–3.) becoming phonemic, as in Table 10.1. (1.) The exceptions to any attempt to reduce [g/ŋ] allophony to rules suggest there is actually a phonemic /g/ŋ/ distinction word-medially. However, this distinction has never been made in writing (a <g> spelling for both), and the innovating variety if anything is generalizing [g]. (2.) /ϕ/ ~ [ϕ<sup>w</sup>] *f* and /ts/ *ts* now occur before any vowel. (3.) /ɕ/ *sh*,

**TABLE 10.1 JAPANESE CONSONANTS**

	Bilabial	Dental/ Alveolar	(Alveolo-) Palatal	Velar	Glottal
Plosive	p b	t d		k g	
Affricate		ts z	ch j		
Fricative	f {v}	s	sh		
Nasal	m	n		(g)	h
Tap		r			
Approximant	w		y		

/tɛ/ *ch*, /dz/ *j* now occur *\_e/*, while /t/ *t* and /d/ *d* now occur *\_i/*. This phonemicization of /s/ vs /ç/, /t/ vs /tɛ/, and /d/ or /dz/ vs /dz/ means that, for example, [tea] cannot be treated any more as /t/ + /ja/. Although [ç] appears not to have phonemicized, some Tokyo speakers show a merger of [ç] with /ç/, e.g. *hito* /hito/ [çito] > [çito], as though *shito*.

A /v/ phoneme is supposedly available to distinguish /v/ from /b/ in Western loans at least in writing (ヴア /va/, ヴイ /vi/, ヴウ /vu/, ヴエ /ve/, ヴオ /vo/), but in practice only /b/ is found (Vance 1987: 32).

/w/ and /j/ in the most conservative variety occur before /a/ and /a/o/u/ respectively. The most innovating speakers also produce /we/, /wi/, /wo/ and /je/ in loans. However, many speakers produce two-mora /ue/, /ui/, /uo/ and /ie/ instead.

### 10.2.1.2 Moraic consonants

In addition to consonant segments, Japanese has moraic /Q/ and /N/. /Q/ represents gemination of a following consonant. In most conservative Japanese only /p/t/s/k/ can be geminated. In most innovating Japanese, in addition to former allophones of the above (/tɛ/ts/ç/), all other consonants in loanwords can be geminated except /w/j/t/. <Qr> occurs in a few loans, e.g. *aqarā* ‘Allah’, which it seems speakers pronounce as [Ql], [Qr] or just [r]. It will seem odd to specialists in other languages that gemination is often treated as a separate phoneme, which is in part because kana uses a distinct symbol, but more validly because /Q/ is phonetically not a segment but a mora: /Qt/ is not twice the length of /t/, but /Q.te/ is twice the length of /te/ (Vance 1987: 40).

/N/ is a moraic homorganic nasal before a consonant, and variously but typically as [ŋ] otherwise. Moraic /N/ contrasts with segmental /n/, so /ka.ni/ ‘crab’ = two, /ka.N.i/ ‘easy’ = three, and /ka.N.ni.N/ ‘patience’ = four morae.

## 10.2.2 Vowels

Japanese vowels display a five-way quality distinction (Table 10.2), though lip rounding of ‘rounded’ vowels is minimal; /u/ is also somewhat fronted [u̠]. /i/u/ are devoiced between voiceless consonants, e.g. *kita* ‘north’ [k̠ita], unless the previous mora contains a devoiced vowel, e.g. *tsukusu* ‘use up’ [ts̠ukusu]. Devoicing is also found word-finally with *-su*, especially in polite forms *-masu* [mas̠u], *-desu* [des̠u].

Long vowels are one mora longer, and may be treated as sequences of identical vowels, e.g. *ojisan* /o.dzi.sa.N/ ‘uncle’ vs *ojisan* /o.dzi.i.sa.N/ ‘grandpa’. /ee/ is written <ei> *ei* in Sino-Japanese (SJ) morphemes, <ee> *ē* in native (NJ) morphemes, and either way in loan morphemes. There are also two- or three-mora vowel sequences, e.g. *aoi* /a.o.i/ ‘be blue’. There are no diphthongs, and spectrographs show minimal transition between vowels in sequence. Certain word-final sequences of /a.e/ and /a.i/, however, can assimilate to /e.e/ in particularly colloquial/macho style, especially *omae* /o.me.e/ and *temae* /te.me.e/, and *i*-adjectives, e.g. *urusai* /u.ru.se.e/ ‘be noisy’.

Pre-1946 orthographic <iu>/<if> were phonemically /juu/, and post-war reforms regularized the orthography to <yuu> in NJ/SJ forms, e.g. <giuniu> ぎゅうこう ‘milk’ >

TABLE 10.2 JAPANESE VOWELS

i				u
	e		o	
		a		

*gyūnyū* ぎゅうにゅう, <sanjifu> さんじふ ‘30’ > *sanjū* さんじゅう. <iu> is preserved in *iu* いう /juu/, nonpast of *iw-* ‘say; be called’ and loanwords, e.g. *heriumu* ヘリウム /herjuumu/ ‘helium’. The sequence across a compound morpheme boundary is /iu/ in slow speech, but can tend towards /juu/ in fast speech, e.g. *yomi-uri* /jomiuri/ ~ /jomjuuri/ (old word for ‘newspaper’, now a national newspaper’s name). Note also that *kiwi* ‘kiwi(-fruit)’ pronounced by speakers who substitute /u/ for /w/ can result in /kjuui/.

### 10.2.3 Suprasegmentals

Every phonological word in Japanese has a phonemic pattern (‘accent’) of high- (H) and low- (L) pitched syllables. This brief discussion is based on Higurashi (1983), Martin (1987) and Vance (1987) amongst others. In the Tokyo standard: (1.) There is one and only one unbroken stretch of high-pitched morae in a phonological word. (2.) At *surface* level, /Q/ and /N/ can never be the last high-pitched mora in a phrase, and /V/ (vowel length) can only be the last high-pitched mora if immediately followed by a ‘preaccented’ particle. (3.) The first and second morae of a phrase are different pitches, i.e. there is either a rise or a drop after the first mora (unless delayed by (2.)). SJ morphemes’ accent bears no relationship with Chinese tone. Accent is not represented orthographically.

Nouns exhibit the most varied accent possibilities, and accent is lexically determined. The end of an H stretch (‘accent’, marked ´ in Western works) may fall on any permissible mora (‘accented’), or may not fall at all (‘unaccented’, no ´), high pitch continuing onto a following particle such as *-wa*.

- |               |                      |    |                |      |             |
|---------------|----------------------|----|----------------|------|-------------|
| (1) Accented: | <i>hána</i> ‘outset’ | HL | <i>hána-wa</i> | HL-L | (class 2.4) |
|               | <i>haná</i> ‘flower’ | LH | <i>haná-wa</i> | LH-L | (class 2.3) |
| Unaccented:   | <i>hana</i> ‘nose’   | LH | <i>hana-wa</i> | LH-H | (class 2.1) |

Verbs and adjectives exhibit only ‘unaccented’ and a single ‘accented’ pattern, e.g. *atsui* ‘be thick’ vs *atsúi* ‘be hot’. Accented adjectives are currently undergoing change in the position of their accent in all inflections but the citation form (e.g. *átsuku* ~ *atsúku* < ‘be hot’). Despite minimal pairs, accent has relatively low functional load as there are surprisingly few accent minimal pairs. Moreover, confusable homophones often have the same accent, e.g. *kágaku* 科学 ‘science’ or 化学 ‘chemistry’, *ryōshi* 獵師 ‘hunter’ or 漁師 ‘fisherman’. Accent differs greatly by dialect, and that in Osaka is largely the opposite of that in Tokyo, e.g. *haná* ‘outset’, *hána* ‘flower; nose’.

In compounds, the first element’s accent is irrelevant; it is the second element that determines the whole word’s accent. (Below, <sup>x</sup> indicates deletion of the first element’s accent.) There are typically three patterns in compounds: the whole word is unaccented (deaccentuation); the accent falls on the second element’s first mora (initial accentuation); the accent falls on the first element’s last mora (preaccentuation). Which pattern applies generally cannot be predicted phonologically, but it can be lexically: a particular morpheme will impose the same pattern on every compound of which it is the second element (Table 10.3).

Inflectional morphemes (including ‘particles’) are more varied. Some display the same patterns as compounds, betraying their independent origins, i.e. the inflectional morpheme determines the pattern for the whole phrase, the preceding morpheme’s accent being irrelevant.

Most inflectional morphemes, however, have an inherent accent that only manifests depending on the accent type of what precedes, giving four categories: unaccented,

TABLE 10.3 ACCENT IN COMPOUNDS IN JAPANESE

Alone	Compounded	Examples
<i>kaze</i>	<sup>x</sup> - <i>'kaze</i> (preaccentuation)	<i>asákaze</i> 'morning breeze', <i>kamikaze</i> 'divine wind'
<i>sóra</i>	<sup>x</sup> - <i>zóra</i> (initial accentuation)	<i>akanézóra</i> 'red sky', <i>fuyuzóra</i> 'winter sky'
<i>ichi</i>	<sup>x</sup> - <i>'ichi</i> (preaccentuation)	<i>asáichi</i> 'morning market', <i>yasaiichi</i> 'vegetable market'
<i>yamá</i>	<sup>x</sup> - <i>yama</i> (deaccentuation)	<i>fuyuyama</i> 'wintry mountain'

accented, preaccented, accent-shifting. After accented stems, the first three categories are not distinct, as the only permissible accent drop for the phrase has already occurred; after unaccented stems, however, their inherent accent becomes manifest. Accent-shifting morphemes, on the other hand, move the accent of a preceding accented stem rightwards (¿, e.g. *tábe*- 'eat' + *-ru* > *tabéru*); their own inherent accent becomes manifest only after unaccented stems. *-no* behaves differently and irregularly after single nouns, but regularly unaccented after larger noun phrases or expressions of quantity or sequence. Typical examples are as follows. (Alternative patterns exist for <sup>x</sup>-*nágara* and <sup>x</sup>-*só*.)

Compounding:

**Noun+:** <sup>x</sup>-*daké/dake*, <sup>x</sup>-*góro*, <sup>x</sup>-*gúrai*, <sup>x</sup>-*kúrai*, <sup>x</sup>-*rashii*; **Verb+:** <sup>x</sup>-*másu*, <sup>x</sup>-*nágara*;  
**Adjective+:** <sup>x</sup>-*sugiru* 'be too . . .'; **Verb/Adjective+:** <sup>x</sup>-*só*.

Inflectional:

**Noun+:** *-bákari*, *-dá*, *-daró*, *-dé*, *-deshó*, *-désu*<sub>1,2</sub>, *-é*, *-ga*, *-hodo*, *-kará*, *-kóso*, *-máde*, *-mitai*, *-mo*, *-na*, *-nádo*, *-ní*, (phrase/quantifier+) *-no*, *-o*, *-ra*, *-sáe(mo)*, *-shika* (Vance (1987: 84–5) observes *-shika* when /i/ is not devoiced), *-tachi*, *-tó*, *-wa*, *-yóri*.

**Verb+:** quinquagrade *-u*, monograde *-ru*; quinquagrade *-e*, monograde *-ro/yo*; quadrigrade *-ó*, monograde *-yó*; quinquagrade *-éba*, monograde *-réba*; quinquagrade *-eru*; quinquagrade *-seru*, monograde *-saseru*; quinquagrade *-reru*, monograde *-rareru*; *-nai*, *-tai*, *-ta*, *-tára*, *-tári*, *-te*, *-zu*.

**Sentence/te-form+:** *'ga*, *'ka*, *'kara*, *'mo*, *'ne*, *'no* (sentence particle or dummy noun), *'sa*, *'shi*, *'wa* (sentence particle づ), *-wa* (topic particle ば), *'yo*, *'yori*, *'zo*.

**Irregular:** Noun+: <sup>x</sup>-*nó* after (underlying) ultima-accent nouns, *-no* otherwise.

### 10.2.4 Morphophonology

Two phenomena that apply to compound words are assimilation (SJ) and *rendaku* (連濁, predominantly NJ). Both have their roots in the earliest Japanese and apply to 'sei' consonants (the descendants of OJ *p/t/s/k* = CJ *f/t/s/k*). A third, the apophony of OJ/CJ, exists in a small number of roots, though it is of minimal productivity: *ame* 'rain' > *amado* 'raindoor', *fune* 'boat' > *funabin* 'sea mail', *sake* 'saké' > *sakaya* 'alcohol shop'.

Assimilation changes  $-C^1u+C^2-$  or less commonly  $-C^1i+C^2-$  to  $-Q+C^2-$ , where  $C^2$  is 'sei',  $-C^1u$  is *-tsu/-ku* and  $-C^1i$  is *-ki/-chi*. A few examples reflect OJ  $-pu+C^2-$ ; sound changes have meant that the vowels of the unassimilated and assimilated variants are now different, e.g. 合 *gō* (< OJ \**gapu* = CJ *gafu*) vs 合唱 *gaQ+shō* 'chorus', and the process is no longer productive. SJ *jū* (< OJ \**zipu*) '10' still assimilates to a *sei*-initial classifier, e.g. *jiQ+pāsento* '10 per cent', but standard usage 'regularizes' the vowel: *juQ+pāsento*.

Rendaku involves the voicing of the initial consonant of a compound word's second morpheme: *sei* consonants to their 'daku' equivalents (OJ/CJ *b/d/z/g*), which in our current analysis means *tV, tsu, ch, k, h/f, sV, sh > dV, zu, j, g, b, zV*. It predominantly affects NJ morphemes in noun-noun non-dvandva compounds, and in noun-verb/adjective compounds. In the latter case, the relationship of noun to verb is typically that expressed by the case particles *-ga* or *-ni*, and there are examples in which either a particle or rendaku is possible: *ki-ga tsuku* or *kizuku* 'notice'. Whether a morpheme undergoes rendaku is largely lexically determined: some always undergo rendaku, some never undergo rendaku, and some undergo rendaku in some compounds but not in others. There are some phonological, morphosemantic and morphological factors: morphemes already containing a *daku* consonant never undergo rendaku (Lyman's Law, e.g. *kami+kaze* 'divine wind' not \**kami+gaze*); dvandva noun-compounds do not undergo rendaku (*oya+ko* 'parent and child' not \**oya+go*). Also, rendaku is less common if one element is itself already a compound.

Loanwords are not normally subject to rendaku; a common exception is *karuta* 'playing cards': *hana+garuta* 'flower playing cards'. Some SJ morphemes appear to display rendaku, perhaps a reinterpretation of the two SJ lexical layers (in which the go'on layer preserves Early Middle Chinese voicing and the kan'on layer reflects Late Middle Chinese loss of voicing) as a rendaku relationship.

## 10.3 ORTHOGRAPHY

### 10.3.1 Script distribution

Japanese is written in a mixed script of Chinese characters (*kanji*), two formally distinct syllabaries (*hiragana* and *katakana*, collectively *kana*), as well as use of non-Asian scripts, particularly roman letters and Western numerals. There are also many pseudo-Chinese characters invented in Japan (*kokuji*, or *wasei kanji*), e.g. 峠 *tōge* 'mountain pass', 働く *hataraku* 'to work'; most represent species and are no longer in general use. The system has a history of open-endedness (Backhouse 1984: 221), in that there is theoretically no limit on the number of characters available, e.g. the larger number of hentaigana symbols (see 10.3.3) than current hiragana ones. This has facilitated the adoption of other scripts.

Reforms reduced the number of kanji available to a *prescribed* 1,850 (*tōyō kanji* 当用漢字) in 1946 enforced in government and educational publications, with a 1948 subdivision of 881 kanji (*kyōiku kanji* 教育漢字) for teaching in primary education, all the Ministry of Education's remit. As many people's names used non-tōyō kanji, a new list of initially 90 extra or variant kanji for registering names (*jinmeiyō kanji* 人名用漢字) was issued in 1951 (Seeley 1991: 157), the remit of the Ministry of Justice. The prescriptiveness of these proved problematic, not least because certain prefecture/city names e.g. 奈良 *nara* or 大阪 *ōsaka*, and even 韓国 *kan'koku* 'South Korea' included off-list kanji, so in 1981 these lists were converted into a supposedly set-in-stone 1,945 *recommended* rather than prescribed kanji (*jōyō kanji* 常用漢字, including 1,006 *kyōiku kanji* from 1982) from the Ministry of Education, plus 166 *jinmeiyō kanji* from the Ministry of Justice. The split between two ministeries was problematic, and in 2004 the MoJ proposed to expand its own list, apparently also intending to slip in some non-name characters in recognition of their frequency of occurrence in print. Unfortunately, these included 癌 *gan* 'cancer', 尻 *shiri* 'buttocks' (also used in e.g. 目尻 *mejiri* 'corner of eye', 尻取り *shiritori* a wordgame) and 糞 *kuso* 'shit' (also used in acceptable 鳥糞

*chōfun* ‘guano’). This bureaucratic misstep was widely ridiculed in Japan and abroad and the nine worst were not adopted. By 2010, though, the MoJ list had reached almost a thousand. Resolution of the issue was attempted in 2010, with the first revision of the jōyō kanji to 2,136 including 奈 阪 韓 尻 but not 癌 糞 (Bunka Shingikai 2010) and confirmation of 861 jinmeiyō kanji.

The system of multiple scripts co-occurring in the same text means that individual words may be written in more than one way, and so, despite a range of government prescriptions or recommendations regarding writing, there is a great variation in usage in practice, and so the term ‘orthography’ in its strictest sense barely seems appropriate. The following outline (based on Backhouse 1984 and Tranter 2008) reflects tendencies in actual use rather than official guidelines.

- (1.) Regarding SJ and NJ words excluding names of species, where an appropriate kanji is included among the jōyō kanji:
  - (a.) SJ words (including those created in Japan from Chinese morphemes) are written in kanji.
  - (b.) NJ content words are written in kanji.
  - (c.) Numerals are written either in Western numerals or kanji, but even when written in kanji page numbers, telephone numbers, statistics and dates generally have a Western-style decimal system with 〇 as zero, e.g.: 1980年 or 一九八〇年 (not \*千九百八十年) *sen kyūhyaku hachijū-nen* ‘(the year) 1980’.
- (2.) Regarding SJ and NJ words excluding names of species, where an appropriate kanji is not included among the jōyō kanji:
  - (a.) Chemical elements use katakana (but only for the affected morpheme, not the whole word), presumably because terms for most other elements are loans from Dutch and thus written in katakana: ホウ素 *hōso* ‘boron’, ケイ素 *keiso* ‘silicon’, ヒ素 *hiso* ‘arsenic’, ヨウ素 *yōso* ‘iodine’.
  - (b.) Other words use hiragana or a non-jōyō kanji, e.g. がん/癌 *gan* ‘cancer’. In the case of bimorphemic words of which only one morpheme has a jōyō kanji, there are three solutions: all kana, all kanji, or half-half, e.g. ゆうつ/憂鬱/憂うつ *yūutsu* ‘depression’. (鬱 became jōyō in 2010.)
- (3.) Only a handful of NJ/SJ species names have jōyō kanji, e.g. 蛇 *hebi* ‘snake’, 杉 *sugi* ‘cedar’, so in non-scientific writing the rest tend to be written in hiragana, e.g. たい/鯛 *tai* ‘bream’, さかき/榊 *sakaki* ‘sakaki-tree’. In scientific writing (including field guides) all species are written in katakana: ヘビ *hebi*, スギ *sugi*, タイ *tai*, サカキ *sakaki*. Even 人 *hito* ‘person, human’ is written in katakana when presented as a species, e.g. ヒトゲノム *hito genomu* ‘the human genome’.
- (4.) Loans from languages other than Chinese are written in katakana. Just a few old loans still retain pre-twentieth-century kanji spellings, though these are less common than katakana, e.g. タバコ/煙草 *tabako* ‘cigarette’, ページ/non-jōyō 頁 *pēji* ‘page’, クラブ/倶楽部 *kurabu* ‘club’. *tenpura* ‘tempura’ and *tabako* can also have hiragana spellings.
- (5.) Emotive forms such as interjections or phonaesthesia are written in kana. High-brow writing sometimes uses katakana for true onomatopoeia and hiragana for other phonaesthesia, but in practice most writing makes no clear distinction, and popular writing tends to use katakana in line with (6.).

- (6.) Emphatic forms in popular writing are written in katakana, including interjections, emphatic particles, emphatic lengthening of final syllables, and slang or taboo vocabulary.
- (7.) NJ grammatical forms are written in hiragana. The borderline between content word and grammatical form is fuzzy, so certain grammaticalized content words may be written in either kanji or hiragana, e.g. 時/とき *toki* ‘when’.
- (8.) ‘Ruby’ (*rubī*), auxiliary text written to the right of the main text in vertical script and above it in horizontal, is common. Conventionally this takes the form of ‘*furigana*’, i.e. ruby hiragana, which indicate the pronunciation of the kanji that they accompany, e.g. <sup>とき</sup>時. The amount of ruby depends on genre and assumed literacy level of the readership. Newspapers and scientific writing use it rarely, while there is a higher proportion in popular fiction. The younger the reader, the more frequent the device is, so that young children’s books may have *furigana* on every kanji except numerals, and those for the youngest even have it on katakana words (Backhouse 1984: 220).

### 10.3.2 Kana

The basic kana signs are presented in Table 10.4. <wi>, <we>, <wo> are pronounced *i*, *e* and *o*; since 1946 <wo> has been restricted to writing only the object particle *-o* and *kiotsuke* ‘standing to attention’ < *ki-o tsuke(-ro)*! ‘attention-ACC pay(-IMP)’ whereas <wi> and <we> have been officially obsolete. However, <wi> and <we> still occur in all Japanese fonts, are consistently used in publishing CJ texts, and are sporadically used to impart a traditional feel particularly to products: the lucky god Ebisu (CJ *webisu*) is often written 恵びす <*webisu*> at shrines, etc., ‘Yebisu’ beer is written エビス <*webisu*>, and Nikka ‘Whisky’ has used キスキイ <*wisukii*>.

Other morae not in the chart are derived either by diacritics or by compound symbols. There are two diacritics: <sup>˘</sup> added to the *k*-, *t*-, *s*-, *h*- columns and to <u> to derive *g*, *d*, *z*, *b* and <vu>, e.g. て *te* > で *de*. Though <vu> has been used for over a century, official approval is relatively recent and so it has still to gain ground in more formal writing. <sup>˚</sup> added to the *h*- column derives *p*, e.g. ひ *hi* > ひ<sup>˚</sup> *pi*. /q/ is represented by a smaller version of the symbols for *tsu*: か<sup>˚</sup>つて /ka.tsu.te/ ‘formerly’ vs か<sup>˚</sup>つて /ka.q.te/ ‘wilful’.

There are only two ‘irregular’ spellings, where the spelling suggests a different pronunciation: へ <*he*> for *-e* ‘to’ and は <*ha*> for *-wa* Topic. To this we may add several cases in which the same sounds may be written in more than one way: (1.) <o> and <wo> for *o* (see above). (2.) <oo> and <ou> for *ō* (next paragraph). (3.) <ee> and <ei> for /eɛ/, a distinction maintained in Hepburn transcription as *e* vs *ei*. (4.) The system still distinguishes づ (<*tsu*> + <sup>˘</sup>) vs ず (<*su*> + <sup>˘</sup>) for /dzu/ *zu*, and ぢ (<*chi*> + <sup>˘</sup>) vs じ (<*shi*> + <sup>˘</sup>) for /dzi/ *ji*, as well as their katakana equivalents, even though there is no longer any phonological difference. Before 1946, the spellings reflected etymology, and were surprisingly consistent even in pre-Meiji manuscripts, but since 1946 the distribution is predictable on a mix of phonological and morphological grounds: づ and ぢ are used: (a.) following <*tsu*> and <*chi*> respectively in the same NJ morpheme, e.g. つづく *tsuzuku* ‘to continue’, ぢぢむ *chijimu* ‘to shrink’; (b.) morpheme-initially as the outcome of *rendaku*, e.g. はなぢ *hanaji* ‘nosebleed’ < *hana* ‘nose’ + *chi* ‘blood’. ず and じ are used everywhere else, except for archaicizing purposes (e.g. ぢ ‘haemorrhoids’ in pharmacy advertisements).

TABLE 10.4 HIRAGANA | KATAKANA SYLLABARIES

	Ø-	k-	s-	t-	n-	h-	m-	y-	r-	w-
-a	あ ア	か カ	さ サ	た タ	な ナ	は ハ	ま マ	や ヤ	ら ラ	わ ワ
<i>a</i>	<i>ka</i>	<i>sa</i>	<i>ta</i>	<i>na</i>	<i>ha</i>	<i>ma</i>	<i>ya</i>	<i>ra</i>	<i>wa</i>	
-i	い イ	き キ	し シ	ち チ	に ニ	ひ ヒ	み ミ	—	り リ	ゐ ヰ
<i>i</i>	<i>ki</i>	<i>shi</i>	<i>chi</i>	<i>ni</i>	<i>hi</i>	<i>mi</i>		<i>ri</i>	<i>i</i>	
-u	う ウ	く ク	す ス	つ ツ	ぬ ヌ	ふ フ	む ム	ゆ ユ	る ル	—
<i>u</i>	<i>ku</i>	<i>su</i>	<i>tsu</i>	<i>nu</i>	<i>fu</i>	<i>mu</i>	<i>yu</i>	<i>ru</i>		
-e	え エ	け ケ	せ セ	て テ	ね ネ	へ ヘ	め メ	—	れ レ	ゑ ヱ
<i>e</i>	<i>ke</i>	<i>se</i>	<i>te</i>	<i>ne</i>	<i>he</i>	<i>me</i>		<i>re</i>	<i>e</i>	
-o	お オ	こ コ	そ ソ	と ト	の ノ	ほ ホ	も モ	よ ヨ	ろ ロ	を ヲ
<i>o</i>	<i>ko</i>	<i>so</i>	<i>to</i>	<i>no</i>	<i>ho</i>	<i>mo</i>	<i>yo</i>	<i>ro</i>	<i>o</i>	
<i>N</i>	ん ン									

Vowel length in katakana is represented by a line after the appropriate kana, vertical in vertical script and horizontal in horizontal script. In hiragana the line is replaced by vowel kana: <a> to lengthen /a/, <i> to lengthen /e/ and /i/, <u> to lengthen /u/ and usually /o/, and <o> to lengthen /o/ in just a handful of NJ roots (e.g. *ō-* ‘big’, *ōkii* ‘be big’, *ōi* ‘be many’, *tōru* ‘go through’).

Compound symbols take the form of two symbols representing a single mora. To indicate that they are not separate morae, the second is written smaller than the first, though with combinations used in loanwords errors are sometimes found. In light of the phonological analysis presented above, we should represent the spellings as:

- (a.)  $C = k/g/n/h/b/m/r: \langle Ci \rangle + \langle y_a \rangle = Cy_a, \langle Ci \rangle + \langle y_o \rangle = Cy_o, \langle Ci \rangle + \langle y_u \rangle = Cy_u$   
 (b.)  $C = sh/ch/j: \langle Ci \rangle + \langle y_a \rangle = Ca, \langle Ci \rangle + \langle y_o \rangle = Co, \langle Ci \rangle + \langle y_u \rangle = Cu$   
 (c.)  $C = sh/ch/j: \langle Ci \rangle + \langle e \rangle = Ce$   
 (d.)  $C = ts/f/v: \langle Cu \rangle + \langle a \rangle = Ca, \langle Cu \rangle + \langle i \rangle = Ci, \langle Cu \rangle + \langle e \rangle = Ce, \langle Cu \rangle + \langle o \rangle = Co$   
 (e.)  $C = t/d: \langle Ce \rangle + \langle i \rangle = Ci$   
 (f.)  $\langle i \rangle + \langle e \rangle = ye, \langle u \rangle + \langle a \rangle = wa, \langle u \rangle + \langle i \rangle = wi, \langle u \rangle + \langle e \rangle = we, \langle u \rangle + \langle o \rangle = wo$

Of these, (a.–b.) are inherited from the earliest kana spellings of Chinese-derived words, though the second element’s reduced size was standardized only in 1946. (c.–f.), on the other hand, are more recent combinations used largely for writing twentieth- and twenty-first-century loanwords or interjections, e.g. *ちえ* *che* ‘tut!’, *ファン* *fan* ‘fan’, *ヴェネツィア* *venetsia* ‘Venice’, *パーティー* *pātī* ‘party’, *イエロー* *yerō* ‘yellow’, *ウィンドー* *windō* ‘window’. Digraphs with <tsu>, <vu>, <u> and <i> are not fully accepted, as many Japanese find them difficult to pronounce, and alternative spellings are still common: *ベネチア* *benechia* ‘Venice’, *イエロー* *ierō* ‘yellow’, *ウィンドー* *windō* ‘window’.

### 10.3.3 Script change

Pre-c. 1880 writing differed from current usage variously. (a.) The kana component of a sentence was often katakana, especially formal writing. (b.) Hentaigana was used, with multiple graphemes for each mora, e.g. し ~ 志 *shi*, ふ ~ 瀧 *fu*, れ ~ 累 *re*. (c.) Some kanji were written variously, e.g. *machi*–*chō* ‘town(ship)’ 町 ~ 早 ~ 町. (d.) The more formal the writing, the larger the proportion of kanji, which also wrote certain particles

or endings with Chinese-influenced inversions, e.g. 迄 *-made* ‘until’ or 不  $V = V$ -zu negative. (e.) Okurigana were used minimally. (f.) Punctuation/paragraphing was unused or limited to sentence/paragraph-ending ○ or ● or \ (Fujii 1984). (g.) Special symbols from grass-written characters or fused kana sequences occurred, e.g. ㄱ *koto* ‘fact; words’, ㄱ *shite te*-form of *su(ru)* ‘do’, ㄱ *toki* ‘time’, ㄱ *-domo* ‘although’. ㄱ in play-scripts or fiction already indicated the start of speech. (h.) Loanwords, including people’s names, used kanji. (i.) A range of repetition symbols existed: 々 or 仝 or 々 for kanji, 々 for hentaigana, 々 for katakana, and 々 for several kana. (j.) Writing was strictly vertical, with horizontal writing read right-to-left originating in vertical single-character columns e.g. above gates, etc.

During the Meiji period, movable type printing entailed the discarding of multiple kana forms for a single mora, i.e. hentaigana became hiragana as we now know it, and standardization of kanji forms. Paragraphing was adopted, and native punctuation marks developed distinct functions in imitation of the similar-looking Western marks: 〃 marked the end of a sentence, and 〃 became a comma. Special symbols (g.) were abolished, though the use of ㄱ at the start of a direct quotation was developed into the pair of inverted commas as currently used. Loanwords were written in katakana.

The distribution of symbols has changed little since. Though the 1946 guidelines allowed a range of repetition symbols, only the kanji-repetition 々 survives. Certain older symbols survive on the periphery: 々 *shime* is used widely on doors to indicate that they will not open (*shimete aota* ‘is shut’); and some newspaper columns save space by replacing paragraphing with variants of paragraph-marking ●, e.g. ▼ in *Asahi Shinbun*’s ‘Tensei jingo’. From 1947, horizontal writing changed from right-to-left to left-to-right (2):

- (2) ← 望展と顧回の動運働勞 (Asahi Shinbun, 31 Dec 1946, p.1)  
*rōdō undō-no kaiko-to tenbō*  
 ‘Labour unions looking back and looking ahead’
- マ元帥・年頭の言葉 (Asahi Shinbun, 1 Jan 1947, p.1)  
*ma[okāsā] gensui – nentō-no kotoba*  
 ‘General MacArthur – New Year’s message’

Modern writing has developed a range of non-conventional devices and options offered by technology primarily in manga, advertising and commercial use, and SMS and online discussions. Manga display e.g. prolonged vowels by means of small-sized katakana; ruby to write two words simultaneously, typically an English-derived term written in katakana on an NJ/SJ word written in kanji, or vice versa; excessive use of !, ?, !?; and emotional symbolizations, such as ♥, ‘pleuds’ (beads of sweat), etc., which are drawn beside a character’s head, but if the character’s head is out of frame they are drawn in the speech bubble almost like sentence-particles (Gravett 2004: 79; Tranter 2008: 141). These devices have spread into young (especially female) casual handwriting. Some, especially the use of small katakana and unusual ruby, have spread into certain lowbrow fiction (Tranter 2008: 139–40). Advertising, etc. makes particularly heavy use of English or the roman alphabet (Inoue 2005), because they are eye-catching and because English is perceived as cool; popular magazines tend to write certain potentially taboo loanwords in the roman alphabet, e.g. SEX. By ‘options offered by technology’ is meant the sudden availability of non-jōyō kanji, other scripts, non-linguistic symbols, etc. that modern word-processing and SMS texting offer to the user. Finally, modern word-processing readily allows users to input characters that they never would have used naturally, such as more obscure

kanji; moreover, it has allowed young, mostly female, users to develop a code (*gyarumoji*) in texting, etc. (Miller 2004), whereby Japanese symbols are replaced by similar-looking symbols, e.g. |£ for hiragana り.

Finally, it should be noted that not only has script changed, but also handwriting, which now tends to be fairly square-script in style, in contrast with cursive- or semi-grass-script before 1945.

## 10.4 MORPHOLOGY

### 10.4.1 Nouns and particles

The noun class includes equivalents to pronouns in other languages. There is little morphological difference between pronouns and other nouns. Some pronouns derive from earlier true pronouns and end in *-re*: demonstratives *kore* ‘this’, *sore* ‘that’ (near you), *are* ‘that’ (near neither of us), *dore* ‘which?’ and in very restricted use *ware* ‘I’. *ware* is usually encountered reduplicated as *wareware* ‘we’ (inclusive). Historical *re*-less forms occur only with *-no* forming modifier demonstratives *kono/sono/ano/dono*, and in e.g. *wa-ga-kuni* ‘our country’, *wa-ga-mama* ‘(my own manner >) selfish’.

Pronoun equivalents include: (1.) ‘I’: *watashi* and variants (formal *watakushi*, female *atashi*, *aoshi*), male familiar *boku*, male casual *ore*, highly formal *kochira*; (2.) ‘you’: *anata* and variants (female *anta*), male familiar *omae*, intimate *kimi*, male confrontational *kisama*. (3.) *kare* ‘he’, *kanojo* ‘she’, formal *ano hito*, honorific *ano kata* ‘he/she’. *anata* is the least marked and most acceptable for ‘you’, but it does have connotations, used to address people of whose rank the speaker is unaware (hence its frequency in advertising), and by wives to their husbands. Third-person *kare/kanojo* were adopted by writers imitating male/female distinctions in Western pronouns and spread into general use; supposedly used only of equals in familiar use or of inferiors, they are nevertheless widespread. Instead of second/third-person pronouns, names, titles or no form at all are preferred.

Japanese has a three-way demonstrative system (see 10.5.2.2) based on the roots *ko-*, *so-*, *a-*; a fourth, interrogative set *do-* ‘which’ completes the system. The four roots regularly create pronouns (*-re*; choice-of-two or polite *-chira* ~ colloquial *-ochi*; place *-ko*; human pejorative *-itsu*), modifiers (*-no*; choice-of-two or polite *-chira-no* ~ colloquial *-ochi-no*; adjective of ‘type’ *-nna* ~ *-iu* ~ *-iota*) and adverbs (manner *-*; direction *-chira* ~ colloquial *-ochi*); the only irregularity is distal *asoko* ‘there’. *-chira/-ochi*, it will be noticed, has a range of functions, to which we can add very formal *kochira* ‘I’, *sochira* ‘you’.

Nouns are not generally marked for number, though there are pluralizing suffixes: only *-tachi* is productive; lexically restricted are *-ra*, honorific *-gata*, and humble/pejorative *-domo*. They are only attached optionally to nouns with human reference, or through anthropomorphism to those with animal or inanimate reference in folktales or children’s literature. Attached to nouns with unique reference, typically names or pronouns, they express a group that has the referent at the centre. There is a degree of collocation between pronoun and suffix: *kare* and *kanojo* take *-ra*; *watashi* takes *-tachi/-domo*, *anata* takes *-tachi/-gata*.

*Case particles and postpositions*: *-ga* subject; *-o* object; *-ni* indirect object, agent of passive, static location, direction, adjunct to *naru* ‘become’/ *suru* ‘make (into)’; *-no* possessive and general noun-linking particle; *-de* dynamic location, instrument; *-e* direction; *-kara* ‘from’; *-made* ‘until, as far as’; *-yori* ‘than’ (or ‘from’ in highly formal use); *-to* ‘with’, adjunct (esp. abstractions) to *naru* ‘become’/ *suru* ‘make (into)’; plus zero particle.

Zero particle in formal Japanese expresses time/distance covered; colloquially it commonly substitutes for *-ga*, *-o*, and sometimes *-ni*. CJ *-ni-te* is largely obsolete, but is still used e.g. on events posters: PLACE-*ni-te* ‘venue: PLACE’.

*Coordinating particles*: (a.) between nouns: *-to* ‘and’, *-ya* ‘and, or’ (inexhaustive listing), *-ka* ‘or’; (b.) otherwise: *-nado* ‘etc.’, *-nante* ‘or such like’, *-nanka* ‘or such like’, *-to* as in *X-to Y-to = X-to Y*.

*Delimiting particles*: *-dake* ‘only’, *-shika* ‘[nothing] but’, *-bakari* ‘only; approximately’, *-gurai* ‘approximately’, *-goro* ‘approximately’, *-atari* ‘around; per’, *-hodo* ‘approximately; [not as . . .] as’.

*Discourse particles*: *-wa* topic, *-mo* ‘also, even’, *-sae(mo)* ‘even’, *-sura* ‘even’, *-made(mo)* ‘even’, *-demo* ‘or such like’, *-koso* focus (highly formal only).

*-made* occurs twice in the list with different functions. *-to* also occurs twice. Particles do not display allomorphy, except *-gurai* ~ *-kurai* and *-goro* ~ *-koro*, the latter of each being slightly more formal. In very colloquial language, *-no* can contract to *-n*, especially in set phrases, e.g. *boku-n chi < boku-no uchi* ‘my house’.

Particles may combine usually in a fixed order (Table 10.5). Other than semantic restrictions, there are restrictions on combinations with *-ga*, *-o* and *-wa*, and perhaps *-ni*. In practice, combinations of three or more particles are uncommon, though not impossible.

In colloquial Japanese (*-ga*), *-o*, *-wa* are commonly ellipted.

To the list of particles we may add the copula in slot 5, or a number of forms that typically follow verbals, e.g. feedback-requesting *ne* or the polite *-desu<sub>2</sub>*.

#### 10.4.2 Verbs and adjectives

Morphologically, Japanese verbals may be divided into verbs proper and adjectives. Verbals may also be divided into three semantic-based classes that affect morphology and syntax: durative, punctual and stative. Stative verbals include all adjectives, the copula, *aru* and *iru* ‘to exist, be (located), have’, *iru* ‘to need’, *wakaru* ‘to understand’, *dekiru* ‘can do/speak’, and the *-[i]tai* and potential forms of verbs. Morphologically, statives generally lack progressive (see 10.5.3.2) or passive forms. Syntactically, they mark objects with *-ga* rather than *-o*; the *-[i]tai* and potential forms consist of typically a *non*-stative verb + a *stative* morpheme, and so tend to be torn between *-o* and *-ga*. *dekiru* is a suppletive rather than suffixal potential, and so is not torn, always taking *-ga*.

Durative and punctual verbs express actions that have duration and actions that are instantaneous respectively. Both form progressive forms, but with punctual verbs they express ongoing *result*. Punctuals include verbs that express change of state (e.g. *naru* ‘to become’, *nureru* ‘to become wet’, *shinu* ‘to die’, *ke@kon suru* ‘to get married’), the intransitive partner of a transitive/intransitive pair (e.g. *aku* ‘to open’ (intr.) vs *akeru* ‘to open’ (tr.)), intransitive verbs of posture (e.g. *suwaru* ‘to sit down’, *tatsu* ‘to stand up’), or motion verbs that express endpoint or path (e.g. *iku* ‘to go’, *kuru* ‘to come’, *kaeru* ‘to return’, *tōru* ‘to go through’). Motion verbs that express manner are durative (e.g. *oyogu* ‘to swim’, *aruku* ‘to walk’). The transitive partner of a transitive/intransitive pair is usually durative. Verbs of putting on or removing clothing and accessories are punctual, so their progressive form is equivalent to ‘is wearing’ or ‘is no longer wearing’. *dekiru* is stative when it means ‘can do/speak’ but punctual when it means ‘be completed’.

TABLE 10.5 JAPANESE PARTICLES

1.		-atari	+ -ga/-o/-wa
		-bakari	± -ga/-o/-wa
		-dake	+ -ga/-o/-wa
		-goro (koro)	n/a
		-gurai (kurai)	- -ga/-o, + -wa
		-hodo	- -ga/-o/-wa
2. a.	Coordinating:	-ka	
		-to	
		-ya	
b.	Coordinating:	-nado	+ -ga/-o/-wa
		-to	+ -ga/-o/-wa
		-nanka	- -ga/-o/-wa
		-nante	- -ga/-o/-wa
3.	Case/Postposition:	-∅	
		-de	
		-e	
		-ga	
		-kara	
		-made	
		-ni	
		-o	
		-to	
		-yori	
4.	Delimiting:	-shika	- -ga/-o/-wa
5. a.	Discourse:	-demo	- -ga/-o
		-koso	- -ga/-o
		-made(mo)	- -ga/-o
		-mo	- -ga/-o
		-sae(mo)	- -ga/-o
		-sura	- -ga/-o
		-wa	- -ga/-o
b.	Noun-Modifying:	-no	

#### 10.4.2.1 Adjectives

Inflected adjectives distinguish *i*-adjectives (*keiyōshi* 形容詞) and *na*-adjectives (*keiyō dōshi* 形容動詞 ‘adjective-verbs’, in the West often called ‘nominal adjectives’ or ‘adjectival nouns’). *na*-adjectives (< CJ *nari*-adjectives) mostly inflect like the copula. A few adjectives are irregular: *onaji* ‘to be the same’ is mostly *na*-type, while *ōkii* ‘to be big’, *chiūsai* ‘to be small’ are *i*-adjectives with alternative *na*-adjective attributives. *beki* ‘should’ is limited in all but the most formal language to nonpast *beki-da* and attributive *beki*. *i*-adjectives have lost their earlier *ō*-form (*-karō*, now highly literary) in favour of the *na*-adjective construction nonpast + *-darō*. In Table 10.6, # indicates places in which a discourse particle, e.g. *-wa* or *-mo*, may occur.

*ii ~ yoi* ‘to be good’ uses *yo-* as the root for all forms, except the nonpast, in which both occur. The *sō*-forms of *ii ~ yoi* and of *nai* ‘not to exist, not to be (located); not to have’ consist of suffixed *-sasō*, not *-sō*. Older forms of *i-* and *i/na*-adjectives that are largely obsolete include attributive *-ki* (sometimes used in book/film titles for gravitas, e.g. *subarashiki aru hi* ‘One Fine Day’), nonpast *-shi* (cf. *jiko ōshi* ‘there are many accidents’, i.e. ‘accident blackspot’ on road-signs), and *ō*-form *-karō*.

**TABLE 10.6 JAPANESE ADJECTIVE MORPHOLOGY**

		<i>i</i> -Adj.	<i>i/na</i> -Adj.	<i>na</i> -Adj.	<i>onaji</i>
Finite:	Nonpast:	-i	-i	-da	-da
	Attributive:	-i	-na or -i	-na	-Ø
	Past:	-kaŋta	-kaŋta	-daŋta	-daŋta
	<i>ō</i> -form:	-i-darō	-i-darō	-darō	-darō
Conjunctive:	Adverb:	-ku	-ku	-ni	-ku
	<i>te</i> -form:	-kute	-kute	-de	-de
	<i>tari</i> -form:	-kaŋtari	-kaŋtari	-daŋtari	-daŋtari
	<i>tara</i> -form:	-kaŋtara	-kaŋtara	-daŋtara	-daŋtara
	<i>ba</i> -form:	-kereba	-kereba	-nara	-nara
Derivatives:	<i>sō</i> -form:	-sō-na	-sō-na	-sō-na	...
	Polite:	PARTICLE	PARTICLE	-desu <sub>1</sub>	-desu <sub>1</sub>
	Negative:	-ku # nai	-ku # nai	-de # nai -janai	-de # nai -janai

**TABLE 10.7 POLITE FORMS OF JAPANESE ADJECTIVES**

		<i>i</i> -Adjectives	<i>na</i> -Adjectives
Affirmative:	Nonpast:	-i-desu <sub>2</sub>	-desu <sub>1</sub>
	Past:	-kaŋta-desu <sub>2</sub>	-deshita
	<i>ō</i> -form:	-i-deshō	-deshō
Negative:	Nonpast:	-ku # arimasen	-de # arimasen
		-ku # nai-desu <sub>2</sub>	-jaarimasen
	Past:	-ku # arimasen-deshita	-de # arimasen-deshita
		-ku # nakaŋta-desu <sub>2</sub>	-jaarimasen-deshita
<i>ō</i> -form:	-ku # nai-deshō	-de # nai-deshō -janai-deshō	

*i*- (and *i/na*-)adjectives do not have a true polite form. Whereas the polite forms of *na*-adjectives or verbs are inflected morphemes, *i*- and *i/na*-adjectives use the polite particle *desu*<sub>2</sub>, homophonous with the copula or *na*-adjective polite form but not inflected (Table 10.7). Substandard language sometimes treats the polite particle as the copula, giving !*takai-deshita* ‘it was expensive’ for standard *takakaŋta-desu*<sub>2</sub>.

10.4.2.2 Verbs

Most verbs are quinquagrade (*godan dōshi* 五段動詞) or monograde (*ichidan dōshi* 一段動詞). Quinquagrade verb stems end in *t*, *r*, *w*, *k*, *g*, *m*, *b*, *n*, and the nonpast form adds a further *-u*; monograde verb stems end in *e* or *i*, and the nonpast adds a further *-ru*. The conjugation of verbs whose citation (‘dictionary’) form ends in *-ru* is only partly predictable: if *-ru* is preceded by *a*, *o* or *u*, it is quinquagrade; if preceded by *e* or *i*, most are monograde, some quinquagrade. Of the latter, only *hairu* ‘to enter’, *hashiru* ‘to run’, *heru* ‘to decrease’, *iru* ‘to enter’, *kaeru* ‘to return’, *kiru* ‘to cut/slice’, *mairu* ‘to go/come (HUM)’, *shiru* ‘to come to know’ occur in this chapter (cf. monograde *iru* ‘to exist’, *kaeru* ‘to change’, *kiru* ‘to put on [clothes]’). Inflection is shown in Table 10.8.<sup>3</sup> Quinquagrade root-consonants reflect the most conservative consonant inventory, i.e. *mat-* ‘to wait’, *das-* ‘to take out’, *kaw-* ‘to buy’ + *-u*, *-i*, *-anai* = *matsu*, *machi*, *matanai*; *dasu*, *dashi*, *dasanai*; *kaŋu*, *kaŋi*, *kawanai*.

TABLE 10.8 JAPANESE QUINTUAGRADE AND MONOGRADE VERBS AND *-[a]nai*

		Quinquagrade	Monograde	<i>-[a]nai</i> Negative
Finite:	Nonpast:	-u	-ru	-[a]nai
	Past:	-Ta	-ta	-[a]nakaota
	<i>ō</i> -form:	-ō	-yō	...
	<i>mai</i> -form:	-umai	-mai	...
	Imperative:	-e	-ro	...
	Prohibitive:	-una	-runa	...
	Conjunctive:	Infinitive:	-i	-Ø
<i>te</i> -form:		-Te	-te	-[a]nakute -[a]naide
<i>tari</i> -form:		-Tari	-tari	-[a]nakaotari
<i>tara</i> -form:		-Tara	-tara	-[a]nakaotara
<i>eba</i> -form:		-eba	-reba	-[a]nakereba
<i>zu</i> -form:		-azu	-zu	...
Derivatives:		<i>tai</i> -form:	-itai	-tai
	<i>sō</i> -form:	-isō-na	-sō-na	...
	Polite:	-imasu	-masu	...
	Negative:	-anai	-nai	...
	Potential:	-eru	-rareru (-reru)	...
	Passive:	-areru	-rareru	...
	Causative:	-aseru	-saseru	...

Verb inflection is agglutinative, e.g. *tabe-sase-rare-mashi-ta* ‘eat’-CAUS-PASS-POL-PAST or *tabe-sase-rare-na-kattara* ‘eat’-CAUS-PASS-NEG-*tara*. The sequence is causative – potential/passive/*-[i]tai/-[i]sō-na* – polite – negative – finite/conjunctive. The potential and passive do not co-occur, while causative-passive *-[s]ase-rareru* occurs, but not passive-causative. The monograde potential is formally identical to the passive, though current colloquial usage also has an analogical potential in *-reru*. Potential, passive and causative forms inflect as monograde verbs, and *-[i]tai*, *-[a]nai* and *-[i]sō-na* as adjectives. The polite form is irregular and defective (Table 10.10); its negative *-[i]masen* uses the polite copula in the past: *-[i]masen-deshita*. Quinquagrade verbs display sandhi with endings beginning with *t*, marked in Table 10.8 by ‘T’, illustrated in Table 10.9.

Irregular verbs include quinquagrade superpolite *gozaru* ‘to exist/be/have’, and honorific *iraqsharu* ‘to exist/go/come’, *kudasaru* ‘to give’, *nasaru* ‘to do’, *oqsharu* ‘to say’. Imperative and polite forms lose the *-r-*, while the imperative ends in *-i*, not *-e*, e.g. *nasaru* (honorific) > IMP *nasai*, POL *nasaimasu*. *iraqsharu* has alternative forms before *T*: *iraqshita* alongside *iraqshaota*, etc. Table 10.10 gives the inflection of other irregular verbs. The bound variant of *suru*, *-zuru*, occurs in formal language, but otherwise is now a regular monograde *-jiru*, e.g. *nenzuru* > *nenjiru* ‘to keep in mind’. There are several copulas: least marked *-da* is the same as the ending of *na*-adjectives (Table 10.6) except that the attributive is *-na* only before dummy noun *no*, and otherwise is *-no*; polite is *-desu*, (as in Table 10.7); formal style, typical of academic-type writing, is *-de # aru*, inflected just like *aru*; its polite equivalent, typical of formal speech such as speech-making, is *-de # arimasu*; and honorific and humble versions are *-de # iraqsharu* and *-de # gozaru*, inflected like *iraqsharu* and *gozaru* respectively. The negative of the last four often incorporates a ‘topic’ *-wa*, e.g. *-de-wa nai*.

**TABLE 10.9 JAPANESE VERB SANDHI**

		e.g.			
		Nonpast	Stem		Past
<i>w</i> -Stem:	w + T > qt	kau	kaw-	‘to buy’	kaqta
<i>r</i> -Stem:	r + T > qt	karu	kar-	‘to mow’	kaqta
<i>t</i> -Stem:	t + T > qt	katsu	kat-	‘to win’	kaqta
<i>m</i> -Stem:	m + T > nd	kamu	kam-	‘to bite’	kanda
<i>b</i> -Stem:	b + T > nd	yobu	yob-	‘to call’	yonda
<i>n</i> -Stem:	n + T > nd	shinu	shin-	‘to die’	shinda
<i>s</i> -Stem:	s + T > shit	dasu	das-	‘to take out’	dashita
<i>k</i> -Stem:	<i>iku</i> : k + T > qt	iku	ik-	‘to go’	iqta
	Other: k + T > it	shiku	shik-	‘to spread’	shiita
<i>g</i> -Stem:	g + T > id	kogu	kog-	‘to row’	koida

**TABLE 10.10 JAPANESE IRREGULAR VERBS**

		<i>suru</i>	<i>kuru</i>	<i>-[i]masu</i>
		‘to do’	‘to come’	Polite
Finite:	Nonpast:	suru	kuru	<i>-[i]masu</i>
	Past:	shita	kita	<i>-[i]mashita</i>
	<i>ō</i> -form:	shiyō	koyō	<i>-[i]mashō</i>
	<i>mai</i> -form:	shimai/surumai	komai/kurumai	...
	Imperative:	shiro	koi	<i>-[i]mase</i>
Conjunctive:	Prohibitive:	suruna	kuruna	RARE
	Infinitive	shi	ki	...
	<i>te</i> -form:	shite	kite	<i>(-[i]mashite)</i>
	<i>tari</i> -form:	shitari	kitari	RARE
	<i>tara</i> -form:	shitara	kitara	RARE
	<i>eba</i> -form:	sureba	kureba	RARE
	<i>zu</i> -form:	sezu	kozu	...
Derivatives:	<i>tai</i> -form:	shitai	kitai	...
	<i>sō</i> -form:	shisō-na	kisō-na	...
	Polite:	shimasu	kimasu	...
	Negative:	shinai	konai	<i>-[i]masen</i>
	Potential:	dekiru	korareru	...
	Passive:	sareru	korareru	...
	Causative:	saseru	kosaseru	...

10.4.2.3 Changes

Kōgo standardization resulted in various changes in inflection. Older inflection included several sub-systems, with, for example, their own conditional or concessive forms: a separate past sub-system, and two parallel negative systems (Table 10.11, based on Plaut (1905: 68, 76), Chamberlain (1907: 166–9), McGovern (1920: 117–37)). Negative (2.) *-nai* is likely cognate with EOJ *-nappye/nopye*. Originally uninflected, confusion in the spoken language with *nai* ‘not to exist’ resulted in it being inflected as an *i*-adjective. The process of standardization in education and literature produced the modern simpler system. The older system is shown in Table 10.11, using the quinquagrade verb *dasu*

TABLE 10.11 OLDER INFLECTION OF *DASU* 'TO TAKE OUT'

		Affirmative	Negative (1.)	Negative (2.)
Nonpast:	Finite:	dasu	<sup>P</sup> dasanu/ <sup>P</sup> dasan	dasanai
	Tentative:	dasō	dasumai	<sup>P</sup> dasanakarō
	Frequentative:	...	...	...
	Conditional:	daseba	<sup>P</sup> dasaneba	dasanakereba
	Concessive:	†dasedo(mo)	†dasanedo(mo)	†dasanakeredo(mo)
Past:	Finite:	dashita	†dasananda	dasanakaota
	Tentative:	<sup>P</sup> dashitarō	†dasanandarō	<sup>P</sup> dasanakaotarō
	Frequentative:	dashitari	†dasanandari	dasanakaotari
	Conditional:	dashitara(ba)	†dasanandara(ba)	dasanakaotara(ba)
	Concessive:	†dashitaredo(mo)	†dasanandaredo(mo)	†dasanakaotaredo(mo)
Infinitive:	dashi	...	dasanaku	
Gerund:	dashite	<sup>P</sup> dasazu	dasanakute	
Imperative:	dase	dasuna	...	

<sup>P</sup>: literary or somewhat dialectal. †: now largely obsolete

'to take out'. Forms marked with <sup>P</sup> are somewhat dialectal or highly literary; those marked with † are essentially obsolete.

### 10.4.3 Numerals and classifiers

#### 10.4.3.1 Numerals

Japanese uses NJ numbers (*hito*- '1', *futa*- '2', *mi*(*o*)- '3', *yo*(*o*/*N*)- '4', *itsu*- '5', *mu*(*o*)- '6', *nana*- '7', *ya*(*o*)- '8', *kokono*- '9', *tō*(-)- '10', *hats*-/*hatachi* '20', *iku*- 'how many'), SJ numbers (*ichi* '1', *ni*(*ji*) '2', *san* '3', *shi* '4', *go* '5', *roku* '6', *shichi* '7', *hachi* '8', *kyū*/*ku*- '9', *jū*/*ju*/*ji**o*- '10', *hyaku* '100', (-)*sen* '1,000', -*man* '10,000', -*oku* '100 million', -*chō* 'trillion', and nine higher terms; *nan*- (etymologically NJ!) 'how many', *sū*- 'several' and *ryō*- 'both'), and sporadic use of Anglo-Japanese numbers, especially '1' *wan*, '2' *tsū*, '3' *surī*. SJ '9' is normally *kyū*, only a few classifiers selecting *ku*-, e.g. *kuji* '9 o'clock', *kugatsu* 'September'. For '2', *ji*- is mostly limited to -*jō* '-th power', i.e. *jijō* 'squared'. Only the SJ set is complete. Compound numbers use no linking particle. Multiples of decimal places are [unit × decimal place] combinations, hence *jū* '10', *nijū* '20', *sanjū* '30', *sūjū* 'several tens', vs *jū ni* '12', *jū san* '13', *jū sū* '10 odd'. '1,000' is usually *iōsen* for '1,000', except in dates (*sen kyūhyaku kyūjūnen* '1990'), while '10,000' upwards require *ichi*-: *ichiman* '10,000' (not \**man*), etc.

NJ and SJ numbers generally combine with classifiers. Most classifiers are SJ-collocating, while only a few are NJ-collocating, but among SJ-collocating classifiers most combine not with *shi* '4' and *shichi* '7' but with native *yo*(*o*/*N*)- and *nana*-, with only a few collocating with *shi* and *shichi*. Two common classifiers (-*ji* 'o'clock' and -*jikan* 'hours') combine with *yo*- for '4' but *shichi* for '7'. The prevalence of *yo*(*N*)- over *shi* is attributed to homophony with *shi* 'death' (cf. *yoji* '4 o'clock' and *shiji* 'hour of death'); homophony with *ku* 'pain' may explain the prevalence of *kyū* over *ku* '9'. Moreover, most of the few NJ-collocating classifiers do not collocate with the full NJ set. -*ma* (rooms) combines only with NJ *hito*- and *futa*-, and thereafter with SJ numbers. -*ri* (people) combines only with *hito*- '1' and *futa*- '2', and thereafter is suppletively replaced by SJ numbers + -*niN*; -*ri* and -*niN* are written with the same

kanji. The only classifiers used with (almost) the full NJ set are *-tsu/Ø* (default classifier) and *-(u)ka* (days).

There are also specific forms used for *counting* in sequence (e.g. ‘1, 2, 3 . . .’ or ‘. . . 3, 2, 1’) or for *citing* telephone numbers, etc. digit-by-digit, and are bimoraic: *ichi, nī, san, shi* (counting) and *yōN* (citing), *gō, roku, nana, hachi, kyū*. ‘0’ in citing numbers is *maru/rei/zero*.

There have been changes during the nineteenth and twentieth centuries, e.g. before standardization *shi-* ‘4’ and *ku-* ‘9’ were more common (e.g. *shisen* ‘4,000’, *kusen* ‘9,000’, now *yōnsen, kyūsen*). Counting forms based on the NJ set (*hi, fū, mi, yō*, etc. (Chamberlain 1907: 101)) are now obsolete. *yōgtari* ‘4 (people)’ and *ikutari* ‘how many (people)’ were still used in the twentieth century for *yōnnin* and *nannin*.

#### 10.4.3.2 Classifiers

Japanese is rich in classifiers. The number + classifier combination forms a single word, e.g. *ichi + -hon > iipon*. Endocentric classifiers, e.g. time (*-nen* ‘years’, *-nichi/-(u)ka* ‘days’, *-byō* ‘seconds’), currencies (*-en* ‘yen’, *-doru* ‘dollar’), measurement (*-senchi* ‘cm’, *-mētoru* ‘metre’, *-ritoru* ‘litre’), and *-do* ‘degrees’, *-wari* ‘10 per cent’, *-pāsento* ‘per cent’, are their own referents, e.g. *toshi* = ‘year’, but *sannen* = ‘three years’, not \**sannen no toshi*. The number of exocentric classifiers in use depends heavily on register. The most formal writing will use many classifiers that are rarely used elsewhere, whereas everyday colloquial speech makes use of a relatively small set.

Unlike South-East Asian languages, then, flexibility in choice is more register-based than semantic. For example, *-tō* vs *-hiki* (animals) denotes *prototypically* large vs small, regardless of the specific referents’ actual size. There are few shape-based classifiers, namely *-hon* (long-thin) and *-mai* (sheet-like), and perhaps the default classifier *-ko* (round). *-hon* has multiple uses: extended objects; recordings; arrivals and departures.

Though SJ and loan classifiers collocate with SJ numbers and NJ classifiers with NJ numbers, there are numerous exceptions: NJ numbers (to 4) + SJ *-ban* (nights), SJ numbers + NJ *-wari* (10 per cent).

*Default classifiers:* Default *-ko* is restricted to tangible referents, while *-tsu/Ø* is used also with abstractions, and to mean ‘years old’, and is morphologically complex: ‘1–9’ = NJ + *-tsu*, ‘10’ = NJ + Ø, ‘20’ = NJ + Ø (age) and SJ + Ø (otherwise), and all other numbers = SJ + Ø (age, mostly): *hitotsu, futatsu, miōtsu, yōtsu, itsutsu, muōtsu, nanatsu, yaōtsu, kokonotsu, tō, hatachi* (age) and *nijū* (otherwise), etc.

*Day classifiers:* The endocentric classifier counting days (‘seven days’) or naming days (‘seventh [of the month]’) is a suppletive paradigm of irregular NJ *-(u)ka*, SJ (go’on) *-nichi* and SJ (kan’on) *-jitsu*, and uses an originally non-numeric term for ‘first day’. *-jitsu* is used after quasi-numbers *sū-* and *ryō-* only, *-(u)ka* after ‘two-ten’ and ‘twenty’, *-nichi* after other numbers and *nan-*; larger numbers ending in ‘four’, though otherwise SJ, use *yōka*: *tsuitachi* (‘first’) and *ichinichi* (‘one day’), *futsuka, miōka, yōka, itsuka, muika, nanoka, yōka, kokonoka, tōka, jūichinichi, jūninichi, jūsannichi, jūyōka, jūgonichi, jūrokunichi, jūshichinichi, jūhachinichi, jūkunichi, hatsuka*.

#### 10.4.3.3 Ordinals, fractions, decimals

Ordinal numbers consist of either *dai-* + SJ number or number–classifier combination + *-me*. Various classifiers are inherently ordinal and do not take *dai-* or *-me*, e.g. *-kai*

‘-th floor’. Fractions consist of numerator + *-bun* + *-no* + denominator, e.g. *sanbun-no ichi* ‘1/3’. Decimals are expressed as integer + *ten* ‘point’ + decimal places, the latter digit-by-digit, e.g. *jūsan ten gō gō* ‘13.55’.

## 10.5 SYNTAX

Japanese is an SOV language, and is often cited as a canonical verb-final or left-branching language: heads are rightmost within their phrases and clauses, and therefore all elements selected by these heads are to the left. For example, verbs are rightmost in verb phrases, verb phrases are rightmost in the clause, noun heads are rightmost in noun phrases, and postpositions are rightmost in postpositional phrases. Though a canonical SOV language, Japanese does display some exceptions: the positions of phrases other than the verb can be changed (known as ‘scrambling’), expressions of quantity modifying a noun may ‘float’ to the right of the entire noun phrase; internally headed relative clauses also occur; and ‘afterthoughts’ in colloquial speech may alter the SOV order of the sentence. Another noticeable syntactic feature is that Japanese is a ‘pro-drop’ language, but is not exactly the same as in Romance languages. In Japanese, sentential elements, including not only subjects but also objects and other grammatical functions, can be omitted as long as the missing elements are interpretable from their predicate forms or contexts.

### 10.5.1 Noun phrase structure

Noun phrase structure consists of modifier + head. Modifiers are headed either (1.) by the possessive/noun-modifying particle *-no*, or (2.) by an attributive verb or adjective. (1.) includes possession (*watashi-no kuruma* ‘my car’), a nominal attribute of the following noun (*eigo-no shinbun* ‘an English-language newspaper’), an appositive noun (*gakusei-no musuko* ‘my son who is a student’), modification by a postpositional phrase (*kōkō-kara-no shinkyū* ‘my best friend from high school’), quantification (*sannin-no gakusei* ‘three students’), and demonstratives (*sono hito* ‘that person’). Modifying clauses headed by a verb in (2.) are generally known as ‘relative clauses’, but in Japanese, modifiers headed by an adjective can be considered as relative clauses as well; in such clauses, tense morphology and its complement may appear with the adjective and form a clause (e.g. [*atama-ga ōki-kaōta*] *akachan* ‘the baby whose head was big’) just like a verbal relative clause. Here we consider both verbal and adjectival modifiers as relative clauses.

The past and – in most cases – the nonpast attributive forms are identical to the sentence-ending predicate forms in Modern Japanese. The nonpast exceptions are the copula (*-da* > *-no*), *na*-adjectives (*-da* > *-na*), *onaji* ‘the same’ (*onaji-da* > *onaji*), and *beki* ‘should’ (*beki-da* > *beki*). A few *i*-adjectives have *na*-adjective alternatives in attributive position only (see ‘*i/na*-adjectives’, above).

Not all predicates can form a noun-modifying clause. The volitional *ō*-form or the hearsay pattern V/A (dictionary form) + *sō-da* have no attributive equivalents. They are either avoided in noun modification, or require a linking marked by *to iu* (lit. ‘saying that . . .’).

Two alternative patterns are employed for noun-phrase structures with quantification and relative clauses. Quantification patterns are expressed as either quantifier + *-no* + head-noun + particle(s) (e.g. *ni-satsu-no hon-o (yomu)* ‘(will read) two books’), or head-noun + particle(s) + quantifier (e.g. *hon-o ni-satsu (yomu)* ‘(will read) two books’). The

meanings of the two patterns are more or less identical; however, one may find that the quantifier + *-no* + head-noun pattern is interpreted either with a collective or a distributive reading whereas the head-noun + particle(s) + quantifier pattern is interpreted distributively. The quantifier, which covers numeral + classifier combinations (e.g. *ni* ‘two’ and *-satsu* = a counter for books, magazines, etc.) as well as indefinite forms such as *takusan* ‘lots’, ‘floats’ off after the NP and any accompanying case particles.

The common type of relative clause is ‘externally’ headed, meaning that the modified noun is located right to the modifying clause. Relative clauses in colloquial or narrative usage may be internally headed as in CJ. For instance, unlike usual relative clauses where the modified head follows its relative clause (e.g. [*nure-ta*] *taoru-o kawakashita* ‘I dried the towel which was wet’), an internally headed relative clause contains its head inside the clause (e.g. [*taoru-ga nureta*] *no-o kawakashita* ‘I dried the towel which was wet’), the clause being marked by *no*.

Noun-modifying clause structures are the basis for a range of other constructions in Japanese: (1.) many subordinate clauses such as those marked by *ga/bedo* ‘but’, and *kara/node* ‘because’; (2.) all noun clauses where the relative clause modifies a dummy noun *koto* or *no* (distinct from the possessive particle *-no*); (3.) clauses inside various modal or quasi-modal predicates such as *n(o)-da*, *wake-da*, *mono-da* (explanatory) and *tsumori-da* ‘I intend to’.

## 10.5.2 Pronouns and anaphora

### 10.5.2.1 Personal pronouns

As shown in 10.4.1, the number and use of personal pronouns in Japanese are distinct from those in English. Japanese has a variety of personal pronouns because the language is more controlled by gender difference and speech level. Considering the use of pronouns at the syntactic level, it is said that the major and peripheral personal pronouns are not used like personal pronouns in English due to the ‘pro-drop’ nature of the language. In many cases where a personal pronoun must appear in English, it is normally elided in Japanese when identifiable from sentential or extra-sentential contexts. Kinship terms, workplace ranks or titles often function as ‘personal pronouns’.

Anaphoric personal pronouns or ‘reflexive’ pronouns include *jibun* ‘self’, *jibun-jishin* ‘self-self’, *kare/kanajo-jishin* ‘him-/herself’ and *otagai* ‘each other’. The syntactic behaviour of such pronouns is not exactly the same as that of English reflexive pronouns such as *my-/him-/herself* and *each other*; for instance, *jibun* is known as a long-distance anaphor since its referent does not have to appear within the same clause, unlike English reflexives. Moreover, it should be pointed out that whether or not those pronouns are indeed considered reflexive pronouns has been controversial in previous research in theoretical linguistics because they can be used logophorically.

### 10.5.2.2 Demonstrative words

The three demonstrative roots, *ko-* (proximal to the speaker), *so-* (distal from the speaker but proximal to the hearer) and *a-* (distal from the speaker and the hearer), have anaphoric (or discourse deixis) as well as deictic (or spatial deixis) uses. See their morphology in 10.4.1.

Cataphoric uses (referring to something upcoming in the discourse) are avoided for the *so-* and *a-* series. The unmarked choice for the anaphoric use is the *so-* demonstrative,

and this in turn forms the basis of a large number of sentence-initial conjunctions: *soshite* ‘then’, *sore-de* ‘that’s why, after that’, *sō suru-to* ‘then, after doing that’, etc.

### 10.5.2.3 Interrogative and indefinite pronouns

Interrogative pronouns are *dare* ‘who?’, *nani* ‘what?’, *naze* ‘why?’, *itsu* ‘when?’, *nan-* + classifier ‘how many?; which?’, and a series of *do*-forms parallel to the *ko/so/a*-demonstratives: *dore* ‘which (one)?’, *dono* ‘which?’, *dochira/dogochi* ‘which [of two]?; which way?’, *doko* ‘where?’, *dō* ‘how?’, *donna/dōiū/dōiōta* ‘what sort of?’, *dōyaote* ‘how?’ (colloquial), *dōshite* ‘why?’. *nani* reduces to *nan* before the copula or quotative *to*. Unlike English, those pronouns in interrogative sentences do not have to appear at the beginning of the sentence, but they may stay in their logical positions equivalent to those in declarative sentences. Therefore, Japanese is known as a so-called ‘*wh*-in-situ language’.

Indefinite pronouns consist of interrogative pronoun + *-ka* (question marker). Some instances are *dareka* ‘some/anyone’, *nanika* ‘some/anything’, *itsuka* ‘some time’, *nan-* + classifier + *-ka* ‘some number/amount of’, *doreka* ‘some of those’, *dokoka* ‘somewhere’ and so forth. *-ka* does not attach to *dono* ‘which’, however, and *nanika* can reduce to *nanka* in colloquial speech.

## 10.5.3 The basic sentence

Japanese is a topic-oriented language, where many sentences are structured Topic (marked by *-wa*) + Comment. The topic normally occurs at the beginning of the sentence, and the comment can include a full sentence where the word order in between is relatively free as long as a main verbal/adjectival/nominal + copula element comes at the end, unmarked order being subject > time expression > other adverbials > object-type element. ‘Object-type element’ includes a range of noun phrases that belong in the subcategorization frame of the verb, including objects marked by *-o*, direction/destination phrases marked by *-ni/-e/-made* with verbs of motion, and ‘*ga*-marked’ objects of adjectives of emotion or wish or of stative verbs: *wakaru* ‘to understand’, *iru* ‘to need’, *aru/iru* ‘to be, exist; have’, *dekiru* ‘can do’.

SOV word order is the default in prescriptive grammar. Informal conversation is characterized by frequent right-movement of phrases, usually subject or object (with or without particles) as shown in (3) and (4), which create apparent OVS or SVO orders. Other phrases such as postpositional phrases can also be moved rightward (see (5)).

- (3) *kawaō-teru ne, ano hito-mo*  
 odd-PROG SFP that person-also  
 ‘That person is an odd fish, isn’t (s)he?’
- (4) *shiōkari mi-te.kudasai, kachō-o*  
 well watch-please section.chief-ACC  
 ‘Please keep a close eye on the section chief.’
- (5) *yoku deki-mashita ne, ano awatadashisa-no naka-de*  
 well can.do-POL.PAST SFP that flurry-GEN middle-in  
 ‘You did well even in a whirl of activity!’

Such right-movement is marked by *pausa* and intonation break after the main predicate, which in writing is indicated invariably by a comma, identifying such structures as marked. The right-moved element functions either as emphasized, or merely as an afterthought.

The basic sentence can be followed by a range of sentence-final particles: (a.) question markers such as *ka*, *no*,  $\emptyset$  (+ rising intonation), (b.) epistemic modal markers such as *kana*/(female speech) *kashira* (uncertainty), *ote* (hearsay), *oke* (recollection); (c.) interactional markers such as *ne* (inviting the hearer's uptake, confirmation), *yo* (presenting the speaker's conviction) and other similar markers such as *wa* (female speech), *zo*, *ze* or *sa* (male speech).

### 10.5.3.1 Declarative

Declarative sentences are unmarked in the sense of the word order of the basic sentence structure. Verbals may be transitive or intransitive. The object of a durative/punctual verbal is mostly marked by *-o*, but the object of a stative verbal is marked by *-ga*. This in theory means that a stative verbal can have two participants (i.e. subject and object) marked by *-ga*, and the construction is indeed known as the double-subject construction, and there is much discussion in linguistic literature as to which is the underlying subject and which is the underlying object (or both are subjects). In practice, however, two *-gas* in a simple, mono-clausal sentence may sound awkward and rare, as speakers tend to mark the first with *-wa* or *-ni* instead to avoid two *-gas* or express the grammatical relations of the two explicitly.

A subtype of such transitive statives is that of adjectives of emotion/desire or experience, e.g. *ureshii* 'to be happy', *-[i]tai* 'to want to . . .', *kowai* 'to be afraid', *tanoshii* 'to enjoy'. These use the 'double-subject' construction to mark both experiencer and cause. These adjectives are distinctive in that by themselves, they are interpreted only as describing a sensation of which the speaker can be directly certain, and so are used in statements only of a first-person experiencer and in questions only of a second-person experiencer. To express another person's sensation, a derivative verb is formed. In most cases this is root + *-garu*, e.g. *kowai* > *kowagaru*; *ureshii*, however, is suppletive and *yorokobu* is used rather than *ureshigaru* for usual/neutral meaning. The derivative is not a stative verb, and so the cause of the sensation is marked by *-o* rather than by *-ga*, and in practice the verb appears usually in its progressive form (*kowagaoteiru*, *yorokondeiru*) to describe the present emotion of the third party. Alternatively, when the speaker's knowledge of the sensation is based on what the experiencer has told him/her, adjectival clause + quotation structure (e.g. *to ioteita* '(s)he said that') or hearsay markers (*sō-da*, *ote*) are found.

- (6) a. first person: *inu-ga kowai*  
dog-NOM be.afraid 'I'm afraid of dogs'
- b. third person: *inu-o kowagaoteiru* '(s)he's afraid of dogs'  
*inu-ga kowai sō-da* '(s)he's afraid of dogs (and told me so)'  
*inu-ga kowai to ioteita* '(s)he told me (s)he's afraid of dogs'

Noun and *na*-adjective predicates, both of identification and description, make use of the copula. Whenever change of state is intended, the copula is replaced by *-ni* (or *-to*) + *naru* 'become'. Note the difference between:

- (7) *kachō-da*  
section.chief-COP '[I/you/someone else] is section chief.'
- (8) *kachō-ni naru*  
section.chief-into become '[I/you/someone else] becomes section chief.'

When the predicates above are followed by *-[i]tai* ‘to want to’, the form of the copula *-da* above changes to *-dearu(u)*.

- (9) *kachō-deari-tai*  
 section.chief-COP-want ‘I want to remain section chief.’
- (10) *kachō-ni nari-tai*  
 section.chief-into become-want ‘I want to be section chief.’

Location and existential predicates make use of the verbs *aru* and *iru* ‘to be, exist, have’. These verbs can express possession too, with the possessed marked by *-ga* and the possessor marked by *-ni* (or *-(ni)wa*). The choice of verb is largely determined by the animacy of the subject: animate subjects collocate with *iru*, inanimate ones with *aru*. The major exceptions are kin-terms in existential statements in formal contexts, or the starts of folk- or children’s tales, where *aru* is common despite the animacy of the subject:

- (11) *go-kyōdai-ga ari-masu ka*  
 HON-siblings-NOM be-POL Q  
 ‘Have you any brothers or sisters?’
- (12) *mukashi mukashi binbōna o-samurai-ga aq-ta*  
 long.ago long.ago poor HON-samurai-NOM be-PAST  
 ‘Once upon a time there was a poor samurai.’

Word order is influenced by the definiteness of the subject or the topic phrase marked by *-wa* in location/existential sentences. When the subject is definite, the order is Topic (Subject) – Location – *aru/iru* as shown in (13) whereas when the subject is indefinite, the order can be Location – Subject – *aru/iru* as in (14).

- (13) *kagi-wa hikidashi-no naka-ni ari-masu*  
 key-TOP drawer-GEN inside-in be-POL  
 ‘The key is in the drawer.’
- (14) *hikidashi-no naka-ni kagi-ga ari-masu*  
 drawer-GEN inside-in key-NOM be-POL  
 ‘There is a key in the drawer.’

For other types of declarative sentences, see the following subsections.

### 10.5.3.2 Tense and aspect

The main markers of tense are the nonpast and the past. The nonpast form expresses present and future meaning, depending on the predicate to which it attaches or the context. The major aspect marker is the progressive, *-teiru*, derived from the *te*-form + auxiliary *iru*. In spoken language, even in fairly formal speech, the root /i/ of *iru* is dropped: *kawaqteiru* ~ *kawaqteru* ‘is unusual, has changed’, *shiteita* ~ *shiteta* ‘was doing’, *haiqteinai* ~ *haiqtenai* ‘has not entered yet; there’s nothing in there’.

The progressive expresses continuation in the present (*-te(i)ru*) or in the past (*-te(i)ta*). It cannot be used for future reference unless it is followed by some modal marker to express the speaker’s conjecture, and so with most verbs it helps distinguish present from future: *hanas-u* (speak-NONPAST) ‘I will speak’ (future) or ‘I speak’ (habitual) vs *hanashi-teiru* (speak-PROG) ‘I’m speaking’ (progressive). The progressive construction is incompatible with most stative verbs; however, *wakaru* ‘to understand’ may form a

progressive with the meaning that durative verbs have, and *mieru* ‘can see’ and *dekiru* ‘can do’ can go with *-te(i)ru* to indicate a temporal state at the time of speech. The progressive of the copula is extremely specialized and such *-iru* in *-deiru* does not really change the aspect but could be considered as an existential verb (e.g. *heiki-de iru* ‘to be/stay self-possessed,’ *gakusei-no mama-de iru* ‘to keep being a student’). Durative and punctual verbs, on the other hand, all regularly form the progressive, V + *-te(i)ru*, but because of the nature of the two types of verb, the construction expresses ongoing action with duratives but continuing resultant state with punctuals, as shown in (15) and (16).

(15) *inu-ga hashiq-teiru*  
dog-NOM run-PROG ‘The dog is running.’

(16) *mado-ga koware-teiru*  
window-NOM break-PROG ‘The glass is broken.’

The simple past form expresses both action in the past and resultant state with duratives in relative clauses such as *kiQ-ta pan* ‘sliced bread’ and *yude-ta tamago* ‘boiled egg’. Punctuals use the simple past form instead of the progressive, either in relative clauses or for effect at the end of sentences, *nureteiru ~ nureta* ‘it’s wet’, to express the resultant state. With verbs of putting on or removing, their progressives by themselves express resulting state (*kiteiru* ‘is wearing’), but combining with *tokoro* + copula to express continuing action (*kiteiru tokoro-da* ‘is putting on’) or with adverbials (e.g. *ima* ‘now’) they can indicate an ongoing action clearly.

There are some other aspectual markers; one is the aforementioned *tokoro* + copula. This is used also to express an immediate future or immediate past event, the latter interchangeable with *bakari* + copula: *yom-u tokoro-da* ‘is about to read’, *yON-da tokoro-da* or *yON-da bakari-da* ‘has just read’.

Other broadly aspectual markers take the form of *-te* + auxiliary verb. *-te shimau* indicates that an action is complete and typically irrevocable, at least in the short term. As a result, it is normal when dealing with unfortunate events that cannot be easily rectified, e.g. breaking something important. *-te shimau* can also be used of actions completed in a neutral sense (e.g. *shukudai-o yaQ-te shimau* ‘to finish doing homework’). *-te oku* indicates advance preparation for a logical follow-up action at a later point (e.g. *shukudai-o yaQ-te oku* ‘to complete homework in advance’). *-te iku/kuru* basically expresses a directed action; if the action takes place away from the speaker, *-te iku* is used, whereas when the action is towards the speaker, *-te kuru* is adopted. Extending these directional meanings, *-te iku/kuru* can function as aspectual markers to show change of state or event, e.g. *kumoQ-te ki-ta* ‘It’s getting cloudy’ and *nakunaQ-te iQ-ta* ‘It was disappearing’.

Two more *-te* + auxiliary verb forms are found in Japanese: *-te aru* and *-te miru*. *-te aru* could be regarded as an aspectual expression since it may create a durative interpretation (e.g. *mado-ga ake-te aru* ‘the window was opened and it is still open’); however, this will be discussed more in relation to voice in 10.5.5. *-te miru* is less aspectual in the sense that it does not necessarily show an internal temporal contour; it merely indicates trying out (e.g. *yaQ-te miru* ‘I’ll try doing it’).

The auxiliary verbs above have independent lexical meaning: *shimau* ‘to finish’, *oku* ‘to put’, *iku/kuru* ‘to go/come’, *aru* ‘to be, exist’ and *miru* ‘to see’. But their grammaticalizations in such combinations are seen in their tendency to adopt inflectional-style accent and for *-te shimau* and *-te oku* to contract to (*-chimau >*) *-chau* and *-toku* in casual speech.

Consequently, the grammaticalized form has come to be written in kana alone and not in kanji in post-war writing.

Other instances of aspectual expression are experiential forms expressed through a tense form + *koto-ga aru* + tense form. The combination of two tense slots within the construction allows the expression of a wide range of experiential meanings: *oyogu koto-ga aru* ‘I sometimes swim’, *oyoida koto-ga aru* ‘I have swum (in the past)’, *oyogu koto-ga aqta* ‘I sometimes used to swim’, *oyogu koto-wa nai* ‘I never swim’, *oyoida koto-wa nai* ‘I have never swum’, *oyoganai koto-ga aru* ‘sometimes I don’t swim’, etc.

### 10.5.3.3 Negation

Negation is expressed morphologically, as described above in section 10.4. Negation of *aru* ‘be (located), have’ is suppletive, i.e. *nai* (not *\*aranai*), though in highly formal registers *zu-* and *mai-*forms of *aru* occur, regular in formation (*arazu*, *arumai*). The copula + negation, *-janai*, is ultimately a contraction of *-de-wa nai*, and is used frequently for noun + copula forms and *na-*adjectives in casual speech. Negative main clauses frequently attract a topic marker to the element in question.

Japanese has negative polarity items, such as *-shika* + *-nai* ‘only’, *zenzen/maqtaku* + *-nai* ‘not at all’, *meqtani* + *-nai* ‘rarely’ and *kanarazushimo* + *-nai* ‘not necessarily’. These items preceding *-nai* require a negative expression, while it should be noted that in recent years *zenzen* is sometimes used with an affirmative predicate (e.g. *zenzen daijōbu-da* ‘(I) am totally fine’). When a *wh*-word or the numeral expression *ichi* ‘one’ + classifier is marked by *-mo* ‘also’ and is followed by a negative predicate, it functions as a negative polarity item too.

- (17) *dare-mo ko-nai*  
 who-also come-NEG ‘Nobody comes.’
- (18) *ichi-do-mo ik-anai*  
 one-time-also go-NEG ‘(I/you/someone) don’t go at all.’

### 10.5.3.4 Modality

Roughly speaking, modality expressions in Japanese can be divided into two types: modality of judgement, which are realized by epistemic and deontic (or evaluative) modal expressions, and modality of utterance (Masuoka 2009). The modality of judgement is expressed largely by verb-following expressions that include an auxiliary verb or adjective. Typical markers of epistemic and deontic (or evaluative) modality are given in Table 10.12. Subscript <sub>τ</sub> indicates which item undergoes inflection for tense/aspect; subscript <sub>p</sub> indicates which item is marked for politeness.

Epistemic expressions appear with verbs, adjectives or noun + copula constructions. With the latter, the form that the copula takes varies:  $\emptyset$  before *ka-mo shirenai* or *rashii*, *-no* before *yō-da*, *-da* before hearsay *sō-da* or *qte*, etc. *rashii*<sub>1</sub> preceded by noun (+ copula) is similar to denominal derivative *-rashii*<sub>2</sub>, but the two uses have different morphology, particularly in negation: *onna- $\emptyset$  rashii*<sub>1</sub> ‘(that person) looks a female’ (I know she is) > *onna-janai rashii* ‘(that person) doesn’t look a female’, vs *onna-rashii*<sub>2</sub> ‘she’s feminine’ > *onna-rashiku nai* ‘she’s not feminine’. The pitch accent patterns are different too; each of the preceding noun and *rashii*<sub>1</sub> has its own accent whereas the preceding noun + *-rashii*<sub>2</sub> is considered as one phonological unit with one accent. Noun-*no yō-da* can be replaced in informal language with Noun- $\emptyset$  *mitai-da* or (lexically restricted)

**TABLE 10.12 JAPANESE EPISTEMIC AND DEONTIC/EVALUATIVE MODAL EXPRESSIONS (FOR YOMU ‘TO READ’)**

Epistemic:	‘perhaps’:	yomu <sub>T</sub> ka-mo shirenai <sub>P</sub>
	‘probably’:	yomu <sub>T</sub> -darō <sub>P</sub>
	‘certainly’:	yomu <sub>T</sub> ni chigai nai <sub>P</sub>
	Deduction:	yomu <sub>T</sub> hazu-da <sub>P</sub>
	Evidential:	(a.) yomu <sub>T</sub> yō-da <sub>P</sub>
	(b.) yomu <sub>T</sub> rashii <sub>P</sub>	
	(c.) yomi-sō-da <sub>T/P</sub>	
	(d.) yomu <sub>T</sub> sō-da <sub>P</sub> (hearsay)	
Deontic:	‘have to’:	yomanakereba naranai <sub>T/P</sub> (etc.)
	‘don’t have to’:	yomanakute-mo ii <sub>T/P</sub> (etc.)
	‘can/may’:	yonde-mo ii <sub>T/P</sub> (etc.)
	‘mustn’t’:	yonde-wa naranai <sub>T/P</sub> (etc.)

Noun-*opoi*: *kankokujin-Ø mitai-da* ‘(s)he looks Korean’ (I don’t know/I know she isn’t), *onna-opoi* ‘he’s effeminate, she’s feminine’.

The meanings or functions of evidential modal expressions are similar but not the same. *yō-da* in (a.) in Table 10.12 is used to show that the event or state is presented based on the speaker’s own observation or the speaker’s conjecture on the basis of someone else’s observation. It expresses something the truth of which is not known or is counterfactual. *rashii* in (b.) expresses something known to be the case from the speaker’s point of view, where the speaker’s judgement is usually based on his/her own observation or on information obtained from someone else. The *-sō* form in (c.) expresses something that the speaker is relatively certain of based on what (s)he can see (or hear, smell where relevant): with verbs something in the imminent future (*ame-ga furi-sō-da* ‘it looks like it’s going to rain’); with adjectives the current state (*oishi-sō-da* ‘it looks tasty’). The hearsay *sō-da* in (d.) specifically indicates that the source of information comes from someone other than the speaker him/herself.

Epistemic modals may be pre-empted by adverbs; for instance, *tabun* or *kioto* are used to pre-empt *darō*. These adverbs only indicate what is to come, and do not normally constitute epistemic expressions in their own right unless uttered in casual speech or in the case of affirmative (zero-marked) predicates which concord with adverbs such as *kanarazu* ‘surely’.

Deontic modals in Japanese are different from those in English. It is common that a deontic modal is followed by an epistemic modal expression, such as *ka-mo shirenai* or *-darō*. Deontic expressions are generally conditional expressions combining *-te-wa* ‘if’, *-nakute-wa* ~ *-nakereba* ‘if . . . not’ and *-te-mo* ‘even if’ with *naranai* ‘it won’t do’ (or synonyms *ikenai*, *dame-da*) or *ii* ‘it’s all right’ (or synonym *kamawanai*). *yomanakute-wa naranai*, therefore, literally means ‘it won’t do if I don’t read it’, i.e. ‘I have to read it’. Colloquial language uses a contracted form of *-nakute-wa* without a following *naranai*, i.e. *-nakya*, as a shorter alternative.

Earlier modal adjective *beshi* exists in Modern Japanese, but in all but the most formal use exists only as uninflected *beki*, and in this use expresses ‘should, ought to’. When *suru* ‘to do’ comes before the modal, it usually takes its CJ conclusive form, i.e. *su*. *beki* is particularly common in fixed expressions equivalent to ‘-able’, in which monograde verbs tend to take their classical forms, e.g. *osoru-beki* ‘terrible (cf. *osoreru*, CJ *osore-* ‘to be afraid’); however, these are best treated as derivatives rather than as

modal expressions. Very formal language preserves the CJ inflection of *beshi*, especially the negative *bekarazu* ‘can’t’, but this expression can be seen in a limited context only (e.g. a written notice to show some prohibition).

In addition to the epistemic and deontic/evaluative modal expressions, Japanese demonstrates a rich system of modality of utterance expressions. It has been controversial whether or not these expressions should be regarded as modality items; however, the majority of ‘*modariti*’ research in Japanese considers them in modality studies. Typical examples, which are found as sentence-final particles, are listed in Table 10.13.

These expressions typically appear at the end of the sentence in spoken Japanese. They are known as interactional discourse markers and play an important role in co-constructing an ongoing conversation.

There are also some modality items called explanatory modality. They include *no-da*, *wake-da*, *mono-da* and *koto-da*, of which *no-da* ‘it is that’ is frequently used and is known as a typical instance that contains both the functions of modality of judgement and modality of utterance. In colloquial speech, *o* in *no-da* is often dropped. The explanatory modality items basically show the hearer how the sentence marked by an explanatory modal is related to the previous discourse and also indicate that the speaker takes in the cause or corollary of the event/state uttered or perceived in the previous discourse. Due to this basic function, *no-da*, for instance, is used when the speaker wants the hearer to be aware of a fact unknown by the hearer or when the speaker takes in a fact unknown to the speaker him/herself. Some examples with *no-da* are shown below.

- (19) *ashita-wa keqseki-shi-masu. byōin-ni iku n.desu*  
 tomorrow-TOP absence-do-POL hospital-to go EXP.MOD.POL  
 ‘I’ll be absent tomorrow. I’ll go to a hospital.’
- (20) (when someone doesn’t show up)  
*kioto michi-ga konderu n.da*  
 more.than.likely roads-NOM be.crowded EXP.MOD  
 ‘I think the roads are congested.’
- (21) (when the speaker realized that the hearer misunderstood something)  
*iie, sō-janai n.desu*  
 no so-COP.NEG EXP.MOD.POL  
 ‘No, that’s not the case.’
- (22) (when the speaker realized that his/her friend had gone)  
*mō kono kuni-o de-ta n.da*  
 already this country-ACC leave-PAST EXP.MOD  
 ‘(S)he had already left this country.’

**TABLE 10.13 JAPANESE MODALITY OF UTTERANCE EXPRESSIONS**

Conveyance:	<i>yo</i>
(male)	<i>zo, ze, sa</i>
(female)	<i>wa</i>
Confirmation, exclamatory:	<i>ne(e), yone, na(a)</i>
Justification:	<i>mon(o)</i>
Recollection:	<i>oke</i>

## 10.5.3.5 Non-declarative sentence types

Questions of all types including *wh*-questions and *yes/no* questions are expressed by a change of sentence intonation and the addition of a particle at the end of the sentence: familiar-style *-no*, *-Ø* or polite-style *-ka*. Soliloquy-type questions that do not necessarily require a response from a hearer use *-kana(a)* or *-kashira* (female). There is no change in word order, and no fronting in *wh*-word questions. Questions that present alternatives, including those that ask a question in an affirmative form with a negative alternative of the same situation, have the structure of two separate questions. The second part of the question tends to include the verbal phrase, or a copular phrase to avoid repeating the same verbal phrase as in the first part of the question.

- (23) *rondon-e iki-mashita ka, (soretomo) iki-masendeshita ka*  
 London-to go-POL.PAST Q or go-POL.NEG.PAST Q  
 ‘Did you go to London (or not)?’

- (24) *kono hon-o yomi-mashita ka (soretomo) ano hon-desu ka*  
 this book-ACC read-POL.PAST Q or that book-COP.POL Q  
 ‘Did you read this book or that book?’

Affirmative-leading questions to ask for the hearer’s agreement use a negative verbal, while purely non-leading questions are expressed as an affirmative question or as alternative questions, typically the first with the verbal in the affirmative, the second with it in the negative.

- (25) *kirei-jaarimaseN ka*  
 pretty-COP.POL.NEG Q ‘Isn’t it pretty?’

- (26) *kirei-desu ka*  
 pretty-COP.POL Q ‘Is it pretty?’

- (27) *kirei-desu ka, kirei-jaarimaseN ka*  
 pretty-COP.POL Q pretty-COP.POL.NEG Q ‘Is it pretty (or not)?’

Responses are normally expressed by reduced sentences. *Yes/no* question responses are *hai/ē/un* + repeated verbal (agreement) and *iie/ya/uun* + repeated verbal (disagreement); when the *yes/no* question has the structure of Noun + copula, the answer can be *hai/ē/un* + *sō-desu* for positive and *iie/ya/uun* + *sō-jaarimaseN*. Question-word responses either repeat the verbal of the original or use a copula to replace the verbal of the original. It should be noted that the choice of *yes* or *no* in Japanese is dependent on whether the speaker agrees with the propositional content (i.e. [*nani-mo yom-anakaqta*] in (28) below) as a whole in question.

- (28) Q: *nani.mo yom-anakaqta-N.desu ka*  
 anything read-NEG.PAST-EXP.MOD.POL Q  
 ‘Didn’t you read anything?’ (lit. ‘Is it the case that you did not read anything?’)

- (29) A: *ee, yomi-masendeshita / yom-anakaqta-N.desu*  
 yes read-POL.NEG.PAST read-NEG.PAST-EXP.MOD  
 ‘No, I didn’t.’ (lit. ‘Yes, you are right.’)

- A: *iie, yomi-mashita / yon-da-N.desu*  
 no read-POL.PAST read-PAST-EXP.MOD  
 ‘Yes, I did.’ (lit. ‘No, you are not right.’)

The closest equivalent to tag questions in Western languages are sentences with a sentence-final particle such as *ne* or *yone*.

- (30) *tanaka-san-desu (yo)ne*  
Tanaka-Mr-COP SFP ‘You are Mr Tanaka, aren’t you?’

Other expression types such as commands and requests are expressed by a range of devices. Common examples are summarized in Table 10.14. Subjects of the sentences with the expression types in Table 10.14 are normally omitted since they are interpretable from the form of the predicate and context. The choice of expression depends on the level of speech. See 10.5.6 for details and examples of different levels.

Exclamations are expressed by various sentence patterns with distinct intonations: (1.) by ending with a noun, (2.) with *nante/nanto* ‘how!’, (3.) by ending with *N-darō/-deshō* or *V/A n(o)-darō/-deshō*, (4.) by ending with *-to wa*, and combinations of (1.)–(4.).

- (31) *sutekina kutsu!*  
nice shoes ‘Nice shoes!’
- (32) *nante oishii sakana-dar-ō!*  
how delicious fish-COP-ō ‘How delicious this fish is!’
- (33) *kono ie, nante hiroi-N.desh-ō!*  
this house how spacious-EXP.MOD.POL-ō ‘How spacious this house is!’
- (34) *koko-ni aq-ta to-wa (shir-anakaqta)!*  
here-in exist-PAST QUOT-TOP know-NEG.PAST  
‘(I didn’t know) it was here!’

#### 10.5.4 Topic, focus and emphasis

Japanese is a topic-oriented language, and explicitly marks the topic of a sentence with a pragmatic particle. When the topic is a word or concept for which the following comment provides a definition, *to iu no-wa* or (colloquially) *-qte* can mark the topic; otherwise the topic is directly marked by *-wa* or *-mo* ‘also’.

*-wa* has two basic functions: (1.) it marks backgrounds or the theme of the utterance, which typically appears at the beginning of the sentence; and (2.) it marks a contrast

**TABLE 10.14** EXAMPLES OF OTHER JAPANESE EXPRESSION TYPES USING *TORU* ‘TO TAKE’

Imperative (Do . . .)	<i>tore, tori-nasai</i>
Request (Please . . .)	<i>toqte kure, toqte kudasai, toqte kurenai/moraenai ka, toqte kureru/moraeru ka, toqte, toqte hoshii, toqte moraitai</i>
Cohortative (Let’s . . . , Shall I/we . . .)	<i>torō (ka), tori-mashō (ka), toranai ka, tori-masen ka</i>
Prohibition (Don’t . . .)	<i>toruna</i>
Suggestion (You should . . . , why don’t you . . .)	<i>toqtara (dō (desu) ka), toru-to ii, toqta hō-ga ii</i>
Intention (I will . . . , I intend to . . .)	<i>torō, toru, toru tsumori-da; torumai (negative intention)</i>

with or without phonological emphasis on the contrastive item(s), and typically two items will be contrasted with each other with a *-wa*, e.g. in separate clauses linked by *ga/keredo(mo)* ‘but’.

- (35) *ikeda-san-wa daigaku-no sensei-desu*  
Ikeda-Mrs-TOP university-GEN teacher-COP.POL  
‘Speaking of Mrs Ikeda, she is a university teacher.’
- (36) *tempura-wa tabe-masu ga, sashimi-wa tabe-masen*  
tempura-TOP eat-POL but raw.fish-TOP eat-POL.NEG  
‘I eat tempura but not sushi.’

Thematic topic phrases are normally limited to one occurrence per sentence whereas contrastive topic phrases may appear more than once in a sentence. The thematic topic can be omitted when it is identifiable from the context.

Focus is achieved in Modern Japanese predominantly by changing the construction, particularly a cleft construction: Presupposition + *no(-wa)* + Focus + Copula, movement (with or without phonological emphasis), or delimiting particles such as *-dake* and *-shika* (. . .*-nai*). Examples of cleft constructions are shown in (37) and (38) (focused constituents are underlined).

- (37) *watashi-ga mi-ta no-wa kono eiga-desu*  
I-NOM watch-PAST NMR-TOP this film-COP.POL  
‘What I watched was this film.’
- (38) *yamada-san-ga uchi-e ki-ta no-wa kyonen-no*  
Yamada-Mr-NOM my.house-to come-PAST NMR-TOP last.year-GEN  
*fuyu-deshita*  
winter-COP.POL.PAST  
‘It was last winter that Mr Yamada visited us.’

Moving an element (or not moving it but placing some phonological emphasis on it) may present a focus function, as illustrated in (39) and (40).

- (39) *kono eiga-o, watashi-wa mi-mashita*  
this film-ACC I-TOP watch-POL.PAST  
‘This film, I watched (it).’
- (40) *kyonen-no fuyu-ni, yamada-san-ga uchi-e ki-mashita*  
last.year-GEN winter-in Yamada-Mr-NOM my.house-to come-POL.PAST  
‘Last winter, Mr Yamada came to our house.’

Some of the delimiting particles also express focus, as shown in (41) and (42).

- (41) *kono hito-dake-ga iki-noko<sub>o</sub>-ta*  
this person-only-NOM live-remain-PAST  
‘Only this person survived.’
- (42) *watashi-wa nihongo-shika hanas-e-masen*  
I-TOP Japanese-only speak-can-POL.NEG  
‘I can speak only Japanese.’

There are other particles that do not necessarily express focus but are used as emphatic markers. These include *-sae(mo)* ‘even’, *-mo* ‘also, even’, *-sura* ‘even’, and *-made(mo)* ‘even’. Very formal language preserves the earlier focal particle *-koso*; however, *-koso*

also is now often more a marker of emphasis than of focus. Focal function is maintained in certain set phrases, e.g. *kochira-koso* '[It is] I [that should thank you]' (standard response to thanks when both parties have profited).

### 10.5.5 Passive and causative

The passive construction has two distinct participant-changing roles: (1.) the logical object, which is usually the object marked by *-o* (or *-ni*), becomes the subject of the passive (direct passives); and (2.) someone who is affected by the action (e.g. the owner of the logical object) becomes the subject of the passive ('adversative' or indirect passives).

The agent of direct passives (the original subject) is either marked by *-ni*, slightly less commonly *-ni yoote* (formal, translatese) or *-kara* (especially to avoid any ambiguity of *-ni*), or is unexpressed. The verb is marked by *-[r]are-*.

- (43) *musuko-wa tomodachi-ni tatak-are-ta*  
son-TOP friend-by hit-PASS-PAST  
'My son was hit by his friend.'
- (44) *seika-ga daihyō sēnshu-niyoote tenkas-are-ta*  
sacred.flame-NOM rep. athlete-by light-PASS-PAST  
'The Olympic Flame has been lit by the representative athlete.'
- (45) *kore-wa haha-kara watashi-ni tewatas-are-ta tegami-desu*  
this-TOP mother-from I-to hand-PASS-PAST letter-COP.POL  
'This is the letter that I was given by my mother.'

Unlike direct passives, which must take a transitive verb, indirect passives can in theory involve any non-stative verb including intransitives. However, they are largely formed from transitive verbs, and indirect passives with an intransitive verb are not actually very many; they commonly use verbs such as *furu* 'to precipitate', *kuru* 'to come', and *shinu* 'to die'. The agent must normally be explicit in indirect passives.

- (46) *ame-ni fur-are-ta*  
rain-by fall-PASS-PAST  
'I got rained on.'
- (47) *oya-ni shin-are-ta*  
parent-by die-PASS-PAST  
'My parents died (and this was bad for me).'
- (48) *watashi-wa oqto-ni henna uta-o utaw-are-te*  
I-TOP husband-by strange song-ACC sing-PASS-*te*  
*komaQ-teiru*  
be.troubled-PROG  
'I have been adversely affected by my husband's singing a strange song.'

There are two other constructions that do not use the passive morpheme but which have some broadly 'passive' meaning. The first is the benefactive construction, *-te + morau*, which shows that the subject receives some benefit from the action made by a person expressed by the *-ni* phrase.

- (49) *watashi-wa imōto-ni mise-e iQ-te moraQ-ta*  
I-TOP younger.sister-by shop-to go-*te* receive-PAST  
'I had my younger sister go to the shop.'

- (50) *otōto-wa tomodachi-ni shokuji-o tsukuQ-te moraQ-teiru*  
 younger.brother-TOP friend-by meal-ACC make-*te* receive-PROG  
 ‘My younger brother has his friend cook meals.’

The second is the *-te aru* construction, which indicates a resultant state of someone’s action. The grammatical object becomes the subject of the *-te aru* construction. Whereas the agent of a passive form is understood to exist, even if it is not identified or identifiable in context, an agent cannot even be implicit with *-te aru*; the construction implies the identity of an agent to be irrelevant (though logically all such actions must have one as the verb attached to *-te aru* is transitive).

- (51) *mado-ga shime-te aru*  
 window-NOM close-*te* exist ‘The window is closed.’
- (52) *akari-ga keshi-te aru*  
 light-NOM turn.off-*te* exist ‘The light is off.’

Causative constructions, marked by *-[s]ase-* in the verb, are used to add a participant, the causer, which appears as the subject. The original subject is then marked either by *-o* or *-ni*. If the verb already has an object marked by *-o* (i.e. transitive), the particle of the original subject or the causee must be *-ni* since Japanese does not allow more than one *-o* in positions close to each other. Otherwise, the choice between *-o* and *-ni* for intransitive verbs in causative constructions is influenced by pragmatic and semantic factors.

- (53) *kōchi-wa kodomo-tachi-o oyog-ase-ta*  
 coach-TOP child-PL-ACC swim-CAUS-PAST  
 ‘The coach made the children swim.’
- (54) *itsumo inu-ni niwa-de asob-ase-teiru*  
 always dog-DAT garden-in play-CAUS-PROG  
 ‘(I) always let the dog play in the garden.’
- (55) *musume-ni kādo-o kak-ase-ta*  
 daughter-DAT card-ACC write-CAUS-PAST  
 ‘(I) had my daughter write a card.’

Some of the previous research in syntax argues that *o*-causatives such as (53) have a coerced interpretation while *ni*-causatives such as (54) express permission. However, such interpretations can be overridden by adverbial phrases (e.g. *muriyari* ‘forcedly’ and *jūyū-ni* ‘freely’ may co-occur with either *o*- or *ni*-causatives) or by context.

There are also lexically causative verbs (rather than syntactically causativized verbs by attaching *-[s]ase-* as above), which can be morphologically distinguished by the large number (but closed set) of transitive (causative) and intransitive (inchoative) verb pairs (Table 10.15). Although causative/transitive and inchoative/intransitive forms are morphologically related, lexical causatives are not productive and the form of one cannot automatically be predicted from the other. Unlike English ‘open’, etc., different verbs are used for transitive and intransitive uses as shown in Table 10.15. There are, however, some exceptions, such as *masu/masu* ‘to increase’ and *hiraku/hiraku* ‘to open’, in which the same form is used for transitives (causatives) and intransitives (inchoatives). Intransitives have a subject that typically has no volitional control over the action and which occupies the object slot of the transitive equivalent, e.g. *doa-ga aita* ‘the door opened’ vs *doa-o aketa* ‘he opened the door’. There is no ‘agent’ implied in the action of the intransitive, in contrast with the passive of the transitive: *doa-ga aita* ‘the door opened’ vs *doa-ga akerareta* ‘the door was opened (by . . .)’.

**TABLE 10.15 JAPANESE LEXICAL CAUSATIVES AND THEIR INTRANSITIVE COUNTERPARTS**

Causative/transitive	Inchoative/intransitive	
akeru	aku	'to open'
shimeru	shimaru	'to shut'
kowasu	kowareru	'to break'
nekasu	neru	'to put someone to sleep' : 'to sleep'
herasu	heru	'to reduce/decrease'
otosu	ochiru	'to drop' : 'to fall'

Causative sentences can be passivized as shown below.

- (56) *kachō-ga hishō-ni shorui-o kak-ase-ta*  
 section.chief-NOM secretary-DAT document-ACC write-CAUS-PAST  
 'The section chief made the secretary write a document.'
- (57) *hishō-ga kachō-ni shorui-o kak-ase-rare-ta*  
 secretary-NOM section.chief-by document-ACC write-CAUS-PASS-PAST  
 'The secretary was made to write a document by the section chief.'

Example (57) is a passive sentence based on the causative sentence (56). It should be noted that when a base causative sentence such as (56) involves a transitive verb, the direct object marked by *-o* cannot be the subject of the causative-passive (i.e. \**shorui-ga kachō-ni hishō-ni kak-ase-rare-ta*).

### 10.5.6 Speech levels and respect

Japanese encodes both (1.) addressee honorifics or polite speech levels/styles (typically referred to as 丁寧語 *teineigo*, or 'polite language') and (2.) participant honorifics (敬語 *keigo*, or 'respect language'). The two systems are marked differently and have different triggers: addressee honorifics are triggered by the speaker's relationship with the addressee only, while participant honorifics are triggered by the speaker's relationship with the participant and in which grammatical function (e.g. subject and object) the participant is expressed in the sentence. They are therefore theoretically independent of each other, but contextual triggers are also involved. On the one hand, the higher the addressee honorific is, the more likely a participant honorific will be triggered. On the other, certain formal situations trigger higher honorifics on both scales than normal, and certain formulaic expressions preserve this. On starting a meal, one says *itadakimasu* (*itadaku* is humble and *-masu* is polite) 'I'm given it; I eat it' to the person who provides it even if one would normally use plain language to them.

#### 10.5.6.1 Speech levels

In terms of verb/adjective inflection, Japanese speech levels show a two-way system, referred to in English-language literature as the plain vs polite distinction. The plain form is the morphologically unmarked member, and the polite is marked by the addition of the suffix *-masu* to the verb stem, replacement of the copula *-da* with *-desu<sub>1</sub>*, and the addition of *-desu<sub>2</sub>* after *i*-adjectives. *-masu* and *-desu<sub>1</sub>* are inflected while *-desu<sub>2</sub>*, though identical in form to *-desu<sub>1</sub>*, is an invariable politeness 'particle' (see 10.4.2.1). In practice, the system has more layers. A few verbs/adjectives display a third, superpolite level,

which combines the appropriate politeness marker with a different stem: *ii* ‘to be good/OK’, polite *ii-desu*<sub>1</sub>, superpolite *yoroshii-desu*<sub>2</sub>; *aru* ‘to be; exist; have’, polite *ari-masu*, superpolite *gozaimasu*. Earlier usage also included an even politer form of *i*-adjectives, formed from the descendent of the CJ *-u* infinitive form + *gozaimasu*. Since 1945, this has become obsolete in daily speech, and survives only in fixed expressions: *ohayō gozaimasu* greeting used at the start of the one’s day, *omedetō gozaimasu* ‘congratulations’, *arigatō gozaimasu* ‘thank you’. In questions, the three-way distinction becomes larger, with choice of question particle marking different relationships with the addressee: corresponding to plain *aru* are three forms *aru?*, *aru no*, and *aru (no) ka*. In commands, there are even more levels, e.g. from *matsu* ‘to wait’: imperative *mate!* (military; aggressive), *o-machi-nasai* (brusque), *maōte* (brusque), *maōte kure* (male, blunt), *maōte kudasai* (less blunt a command), *maōte kudasaimasen ka* (polite request), *maōte itadakemasen ka* (polite request), etc.

Verbs are normally only marked for politeness at the end of the sentence, and it is basically the main predicate that is so marked. Before a major break, typically before *ga/keredo(mo)* ‘but’, and, especially in female speech, before *kara* ‘because’, the level of politeness normally reflects that of the end of the sentence. Otherwise, subordinate clauses are no longer marked for politeness. In the most formal language, e.g. speech-giving and letter-writing, certain clause-internal polite forms may occur, e.g. Noun-*ni tsuite* (a postpositional phrase derived from Noun-*ni tsuku* ‘to attach to Noun’) ‘concerning Noun’ > Noun-*ni tsukimashite*. Nevertheless, *-mashite* as an ending occurs in speech only in set formulae, e.g. *hajimemashite* (< *hajimeru* ‘to begin’), used to greet someone that you have never met before. The *-masureba* form is virtually obsolete; it has an archaic feel if it is used in a letter.

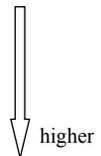
10.5.6.2 Participant honorifics: *keigo*

Participant honorifics (hereafter ‘*keigo*’) with verbs may be divided into (2a.) subject honorifics (尊敬語 *sonkeigo*, or ‘honorific language’) and (2b.) non-subject honorifics (謙讓語 *kenjōgo*, or ‘humble language’), typically object honorifics. Hereafter, ‘honorific’ (HON) will be used to refer to (2a.) and ‘humble’ (HUM) to (2b.). With nouns, *keigo* constitute what may be called possessor honorifics.

There are various devices used to derive honorific and humble forms, listed in Table 10.16. The devices marked † were in use in the earlier twentieth century, but with post-war democratization were considered too subservient in feel and are now obsolete. Unlike politeness, *keigo* can be marked on verbs wherever they occur in the sentence, including in relative clauses. However, excessive use of *keigo* has a subservient, grovelling or even sarcastic feel.

TABLE 10.16 JAPANESE HONORIFIC AND HUMBLE VERBS (E.G. *YOMI-MASU* ‘READ-POL’)

Unmarked	Honorific	Humble
yomi-masu	yom-are-masu	
	o-yomi-ni nari-masu	
		o-yomi shi-masu
	o-yomi nasai-masu	
	†o-yomi asobashi-masu	o-yomi itashi-masu
	†o-yom-are asobashi-masu	



†: now obsolete

The simplest honorific is the V-[r]are- passive form. All other regular forms are derived from an honorific noun, which itself is derived from *o-* + the infinitive. This honorific noun appears in sentences such as *o-yomi kudasai* ‘please read (HON)’, *o-yomi itadaku* ‘has someone read (HUM)’ as well, but this has no independent existence outside of these constructions.

- (58) *sensei-wa sōseki-o o-yomi-ninari-masu*  
 teacher-TOP Sōseki-ACC HON-read-HON-POL  
 ‘The teacher reads Sōseki’s novels.’
- (59) *watashi-wa okyakusama-ni keiyakusho-o o-yomi-shi-mashita*  
 I-TOP customer.POL-to contract-ACC HON-read-HUM-POL.PAST  
 ‘I read the contract for the customer.’

A significant number of verbs have irregular honorific or humble forms. (a.) Two can be analyzed as containing a formative *-[a](o)shar-* (HON), *iraqsharu* (for *iru*, ‘to be; exist’, *iku* ‘to go’, *kuru* ‘to come’), *oqsharu* (for *iu* ‘to say; be called’).<sup>4</sup> (b.) Some use suppletive verbs, e.g. *iru* ‘to be; exist’ > *oru* (HUM), *iu* ‘to say; be called’ > *mōsu* (HUM), *suru* ‘to do’ > *nasaru* (HON), *itasu* (HUM), *iku* ‘to go’/ *kuru* ‘to come’ > *mairu* (HUM). (c.) Some use euphemistic phrases, e.g. *taberu* ‘to eat’/ *nomu* ‘to drink’ > *meshi-agaru* (HON) and *itadaku* (HUM). (d.) Some use SJ forms combined with *suru*, *-da* or *-ni naru*, e.g. *miru* ‘to see’ > *goran-ni naru* (HON), *haiken suru* (HUM), *shiotte iru* ‘to know’ > *go-zonji-da* (HON). (e.) Some use suppletive stems in otherwise regular patterns, e.g. *iku* ‘to go’/ *kuru* ‘to come’ > *o-ide ni naru* (HON), a regular formation from CJ *ide-* (now *deru*) ‘to go out’. Most verbs with irregular honorific forms can also use the passive form; some also produce (certain) regular forms; some have multiple irregular forms. For example, *kuru* ‘to come’ has the honorific forms *iraqsharu*, *ko-rareru*, *o-ide ni naru*, among others. A definitive list of such forms is difficult to establish, as the more formal language tends to use rarer forms.

*oru* and *iru* in earlier Japanese (CJ *wor<sup>R</sup>-* and *wi<sup>M</sup>-* ‘sit; be’) were roughly synonymous. *oru* remains suppletive in the paradigm of *iru*. In coordinating function, the infinitive *i* is too short, particularly since the /i/ tends to be dropped in the *-te iru* form, and so *ori* is used instead. Similarly, the *zu*-form *orazu* replaces *\*irazu*, and the passive (honorific) of *iru* is either *irareru* or *orareru*. Otherwise, *oru* is now the usual humble equivalent of *iru*.

Keigo nouns are typically formed by the prefix *o-* or *go-*. Strictly, *o-* is used with NJ nouns, and *go-* is used with SJ nouns, but even in standard language there is some etymological inconsistency: SJ *denwa* ‘telephone call’ > *o-denwa*, and NJ *yuokuri* ‘at ease’ (not even a noun) > *go-yuokuri*. In more colloquial use, *o-* is also widely used on loanwords, particularly in connection with foodstuffs, and is considered a female affectation: *o-bīru* ‘beer’. In practice, such nouns are often not keigo at all, as they do not show respect to a possessor. They are what is known as *bikago* (美化語 ‘beautifying words’), or euphemism. These include forms referring to foodstuffs and to taboo concepts.

### 10.5.6.3 Pejorative verbs

Finally, it should be observed that Japanese has a form of anti-keigo in the form of pejorative verbs. As with honorifics, they are triggered by the grammatical subject, but instead show disapproval or dislike. They are typical of macho male language. Pejorative verbs are formed from infinitive + *-yagaru*, e.g. *yomu* ‘to read’ > *yomi-yagaru*.

### 10.5.7 Adverbials

Morphologically, adverbs may be derived from the infinitive of adjectives by changing *-i* to *-ku* for *i*-adjectives and *-na* to *-ni* for *na*-adjectives: e.g. *haya-i* ‘to be early’ > *haya-ku* ‘early’ and *kantan-na* ‘to be easy’ > *kantan-ni* ‘easily’. A number of adverbial expressions are formed from the *te*-form of verbs: *isoide* ‘hurriedly’, *kanete* ‘formerly’, *hajimete* ‘for the first time’, and *aratamete* ‘anew’. There are many other adverbs that are not derived from adjectives or verbs such as *totemo* ‘very’, *yuokuri* ‘slowly’, *tokidoki* ‘sometimes’, and *mochiron* ‘of course’.

Some types of adverbial expressions show concordance with modality items and other related items at the end of the clause. Examples are given in Table 10.17.

Onomatopoeic expressions or mimetics also function as adverbials. There are three types: *giongo* 擬音語 ‘phonomimes’ such as *gata-gata* ‘clattering’; *gitaigo* 擬態語 ‘phenomimes’ such as *gura-gura* ‘wobbly’; and *gijōgo* 擬情語 ‘psychomimes’ such as *ira-ira* ‘nervously’. Mimetics are often structured by reduplicating as shown above; however, non-reduplicated mimetics are also very common (e.g. *saō* ‘quickly’, *koqsori* ‘stealthily’, *guō* ‘really (get to)’). They can be followed by the particle *-to* or *-ni* with some restriction.

Postpositional phrases (PPs) function as ‘adverbial’ phrases (in the sense that they can be optional in syntax) and at their simplest consist of noun + particle, e.g. *expitsu-de* ‘with a pencil’ or *yoji-made* ‘till 4 o’clock’. More complex structures make use of a range of quasi-nouns, e.g. *mae* ‘before’, *naka* ‘inside’, *mawari* ‘around’, *aida* ‘[space] in-between’. These take the form [noun + *-no* + quasi-noun] + particle.

- (60) *butai-no ue-de odoō-ta*  
stage-GEN top-on dance-PAST  
‘(Someone) danced on the stage.’
- (61) *doa-no shita-kara tegami-o ire-ta*  
door-GEN beneath-from letter-ACC put-PAST  
‘(Someone) left a letter under the door.’

### 10.5.8 Coordination

Strictly speaking, verb phrase coordination is formed by (a.) the *te*-form (e.g. *aruite* ‘walk and’) or the infinitive (e.g. *aruki* ‘walk and’). There are other expressions, (b.)

TABLE 10.17 CONCORDANCE BETWEEN ADVERBIAL AND PREDICATE IN JAPANESE

Negation (e.g. <i>-nai</i> ):	<i>amari</i> ‘(not) really’, <i>meqtani</i> ‘rarely’
Perfect/past (e.g. <i>-ta</i> ):	<i>sakihodo</i> ‘a little while ago’, <i>taōta ima</i> ‘just now’
Affirmative (no apparent morphology):	<i>kanarazu</i> ‘certainly’, <i>mochiron</i> ‘of course’
Suppositionals (e.g. <i>-darō</i> ):	<i>tabun</i> ‘probably’, <i>osoraku</i> ‘perhaps’
Evidentials (e.g. <i>sō-da</i> , <i>yō-da</i> ):	<i>marude</i> ‘as if’, <i>imanimo</i> ‘at any time’
Exclamatory (e.g. <i>-darō</i> , <i>koto-da</i> ):	<i>nanto</i> ‘how . . . !’, <i>dorehodo</i> ‘how much . . . !’
Request (e.g. <i>-te kudasai</i> ):	<i>dōka</i> ‘(lit.) somehow, please’, <i>dōshitemo</i> ‘anyhow’
Desire (e.g. <i>-[i]tai</i> ; <i>-te hoshii</i> ‘I want X to . . .’):	<i>zehi</i> ‘at all costs’, <i>semete</i> ‘at the very least’
Conditionals (e.g. <i>-tara</i> , <i>nara</i> ):	<i>moshi(mo)</i> ‘if’, <i>karini</i> ‘supposedly’
Others 1 ( <i>-te oku</i> ):	<i>arajakime</i> ‘in advance’, <i>sono mama</i> ‘as-is’
Others 2 ( <i>-te shimau</i> ):	<i>tsui</i> ‘accidentally’, <i>uokari</i> ‘due to my carelessness’

infinitive + *-nagara* and (c.) the *tari*-form, but they contain nuances other than simple coordination (see below). The negative equivalents of (a.) are *-[a]naide* and *-[a]zu-ni* for verbs and *-nakute* and *-naku* for *i*-adjectives. *-nagara* may not be preceded by a negative form of a verb so long as it functions as a coordination marker (*-nagara(-mo)* is also a formal register subordinator ‘despite’ that does not share this restriction.). Affirmative verbs may be linked by *-tari* (e.g. *tabe-tari non-dari suru* ‘eat, drink and so on’), and affirmative and negative verbs expressing the same action ‘on and off’ can be linked by *-tari* usually in affirmative-negative order (e.g. *tabe-tari tabe-nakaQ-tari suru* ‘sometimes eat but sometimes don’t’, ‘keep eating and pausing’).

Of a sequence of verbs linked, for example, with the *te*-form, only the final verb is marked explicitly for tense, modality, etc.; the other verbs may be marked for aspect but otherwise take no inflection other than *-te*: A-*te* B-*te* C<sub>T</sub>. In the case of *-tari*, the final predicate may also take *-tari*, requiring the addition of an auxiliary verb, *suru*, to take tense, modality and/or politeness inflection: for instance, A-*tari* B-*tari* C(-*tari suru*)<sub>T</sub>.

- (62) *ongaku-o kii-te shinbun-o yon-da*  
 music-ACC listen-*te* newspaper-ACC read-PAST  
 ‘I listened to music and then read a newspaper.’
- (63) *ongaku-o kiki-nagara shinbun-o yon-da*  
 music-ACC listen-*nagara* newspaper-ACC read-PAST  
 ‘I read a newspaper while listening to music.’
- (64) *ongaku-o kii-tari shinbun-o yon-dari shi-ta*  
 music-ACC listen-*tari* newspaper-ACC read-*tari* do-PAST  
 ‘I listened to music and read a newspaper [alternating between the two or doing other things as well].’

Semantically, (a.), (b.) and (c.) differ significantly. (a.) acts as simple coordination. Linking dynamic verbs, however, (a.) expresses a sequence of events, and (b.) expresses simultaneity of events (the *nagara*-marked clause marking the accompanying as opposed to the main action). (c.) is used in circumstances in which the events cannot be sequenced, typically when different agents, represented by a subject of the *tari*-clause, are performing different actions at around the same time, or when the subject is alternating between two actions. The latter often involves the same verb repeated in the affirmative and negative forms.

- (65) *hiro-kute akaru-i heya*  
 be.wide-*te* be.bright room  
 ‘a large bright room’
- (66) *ame-ga fuQ-tari fura-nakaQ-tari shi-ta*  
 rain-NOM precipitate-*tari* precipitate-NEG-*tari* do-PAST  
 ‘It rained on and off.’

Other coordinating devices are *ga* ‘but; and’, *ke(re)do(mo)* ‘but’ and *shi* ‘and’, which follow finite verb forms. *ga* is not necessarily adversative in meaning, as it is an alternative to the *te*-form particularly when syntactically larger components are linked with different subjects. *shi* has the effect of listing situations the impact of which builds up and leads to an explicit or implicit conclusion. *shi* can be interpreted as similar to *node/kara* ‘because’ to indicate a reason of the following clause; this type of *shi* is categorized as a subordinate clause marker here although some researchers may include it as one of coordination.

### 10.5.9 Subordinate clauses

Subordinate clauses include (1.) conditional clauses, (2.) nominal clauses, (3.) conjunctive (reason) clauses, and (4.) quotation clauses. The distinction is not only made syntactically or morphologically but also from types of modal expressions and other related items appearing within the clause.

#### 10.5.9.1 Conditional clauses

Conditional clauses precede the main clause unless they are dislocated rightward for a stylistic reason. They are formed by a morphological marker, such as *-eba* 'if', *-tara* 'if; when', nonpast + *-to* 'when', tense + *nara* 'if', and the *te*-form in combination with a particle: *-te-mo* 'even if/though', *-te-wa* 'if'. The use of *-te-wa* is rather limited; it is mostly followed by expressions such as *dō-desu ka* (e.g. *kono hon-o yon-de mi-te-wa dō-deshō ka* 'How about reading this book?').

- (67) *iki-ta-kereba, iō-te kudasai*  
 go-want-COND go-te please  
 'If you want to go, please go.'
- (68) *shujin-ga modoō-tara, renraku-shi-masu*  
 husband-NOM return-COND inform-do-POL  
 'When my husband comes back, I will inform you.'
- (69) *ima ie-o deru-to, densha-ni maniau*  
 now home-ACC leave-COND train-DAT arrive.in.time  
 'If you leave home now, you can catch a train.'
- (70) *ganbaō-temo, tensū-wa kawar-anai*  
 do.best-COND mark-TOP change-NEG  
 'Even if I do my best, the mark does not change.'
- (71) *ima o-kyakusan-ni ko-rare-tewa, komaru*  
 now POL-guest-DAT come-PASS-COND be.troubled  
 'It's not convenient if a guest comes in now.'

The subject in the conditional clause does not have to be the same subject as in the main/following clause, as shown in (68), (70) and (71) above. The only conditional marker that allows morphological tense (either past or nonpast) before it is *nara* 'if'. Semantically, however, the tense of the verb at the end of the conditional clause is relative to that of the main verb. Those conditional clauses that express a situation that is dependent on ('if') or independent of ('even if/though') the situation of the main clause can be pre-empted by clause-initial or internal adverbs, *moshi* and *tatoe* respectively.

#### 10.5.9.2 Nominal clauses as adverbials

Basic structures for nominal or noun-modifying clauses (including relative clauses) are discussed in 10.5.1. This subsection deals with nominal clauses which function as sentential adverbials with the pattern of [relative clause] + grammatical/relational noun (+ particle). Examples with nominal clauses indicating a time are shown below.

- (72) [*kankoku-ni ik-u*]                    **mae-ni**, *chizu-o kai-mashita*  
 Korea-to go-NONPAST before-at map-ACC buy-POL.PAST  
 ‘Before I went to Korea, I bought a map.’
- (73) [*shokuji-o shi-ta*]            **ato-de**, *kēki-o tabe-mashita*  
 meal-ACC do-PAST after-at cake-ACC eat-POL.PAST  
 ‘After I had a meal, I ate a piece of cake.’
- (74) [*densha-o maQ-teiru aida(-ni)*] *hon-o yomi-masu*  
 train-ACC wait-PROG while(-at) book-ACC read-POL  
 ‘While I wait for a train, I read a book.’

Morphologically, *mae-ni* ‘before’ is preceded by the nonpast (which has future meaning/function), *ato-de* ‘after’ by the past, and *aida* ‘while’ by the progressive (or the nonpast) of verbs that can form it. *ato-de* can be replaced by *-te kara*; however, *-te kara* takes a *-te* form rather than the past tense form as in (73). The tense of the nominal clause is governed by and relative to that of the main predicate.

Another example is a nominal clause followed by *toki* ‘time > when’.

- (75) [*watashi-ga ie-ni i-ta*]            **toki**, *haha-ga kaeQ-te*  
 I-NOM home-in stay-PAST when mother-NOM return-*te*  
*ki-mashita*  
 come-POL.PAST  
 ‘When I was at home, my mother came back.’

*toki* can take either a nonpast or a past verb in its clause, and when it is a stative verb, the meaning of the sentence does not really change whether the verb is nonpast or past (i.e. [*watashi-ga ie-ni i-ru*] *toki*, *haha-ga kaeQ-te ki-mashita* has essentially the same meaning as (75)). However, the meaning of the sentence changes depending on the tense of the verb in the nominal clause when the verb is an activity or change of motion verb.

- (76) [*yubune-ni hairu*] *toki*, *karada-o arai-masu*  
 bathtub-to enter when body-ACC wash-POL  
 ‘Before I get in a bath, I wash myself.’
- (77) [*yubune-ni haiQ-ta*] *toki*, *karada-o arai-masu*  
 bathtub-to enter-PAST when body-ACC wash-POL  
 ‘After I get in a bath, I wash myself (in the bath).’

On the one hand, in (76) where the verb in the nominal clause is the nonpast, the action in the clause takes place right before the action expressed in the main clause (i.e. washing before getting into a bath to soak, which is traditional in Japan). On the other, the action expressed by the past tense form of the verb in the nominal clause in (77) should precede the action expressed in the main clause (i.e. washing after getting into a bath, meaning washing in a bath).

### 10.5.9.3 Conjunctive (reason) clauses

There are mainly three instances of conjunctive clauses which express a reason or a cause: *node*, *kara*, and (reason) *shi*.

- (78) *shigoto-ga aru node*, *ik-e-masen*  
 work-NOM exist because go-POT-POL.NEG  
 ‘Because I have to work, I can’t go.’

- (79) *ame-ga furu kara, kasa-o moQ-te iki-masu*  
 rain-NOM fall because umbrella-ACC hold-te go-POL  
 ‘Because it will rain, I’ll bring my umbrella.’
- (80) *dezāto-mo tabe-ta shi, mō kaeri-mash-ō*  
 dessert-also eat-PAST because already return-POL-ō  
 ‘We had the dessert, so let’s go home now.’

The use of *node* clauses is restricted more than that of *kara* and *shi* clauses; the latter are able to take a modal expression (and possibly a polite form) within their clauses and be followed by a command, request, cohortative, etc., while the former is not allowed to appear in such syntactic environments.

#### 10.5.9.4 Quotation clauses

Quoting a statement is expressed by quoted speech + *to* + quoting verb. Colloquial language can replace *to* with *ote*.

- (81) *tomodachi-wa [ashita denwa-suru] to ii-mashita*  
 friend-TOP tomorrow phone-do QUOT say-POL.PAST  
 ‘My friend said, “I’ll phone you tomorrow.”’ or  
 ‘My friend said that (s)he would phone me tomorrow.’

Where the main clause verb is a verb of speaking, (81) is the pattern chosen for indirect or direct quotation; however, the difference between direct and indirect quotations is not obvious; to make it clearer that the sentence involves a direct quotation, 「 」 is used in Japanese script for the quoted sentence, in which even sentence-final particles are possible. No other sentence-final particle can precede the quoting particle in indirect statements, even if one was used in the original words. The quoting particle is *ka* or *ka dō ka* ‘whether or not’ (never *no* or  $\emptyset$ , except for some dialects) for questions. With commands and requests the simplest structure is nonpast + *yō-ni*. Examples of *ka*, *ka dō ka* and *yō-ni* are shown in (82)–(84) below.

- (82) *chichi-wa [kinō doko-e iQ-ta] ka(\*-dōka) kii-ta*  
 father-TOP yesterday where-to go-PAST QUOT ask-PAST  
 ‘My father asked (me) where I went yesterday.’
- (83) *[shukudai-o yaQ-ta] ka(-dōka) oboe-tei-nai*  
 homework-ACC do-PAST QUOT remember-PROG-NEG  
 ‘I don’t remember whether I did homework or not.’
- (84) *ueitoresu-ni [hayaku kōhī-o moQ-te kuru] yō-ni tanon-da*  
 waitress-to quickly coffee-ACC hold-te come QUOT ask-PAST  
 ‘I asked the waitress to bring coffee quickly.’

## 10.6 LEXICON

In contrast with the usual native vs loan two-way distinction, it is conventional in Japan to distinguish three lexical strata: NJ, SJ and ‘*gairaigo*’ (外来語, loanwords). Strictly, SJ morphemes are also ultimately loans, but their special status in Japanese makes their separate classification valid (see Chapter 1 this volume, 1.2.4). Although typically morphemes combine with morphemes of the same stratum in first-level compounds, there

are many exceptions, e.g. NJ *mi* + SJ *bun* > *mibun* 身分 ‘identity’, SJ *ni* + NJ *kata* > *nigata* 二型 ‘type 2’. In the Edo period there was increased neologism in the areas of popular culture and these frequently mixed NJ and SJ morphemes. More recently, popular forms have been created that deliberately mix strata, e.g. *ganguro* ガングロ urban girl ‘tribe’ known for dark face-paint < SJ {*gan*} 顔 ‘face’ + NJ *kuro* 黒 black’. { } indicates morphemes that only occur bound. Within the NJ layer, onomatopoeic/mimetic/phonaesthetic forms constitute a distinct sub-layer.

### 10.6.1 Onomatopoeia and phonaesthesia

Japanese makes extensive use of phonaesthetic forms – some onomatopoeic but most not – as adverbs (see 10.5.7 for their syntax), and dictionaries are dedicated to them. Phonologically, the phonaesthetic layer of NJ is notable for the occurrence of /p/ morpheme-initially; otherwise NJ /p/ only occurs after /N/ or /Q/. Some phonaesthetic forms are derivable from content words, but most are not. Formal non-fiction writing makes very limited use of them, but they are common in casual speech. The commonest pattern is a reduplication, C<sup>1</sup>V<sup>1</sup>C<sup>2</sup>V<sup>2</sup>-C<sup>1</sup>V<sup>1</sup>C<sup>2</sup>V<sup>2</sup>, such as *fuwa-fuwa* ‘very lightly, very soft’, *neba-neba* ‘stickily’, *pika-pika* ‘shiningly, flashing’ (cf. *hikar-u* ‘to shine’ < OJ *pikar-*), *jiwa-jiwa* ‘stealthily’, *giri-giri* ‘in the nick of time’. There is no *rendaku* normally, but exceptions occur, indicating a different etymological pattern, e.g. *kowa-gowa* ‘nervously’ (cf. *kowai* ‘to be scared’). More sudden or semelfactive actions take other patterns, especially C<sup>1</sup>V<sup>1</sup>Q<sup>2</sup>V<sup>2</sup>ri, C<sup>1</sup>V<sup>1</sup>C<sup>2</sup>V<sup>2</sup>N-to, or CVQ-to. Other than shifting between patterns, subtle nuances can be expressed by consonant voicing, e.g. *fura-fura* ~ *bura-bura* ‘walking casually’, or less commonly by vowel changes, e.g. *muku-muku* ~ *moku-moku* ‘[smoke/cloud] rising’. Though adverbs, some are commonly verbalized, especially *biokuri suru* ‘to be surprised’.

### 10.6.2 Neologism

The modern period has been characterized by a vast quantity of neologism. The 1868 Meiji Restoration and subsequent abolition of the feudal system and rapid modernization of Japan led to large-scale neologism to express new scientific, medical, military, political, economic and sociological concepts, and these overwhelmingly took the form of neo-classical (SJ) compounds. See Chapter 1 (this volume, 1.2.3.2) for discussion of neo-classical forms in Japan and subsequent graphic borrowing within East Asia. By 1900, most of this modern vocabulary had been settled. Since then, SJ-based neologisms have tended to be not first-level but higher-level compounds, e.g. *onshitsu* ‘greenhouse’ + *kōka* ‘effect’ > 温室効果 *onshitsu kōka* ‘greenhouse effect’. Since 1945, neologism has been heavily based on loanwords from English.

There are morphological processes other than neo-classical and loan forms. SJ words predominantly consist of two morphemes = two characters. In Chinese, many of these were created from near-synonyms to create a bimorphemic word to replace an earlier, perhaps ambiguous, monomorphemic word. However, when these were the input of further compounds, it was the underlying monomorphemic word that was used, thereby preserving a two-morpheme pattern in Chinese. Japanese adopted the same principle. As most SJ morphemes are bimoraic, and NJ nouns also tend to be bimoraic, new Japanese compounds came to prefer four morae. This has resulted in double truncations of SJ, NJ, loan- and hybrid compounds of the pattern AB-CD > A-C or (if B or D are more distinctive to the meaning or less ambiguous than A or C) B-C or A-D: 東京大学

*tōkyō daigaku* ‘Tokyo University’ > 東大 *tōdai*. The preference is to replace NJ readings of characters with SJ ones, so \*京都大阪 *kyōto-ōsaka* > 京阪 *keihan* ‘Osaka [to] Kyoto’, \*大阪神戸 *ōsaka-kōbe* > 阪神 *hanshin* ‘Osaka [to] Kobe’. The retention of NJ readings only occurs when the SJ reading is shared by too many morphemes for it to be distinctive, e.g. 静岡大学 *shizuoka daigaku* ‘Shizuoka University’ > 静大 *shizudai*. The device has been used for neologisms by companies and by youth slang, and loanwords are particularly prone due to their length, e.g. ゴスリック・ロリータ *gosurīku rorīta* ‘Gothic Lolita’ > ゴスロリ *gosurori*, シヤープ・ペンシル *shāpu penshiru* ‘sharp pencil, i.e. telescopic pencil’ (originally a tradename) > シヤープペン *shāpen*. Long vowels are generally shortened in loanwords, e.g. パーソナル・コンピューター *pāsonaru konpyūtā* ‘personal computer’ > パソコン *pasokon* (not \**pākon*).

### 10.6.3 Slang

Youth and street language are notable for their innovation. Some forms naturally enter into common usage, but most are vogue words that are doomed to obsolescence. Some use devices outlined above, and industry sometimes even adopts such forms or uses the same devices, e.g. double truncation 着信メロディー *chakushin merodī* > 着メロ *chakumero* ‘ringtone’. Similar moraic manipulation characterizes one argot that a century ago produced a variety of words that, at least originally, were opaque, whereby the first and last syllables are reversed, and any intervening syllables are dropped. Two surviving examples of this long-obsolete device are *fuda* ‘card’ > *dafu*- ‘ticket’ in *dafuya* ‘ticket tout’, and †*kakusode* ‘policeman’ (named after their late nineteenth-century uniform) > *deka* ‘cop’.

In standard Japanese, NJ verbs and *i*-adjectives form closed sets. When a new verb or adjective is created or borrowed, it takes the auxiliary verb *suru* ‘to do’ or the particle *no* (or sometimes becomes a *na*-adjective) respectively, e.g. neo-classical *shuōpan suru* ‘to publish’, English loan *haikingu suru* ‘to hike’. In colloquial usage, there is a long history of creating new verbs or *i*-adjectives, few of which have entered standard use, and most tend to become obsolete eventually. The preferred pattern is two morae + *-ru* and two morae + *-i*. Two that have survived many decades are *daburu* ‘to repeat [a course]’ < E *double* and *saboru* ‘to play truant’ < Fr *sabotage* ‘sabotage’ (or Fr *sabot* ‘clog; sabotage’?). Most have become obsolete, e.g. *takuru* ‘to take a taxi’ < *takushī*. *nai* < E *now* and *imai* < *ima* ‘now’, both meaning ‘to be trendy’, appeared in the 1980s and have largely vanished. However, despite the short-term existence of most such forms, they continue to appear. Common current forms used in youth slang include *kimoi* ‘to be weird’ < *kimochi(-ga) warui* idem, and *kishoi* ‘to be gross; be creepy’ < *kishoku(-ga) warui* idem. The latter tend to be used as interjections and are rarely inflected.

### NOTES

- 1 The syntax sections were written by Mika Kizu, and the preceding and following sections by Nicolas Tranter.
- 2 Full Hepburn transcription is used for publications, people and terms; our modified transcription using *N* and *Q* is used for language data.
- 3 We keep the paradigmatic presentation simple rather than incorporate reference to traditional verb stems as is common in Japan. Moreover, *ō-*, *te-*, *tari-*, *tara-*, *eba-* forms, etc. are referred to as such, though Western works tend to use terms such as ‘tentative’, ‘gerund’, ‘frequentative’, ‘conditional’, ‘provisional’, etc. We use the latter terms only in Table 10.11 where a different and older system is presented, and ‘conditional’ is otherwise used in this chapter to describe function rather than to name a form.

- 4 The Edo-period ending *-[a](o)shar-* derived from LMJ honorific *-[s]aserare-* (Irwin and Narrog, this volume, 9.5.6) and was more productive. Strictly *iraqsharu* derives from quinquagrade *iru* ‘to enter’ even though it now substitutes for mono-grade *iru* ‘to exist’, while *oqsharu* derives from CJ *ofose-* (‘say’ HON) + LMJ honorific *-[r]are-* > *ōse-rare-* (probably reanalyzed analogically as *ō-serare-*) > *ōqshar(e)-*.

## APPENDIX: STYLE DIVERSITY IN 1903/1904 NEWSPAPERS

Transcription key: distinctly bungo/sōrōbun forms (a–d), reading no longer used (e); kanji simplified or abolished 1946–1949; kanji spellings no longer used.

- a. Front page editorial: bungo, kanji + hiragana, furigana on most kanji, no punctuation.

とく はくらんくわい 開會式 の とうじつ おま とくしや わか  
 特に博覽會開會式の當日において普ねく讀者に頒たんとす  
 ... *toku-ni hakurankai kaikaishiki-no tōjitsu-ni oite amaneku tokusha-ni wakatan to su*  
 ‘We shall ... specially distribute it to every reader on the day of the opening of the exhibition.’

- b. Review article: sōrōbun, kanji + hiragana, no punctuation.

本日は接待所々見なるものを御覽に入れ申候  
 ... *honjitsu-wa seotaijo shōken-naru mono-o goran-ni ire-mōshi-sōrō*  
 ‘... today I humbly submit for your reading what I observed at the reception.’

- c. Kimono shop advert: sōrōbun, kanji + minimal katakana, no punctuation.

四月一日ヨリ六月下旬迄左ノ諸展覽會開催致候  
 ... *shigatsu tsuitachi-yori rokugatsu gejun-made sa-no sho-tenrankai kaisai itashi-sōrō*  
 ‘... we humbly open each display listed on the left from 1 April to late June.’

- d. Advert: bungo, kanji + katakana, no punctuation.

專賣特許粉磨機麥磨機ノ二機ハ其目的ヲ達スル最新發明ノ農具ナリ  
 ... *senbai tokyo momimigaki-ki mugimigaki-ki-no niki-wa sono mokuteki-o taosuru saishin hatsumei-no nōgu-nari*  
 ‘... the patent rice and buckwheat threshing machines are both recently invented farm tools that fulfil this purpose.’

- e. Serialized story: polite-style kōgo, kanji + hiragana + katakana, furigana on kanji, punctuation (but no full stops).

女「ハイ」女中は膨れ面をして見世へ参りました  
 “hai” *nyochū-wa fukuretsura-o shite mise-e mairimashita*  
 ‘Maid: “Yes sir.” The maid went to the shop with a sulky face.’

## REFERENCES AND FURTHER READING

- Aikawa, Takako (1994) ‘Logophoric use of the Japanese reflexive *zibun-zisin* “self-self”’, *MIT Working Papers in Linguistics*, 24, 1–22.  
 Backhouse, A. E. (1984) ‘Aspects of the graphological structure of Japanese’, *Visible Language*, 18: 219–28.  
 Bloch, Bernard (1950) ‘Studies in colloquial Japanese IV: Phonemics’, *Language*, 26(1): 86–125.  
 Bunka Shingikai (2010) *Kaitei Jōyō Kanjihyō* [The revised jōyō kanji list]. ([http://www.bunka.go.jp/kokugo\\_nihongo/pdf/jouyoukanjihyou\\_h22.pdf](http://www.bunka.go.jp/kokugo_nihongo/pdf/jouyoukanjihyou_h22.pdf))  
 Carroll, Tessa (2001) *Language Planning and Language Change in Japan*. Richmond, UK: Curzon.

- Chamberlain, Basil Hall (1907) *Handbook of Colloquial Japanese*, 4th ed. revised. London, Yokohama, Shanghai, Hongkong, Singapore: Crosby Lockwood/Kelly & Walsh (1st ed.: 1888, 2nd ed.: 1889, 3rd ed.: 1898).
- Fujii, Noriko (1984) 'History of the use of punctuation marks in Japanese and the influence of Western languages', *6th International Symposium on Asian Studies*, 677–85.
- Georgopoulos, Carol and Ishihara, Roberta (1990) *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*, Dordrecht: Kluwer.
- Gottlieb, Nanette (1995) *Kanji Politics: Language Policy and Language Script*, London and New York: Kegan Paul International.
- Gravett, Paul (2004) *Manga: Sixty Years of Japanese Comics*, London: Laurence King.
- Hepburn, J. C. (1887) *A Japanese–English and English–Japanese Dictionary*, 2nd revised and enlarged ed., Tokyo: Z. P. Maruya & Co.
- Higurashi, Yoshiko (1983) *The Accent of Extended Word Structures in Tokyo Standard Japanese*, Tokyo: Educa.
- Hinds, John and Howard, Irwin (1978) *Problems in Japanese Syntax and Semantics*, Tokyo: Kaitakusha.
- Inoue, Fumio (2005) 'Econolinguistic aspects of multilingual signs in Japan', *International Journal of the Sociology of Language*, 175/176: 157–77.
- Kuniyoshi, Kataoka (1997) 'Affect and letter-writing: unconventional conventions in casual writing by young Japanese women', *Language in Society*, 26(1), 103–36.
- Kuno, Susumu (1973) *The Structure of the Japanese Language*, Cambridge, MA: MIT Press.
- Kuroda, S.-Y. (1992) *Japanese Syntax and Semantics: Collected Papers*, New York: Springer.
- McGovern, William Montgomery (1920) *Colloquial Japanese*, London: Kegan Paul, Trench, Trubner & Co.
- Makino, Seiichi and Tsutsui, Michio (1989) *A Dictionary of Basic Japanese Grammar*, Tokyo: Japan Times.
- (1995) *A Dictionary of Intermediate Japanese Grammar*, Tokyo: Japan Times.
- (2008) *A Dictionary of Advanced Japanese Grammar*, Tokyo: Japan Times.
- Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.
- Masuoka, Tadashi (2009) 'Modality from a Japanese perspective', in Pizziconi, Barbara and Kizu, Mika (eds) *Japanese Modality: Exploring its Scope and Interpretation*, Basingstoke, UK: Palgrave MacMillan, 36–55.
- Maynard, Senko (1990) *An Introduction to Japanese Grammar and Communication Strategies*, Tokyo: Japan Times.
- Miller, Laura (2004) 'Those naughty teenage girls: Japanese Kogals, slang, and media assessments', *Journal of Linguistic Anthropology*, 14: 225–47.
- Miyagawa, Shigeru (1989) *Syntax and Semantics: Structure and Case Marking in Japanese*, vol. 22, New York: Academic Press.
- Miyagawa, Shigeru and Saito, Mamoru (2008) *The Oxford Handbook of Japanese Linguistics*, New York: Oxford University Press.
- Nakanishi, Kimiko (2004) 'Domains of Measurement: Formal Properties of Non-Split/Split Quantifier Constructions', Ph.D. dissertation, University of Pennsylvania.
- Pizziconi, Barbara and Kizu, Mika (2009) *Japanese Modality: Exploring its Scope and Interpretation*, Palgrave Macmillan.
- Plaut, Hermann (1905) *Japanese Conversation-Grammar with Numerous Reading Lessons and Dialogues*, London, Heidelberg: Julius Groos.
- Seeley, Christopher (1991) *A History of Writing in Japan*, Honolulu: University of Hawai'i Press.
- Sells, Peter (1987) 'Aspect of logophoricity', *Linguistic Inquiry*, 18: 445–81.
- Shitabani, Masayoshi (1976) *Syntax and Semantics*, vol. 5 (Japanese Generative Grammar), New York: Academic Press.

- (1990) *The Languages of Japan*, Cambridge: Cambridge University Press.
- Smith, Janet S. and Schmidt, David L. (1996) 'Variability in written Japanese: towards a sociolinguistics of script choice', *Visible Language*, 30: 46–71.
- Tranter, Nicolas (2008) 'Non-conventional script choice in Japan', *International Journal of the Sociology of Language*, 192: 133–51.
- (forthcoming) 'Graphic loans: East Asia and beyond', *Word*, 58.
- Tsujimura, Natsuko (1999) *The Handbook of Japanese Linguistics*, Cambridge, MA: Blackwell Publishing.
- (2007) *An Introduction to Japanese Linguistics [second edition]*, Cambridge, MA: Blackwell Publishing.
- Twine, Nanette (1984) 'The adoption of punctuation in Japanese Script', *Visible Language*, 18(3): 229–37.
- (1991) *Language and the Modern State: The Reform of Written Japanese*, London, New York: Routledge.
- Vance, Timothy (1987) *An Introduction to Japanese Phonology*, Albany: State University of New York Press.

# JAPANESE DIALECTS: FOCUSING ON TSURUOKA AND EI

*Akiko Matsumori and Takuichiro Onishi*

## 11.1 INTRODUCTION

Japanese dialects are a rich source of topics in phonology, morphology, and grammar, as well as many sub-areas of linguistics. Such regional variety will continue to be an excellent testing ground for any theoretical and typological generalizations made in studies on the Japanese language, which have so far been heavily based on Standard Japanese (ModJ). Among the various Japanese dialects, this chapter focuses on two dialects in the main islands of Japan. One of them is the dialect of Tsuruoka, located in Yamagata prefecture in the Tōhoku region, while the other one is Ei, which is located in Kagoshima prefecture, the southernmost prefecture in the Kyūshū region.<sup>1</sup>

The purpose of this chapter is not so much to give an exhaustive description of the features of these two dialects as to illustrate how dialectal differences may provide some illuminating topics for future research on Japanese. In order to pursue this goal, we will describe these two dialects, particularly focusing on how they differ from ModJ. Thus, comparisons with ModJ will be made whenever necessary throughout this chapter.

The Tsuruoka and Ei dialects were chosen not only because they have been relatively thoroughly described by earlier scholars, especially due to the efforts made by the National Institute for Japanese Language and Linguistics (Kokuritsu Kokugo Kenkyūjo: hereafter NINJAL), but also because they are ideal examples of some of the most well-debated features of Japanese dialects, most of which are distinctively different from ModJ, and thus are suitable vehicles to introduce the major achievements of past studies on Japanese dialects.

Among the various dialect-related projects so far conducted by NINJAL, of especial note is the creation of *Nihon Gengo Chizu* (Linguistic Atlas of Japan: hereafter LAJ), which was published from 1966, and completed in 1974. The LAJ, which consists of six volumes, containing maps for 300 vocabulary items, has provided a bird's eye view of the distribution of vocabulary and grammatical items throughout the Japanese archipelago including Ryukyuan varieties, and has significantly contributed to the advancement of subsequent research on Japanese dialects. Soon after the completion of this atlas, NINJAL planned another project to publish atlases focusing on grammatical aspects of Japanese dialects, starting in 1976. They subsequently conducted research from 1979 to 1982, publishing the results under the title *Hōgen Bunpō Zenkoku Chizu* (Grammar Atlas of Japanese Dialects: hereafter GAJ). The GAJ also consists of six volumes, the first one of which was published in 1989, the last one (i.e. vol. 6) completed in 2006 (Onishi 2006).

Such projects by NINJAL have revealed that many of the greatest regional differences exist between eastern and western parts of mainland Japan, which is typically illustrated by the map of the negative suffixes of verbs in Onishi (2008: 68). Our choice of one dialect from the Tōhoku region (Tsuruoka) and the other one from the Kyūshū region (Ei) is first intended to delineate such east–west contrasts, based on the assumption that these two systems typically represent eastern and western dialects, respectively.

## 11.2 TSURUOKA DIALECT

The dialect of Tsuruoka in Yamagata prefecture is one of the best-described systems in Japan. So far, NINJAL has conducted three series of fieldwork projects in this dialect, the first of which was conducted in 1950, the second in 1971, and the third in 1991–1992. The fieldwork reports were published by NINJAL (1953, 1974, 1994, 2006b).

### 11.2.1 Phonology

#### 11.2.1.1 Vowels

The most significant feature of vowels in this dialect, also shared with many other dialects in the Tōhoku region, lies in the vowel length. In ModJ, the phonological distinctions between **obasan** ‘aunt/middle-aged woman’ and **obāsan** ‘grandmother/old woman’, or between **chizu** ‘map’ and **chīzu** ‘cheese’ (hereafter, the ModJ words will be denoted in bold font) are made by the clear durational differences of the vowels, in that the vowels in the latter words are one mora longer than in the former ones. Such length differences of vowels are more subtle in the Tsuruoka dialect, so that in this dialect long vowels are phonetically shorter than those in ModJ.

The vowel system of Tsuruoka consists of six vowels /i e ε a o u/ (Inoue 1994), of which /ε/ is derived from the contraction of the vowel sequences /ai/, /oi/, or /ae/: [ke:go] ‘silkworm’ (**kaiko**), [de:gon] ‘radish’ (**daikon**), and [sa<sup>n</sup>ze:] ‘turban shell’ (**sazae**) (NINJAL 1994: 52). Although Inoue (1994, 2000) reported that the six-vowel system is now becoming archaic and that recently the older /ε/ has merged with /e/, which, as a result, made the whole system into one consisting of five vowels /i e a o u/ (NINJAL 1994: 52), this chapter will describe the Tsuruoka dialect based on the older generation’s system of six vowels.

The phonetics of the phoneme /i/ in this dialect are different from ModJ in that it is much more centralized than that of ModJ, i.e. [i]: e.g. [kagī] **kaki** ‘persimmon’, [hana<sup>m</sup>bi] **hanabi** ‘firework’, [tori] **tori** ‘bird’, [kani] **kani** ‘crab’, [tʃi] **chi** ‘blood’, [iʃi] **ishi** ‘stone’, [aʃi] **ashi** ‘leg’, and [sa<sup>n</sup>dʒi] **saji** ‘spoon’.<sup>2</sup> Moreover, the vowel corresponding to ModJ /u/ is (or used to be; see Note 3) fronted to [i] after the coronal consonants *c*, *z*, and *s*. Thus, for example, the ModJ words **tsue** ‘cane’ and **chizu** ‘map’ are pronounced as [tʃiε] and [tʃi<sup>n</sup>dʒi] respectively in this dialect. /i/ and /u/ also devoice in the same circumstances as in ModJ.

As a result, the merger of /cʃi/ and /cu/, /zʃi/ and /zu/, and /sʃi/ and /su/ took place, so that the ModJ words **machi** ‘town’ and **matsu** ‘pine tree’ are both pronounced as [maʃi], **chichi** ‘father’ and **tsuchi** ‘soil’ as [tʃiʃi], **chizi** ‘governor’ and **chizu** ‘map’ as [tʃi<sup>n</sup>dʒi]. This type of dialect, found widely in the Tōhoku region, is known as a *hitotsugana* dialect, literally translated as ‘one-letter dialect’, based on the fact that the older distinction of EMJ ぢ \**di*, づ \**du*, じ \**zi*, ず \**zu* (see Irwin and Narrog, this volume, 9.2.1) is lost, resulting in one pronunciation of all four, i.e. [i<sup>n</sup>dʒi] in the case of Tsuruoka.<sup>3</sup>

11.2.1.2 Consonants

In Tsuruoka, a number of voiced obstruents corresponding to ModJ **b**, **d**, **j**, and **z** are prenasalized,<sup>4</sup> as illustrated in the following examples:

(1)	Tsuruoka		ModJ	Tsuruoka		ModJ	
	[na <sup>m</sup> be]	/na <sup>m</sup> be/	‘pot’	<b>nabe</b>	[ka <sup>m</sup> bi]	/ka <sup>m</sup> bi/ ‘mold’	<b>kabi</b>
	[u <sup>n</sup> de]	/u <sup>n</sup> de/	‘arm’	<b>ude</b>	[ka <sup>n</sup> do]	/ka <sup>n</sup> do/ ‘angle’	<b>kado</b>
	[a <sup>n</sup> dza]	/a <sup>n</sup> za/	‘bruise’	<b>aza</b>	[sa <sup>n</sup> dʒi]	/sa <sup>n</sup> dʒi/ ‘spoon’	<b>saji</b>

As for the voiced velar /g/, it appears as the velar nasal [ŋ] between vowels so ModJ **kagi** ‘key’, **kage** ‘shadow’, **kagami** ‘mirror’, and **ago** ‘chin’ are pronounced as [kaŋi], [kaŋe], [kaŋami], and [aŋo], respectively, by Tsuruoka dialect speakers.

The consonant system of this dialect is summarized in Table 11.1, in which /c, cj, sj, zj/ are phonetically [ts], [tʃ], [ʃ], and [ʒ] respectively.

One of the most salient segmental features of the present Tōhoku dialects, including the Tsuruoka dialect, is the voicing of voiceless obstruents word-internally between vowels, by which the older /t/ and /k/ are changed into /d/ and /g/ respectively.

(2)	Tsuruoka		ModJ	Tsuruoka		ModJ
	<i>hado</i>	‘pigeon’	<b>hato</b>	<i>adama</i>	‘head’	<b>atama</b>
	<i>kagi</i>	‘persimmon’	<b>kaki</b>	<i>sage</i>	‘wine’	<b>sake</b>
	<i>igi</i>	‘breath’	<b>iki</b>	<i>kuRgi</i>	‘air’	<b>kūki</b>
	<i>magura</i>	‘pillow’	<b>makura</b>	<i>odogo</i>	‘man’	<b>otoko</b>

However, this voicing tends to be blocked after devoiced vowels /i, u/, as in *hito* ‘person’, *asjita* ‘tomorrow’, *sjiken* ‘test’, *kucji* ~ *kuzji* ‘mouth’, *cucji* ~ *cuzji* ‘soil’, etc.; moreover, the voicing does not take place when the preceding segments are moraic /N/ or /Q/; thus, word-internal /p/ /t/ /k/, as in *zjiopun* ‘ten minutes’, *bento* ‘lunch’, *sankagu* ‘triangle’, and *gaoko* ‘school’, are not subject to this change.

As a result, the ModJ distinction between **chichi** ‘father’ and **chiji** ‘governor’ is made in Tsuruoka by the absence or presence of the prenasalized quality of the intervocalic consonants, i.e. [tʃiʒi] vs. [tʃi<sup>n</sup>dʒi], rather than a voiceless/voiced distinction. According to Inoue (1994: 54), however, this distinction is lost by younger-generation speakers due to the loss of prenasalization as a feature; for these speakers, **chichi** ‘father’ and **chiji** ‘governor’ now tend to be pronounced the same.

TABLE 11.1 CONSONANTS IN THE TSURUOKA DIALECT

	Bilabial		Dental/ Alveolar		Palatal <sup>5</sup>		Velar		Glottal
Plosive	p	b	t	d			k	g	
Prenasalized		<sup>m</sup> b		<sup>n</sup> d					
Affricate			c		cj				
Fricative			s	z	sj	zj			h
Prenasalized				<sup>n</sup> z		<sup>n</sup> zj			
Nasal		m		n				ŋ	
Flap				r					
Approximant		w				j			

This intervocalic voicing used to be a phonetic quality of the voiceless obstruents; however, recent vocabulary in Tsuruoka includes many words (many of which are Sino-Japanese ones) which contain intervocalic voiceless obstruents: *toRkjo* ‘Tokyo’ (**tōkyō**), *zjikan* ‘time’ (**jikan**), *hoRtai* ‘bandage’ (**hōtai**); thus, the voiced consonants illustrated in (2) are phonemic in the present-day Tsuruoka dialect. (Vowel length in dialects is represented by *R* in this chapter.)

### 11.2.1.3 Suprasegmentals

As for the accent system, nouns in ModJ are well known to belong to an *n+1* type system, in which words with *n* number of morae have *n+1* number of accent patterns (Uwano 1977; Vance 1987). Thus, in ModJ, 1-mora words have two patterns, 2-mora words three, 3-mora words four, and so on. For example, the accentual pattern of a 2-mora noun **hashi** ‘edge’ in ModJ is Low-High (LH) in isolation, while when a suffix such as **-ga** (Nominative), **-o** (Accusative), or **-ni** (Dative) is attached to the noun as in **hashi-ni** it is LH-H, thus the word has a LH~LH-H pattern. On the other hand, **hashi** ‘bridge’ and **hashi** ‘chopstick’ have different accentual patterns in ModJ, the former of which shows a LH~LH-L pattern, while the latter a HL~HL-L pattern. In fact, all the ModJ 2-mora words belong to one of these three types. In the same way, 3-mora words in ModJ belong to either one of the following four types: (1) LHH~LHH-H, (2) HLL~HLL-L, (3) LHL~LHL-L, (4) LHH~LHH-L, while 4-mora words the following five: (1) LHHH~LHHH-H, (2) HLLL~HLL-L, (3) LHLL~LHLL-L, (4) LHHL~LHHL-L, (5) LHHH~LHHH-L. The accentual system of ModJ is illustrated in Table 11.2.

In contrast, the Tsuruoka dialect has an *n+2* system, in which, although 1-mora words have only two patterns, 2-mora words have four patterns, 3-mora words five, and 4-mora words six (Nitta 1994). This is illustrated in Table 11.3.

For example, 2-mora nouns in Tsuruoka consist of the LL~LL-L pattern (*hàna* ‘nose’, *hàna-sà* ‘nose-DAT’), HL~HL-L (*àgi* ‘autumn’, *àgi-sà* ‘autumn-DAT’), LH~LH-L (*kàdà* ‘shoulder’, *kàdà-sà* ‘shoulder-DAT’), and LH~LL-H (*àsji* ‘foot’, *àsji-sà* ‘foot-DAT’). Table 11.4 compares the ModJ and Tsuruoka patterns of 2-mora nouns. Here we see that fairly regular correspondences exist between the three patterns of ModJ and the four of Tsuruoka.

TABLE 11.2 ACCENT SYSTEM OF STANDARD JAPANESE

1-mora words	2-mora words	3-mora words	4-mora words
L(-H)	LH(-H)	LHH(-H)	LHHH(-H)
<b>hi</b> ‘day’	<b>kaze</b> ‘wind’	<b>sakana</b> ‘fish’	<b>niwatori</b> ‘hen’
H(-L)	HL(-L)	HLL(-L)	HLLL(-L)
<b>hi</b> ‘fire’	<b>nabe</b> ‘pot’	<b>awabi</b> ‘abalone’	<b>kamakiri</b> ‘mantis’
	LH(-L)	LHL(-L)	LHLL(-L)
	<b>hana</b> ‘flower’	<b>kokoro</b> ‘heart’	<b>murasaki</b> ‘purple’
		LHH(-L)	LHHL(-L)
		<b>atama</b> ‘head’	<b>tamanegi</b> ‘onion’
			LHHH(-L)
			<b>otōto</b> ‘brother’

**TABLE 11.3 ACCENT SYSTEM IN TSURUOKA DIALECT**

1-mora words	2-mora words	3-mora words	4-mora words
L(-L)	LL(-L)	LLL(-L)	LLLL(-L)
<i>ka</i> ‘mosquito’	<i>hana</i> ‘nose’	<i>sagana</i> ‘fish’	<i>niwadori</i> ‘hen’
L(-H)	HL(-L)	HLL(-L)	HLLL(-L)
<i>e</i> ‘picture’	<i>agi</i> ‘autumn’	<i>ka<sup>m</sup>buto</i> ‘helmet’	<i>mazjidage</i> ‘mushroom’
	LH(-L)	LHL(-L)	LHLL(-L)
	<i>kada</i> ‘shoulder’	<i>usaji</i> ‘rabbit’	<i>asajao</i> ‘morning glory’
	LL(-H)	LLH(-L)	LLHL(-L)
	<i>asji</i> ‘leg’	<i>kogoro</i> ‘heart’	<i>tamaneki</i> ‘onion’
		LLL(-H)	LLLH(-L)
		<i>adama</i> ‘head’	<i>asjido</i> ‘sound of a footstep’
			LLLL(-H)
			<i>inosjisji</i> ‘boar’

**TABLE 11.4 ACCENT PATTERNS OF 2-MORA NOUNS IN STANDARD JAPANESE AND TSURUOKA**

Examples by tone class (Martin 1987)	ModJ	Tsuruoka
2.1: <b>ame</b> ‘candy’, <b>hana</b> ‘nose’, <b>hashi</b> ‘edge’, <b>kaze</b> ‘wind’, <b>mushi</b> ‘insect’	LH~LH-H	LL~LL-L
2.2: <b>hashi</b> ‘bridge’, <b>kami</b> ‘paper’, <b>kawa</b> ‘river’, <b>mune</b> ‘breast’	LH~LH-L	
2.3: <b>ashi</b> ‘foot’, <b>hana</b> ‘flower’, <b>kagi</b> ‘key’, <b>yama</b> ‘mountain’		LH~LL-H
2.4/5: <b>ame</b> ‘rain’, <b>ito</b> ‘thread’, <b>kata</b> ‘shoulder’, <b>nabe</b> ‘pot’, <b>sora</b> ‘sky’		LH~LH-L
2.4/5: <b>aki</b> ‘autumn’, <b>hashi</b> ‘chopstick’, <b>obi</b> ‘sash’, <b>umi</b> ‘ocean’, <b>matsu</b> ‘pine tree’, <b>saru</b> ‘monkey’, <b>shiru</b> ‘soup’	HL~HL-L	HL~HL-L

Diachronically, this system developed from the older *n+1* system, similar to that of ModJ, with the split of the older HL~HL-L pattern into two separate patterns (LH~LH-L and HL~HL-L) depending on the quality of vowels. The new LH~LH-L pattern developed from HL~HL-L by the shift of H tone one mora to the right, when the mora in question consisted of non-high vowels (/a e o/), e.g. \**sòrà-sà* > *sòrá-sà* ‘sky-DAT’; on the other hand, when the vowels were high (/i u/), this shift did not take place (*hásji-sà* ‘chopstick-DAT’, *óbi-sà* ‘sash-DAT’). The synchronic system of this dialect now has some exceptions: e.g. *kasji* ‘scum’, *cjizji* ‘father’, *cjiju* ‘dew’ appear as LH~LH-L, instead of the expected HL~HL-L pattern (Nitta 1994). Thus, the LH~LH-L and HL~HL-L patterns are no longer in complementary distribution, so that both of them are established as different accentual patterns in present-day Tsuruoka dialect. Such a split of an older HL~HL-L pattern into two contrastive patterns is also observed in some of the northern Tōhoku areas, such as Akita and Aomori prefectures, and part of Iwate prefecture.

As with ModJ, verbs and adjectives consist of two accentual patterns, regardless of their length. Table 11.5 shows the accentual patterns of verbs in Tsuruoka, in which underlined letters show predicate-final (intonational) high tones, which is not a part of the accentual patterns of the verbs.

**TABLE 11.5 ACCENT PATTERNS OF VERBS IN TSURUOKA DIALECT (BASED ON NITTA 1994)**

	Non-Past	Past	Negative	Conditional	Negative-Command
A.	<i>og-u</i> ‘put’ LH	<i>oi-da</i> LLH	<i>og-ane</i> LLH	<i>og-e<sup>m</sup>ba</i> LLH	<i>og-una</i> LLH
B.	<i>kag-u</i> ‘write’ HL	<i>kai-da</i> HLL	<i>kag-ane</i> LLH	<i>kag-e<sup>m</sup>ba</i> LHL	<i>kag-una</i> HLL
A.	<i>hadarag-u</i> ‘work’ LLLH	<i>hadara-ida</i> LLLLH	<i>hadarag-ane</i> LLLLH	<i>hadarag-e<sup>m</sup>ba</i> LLLLH	<i>hadarag-una</i> LLLLH
B.	<i>ujogas-u</i> ‘move (tr.)’ LLHL	<i>ujogasj-ita</i> LLHLL	<i>ujogas-ane</i> LLLLH	<i>ujogas-e<sup>m</sup>ba</i> LLLHL	<i>ujogas-una</i> LLHLL

## 11.2.2 Morphology

### 11.2.2.1 Nouns and case marking

As in a number of other Japanese dialects, the accusative, marked by **-o** in ModJ, is not marked by any particle in Tsuruoka, as illustrated in (3).<sup>6</sup>

- (3) a. *gohan-sa sjiō kage-ru*  
rice-ALL salt-Ø sprinkle-NONPAST  
‘(Someone) put some salt on the rice.’
- b. *tora ori-sa ire-ru*  
**tiger-Ø** cage-ALL put.in-NONPAST  
‘(Someone) put the tiger into a cage.’ (Satō 1994: 270–1)

This general tendency of zero-marking for the direct object is especially prevalent in the Tōhoku dialects, in which the accusative marker is generally missing. However, in cases in which the direct object is an animate noun, there is a strong tendency throughout the Tōhoku district to mark it with one of the suffixes *koto*, *godo*, *toko*, *dogo*, etc. (GAJ, vol. 1; Hidaka 2005), of which *dogo* is used in the Tsuruoka dialect. Thus, (3b) is readily paraphrased as *tora-dogo ori-sa ire-ru*. This is illustrated in (4):

- (4) a. *kacjan-dogo mi-da*  
mother-ACC see-PAST  
‘(Someone) saw the mother.’
- b. *inu-dogo kusari-sa cunai<sup>n</sup>-da*  
dog-ACC chain-DAT link-PAST  
‘(Someone) put the dog to the leash.’

According to Hidaka (2005), these new accusative markers are developing throughout the Tōhoku district through grammaticalization of a content word *godo* ‘thing’, or its metathesized form *dogo*, into a grammatical particle (Hidaka 2005: 82–3).<sup>7</sup> In fact, the newly developed direct object marker is now often used for inanimate objects such as *e* ‘house’ or *egi* ‘station’ in Tsuruoka as in *ano e-dogo mi-da* ‘(Someone) saw that **house**’ or *egi-dogo toq-ta* ‘(Someone) passed the **station**’, and so on.

In addition, in a number of dialects in the Tōhoku area, except for many in Iwate prefecture, the nominative, marked with **-ga** in ModJ, is also unmarked, as illustrated in the following examples from the Tsuruoka dialect:

- (5) a. *sanaji cjoR-sa kawaQ-ta*  
**pupa-Ø** butterfly-DAT change-PAST  
 ‘The pupa turned into a butterfly.’ (Satō 1994: 273)
- b. *toRcjan ima e-sa i-da*  
**father-Ø** now house-LOC be-NONPAST  
 ‘Father is now at home.’

In such dialects, both nominative and accusative cases can be zero-marked, and the Tsuruoka dialect is one such dialect.

Characteristically different from ModJ is the dative marker *-sa* in this dialect. The particle, as with the same particle observed throughout the Tōhoku district, developed from a goal-marking (allative) suffix *-sa*, as in (6). This particle is now undergoing a typical path of grammaticalization, i.e. an original noun *sama* ‘direction’ plus a particle cognate with CJ *-fe*, ModJ *-e*, contracted: *sama(-fe)* > *-saN* > *-sa* (Kobayashi 1997; Hidaka 2005).

- (6) *hinasji-no hoR-sa eg-e*  
 east-GEN direction-ALL go-IMP  
 ‘Go to the direction of the east.’

With subsequent semantic bleaching of the original meaning of direction, this suffix is now undergoing a process of semantic extension, by which its original meaning of direction is bleached out, and nondirectional spatial sense including locative ‘at’ or ‘on’ developed as shown in (7a,b):

- (7) a. *curuoga-sa ni.san-nizji i-ru*  
 Tsuruoka-LOC a.couple.of-day stay-NONPAST  
 ‘(Someone) stays in Tsuruoka for a couple of days.’
- b. *kami-sa zji kag-u*  
 paper-LOC letter write-NONPAST  
 ‘(Someone) writes letters on a piece of paper.’

In addition, the suffix is now used to mark non-spatial functions such as dative or benefactive as illustrated in (7c,d), or even purpose following a verb-stem as in (7e):

- (7) c. *hana-sa mi<sup>m</sup>zji kage-ru*  
 flower-DAT water put-NONPAST  
 ‘(Someone) gives some water to the flower.’
- d. *inu-sa esa ker-u*  
 dog-DAT food give-NONPAST  
 ‘(Someone) gives some food to the dog.’
- e. *hana<sup>m</sup>bi mi-sa eg-u*  
 firework see-DAT go-NONPAST  
 ‘(Someone) goes to see the fireworks.’ (Satō 1994)

Such semantic extension of *-sa* is comparable to the ModJ goal-marking suffix **-ni**, which is also used to express various abstract meanings, in addition to denoting a concrete sense of direction. This type of semantic extension of *-sa* is commonly observed in a number of other Tōhoku dialects (Kobayashi 1997; Hidaka 2005), in which the process of their grammaticalization is one of the most debated topics in Japanese dialect research.

## 11.2.2.2 Other particles

Other particles are rather similar to those of ModJ, which include the marker for genitive, which is marked with *-no*:

- (8) a. *ome-no hon* 'your book'  
you-GEN book  
b. *onajo-no heja* 'a woman's room'  
woman-GEN room

Similarly, inclusive 'also' is expressed with *-mo*, as in:

- (9) *ore-no odoRdo-mo isja<sup>n</sup>da*  
I-GEN brother-also doctor-COP  
'My brother is a doctor, too.'

Other particles include *-<sup>n</sup>de* for location, *-gara* 'from', *-ma<sup>n</sup>de* 'until, as far as', *-<sup>n</sup>dage* 'only', *-<sup>n</sup>jure* 'approximately', and *-<sup>m</sup>baqkari* 'only', some of which are exemplified in (10):

- (10) a. *gaqko-gara keq-ta*  
school-ABL return-PAST  
'(Someone) returned from the school.'  
b. *egi-ma<sup>n</sup>de arui-da*  
station-TERM walk-PAST  
'(Someone) walked up to the station.'  
c. *ome<sup>n</sup>dage-sa osjie-ru*  
you-only-DAT teach-NONPAST  
'(Someone) teaches something only to you.'  
d. *zjiq-pun-jure maq-ta*  
ten-minutes-approximately wait-PAST  
'(Someone) waited about ten minutes.'  
e. *ano mazji-sa<sup>m</sup>ba nego<sup>m</sup>baqkari i-ru*  
that town-LOC-TOP cat-only be-NONPAST  
'That town is full of cats.'

## 11.2.2.3 Verbs and adjectives

Similar to ModJ, the dialect has a distinction between consonant-stem verbs such as *kag-* 'write' or *eg-* 'go', and the vowel-stem verbs such as *ogi-* 'get up' or *age-* 'open'. Verb conjugation is shown in Table 11.6, in which (a) and (b) correspond to consonant-stem verbs in ModJ, and (c) to vowel-stem verbs.

Notice that verbs in (a) show a quinquagrade conjugation pattern, while those in (c) show a monograde one. Especially noteworthy here is the behavior of the verbs in (b), consonant-stem verbs the stem of which ends with /r/ (hereafter '/r/-stem verbs'): their negative forms are *to-ne* and *u-ne*, instead of the \**tor-ane* and \**ur-ane* expected based on ModJ **tor-anai** and **ur-anai**. The reason for this may be the similarity of the other patterns of the same paradigm of these verbs with those of vowel-stem verbs in (c) (i.e. they both end with *-ru* (NONPAST), *-re* (IMP), *-ro* (VOL), and *-re<sup>m</sup>ba* (COND)).

**TABLE 11.6 VERB CONJUGATION IN THE TSURUOKA DIALECT (I)**

	Non-Past	Past	Negative	Imperative	Volitional	Conditional	Desiderative
a.	<i>kag-u</i> ‘write’ <i>eg-u</i> ‘go’	<i>kae-da</i> <i>eo-ta</i>	<i>kag-ane</i> <i>eg-ane</i>	<i>kag-e</i> <i>eg-e</i>	<i>kag-o</i> <i>eg-o</i>	<i>kag-e<sup>m</sup>ba</i> <i>eg-e<sup>m</sup>ba</i>	<i>kag-ide</i> <i>eg-ide</i>
b.	<i>tor-u</i> ‘take’ <i>ur-u</i> ‘sell’	<i>toQ-ta</i> <i>uo-ta</i>	<i>to-ne</i> <i>u-ne</i>	<i>tor-e</i> <i>ur-e</i>	<i>tor-o</i> <i>ur-o</i>	<i>tor-e<sup>m</sup>ba</i> <i>ur-e<sup>m</sup>ba</i>	<i>tor-ide</i> <i>ur-ide</i>
c.	<i>ogi-ru</i> ‘get up’ <i>age-ru</i> ‘open’	<i>ogi-da</i> <i>age-da</i>	<i>ogi-ne</i> <i>age-ne</i>	<i>ogi-re</i> <i>age-re</i>	<i>ogi-ro</i> <i>age-ro</i>	<i>ogi-re<sup>m</sup>ba</i> <i>age-re<sup>m</sup>ba</i>	<i>ogi-de</i> <i>age-de</i>

Thus, the negative forms of the /r/-stem verbs in this dialect seem to have become similar to those of the vowel-stem verbs by analogy.

On the other hand, a tendency to merge the conjugational paradigms of /r/-stem verbs and of vowel-stem verbs is also observed on the part of the latter. Notice in (c) that some conjugational forms of the vowel-stem verbs in Tsuruoka are different from ModJ; while the imperative and volitional forms of these verbs in ModJ are **oki-ro** (IMP) ‘Get up!’ and **oki-yō** (VOL) ‘I’ll (Let’s) get up’ respectively, those in Tsuruoka are *ogi-re* (IMP), and *ogi-ro* (VOL). This is also due to the gradual assimilation of vowel-stem verbs (c) into the paradigm of the /r/-stem verbs in (b).

In fact, this tendency, in which the vowel-stem verbs come to conjugate similarly to consonant-stem verbs is not peculiar to the Tsuruoka dialect, but is shared with numbers of other dialects in Japan; i.e. certain dialects in the eastern Tōhoku area as well as a number of dialects in the Kyūshū region also display such a tendency, which is now recognized as a cross-dialectal trend (or ‘drift’) in a number of Japanese dialects, whereby attrition of the vowel-stem verbs is gradually taking place through analogy with the consonant-stem verbs.

Irregular verbs *su-* ‘do’ and *ku-* ‘come’ show the conjugation patterns in Table 11.7. Note that both of these verbs conjugate differently from ModJ, the paradigm of which is listed in Table 11.8. Here, the vowel alternation for the stem of the irregular verb *su-ru* ‘do’ is especially noteworthy. While the corresponding ModJ stem has two allomorphs (**su** or **shi**), its Tsuruoka cognate has five: /sji/ (PAST, DES), /su/ (NONPAST), /sa/ (NEG), /se/ (IMP, COND), and /so/ (VOL), which, as a result, means that this verb conjugates similarly to consonant-stem verbs shown in Table 11.6(a).

**TABLE 11.7 VERB CONJUGATION IN THE TSURUOKA DIALECT (II): IRREGULAR VERBS**

Non-Past	Past	Negative	Imperative	Volitional	Conditional	Desiderative
<i>su-ru</i> ‘do’	<i>sji-da</i>	<i>sa-ne</i>	<i>se (sjire)</i>	<i>so</i>	<i>se<sup>m</sup>ba</i>	<i>sji-de</i>
<i>ku-ru</i> ‘come’	<i>ki-da</i>	<i>ko-ne</i>	<i>koe</i>	<i>ko</i>	<i>koe<sup>m</sup>ba</i>	<i>ki-de</i>

**TABLE 11.8 VERB CONJUGATION OF IRREGULAR VERBS IN STANDARD JAPANESE**

Non-Past	Past	Negative	Imperative	Volitional	Conditional	Desiderative
<b>su-ru</b> ‘do’	<b>shi-ta</b>	<b>shi-nai</b>	<b>shi-ro</b>	<b>shi-yō</b>	<b>su-reba</b>	<b>shi-tai</b>
<b>ku-ru</b> ‘come’	<b>ki-ta</b>	<b>ko-nai</b>	<b>koi</b>	<b>ko-yō</b>	<b>ku-reba</b>	<b>ki-tai</b>

As for adjectives, the above-mentioned monophthongization of /ai/ and /oi/ to /ɛ/ produced the conclusive and attributive forms *aoge* ‘red’ and *sjiRre* ‘white’, corresponding to ModJ **aka-i** and **shiro-i** respectively. These forms *aoge* and *sjiRre* are reanalyzed as stems to produce such conjugational forms as *aoge-ke*, *sjiRre-ke* (past), *aoge-gu-ne*, *sjiRre-gu-ne* (negative), *aoge-gu-de*, *sjiRre-gu-de* (gerund), *aoge<sup>m</sup>-ba*, *sjiRre<sup>m</sup>-ba* (conditional), *aoge-gu-naru* ‘become red’, *sjiRre-gu-naru* ‘become white’. This is in stark contrast to ModJ, in which the root forms such as **aka-** ‘red’ or **shiro-** ‘white’ serve as the stems for all the forms of adjective conjugation, as in **aka-i** ‘red’, **shiro-i** ‘white’, **aka-ka<sup>o</sup>-ta**, **shiro-ka<sup>o</sup>-ta** (past), **aka-ku nai**, **shiro-ku nai** (negative), **aka-ku-te**, **shiro-ku-te** (gerund), **aka-ke-reba**, **shiro-ke-reba** (conditional), and **aka-ku naru** ‘become red’, **shiro-ku naru** ‘become white’. Such a feature of the pattern of adjectival conjugation in Tsuruoka represents the tendency of levelling of the adjectival conjugational paradigms often observed in other dialects in Japan (Onishi 1999: 59). Examples that use the root as in ModJ, probably influenced by ModJ, do occur.

Similar to ModJ, non-past forms for conclusive vs. attributive predicates in this dialect are not distinguished for the predicates so far mentioned. However, in contrast to ModJ, in which **na**-type adjectives maintain the distinction, as in **shizuka-da** (conclusive) ‘quiet’ and **shizuka-na** (attributive), the Tsuruoka dialect does not have such a distinction. Thus, the conclusive forms *sjizjiga<sup>n</sup>-da* ‘quiet’ and *kiRre<sup>n</sup>-da* ‘pretty’ are also used as attributive forms as in *sjizjiga<sup>n</sup>-da dogo* ‘a quiet place’ and *kiRre<sup>n</sup>-da hito* ‘a pretty person’ in Tsuruoka. Such “*da*”-type adjectives in Tsuruoka conjugate as in *sjizjiga<sup>n</sup>-da-ke*, *kiRre<sup>n</sup>-da-ke* (past), *sjizjiga<sup>n</sup>-de-ne*, *kiRre<sup>n</sup>-de-ne* (negative), *sjizjiga<sup>n</sup>-de*, *kiRre<sup>n</sup>-de* (gerund), *sjizjiga<sup>n</sup>-da<sup>m</sup>-ba*, *kiRre<sup>n</sup>-da<sup>m</sup>-ba* (conditional), *sjizjiga-naru* ‘become quiet’, *kiRre-naru* ‘become pretty’.

### 11.2.3 The basic sentence

#### 11.2.3.1 Declarative

As in other Japanese dialects, the basic word order of this dialect is SOV, with tense, aspect, and modality markers attached after verbs and adjectives.

#### 11.2.3.2 Tense and aspect

Auxiliary verbs expressing tense and aspect in the Tōhoku dialects are topics of particular interest. As with ModJ, the non-past tense is expressed by *-ru* for vowel-stem verbs and *-u* for consonant-stem ones, as in *mi-ru* ‘see’, *ne-ru* ‘sleep’, *og-u* ‘put’, *das-u* ‘take out’, *yom-u* ‘read’, *ur-u* ‘sell’, *su-ru* ‘do’, and *ku-ru* ‘come’. Similarly, the past tense is marked by *-ta* (or *-da* due to intervocalic voicing of stem-initial *t*) as in *mi-da* ‘saw’, *ne-da* ‘slept’, *oi-da* ‘put’, *das-i-ta* ‘took out’, *jon-da* ‘read’, *uo-ta* ‘sold’, *sji-da* ‘did’, and *ki-da* ‘came’.

However, what is different from ModJ is that for stative predicates such as adjectives and nominal verbs *-(o)ke*, instead of *-ta* or *-da*, is used to mark the past tense in the Tsuruoka dialect, as in *aoge-ke* ‘red-PAST’, *ho<sup>o</sup>se-ke* ‘thin-PAST’, *sjizjiga<sup>n</sup>-da-ke* ‘quiet-PAST’, or *kiRre<sup>n</sup>-da-ke* ‘pretty-PAST’, which is illustrated below:

- (11) a. *ano umi aoi-**o**ke*  
 that ocean-Ø blue-PAST  
 ‘That ocean was blue.’

- b. *son togi heja sjizjiga-<sup>n</sup>da-ke*  
 that time room-Ø quiet-PAST  
 ‘The room was quiet then.’

The corresponding ModJ forms are all marked with **-ta**, as in **aka-kaq-ta** ‘red-PAST’, **ao-kaq-ta** ‘blue-PAST’, **shizuka-daq-ta** ‘quiet-PAST’, and **kirei-daq-ta** ‘pretty-PAST’.

Among the stative verbs, one, *i-ru* ‘be, exist’, shows especially interesting features in its tense marking. Contrary to other verbs, for which *-ru* ending forms are used for non-past, which includes the present, future, and habitual, the *-ru* form of *i-ru* in this dialect expresses only the future and habitual tense, while the *-da* form expresses both the temporary (non-habitual) present and the past tense; thus, *i-da* expresses both ‘is/are’, meaning temporary existence, and ‘was/were’, while *i-ru* expresses ‘will be’ and habitual ‘is/are’. This is illustrated in (12), in which (12a) indicates a temporary state and (12b) a habitual one:

- (12) a. *ima e-sa nego i-da*  
 now house-LOC cat be-*da*  
 ‘There is a cat in the house now.’
- b. *ano mazji-<sup>n</sup>da<sup>m</sup>ba nego-<sup>m</sup>baqkari i-ru*  
 that town-TOP cat-only be-NONPAST  
 ‘There are only cats (there are so many cats) in that town.’

This distinction may derive from a realis/irrealis distinction. However, it is observed only by the verb *i-ru* ‘be, exist’. Other verbs including the verb *ar-u* ‘be, exist’, the equivalent verb of existence used for non-animate subjects, do not exhibit it; they only show a past/non-past distinction similar to ModJ.

The continuative form is marked by putting *-oda* (or *-ota*) after a verb root, as in *kae-oda* ‘write-CONT’, *mi-oda* ‘see-CONT’, *age-oda* ‘open-CONT’, and *i-oda* ‘be-CONT’. The suffix *-deda* is also possible for continuative as in *kae-deda* ‘write-CONT’ or *mi-deda* ‘see-CONT’. Such forms developed from gerund *-de* + *i-ru* ‘be, exist’ + the tense marker *-da* (*\*-de i-da* > *-deda* > *-oda*). Thus, similar to *i-ru* ‘be, exist’, *-oda* is also used for both temporary present and past tense, e.g. *kae-oda* can indicate both ‘is writing’ and ‘was writing’ in this dialect. Example (13a) shows that the continuative verb forms are used for the present tense even though the past tense marker *-da* is attached.

To express explicitly the past tense, *-ke* can be further added after *-oda* in Tsuruoka as in *kaeq-daq-ke* ‘write-CONT-PAST’, which definitely means ‘was writing’.

- (13) a. *toRcjan e-sa iq-da ga*  
 father house-LOC be-CONT QUESTION  
 ‘Is your father at home?’
- b. *senseR ome-dogo homeq-daq-ke*  
 teacher you-ACC praise-CONT-PAST  
 ‘Your teacher was praising you.’ (National Institute of Informatics 2008)

### 11.2.3.3 Negation

The negative is marked with *-ne*, which corresponds to ModJ **-nai**, as in *kag-a-ne* ‘write-NEG’, *age-ne* ‘open-NEG’, *ogi-ne* ‘get up-NEG’. As shown in Table 11.6, /r/-stem verbs, such as *tor-* ‘take’ and *kir-* ‘cut’, drop the stem-final /r/ in their negative forms, namely *to-ne* ‘take-NEG’ and *ki-ne* ‘cut-NEG’.

The negative of adjectives is expressed by *-gu-ne*, as in *jasu-gune* ‘cheap-NEG’ or *adarasji-gune* ‘new-NEG’, corresponding to ModJ *yasu-ku nai* ‘cheap-NEG’ or *atarashi-ku nai* ‘new-NEG’. The negative forms for adjectival nominals are *-<sup>n</sup>de-ne*, as in *sjizjiga-<sup>n</sup>de-ne* ‘quiet-NEG’, corresponding to ModJ *-janai* as in *shizuka-janai* ‘quiet-NEG’.

#### 11.2.3.4 Modality

Conjecture (epistemic modality) is expressed by a modal auxiliary *-(n)dero* attached after conclusive forms as in *kag-u-<sup>n</sup>dero* ‘may write’, *eg-u-<sup>n</sup>dero* ‘may go’ for verbs, or *taoge-<sup>n</sup>dero* ‘may be expensive’, and *adarasji-<sup>n</sup>dero* ‘may be new’ for adjectives. The auxiliary *-ro* is also possible as in *kag-u-ro*, *eg-u-ro*, *taoge-ro*, *adarasji-ro* ‘may be new’:

- (14) *asjita-wa nara-sa eg-u-ro*  
 tomorrow-TOP Nara-ALL go-NONPAST-MOD  
 ‘(Someone) will go to Nara tomorrow.’ (Onishi 1994: 183)

*-kamosjinne* is also used for conjecture, as in *jom-u-kamosjinne* ‘someone may read (something)’.

Volition is expressed by *-(r)o* as in *kag-o* ‘I intend to write’, *ne-ro* ‘I intend to sleep’, as shown below:

- (15) *asjita haje-sage ne-ro*  
 tomorrow early.NONPAST-because sleep-VOL  
 ‘I will go to bed since I (have to) get up early tomorrow.’ (Onishi 1994: 196)

For non-stative verbs, the auxiliary *-ke* is used to show recollection (and reporting) of some past actions or events observed or experienced by the speaker in the past, as in *su-ru-ke* ‘I remember (and hereby report) someone doing something’, *eg-u-ke* ‘I remember someone going’, and *age-o-ke* ‘I remember someone opening it’ (Shibuya 1994). However, used with stative predicates, as illustrated in 11.2.3.2, *-ke* simply indicates past tense.

In ModJ, the same particle is also used for recollecting the past event as in *shi-ta-oke* ‘I remember doing something’, and *ake-ta-oke* ‘I remember opening it’, *genki yo-kaota-oke* ‘I remember (someone) was in good spirits (or healthy)’. However, ModJ *-oke* almost always follows after a past tense marker *-ta*. In contrast, in the Tsuruoka dialect, simply recollecting the past event is expressed by the non-past conclusive plus the auxiliary *-ke*. On the other hand, this particle attached after the past tense marker *-da*, as in *sjida-ke* ‘I remember and report someone having done something’ in Tsuruoka, indicates that the actions or events observed and reported had already finished at the time of observation (Shibuya 1994). This distinction cannot be made by ModJ *-oke*.

Other evidentiality markers in this dialect include *soR<sup>n</sup>da*, *joR<sup>n</sup>da*, and *doja*. Appearance is expressed by *soR<sup>n</sup>da*, as in *kagi-soR<sup>n</sup>da* ‘It looks like (someone) is going to write’, *taoge-soR<sup>n</sup>da* ‘It looks expensive’, while it is also expressed by *-joR<sup>n</sup>da* as in *taoge-joR<sup>n</sup>da* ‘It seems expensive’, *adarasji-joR<sup>n</sup>da* ‘It seems to be new’. Hearsay is expressed by *-doja* as in *taoge-doja* ‘I hear that it is expensive’ and *adarasji-doja* ‘I hear it is new’.

A deontic modality expression ‘must’ is expressed by attaching *-nemanε* to the same stem that negative *-ne* is attached to:

- (16) a. *rogu-zji-<sup>n</sup>de ogi-nemane*  
 6-o'clock-LOC get.up-must  
 '(Somebody) has to get up at 6 o'clock.'
- b. *iR fugu ki-de ega-nemane*  
 good clothes wear-GER go-must  
 '(Somebody) must go wearing good clothes.'

Other deontic modality markers include gerund *-<sup>n</sup>de + -mo iR*, as in *jon-de-mo iR* 'may write', and gerund *-<sup>n</sup>de + -wa dami-<sup>n</sup>da*, as in *jon-<sup>n</sup>de-wa dami-<sup>n</sup>da* 'must not write'.

Desiderative 'want to do (something)' is expressed by the suffix *-(o)de*, as in *egi-(o)de* 'want to go', *age-(o)de* 'want to open', *sjj-(o)de* 'want to do', and *ki-(o)de* 'want to come', as illustrated below:

- (17) *asjita-mo ko-sa ki-ode no*  
 tomorrow-also here-ALL come-DESID PART  
 'I want to come here again tomorrow.' (Onishi 1994: 170)

### 11.2.3.5 Potential

Potential expressions in Tsuruoka are formed by the auxiliary *-ae-ru*, while its negative form is *-ae-ne*, as illustrated below:

- (18) a. *zjibun zji heda-<sup>n</sup>da-hage kag-ae-ne*  
 myself letter (writing) unskilled-NONPAST-because write-POT-NEG  
 'Since I am not good at writing letters, I cannot write it.'
- b. *agaru-<sup>m</sup>ba kag-ae-ru*  
 bright-if write-POT-NONPAST  
 'If it is bright, (someone) can write (something).' (Onishi 1994: 201–2)

Certain dialects in Japan make a formal distinction between 'ability' potential and 'circumstantial' potential; the former indicating that the agent of a certain action has the ability to achieve the action, the latter that a certain event or action is achievable because the situation allows it. In ModJ, they are both expressed with *-eru* or *-rareru* as in *yom-eru* 'can read' or *ake-rareru* 'can open'. However, a number of dialects in the Tōhoku region are known to make the distinction. For example, in the Akita dialect, *kageru* 'can write' simply indicates that someone has an ability to write something, while *kagunie* 'can write' indicates that the situation allows someone to write something (Hidaka 2000: 103). However, it seems that such a distinction is no longer made in the present-day dialect of Tsuruoka.<sup>8</sup>

### 11.2.3.6 Non-declarative sentence types

Questions are made by attaching *ga*, as in (13a), or *-na*, as in (19):

- (19) *dogo-sa (~ do-sa) eg-u na*  
 where-ALL go-NONPAST QUESTION  
 'Where are you going to?'

Commands are marked by the imperative suffix *-e* or *-re*, as in *kag-e* 'Write!', *tor-e* 'Take it!', *ogi-re* 'Get up!', *mi-re* 'See!', *de-re* 'Go out!'.

Proposal 'let's' is formed with the volitional *-o* or *-ro*, as in *kag-o* 'Let's write', *ogi-ro* 'Let's get up'.

### 11.2.4 Topic and emphasis

Topic can be expressed by zero-marking or *-wa*, as in (20):

- (20) *kjoR<sup>n</sup>de-wa* *ane* *hito-ri-sa* *imodo* *hito-ri-da*  
 siblings-TOP elder.sister one-person-and younger.sister one-person-COP  
 'As for siblings, I have one older sister and one younger one.' (Satō 1994: 273)

Topic may also be marked by emphatic *-<sup>n</sup>da<sup>m</sup>ba* as in (21), in which case a contrastive meaning may be particularly emphasized.

- (21) *ano hito-sa-<sup>n</sup>da<sup>m</sup>ba* *ai-da-gu.ne*  
 that person-DAT-TOP see-DESID-NEG  
 'I don't want to see that person.' (Satō 1994: 271)

Thus, instead of *sono hon-wa e-sa aru* 'The book is at our house', one can say:

- (22) *sono hon-<sup>n</sup>da<sup>m</sup>ba* *e-sa* *ar-u*  
 that book-TOP house-LOC be-NONPAST  
 'As for that book, we have it at our house.'

### 11.2.5 Passive and causative

The passive, including the adversative passive, is marked by the suffixes *-rae-* for vowel-stem verbs and *-ae-* for consonant-stem verbs as in *mi-rae-ru* 'be seen' and *kag-ae-ru* 'be written', as shown in the following examples:

- (23) a. *doro<sup>n</sup>bo-gara* *kane* *tor-ae-da*  
 thief-ABL money take-PASS-PAST  
 '(Someone) had money taken by a thief.' (Satō 1994: 279)
- b. *ano hito-gara* *kaisja* *jame-rae-de* *sumaō-da*  
 that person-ABL company quit-PASS-GER finish-PAST  
 '(I was adversatively affected by the fact that) that person left the company.'

Particularly interesting is the fact that, as illustrated in (23), the agent is marked by the ablative particle *-gara* in this dialect. Both examples in (23) are rendered in ModJ with *-ni* 'by': **dorobō-ni kane-o tor-are-ta** and **ano hito-ni kaisha-o yame-rare-ta**.

The causative is marked with the suffix *-ase-* for consonant-ending verbs, as in *kag-ase-ru* 'cause to write' and *eg-ase-ru* 'cause to go', and *-sase-* for vowel-ending verbs, as in *ogi-sase-ru* 'cause to get up' and *age-sase-ru* 'cause to open'. As illustrated in (24a), the causee is marked by *-sa*. The causative passive is marked by *-(s)ase-rae-* as in (24b):

- (24) a. *toRcjan taroR-sa* *tejami* *kag-ase-da*  
 father Tarō-DAT letter write-CAUS-PAST  
 'Father made Tarō write a letter.'
- b. *sjacjoR-gara* *kaisja* *jame-sase-rae-da*  
 president-ABL company quit-CAUS-PASS-PAST  
 '(Someone) was made to quit the company by the president.' (Satō 1994: 279)

### 11.2.6 Subordination

Similar to ModJ, relative clauses are positioned to the left of the head noun as in (25):

- (25) *toRkjoR-sa eg-u me a-o*  
 Tokyo-ALL go-NONPAST before meet-VOL  
 ‘I will (Let’s) meet before I go to Tokyo.’

Embedded clauses may be marked by a nominalizer, the most frequent of which is *na* in this dialect, which roughly corresponds to ModJ **no**:

- (26) a. *asjita sjikeN a-N na o<sup>m</sup>boi-oda*  
 tomorrow test be-NONPAST COMP remember-CONT  
 ‘Do you know that we will have an exam tomorrow?’  
 b. *kino kaQ-ta na sono hon-<sup>n</sup>da*  
 yesterday buy-PAST COMP that book-COP  
 ‘What (someone) bought yesterday was that book.’

Quotations are marked by *ote* as in:

- (27) *ano hito toRkjoR-sa eg-u ote o<sup>m</sup>boi-oda*  
 that person Tokyo-ALL go-NONPAST QUOT remember-CONT  
 ‘Do you know that that person is going to Tokyo?’

Conditional/hypothetical clauses are marked by *-<sup>n</sup>da<sup>m</sup>ba*, as in (28):

- (28) a. *mi<sup>n</sup>zji-sa cuge-<sup>n</sup>da<sup>m</sup>ba age<sup>n</sup>da iro naQ-ta*  
 water-ALL dip.into-when that color become-PAST  
 ‘When (someone) put it into water, it changed into such a color.’  
 b. *kag-u na-<sup>n</sup>da<sup>m</sup>ba cjan<sup>to</sup> kag-e*  
 write-NONPAST COMP-if properly write-IMP  
 ‘If you write it, write it properly.’

A subordinate clause indicating reason/cause is introduced by *-sage* or *-hage*, which are largely interchangeable:

- (29) *omosji tere<sup>m</sup>bi jaQ-teQ-sage (~-hage) ki-de mi-re*  
 interesting television do-CONT-because come-GER see-IMP  
 ‘Just come and see, since an interesting TV program is now showing.’  
 (National Institute of Informatics 2008)

Concessive clauses are expressed most commonly with *-<sup>n</sup>domo* as in (30a). They can also be expressed with *-nasa*, a combination of the above-mentioned nominalizer *na* and the particle *-sa* as in (30b), which roughly corresponds to ModJ **-no-ni**:

- (30) a. *asjita kuN-<sup>n</sup>domo zjikaN oso-gu.naru zo*  
 tomorrow come-though time late-become PART  
 ‘Though I will come tomorrow, the time will be late.’ (Onishi 1994: 181)  
 b. *mad-e ote ju-nasa, mai-de ke-ne*  
 wait-IMP QUOT say.NONPAST-though wait-GER give-NEG  
 ‘Though I told (someone) to wait, he/she does not wait for me.’  
 (Sato 1994: 274)

## 11.3 EI DIALECT

### 11.3.1 Phonology

#### 11.3.1.1 Vowels

In Ei, as well as most other dialects in Kagoshima prefecture, diphthongs /ai/ and /oi/ are monophthongized and subsequently shortened to /e/. This is illustrated in the following examples, in which the corresponding words in ModJ are shown in bold:

(31)	Ei		ModJ	Ei		ModJ
	<i>degon</i>	‘radish’	<b>daikon</b>	<i>jase</i>	‘vegetable’	<b>yasai</b>
	<i>esaQ</i>	‘greeting’	<b>aisatsu</b>	<i>kure</i>	‘black’	<b>kuroi</b>

However, this monophthongization has many exceptions in the present-day Ei dialect: *hai* ‘lung’ **hai** (vs. *he* ‘ash’ **hai**), *k<sup>w</sup>ai* ‘meeting’ **kai** (vs. *ke* ‘shellfish’ **kai**), *taisoR* ‘physical exercises’ **taisō**, *saiban* ‘trial, justice’ **saiban**, as well as most loanwords, as in *haikingu* ‘hiking’, *saireN* ‘siren’, and *raisu* ‘rice’.

Especially noteworthy in Ei, as well as a number of other dialects in Kagoshima prefecture, is the deletion of word-final high vowels /i/ and /u/, leading to syllable-final consonants, with neutralization of obstruents as /Q/:

(32)	Ei		ModJ	Ei		ModJ
	<i>kaQ</i>	‘persimmon’	<b>kaki</b>	<i>miQ</i>	‘water’	<b>mizu</b>
	<i>kuQ</i>	‘stem’	<b>kuki</b>	<i>kuQ</i>	‘shoe’	<b>kutsu</b>
	<i>kuQ</i>	‘mouth’	<b>kuchi</b>	<i>maQ</i>	‘pine tree’	<b>matsu</b>
	<i>misAQ</i>	‘cape, point’	<b>misaki</b>	<i>esaQ</i>	‘greeting’	<b>aisatsu</b>

Phonetically, the word-final /Q/ is realized as unreleased [t̚], or assimilated to the initial consonant of the following morpheme.

Similar to ModJ, word-internal high vowels /i, u/ are regularly devoiced between voiceless consonants, as in *sjika* [ʃika] ‘deer’, *hito* [hit̚o] ‘person’, *hikai* [hika̠i] ‘light’, *huka-ga* [f̥ukaga] ‘deep’, and *futo-ga* [f̥utoga] ‘big’.

As for words ending in ModJ **shi**, **ji**, **su**, and **zu**, the underlying final vowels are devoiced or often deleted (often accompanied by a phonetic lengthening of the preceding vowel) as in *isji* [iʃi̠] or *isj* [i:ʃ] ‘stone’ **ishi**, *hasj* [haʃi̠] or [ha:ʃ] ‘bridge’ **hashi**, *kosj* [ko:ʃ] ‘waist’ **koshi**, *iwasj* [iwa:ʃ] ‘sardine’ **iwashi**, *k<sup>w</sup>azj* [k<sup>w</sup>a:ʒ] ‘fire’ **kaji**, *sazj* [sa:ʒ] ‘spoon’ **saji**, and *karas* [kara:s] ‘crow’ **karasu**. Devoicing of word-final fricatives results in neutralization between underlying /z/ and /s/ in Ei, e.g. *kaz* ‘number’ **kazu** and *kas* ‘scum’ **kasu** are both pronounced as [ka:s], or *suz* ‘bell’ **suzu** and *sus* ‘soot’ **susu** as [su:s].

Interestingly, the resulting *sj* [ʃ] and *s* [s] are often changed into *i* when followed by other consonants; thus, when the nominative marker *-ŋa* or the genitive marker *-no* are attached to such words as *isj* ‘stone’, *hasj* ‘chopstick’, *iwasj* ‘sardine’, *sus* ‘soot’, *kaz* ‘number’, and *karas* ‘crow’, they can be changed into *ii*, *hai*, *iwai*, *sui*, *kai*, and *karai* as in *ii-ŋa*, *ii-no*, *hai-ŋa*, *hai-no*, *iwai-ŋa*, *iwai-no*, *sui-ŋa*, *sui-no*, *kai-ŋa*, *kai-no*, *karai-ŋa*, *karai-no*, respectively.<sup>9</sup>

When the consonants before the deleted high vowels are nasals /m/ or /n/, they are neutralized as /N/, as illustrated in (33):

**TABLE 11.9 MORPHONOLOGICAL CONSONANT CHANGES IN THE EI DIALECT**

ModJ phonemes	CJ sources	Ei underlying	word-final or before C-initial particles
<b>k/t/ch/ts</b>	<i>k/t</i>	<i>k/t/tj/c</i>	<i>Q</i>
<b>sj/j</b>	<i>s/z</i>	<i>sj/zj</i>	<i>sj ~ i</i>
<b>s/z</b>	<i>s/z</i>	<i>s/z</i>	<i>s ~ i</i>
<b>m/n</b>	<i>m/n</i>	<i>m/n</i>	<i>N</i>
<b>g/b/z</b>	<i>g<sup>h</sup>/b/d</i>	<i>ŋ<sup>h</sup>/b/dj</i>	<i>N ~ Q</i>

- (33) Ei                    ModJ    Ei                    ModJ    Ei                    ModJ  
*mi<sub>N</sub>* ‘ear’    **mimi**    *i<sub>N</sub>* ‘dog’    **inu**    *no<sub>N</sub>* ‘flea’    **nomi**  
*ka<sub>N</sub>* ‘paper’    **kami**    *a<sub>N</sub>* ‘net’    **ami**    *u<sub>N</sub>* ‘ocean’    **umi**  
*ka<sub>ŋa<sub>N</sub></sub>* ‘mirror’    **kagami**    *hasa<sub>N</sub>* ‘scissors’    **hasami**    *ko<sub>jo<sub>N</sub></sub>* ‘calendar’    **koyomi**

What is different from most of the other dialects in Kagoshima prefecture is the fact that for some speakers in Ei, /n/, instead of /Q/, is used when the consonants preceding the deleted high vowels are underlying Ei nasal /ŋ/ or voiced obstruents /b/ or /dj/, which correspond to ModJ **g**, **b**, and **z/j**, respectively.

- (34) Ei                    ModJ    Ei                    ModJ  
*ku<sub>N</sub> ~ ku<sub>Q</sub>* ‘nail’    **kugi**    *mi<sub>N</sub> ~ mi<sub>Q</sub>* ‘water’    **mizu**  
*mi<sub>N</sub> ~ mi<sub>Q</sub>* ‘right (side)’    **migi**    *a<sub>N</sub> ~ a<sub>Q</sub>* ‘taste’    **aji**  
*usa<sub>N</sub> ~ usa<sub>Q</sub>* ‘rabbit’    **usagi**    *aso<sub>N</sub> ~ aso<sub>Q</sub>* ‘play’    **asobi**  
*una<sub>N</sub> ~ una<sub>Q</sub>* ‘eel’    **unagi**    *to<sub>N</sub> ~ to<sub>Q</sub>* ‘fly’    **to<sub>bu</sub>**

For a number of Ei dialect speakers, both members of each pair in (34) seem to be acceptable.

The above morphophonological changes are summarized in Table 11.9, with a contrast with ModJ consonants.

In contrast to the Tsuruoka dialect, as well as other dialects in the Tōhoku region, in which the long vowels are phonetically much shorter than those in ModJ, the long vowels in Ei are not pronounced so short; thus, *koRri* ‘ice’ **kōri**, *kuRki* ‘air’ **kūki**, and *cjiRzu* ‘cheese’ **chizu** are [ko:ri], [ku:ki], and [cɟi:zu], respectively.

### 11.3.1.2 Consonants

In ModJ, the historical distinction between voiced affricates and fricatives before high vowels /u/ and /i/ is no longer kept. Thus, for example, the older distinction of *zi* vs. *di* as in *kazi* ‘fire’ and *kadi* ‘rudder’ in CJ is no longer distinguished in ModJ, where they have merged as **ji** (phonetically realized as [ɟi] ~ [dɟi], as in [kaɟi] ~ [kadɟi]). Similarly, the CJ distinction between *zu*, as in *nezumi* ‘mouse’ or *suzuri* ‘inkstone’, and *du*, as in *aduki* ‘red beans’, *midu* ‘water’, or *tudumi* ‘*tsuzumi*-drum’, is no longer kept in ModJ, where they have merged as **zu** (phonetically [zu] ~ [dzu]).

In contrast, Ei still keeps the CJ *zi di zu du* four-way distinction as /zj/ dji zu dju/. So, CJ *zi* is pronounced *zji* [ɟi] as in *zji* ‘letter’ **ji**, *huzji* ‘Mt. Fuji’ **fujji**, *kuzjira* ‘whale’ **kujjira**, while CJ *di* is pronounced as *dji* [dɟi] as in *dji* ‘hemorrhoid’ **ji**, *hudji* ‘wisteria’ **fujji**, *hanadji* ‘nose-bleed’ **hanajji** (in this case, *cji* ‘blood’ is changed to *dji* due to *rendaku*

voicing), and *cjidjin* ‘shrink’ **chijimu**. Similarly, CJ *zu* is pronounced *zu* [zu] as in *suzume* ‘sparrow’ **suzume**, *suzui* ‘inkstone’ **suzuri**, and *nezun* ‘mouse’ **nezumi**, while CJ *du* is pronounced as *dju* [dʒu] as in *mika-dju* ‘crescent moon’ **mika-zuki** (in this case, *cu* in *cu*Q ‘moon’ becomes *dju* due to *rendaku* sequential voicing), *adju*Q ‘red bean’ **azuki**, *cjudjun* ‘*tsuzumi*-drum’ **tsuzumi**.

The differences between /zji/ and /dji/ or /zu/ and /dju/ are especially clear word-finally, as /zji/ and /zu/ appear as *sj* [ʃ] and *s* [s], respectively, while /dji/ and /dju/ as *o*, due to the word-final deletion of /i/ and /u/ mentioned above. So, underlying *zji* appears word-finally as *sj* as in *kwasj* [kwa:ʃ] ‘fire’ **kaji**, *kisj* [ki:ʃ] ‘pheasant’ **kiji**, *sjoRsj* [sjo:ʃ] ‘shōji screen’ **shōji**, while underlying *dji* as *o*, as in *kaQ* [ka:t] ‘rudder’ **kaji**, *aQ* [a:t] ‘taste’ **aji**, and *hiQ* [çi:t] ‘elbow’ **hiji**. Similarly, underlying /zu/ word-finally appears as [s] as in *kaz* [ka:s] ‘number’ **kazu**, *kiz* [ki:s] ‘cut, hurt’ **kizu**, while /dju/ as /q/ as in *miQ* [mi:t] (or /miN/) ‘water’ **mizu**, and *kuQ* [ku:t] ‘rubbish’ **kuzu**.

We have seen above in 11.2.1.1 that the Tsuruoka dialect is a *hitotsugana* (‘one-letter’) system, in which the distinctions of the pronunciation of all four letters ぢ (CJ *di*), づ (*du*), じ (*zi*), and ず (*zu*) are totally lost. In contrast, all of these distinctions are maintained in Ei. Thus, this type of dialect is often called *yotsugana* (‘four-letter’) dialect, for which the Ei dialect is the archetypical example. This is summarized in Table 11.10.

The *yotsugana* dialects are distributed in many areas in Kyūshū as well as in south-western parts of Shikoku, although at present these dialects seem to be gradually losing the distinction.

Similar to the Tsuruoka dialect, as well as a number of other Tōhoku dialects, what corresponds to ModJ voiced velar stop /g/ appears intervocalically as the nasal /ŋ/. For example, the ModJ words **kage** ‘shade’, **mago** ‘grandchild’, **tamago** ‘egg’, and **kagami** ‘mirror’ are /kaŋe/, /maŋo/, /tamaŋo/, and /kaŋaŋ/ in Ei.<sup>10</sup>

Another similarity to the Tsuruoka dialect is the fact that older voiceless stops also become voiced intervocalically in Ei as illustrated in the following examples:

(35)	Ei		ModJ	Ei		ModJ
	<i>iga</i>	‘squid’	<b>ika</b>	<i>hago</i>	‘box’	<b>hako</b>
	<i>sagura</i>	‘cherry tree’	<b>sakura</b>	<i>magura</i>	‘pillow’	<b>makura</b>
	<i>ida</i>	‘board’	<b>ita</b>	<i>ido</i>	‘thread’	<b>ito</b>
	<i>kada</i>	‘shoulder’	<b>kata</b>	<i>adama</i>	‘head’	<b>atama</b>

The intervocalic voicing is regularly blocked when the preceding vowel /i/ or /u/ is subject to devoicing between voiceless consonants, e.g. *hito* ‘person’, *ki-ta* ‘come-PAST’ and *huto-ga* ‘big’, not \**hido*, \**ki-da*, and \**hudo-ga*.

**TABLE 11.10 THE FOUR-WAY HIRAGANA DISTINCTION IN EI DIALECT (CONTRASTED WITH TSURUOKA DIALECT)**

Ei ( <i>yotsugana</i> dialect)		CJ	EMJ	Pre-1946 hiragana/katakana	ModJ	Tsuruoka ( <i>hitotsugana</i> dialect)
Word-final	Else					
/sj/ [ʃ]	<i>zji</i> [ʒi]	<i>zi</i>	<ji> [ʒi]	じ・ジ	<b>ji</b> [dʒi] ~ [ʒi]	<i>zi</i> [zi]
/q/ [tʰ]	<i>dji</i> [dʒi]	<i>di</i>	<gi> [dʒi]	ぢ・ヂ		
	<i>dju</i> [dʒu]	<i>du</i>	<zzu/dzu> [dzu]	づ・ヅ	<b>zu</b> [dzu] ~ [zu]	
<i>s</i> [s]	<i>zu</i> [zu]	<i>zu</i>	<zu> [zu]	ず・ズ		

**TABLE 11.11 CONSONANTS IN THE EI DIALECT**

	Bilabial	Dental/Alveolar	Palatal <sup>11</sup>	Velar	Glottal
Plosive	p b	t d		k g	
Labialized				k <sup>w</sup>	
Affricate		c	cj dj		
Fricative		s z	sj zj		h
Nasal	m	n		ŋ	
Flap		r			
Approximant	w		j		

As a result, the distinction between /t/ and /d/ is neutralized intervocalically, e.g. both ModJ words **hata** ‘flag’ and **hada** ‘skin’ are *hada*, while both **mato** ‘target’ and **mado** ‘window’ are *mado* in Ei. However, in the present-day Ei dialect, such intervocalic voicing has many diachronic exceptions.

Although such voicing is widely observed in many of the dialects in Tōhoku, this type of voicing is not so common in the western part of the Japanese archipelago. In fact, among other Kyūshū dialects, this is observed only in the areas around the southernmost parts of the Satsuma peninsula where Ei is located. Whether the similarities in such voicing between this dialect and that of many dialects in the Tōhoku region are due to vestiges of an older feature of Japanese or are independent developments is unclear.

Another characteristic feature of this dialect, along with other Kagoshima dialects, is observed in the behavior of older flap /r/, which shows peculiar behavior in the following three respects. First, /r/ is changed into /t/ between a syllable /sji/ and /a/ or /o/; thus, ModJ **ushiro** ‘back’, **mushiro** ‘straw mat’, **hashira** ‘pillar’, **shirami** ‘lice’, and **shiro-i** ‘white’ correspond to *usjito*, *musjito*, *hasjita*, *sjitan*, and *sjito-ga*, respectively in Ei. On the other hand, /r/ preceded by syllables other than /sji/ never changes into /t/, as in *sagura* ‘cherry blossom’, *abura* ‘oil’, *magura* ‘pillow’, *kogoro* ‘heart’, and *iro* ‘color’. Secondly, word-initially, /r/ is strengthened and changed into /d/, as in *danpu* ‘lamp’, *dosog* ‘candle’, *daoppa* ‘trumpet’, and *daokjo* ‘shallot’, while corresponding ModJ words are **ranpu**, **rōsoku**, **raoppa**, and **raokyo**, respectively. Lastly, the /r/ sound is deleted between a vowel and /i/, as the following examples illustrate:

(36)	Ei		ModJ	Ei		ModJ
	<i>toi</i>	‘bird’	<b>tori</b>	<i>hai</i>	‘needle’	<b>hari</b>
	<i>namai</i>	‘lead’	<b>namari</b>	<i>tonai</i>	‘next-door, neighbor’	<b>tonari</b>
	<i>kusui</i>	‘medicine’	<b>kusuri</b>	<i>kemui</i>	‘smoke’	<b>kemuri</b>

Some Sino-Japanese words contain the labialized velar stop /k<sup>w</sup>/ in Ei as in *k<sup>w</sup>aigoR* ‘meeting’, *k<sup>w</sup>annon* ‘the goddess of mercy’, *k<sup>w</sup>aidaku* ‘a willing consent’, or *juk<sup>w</sup>ai* ‘amusement, enjoyment’; the corresponding ModJ words are **kaigō**, **kannon**, **kaidaku**, and **yukai**, respectively.

The consonants in this dialect are summarized in [Table 11.11](#).

### 11.3.1.3 Suprasegmentals

In a number of dialects in the southwestern Kyūshū area, such as Nagasaki, Shimabara, Isahaya, Kashima, etc., as well as most of the dialects in the Amakusa islands, Koshiki-jima, and Yaku-shima, there is a prosodic system in which there are only two accentual

patterns, regardless of the length of the tonal phrase. These two classes of words are generally called Class A and Class B. Although the phonetic realizations of these two tonal patterns are significantly different depending on each dialect, the details of which are beyond the scope of this chapter, the words which belong to each of these two groups are basically the same throughout these dialects. Samples of the vocabulary which belong to these two tonal classes are given in (37), with the forms observed in Ei:

## (37) [Nouns]

Class A: *ame* ‘candy’, *hana* ‘nose’, *hago* ‘box’, *kaze* ‘wind’, *kawa* ‘river’, *tage* ‘bamboo’, *uda* ‘song’, *inaga* ‘countryside’, *minado* ‘port’, *sagura* ‘cherry blossom/tree’

Class B: *ame* ‘rain’, *hana* ‘flower’, *ida* ‘board’, *kome* ‘rice’, *kumo* ‘cloud’, *mame* ‘bean’, *miso* ‘bean paste’, *abura* ‘oil’, *huguro* ‘bag’, *magura* ‘pillow’

## [Verbs]

Class A: *ki*<sub>Q</sub> ‘hear’, *ma*<sub>Q</sub> ‘wind’, *na*<sub>Q</sub> ‘ring’, *ne*<sub>Q</sub> ‘sleep’, *to*<sub>N</sub> ‘fly’

Class B: *ka*<sub>Q</sub> ‘write’, *ma*<sub>Q</sub> ‘sow’, *mi*<sub>Q</sub> ‘see’, *na*<sub>Q</sub> ‘become’, *no*<sub>N</sub> ‘drink’

## [Adjectives]

Class A: *aga-ga* ‘red’, *ama-ga* ‘sweet’, *asa-ga* ‘shallow’

Class B: *sjito-ga* ‘white’, *kara-ga* ‘spicy hot’, *huka-ga* ‘deep’

In Ei, Group A nouns are always accented on the penultimate syllable of each tonal phrase, which consist of a noun and the following particle (or a series of particles), while in Group B it is on its final syllable. Thus, phrases with *-ŋa* nominative, *-mo* ‘also’, *-gara* ‘from’, and *-gara-mo* ‘also from’ display the accentual patterns shown in Table 11.12.

Similar to the Tsuruoka dialect, fairly regular correspondences are observed between the lexical distribution of accent patterns in ModJ and those in Ei, illustrated by the two-syllable nouns in Table 11.13.

Due to high-vowel deletion, the Ei dialect has a number of heavy syllables. In contrast to ModJ, in which the only word-final consonant is *n*, the Ei dialect also permits (at least phonetically) *o* and *sj* word-finally, as in *ka*<sub>Q</sub> ‘persimmon’ or *iwasj* ‘sardine’.

TABLE 11.12 ACCENT PATTERNS OF NOUNS IN THE EI DIALECT (I)

Group A	<i>hána</i>	‘nose’	<i>haná-ŋa</i>	<i>haná-mo</i>	<i>hana-gára</i>	<i>hana-gará-mo</i>
Group B	<i>haná</i>	‘flower’	<i>hana-ŋá</i>	<i>hana-mó</i>	<i>hana-gará</i>	<i>hana-gara-mó</i>
Group A	<i>sagúra</i>	‘cherry tree’	<i>sagurá-ŋa</i>	<i>sagurá-mo</i>	<i>sagura-gára</i>	<i>sagura-gará-mo</i>
Group B	<i>aburá</i>	‘oil’	<i>abura-ŋá</i>	<i>abura-mó</i>	<i>abura-gará</i>	<i>abura-gara-mó</i>

TABLE 11.13 DISTRIBUTION OF NOUN ACCENT PATTERNS IN ModJ AND EI

Examples by tone class (Martin 1987)	ModJ	Ei
2.1: <b>ame</b> ‘candy’, <b>hana</b> ‘nose’, <b>hako</b> ‘box’, <b>kaze</b> ‘wind’	LH ~ LH-H	LH~ LH-L
2.2: <b>hito</b> ‘person’, <b>kawa</b> ‘river’, <b>uta</b> ‘song’	LH ~ LH-L	
2.3: <b>hana</b> ‘flower’, <b>kome</b> ‘rice’, <b>kumo</b> ‘cloud’, <b>yama</b> ‘mountain’		LH ~ LL-H
2.4/5: <b>ame</b> ‘rain’, <b>ito</b> ‘thread’, <b>kata</b> ‘shoulder’, <b>mado</b> ‘window’, <b>miso</b> ‘bean paste’, <b>nabe</b> ‘pot’, <b>sora</b> ‘sky’	HL ~ HL-L	

**TABLE 11.14 ACCENT PATTERNS OF NOUNS IN THE EI DIALECT (II)**

Group A	<i>káQ</i>	‘persimmon’	<i>káQ-ŋa</i>	<i>káQ-mo</i>	<i>kaQ-gára</i>	<i>káQ-djui</i>
Group B	<i>máQ</i>	‘pine tree’	<i>maQ-ŋá</i>	<i>maQ-mó</i>	<i>maQ-gará</i>	<i>maQ-djúi</i>
Group A	<i>kádaQ</i>	‘form’	<i>kadáQ-ŋa</i>	<i>kadáQ-mo</i>	<i>kadaQ-gára</i>	<i>kadáQ-djui</i>
Group B	<i>cjubáQ</i>	‘camellia’	<i>cjubaQ-ŋá</i>	<i>cjubaQ-mó</i>	<i>cjubaQ-gará</i>	<i>cjubaQ-djúi</i>
Group A	<i>káN</i>	‘paper’	<i>káN-ŋa</i>	<i>káN-mo</i>	<i>kaN-gára</i>	<i>káN-djui</i>
Group B	<i>áN</i>	‘net’	<i>aN-ŋá</i>	<i>aN-mó</i>	<i>aN-gará</i>	<i>aN-djúi</i>
Group A	<i>tádaN</i>	‘straw mat’	<i>tadáN-ŋa</i>	<i>tadáN-mo</i>	<i>tadaN-gára</i>	<i>tadáN-djui</i>
Group B	<i>hasáN</i>	‘scissors’	<i>hasaN-ŋá</i>	<i>hasaN-mó</i>	<i>hasaN-gará</i>	<i>hasaN-djúi</i>

However, the dialect of Ei is a syllable language lacking the rhythmic unit ‘mora’; this is also observed in the fact that in contrast to a number of Japanese dialects in which a mora is an accentual unit, Ei is one of the dialects in Japan in which a syllable is the sole unit for its accent placement. This is illustrated in Table 11.14, with the accentual patterns of the nouns *kaQ* ‘persimmon’, *kadaQ* ‘form’, *kaN* ‘paper’, *tadaN* ‘straw mat’, which are Class A nouns, and *maQ* ‘pine tree’, *cjubaQ* ‘camellia’, *aN* ‘net’, and *hasaN* ‘scissors’, which are Class B nouns. The suffix *-djui* (terminative) means ‘up to’. Notice that Class A nouns, e.g. *tadaN* ‘straw mat’, always carry accent on their penultimate syllable, as in *tádaN*, *tadáN-ŋa*, *tadáN-mo* and *tadáN-djui*. That Ei is a syllable-based language is shown by the fact that high tones are always assigned on either the penultimate or the final syllable, not mora.

Verb and adjective conjugation forms also maintain these two clear-cut patterns. In Table 11.15, the verbs *maQ* ‘wind’ and *haimui* ‘begin, start’ belong to Class A, while *maQ* ‘sow’ and *acjumui* ‘gather’ belong to Class B.

The accentual patterns of adjectives are illustrated in Table 11.16.

In terms of adjective accentuation, we could say that this dialect is more conservative than ModJ (in spite of the fact that it has only two patterns for nouns, while ModJ has *n+1* patterns). The old distinction of two adjective accentual patterns is overwhelmingly well preserved in Ei, while in ModJ the distinction between Class A (‘accented’) vs. Class B (‘unaccented’) adjectives is under attrition, i.e. there is no distinction in conclusive forms such as **amá-i** ‘sweet’ and **kará-i** ‘spicy’, even though a distinction

**TABLE 11.15 ACCENT PATTERNS OF VERBS IN THE EI DIALECT**

	Non-Past	Past	Negative	Negative-Past	Imp.	Neg. Imp.
Class A	<i>máQ</i> ‘wind’	<i>mé-da</i>	<i>mág-aN</i>	<i>mag-aN-zjáQta</i>	<i>mág-e</i>	<i>máN-na</i>
Class B	<i>máQ</i> ‘sow’	<i>me-dá</i>	<i>mag-áN</i>	<i>mag-aN-zjaQta</i>	<i>mag-é</i>	<i>man-ná</i>
Class A	<i>háimui</i> ‘begin’	<i>haimé-da</i>	<i>háime-N</i>	<i>haime-N-zjáQta</i>	<i>háim-e</i>	<i>haimún-na</i>
Class B	<i>acjumúi</i> ‘gather’	<i>acjume-dá</i>	<i>acjumé-N</i>	<i>acjume-N-zjaQta</i>	<i>acjum-é</i>	<i>acjumun-ná</i>

**TABLE 11.16 ACCENT PATTERNS OF ADJECTIVES IN THE EI DIALECT<sup>12</sup>**

	Non-Past	Past	Negative	Negative-Past
Class A	<i>agá-ga</i> ‘red’	<i>aga-gá-Qta</i>	<i>ágo-na-ga</i>	<i>ágo-na-ga-Qta</i>
	<i>amá-ga</i> ‘sweet’	<i>ama-gá-Qta</i>	<i>ámo-na-ga</i>	<i>ámo-na-ga-Qta</i>
Class B	<i>sjító-gá</i> ‘white’	<i>sjito-ga-Qtá</i>	<i>sjító-na-ga</i>	<i>sjító-na-ga-Qta</i>
	<i>kara-gá</i> ‘spicy, hot, dry’	<i>kara-ga-Qtá</i>	<i>karó-na-ga</i>	<i>karó-na-ga-Qta</i>

TABLE 11.17 ACCENT PATTERNS OF COMPOUND NOUNS IN THE EI DIALECT

Class A + Class B			
<i>mómo</i>	‘peach’	<i>momobáŋke</i>	~ <i>momobadáge</i> ‘peach field’
<i>tomáto</i>	‘tomato’	<i>tomatobáŋke</i>	~ <i>tomatobadáge</i> ‘tomato field’
<i>kjabécu</i>	‘cabbage’	<i>kjabecubáŋke</i>	~ <i>kjabecubadáge</i> ‘cabbage field’
Class B + Class B			
<i>haná</i>	‘flower’	<i>hanabaŋké</i>	~ <i>hanabadagé</i> ‘flower field’
<i>iŋgó</i>	‘strawberry’	<i>iŋgobaŋké</i>	~ <i>iŋgobadáge</i> ‘strawberry field’
<i>migáN</i>	‘orange’	<i>miganbaŋké</i>	~ <i>miganbadagé</i> ‘orange field’

is retained in attributive forms: **ama-i mono** ‘sweet thing’ vs. **kará-i mono** ‘spicy thing’. Similarly, ModJ past tense forms, **amá-kaŋta** vs. **kára-kaŋta**, or infinitive/adverbial forms, **amá-ku** vs. **kára-ku**, are currently undergoing accentual merger into one pattern: **amá-kaŋta**, **kará-kaŋta**, and **amá-ku**, **kará-ku**.<sup>13</sup>

Accent patterns of compounds in Ei are very regular, in that the first member of the compound determines the tonal patterns of the whole compound; thus, the accentual pattern of a compound *momobáŋke*<sup>14</sup> ~ *momobadáge* ‘a field of peaches’, which is produced by attaching a Class A noun *mómo* ‘peach’ with another noun *hadagé* ‘field’, is Class A, while another compound *hanabaŋké* ~ *hanabadagé* ‘a field of flowers’ is produced by a Class B word *haná* ‘flower’ and *hadagé* is Class B. This is illustrated in Table 11.17.

Compound accentuation in ModJ is exceptional; the above type of compounding rule by which the accent class of the *first* member of a compound decides the accent of the whole compound is widely observed in many other dialects in Japan, such as the dialects in the Kinki (Kyoto-Osaka) region, some of the dialects in the Shikoku and San’in districts, dialects throughout the southwestern Kyūshū areas, as well as a number of Ryukyuan varieties.

### 11.3.2 Morphology

#### 11.3.2.1 Nouns and case marking

It is well known that in a number of the dialects in Kyūshū, such as Kumamoto and Nagasaki, as well as in Ryukyuan varieties, the unmarked nominative marker is *-N*, a shortened version of *-no* (or *-nu* for Ryukyuan varieties). However, in Ei, the nominative marker is *-ŋa* as in *kome-ŋa yo-ga* ‘the rice is nice’, or *mun-ŋa yo-ga* ‘the wheat is nice’; *-no* or *-N* are not used for subjects.

The accusative is typically marked in this dialect by *-o*, as in *hago-o mi-da* ‘saw the box’, or *oge-o mi-da* ‘saw the tub’; but the marker has allomorphs *-u/-nju*, discussed in 11.3.3. Similar to informal ModJ, the accusative may also be zero-marked in this dialect.

The unmarked dative marker in Ei is *\*-i* or *-N*, as illustrated in the following examples, in which *hane* is formed by *hana* ‘flower’ plus *-i* (for details, see 11.3.3):

- (38) a. *hane*            *midj-u*            *jaŋ-ta*  
 flower.DAT    water-ACC    give-PAST  
 ‘(Somebody) gave water to the flower.’

- b. *tage-N*            *midj-u*            *jaQ-ta*  
 bamboo-DAT    water-ACC    give-PAST  
 ‘(Somebody) gave water to the bamboo tree.’

As in ModJ, the dative marker is identical to that expressing direction, as discussed later.

The genitive marker is *-N* or *-no*. When nouns end in vowels, *-N* is used as in *momo-N iro* ‘color of a peach’, *hana-N iro* ‘color of a flower’. However, when suffixed to nouns ending in *-N*, many of which are produced by high-vowel deletion, *-no* is used, as in *kan-no iro* ‘color of paper’, *un-no iro* ‘color of the ocean’. When a noun ends in /Q/, as in *muQ* ‘wheat’, *kaQ* ‘persimmon’, *maQ* ‘pine tree’, or *cjubaQ* ‘camellia’, the final consonant is assimilated to the suffix-initial consonant /n/ of *-no*, i.e. *mun-no*, *kan-no*, *man-no*, and *cjuban-no*.

As for first- and second-person pronouns, individual names, kinship terms, titles referring to a specific individual, such as *senseR* ‘teacher, doctor’, genitive is marked by *-ŋa* instead of *-no*: *omae-ŋa hon* ‘your book’, *wai-ŋa hon* ‘my book’, *joRko-ŋa hon* ‘Yōko’s book’, and *senseR-ŋa je* ‘the teacher’s house’.

### 11.3.2.2 Other particles

Direction can be expressed by means of the dative marker *-i* or *-N*. It can, in addition, be expressed by a marker *-same*, as in *hune-same* ‘into the boat’, or *nabe-same* ‘into the pan’:

- (39) *jenpicu*    *cjukue-same*    *ire-da*  
 pencil    desk-ALL    put.in-PAST  
 ‘(Somebody) put the pencil in the desk.’

Cognate with *-sa* in Tsuruoka dialect, this suffix is probably the grammaticalization of CJ *sama* ‘direction’ + *-ŋe* ‘to’ (ModJ *-e*). However, unlike *-sa*, which is a very common suffix in the Tōhoku dialects expressing various functions such as ‘at’, ‘on’, ‘for the purpose of’ or even dative (see (7)), the particle *-same* in Ei has not expanded its semantic domain beyond the sense of direction.

That such semantic extension of *-same* did not take place in Ei may be due to the fact that this dialect has (an)other (unmarked) suffix(es) by which the notion of direction is expressed, *-i* or *-N* above. The suffixes *-i* or *-N* are used to indicate ‘on’, ‘in’, or ‘onto’ in addition to ‘toward’. When nouns end in /e/ (and sometimes /o/), *-N* is preferred, as in *hune-N* ‘onto the boat’, *oge-N* ‘into the tub’, *nabe-N* ‘into the pot’ (and *togyo-N* ‘to Tokyo’):

- (40) *oge-N*            *karaimo*            *ire-da*.  
 tub-ALL    sweet.potato    put.in-PAST  
 ‘(Somebody) put sweet potatoes in the tub.’

This particle developed as a general dative marker, as illustrated in (38).

Locative is marked by *-zje* ‘at, in’, as in:

- (41) a. *jeQ-zje*            *senseR-to*            *o-da*  
 station-LOC    teacher-with    meet-PAST  
 ‘(Someone) met the teacher at the station.’  
 b. *hadage-zje*    *sjigodj-u*            *su-i*  
 field-LOC    work-ACC    do-NONPAST  
 ‘(Someone) works in the field.’

The same particle is also used as an instrumental marker as *hasan-zje ki-i* ‘cut with scissors’, or *densja-zje iQ* ‘go by train’.

Other basic particles include inclusive *-mo* as in *asjita-mo* ‘tomorrow too’, *odon-mo* ‘my brother as well’, ablative *-gara* ‘from’ as in *jama-gara* ‘from a mountain’, *an-gara* ‘from a net’, and terminative *-djui* ‘up to’, as in *jama-djui* ‘to a mountain’, *un-djui* ‘to the ocean’.

### 11.3.2.3 Verbs and adjectives

Similar to ModJ and the Tsuruoka dialect, forms for conclusive vs. attributive predicates are not distinguished for verbs and adjectives. Verb conjugation in the Ei dialect is illustrated in Table 11.18.

Similar to Tsuruoka dialect, the conjugational forms of historical vowel-stem verbs, e.g. *mi-* ‘see, look’ and *ogi-* ‘get up’ in (b.1) (b.2), are similar to the historical /r/-stem verbs shown in (a.2). First, by analogy with such /r/-stem verbs *nare* ‘Become!’ and *ure* ‘Sell!’, their imperative forms end with *-re* as in *mire* ‘Look!’ and *ogire* ‘Get up!’, instead of *-ro* as in ModJ **mi-ro** or **oki-ro** and many other dialects. Secondly, their negative forms are suffixed with /ran/, as in *miran* ‘does not see’ and *ogiran* ‘does not get up’, by analogy with /r/-stem verbs, such as *naran* ‘does not become’ and *uran* ‘does not sell’.

These imperative and negative forms are another instance of the general tendency in Japanese dialects for vowel-stem verbs to come to conjugate similarly to /r/-stem verbs, also seen in the section on the Tsuruoka dialect. It is noteworthy that such a general tendency of attrition of vowel-stem verbs under the influence of the productive consonant-stem verbs is typically observed in peripheral areas of the Japanese archipelago, such as the eastern Tōhoku area and the Kyūshū region, where it may have developed independently.

In ModJ, vowel-stem verbs which end in /i/ or /e/ conjugate as in Table 11.19. Vowel-stem verbs in OJ and MJ were ‘bigrade’, having two allomorphs, *oku-* and *oki-*, *uku-* and *uke-*, *aku-* and *ake-*, respectively, as in *oku* ‘get up-NONPAST’, *oku-reba* ‘get

TABLE 11.18 VERB CONJUGATION IN THE EI DIALECT (I)

	Non-Past	Past	Negative	Imperative	Conditional
a.1	<i>kaQ</i> ‘write’	<i>ke-da</i>	<i>kag-an</i>	<i>kag-e</i>	<i>kag-eba</i>
	<i>ton</i> ‘fly’	<i>ton-da</i>	<i>tob-an</i>	<i>tob-e</i>	<i>tob-eba</i>
a.2	<i>na-i</i> ‘become’	<i>naQ-ta</i>	<i>nar-an</i>	<i>nar-e</i>	<i>nar-eba</i>
	<i>u-i</i> ‘sell’	<i>uQ-ta</i>	<i>ur-an</i>	<i>ur-e</i>	<i>ur-eba</i>
b.1	<i>mi-i</i> ‘see, look’	<i>mi-da</i>	<i>mi-ran</i>	<i>mi-re</i>	<i>mi-reba</i>
b.2	<i>ogi-i</i> ‘get up’	<i>ogiQ-da</i> <sup>15</sup>	<i>ogi-ran</i>	<i>ogi-re</i>	<i>ogi-reba</i>

TABLE 11.19 CONJUGATION OF VOWEL-STEM VERBS IN ModJ

Non-Past	Past	Negative	Imperative	Conditional	Neg. Imp.
<b>oki-ru</b> ‘get up’	<b>oki-ta</b>	<b>oki-nai</b>	<b>oki-ro</b>	<b>oki-reba</b>	<b>oki-runa</b>
<b>uke-ru</b> ‘receive’	<b>uke-ta</b>	<b>uke-nai</b>	<b>uke-ro</b>	<b>uke-reba</b>	<b>uke-runa</b>
<b>ake-ru</b> ‘open’	<b>ake-ta</b>	<b>ake-nai</b>	<b>ake-ro</b>	<b>ake-reba</b>	<b>ake-runa</b>
<b>hajime-ru</b> ‘begin’	<b>hajime-ta</b>	<b>hajime-nai</b>	<b>hajime-ro</b>	<b>hajime-reba</b>	<b>hajime-runa</b>

up-COND’, *oki-zu* ‘get up-NEG’, *oki-yo* ‘get up-IMP’, etc. Analogical leveling resulted in the loss of this allomorphy, resulting in a ‘monograde’ conjugation in ModJ. This leveling, however, does not take place in all dialects. The bigrade conjugation is still observed in a number of dialects in the Kyūshū region, representative of which is Bungo-Takada in Ōita prefecture, in which the distinction of two allomorphs is still maintained: *oku-ru* ‘get up-NONPAST’, *oku-rja* ‘get up-COND’, *oki-N* ‘get up-NEG’, *oki-jo* ‘get up-IMP’, *aku-ru* ‘open-NONPAST’, *aku-rja* ‘open-COND’, *ake-N* ‘open-NEG’, and *ake-jo* ‘open-IMP’ (Onishi 1998).

Leveling to monograde *partly* took place in Ei, the former bigrade conjugation changing to monograde in the case of the so-called *upper* bigrade (i.e. /i/-stem) verbs as illustrated in (b.2) in Table 11.18. However, the so-called *lower* bigrade (i.e. /e/-stem) verbs still keep the older pattern of two allomorphs in this dialect, as illustrated in Table 11.20.

On the other hand, the Ei cognates of ModJ irregular verbs **suru** ‘do’ and **kuru** ‘come’ show the conjugation patterns in Ei shown in Table 11.21. In contrast to Tsuruoka, in which the the verb ‘do’ conjugates similarly to consonant-stem verbs, it conjugates like a bigrade verb in Ei, in that the verb consists of two allomorphs /su/ and /se/ as shown in Table 11.21, which suggests that the older *lower* bigrade conjugation is still productive in this dialect.

The conjugational pattern for adjectives is illustrated above in Table 11.16, in the section of suprasegmentals. As with ModJ, there is no distinction between conclusive and attributive forms of adjectives. Non-past forms of adjectives seem to consist of two types in Ei, as well as a number of dialects in Kagoshima prefecture: those ending with *-ga* as in *aga-ga* ‘red’ or *huto-ga* ‘big’ and those ending with the vowel /e/ as in *age* ‘red’ or *hute* ‘big’; the latter probably derive from \**aga-i* (cf. SJ **aka-i**) or \**huto-i* (**futo-i**) through vowel coalescence (*ai* > *e*, *oi* > *e*). Thus, for example, ‘red flower’ can be either *aga-ga hana* (red-ATTR flower) or *age hana* in Ei, but the former type (i.e. *aga-ga*) seems to be used more widely than the latter.

The negative form in this dialect also has versions based on *-ga* as in *ama-ga-na-ga* ‘is not sweet’ or *huto-ga-na-ga* ‘is not big’, but those in Table 11.16 (*amo-na-ga* or *huto-na-ga*), built on the adverbial form, seem to be used more widely. Adverbial forms are derived by fusion of root vowel + \**u* (*a-u* > *o*, *o-u* > *o*), as in *amo naqta* ‘became red’ or *huto naqta* ‘became big’.

The non-past form of adjectives, *ama-ga* or *kara-ga*, in this dialect corresponds to *ama-ka* and *kara-ka* observed in the other western and southern Kyūshū dialects, such

TABLE 11.20 VERB CONJUGATION IN THE EI DIALECT (II)

Non-Past	Past	Negative	Imperative	Conditional	Neg. Imp.
<i>agu-i</i> ‘open’	<i>age-da</i>	<i>age-N</i>	<i>age</i>	<i>agu-reba</i>	<i>agu-nna</i>
<i>ugu-i</i> ‘receive’	<i>uge-da</i>	<i>uge-N</i>	<i>uge</i>	<i>ugu-reba</i>	<i>ugu-nna</i>
<i>haimu-i</i> ‘begin’	<i>haime-da</i>	<i>haime-N</i>	<i>haime</i>	<i>haimu-reba</i>	<i>haimu-nna</i>

TABLE 11.21 VERB CONJUGATION IN THE EI DIALECT (III)

Non-Past	Past	Negative	Imperative	Conditional	Neg. Imp.
<i>su-i</i> ‘do’	<i>se-da</i>	<i>se-N</i>	<i>se</i>	<i>su-reba</i>	<i>su-nna</i>
<i>ku-i</i> ‘come’	<i>ki-ta</i>	<i>ko-N</i>	<i>ke</i>	<i>ku-reba</i>	<i>ku-nna</i>

as those in Nagasaki, Kumamoto and a part of Fukuoka prefectures. Diachronically, they are contracted versions of older *\*ama-kari* and *\*kara-kari*, corresponding to CJ *ama-k(u)ar-* and *kara-k(u)ar-*, as in *ama-kar-azu* ‘it’s not sweet’.

Adjectival nouns (ModJ **na**-type adjectives) are conjugated in a similar pattern to ModJ, except that the suffix for conclusive forms is **-da** in ModJ, as in **hade-da** ‘showy, gaudy’, **meiwaku-da** ‘troublesome, bothersome’, **shinsetsu-da** ‘kind’, while it is *-zja* in Ei: *hade-zja*, *mejao-zja*, and *sjinseo-zja*. As in ModJ, the conclusive and attributive forms of nominal adjectives are different in Ei, the suffix for their attributive forms being *-na*, as in *mejao-na hanasj* ‘troublesome story’. Their past tense forms are *hade-zjao-ta*, *mejao-zjao-ta*, and *sjinseo-zjao-ta* in Ei, while their negative forms are *hade-zja-na-ga*, *mejao-zja-na-ga*, and *sjinseo-zja-na-ga*.

### 11.3.3 Morphophonology

Particularly characteristic in the dialect of Ei is that certain particles have a variety of allomorphs, depending on phonological context. Especially interesting behavior is observed in the case of the particles *-o* accusative, *-i* dative and direction, and *-wa* topic.

In ModJ, their cognates **-o**, **-ni**, and **-wa** neither display allomorphy nor fuse with the word-final vowels of the nouns to which they are attached; thus, for example, in a ModJ sentence (42), the underlined noun is interchangeable with any other noun, without any morphophonological change:

- (42) **sono kodomo-wa ame-o mi-ta**  
 that child-TOP candy-ACC see-PAST  
 ‘The child saw the candy.’

For example, **ame-o** in (42) can be replaced with **yama-o** ‘mountain-ACC’, **neko-o** ‘cat-ACC’, **inu-o** ‘dog-ACC’, **kaki-o** ‘persimmon-ACC’, or **mikan-o** ‘orange-ACC’ with no sandhi. This is in sharp contrast to Ei allomorphy/sandhi, in which various particles display a number of interesting ‘declension’ patterns.

Firstly, the topic marker *-wa* has allomorphs *-a* or *-ja*, the most unmarked of which is *-wa*, which is used for nouns ending with vowels /e/, /a/, and /o/: e.g. *ame-wa* ‘candy-TOP’, *jeda-wa* ‘branch-TOP’, *ido-wa* ‘thread-TOP’, *ajame-wa* ‘iris-TOP’, *abura-wa* ‘oil-TOP’, and *tamaŋo-wa* ‘egg-TOP’. When suffixed to nouns ending with /o/sj/s, *-a* is used, and an underlying consonant reemerges: e.g. *tab-a* ‘socks-TOP’ **tabi-wa**, *nasub-a* ‘egg plant-TOP’ **nasubi-wa**, *kag-a* ‘persimmon-TOP’ **kaki-wa**, *cjubag-a* ‘camellia-TOP’ **tsubaki-wa**, *kuŋ-a* ‘nail-TOP’ **kugi-wa**, *miŋ-a* ‘right (noun)-TOP’ **migi-wa**, *usaj-a* ‘rabbit-TOP’ **usagi-wa**, *janaj-a* ‘willow-TOP’ **yanagi-wa**, *is-a* ‘stone-TOP’ **ishi-wa**, *iwas-a* ‘sardine-TOP’ **iwashi-wa**, *karas-a* ‘crow-TOP’ **karasu-wa**, and *suz-a* ‘bell-TOP’ **suzu-wa**.<sup>16</sup> On the other hand, when suffixed to nouns ending with *i*, *-ja* is used, as in *wai-ja* ‘you-TOP’ and *zji-ja* ‘letter-TOP’. The allomorph *-ja* is also used for nouns ending with *N*, with an insert *-n*:- e.g. *miN-nja* ‘ear-TOP’, *non-nja* ‘flea-TOP’, *kaŋan-nja* ‘mirror-TOP’, *iN-nja* ‘dog-TOP’, and *kimon-nja* ‘clothes-TOP’.

The accusative is typically marked by the suffix *-o* in this dialect, as in (43a). However, when nouns end in /o/sj/s derived through high-vowel deletion, this changes to the underlying source consonant + *-u*, as in, *kub-u* from *kuo* ‘neck’ **kubi-o**, *cjubag-u* from *cjubaŋ* ‘camellia’ **tsubaki-o**, *kuŋ-u* from *kuŋ* ‘nail’ **kugi-o**, *kucj-u* from *kuŋ* ‘mouth’ **kuchi-o**, *madj-u* from *maŋ* ‘pine tree’ **matsu-o**, *kucj-u* from *kuŋ* ‘shoe’ **kutsu-o**, *esacj-u* from *esaŋ* ‘greeting’ **aisatsu-o**, *isj-u* for *isj* ‘stone’ **ishi-o**, *iwasj-u* for *iwasj* ‘sardine’ **iwashi-o**, *karas-u* for *karas* ‘crow’ **karasu-o**, and *kaz-u* for *kaz* ‘number’ **kazu-o**. This

is illustrated in (43b) by *miQ* ‘water’. For those ending with *i*, *-ju* is used as in *wai-ju* ‘you-ACC’, *k<sup>v</sup>ai-ju* ‘meeting-ACC’ **kai-o**, *hi-ju* ‘fire-ACC’ **hi-o**. For those ending with nasal *N*, an insert /nj/ is needed before the *-u*, as in *in-nju* from *in* ‘dog’ **inu-o**, *ka<sup>n</sup>an-nju* from *ka<sup>n</sup>an* ‘mirror’ **kagami-o**, and *tadan-nju* from *tadan* ‘straw mat’ **tatami-o**, illustrated in (43c):

- (43) a. *kome-o*    *oge-N*    *ire-da*  
 rice-ACC    tub-ALL    put-PAST  
 ‘(Somebody) put the rice into a tub.’
- b. *oge-N*    **midj-u**    *ire-da*  
 tub-ALL    water-ACC    put-PAST  
 ‘(Somebody) put water into the tub.’
- c. *made*    **kan-nju**    *haQ-ta*  
 window.ALL    paper-ACC    paste-PAST  
 ‘(Somebody) put the paper on the window.’

A more complicated case is observed for the dative (or directional) suffix *-i*, which appears most clearly when attached to *Q/sj*-final nouns, which resulted from the above-mentioned high-vowel deletion, whose underlying consonant reemerges. Thus, *kaQ* ‘persimmon’ **kaki**, *maQ* ‘pine tree’ **matsu**, *usaQ* ‘rabbit’ **usagi**, *usj* ‘cow’ **ushi**, *hasj* ‘bridge’ **hashi**, *kaz* ‘number’ **kazu**, *karas* ‘crow’ **karasu** appear as *kag-i* ‘persimmon-DAT’, *mazj-i* ‘pine tree-DAT’, *usay-i* ‘rabbit-DAT’, *usj-i* ‘cow-DAT’, *hasj-i* ‘bridge-DAT’, *kazj-i* ‘number-DAT’, and *karasj-i* ‘crow-DAT’. However, the suffix *-i* tends to fuse with word-final vowels of nouns to which they are attached, as in *kawe* (< \*kawa-i) ‘to the river’, *jame* (< \*jama-i) ‘to the mountain’, *ka<sup>n</sup>oime* (< \*ka<sup>n</sup>oima-i) ‘to Kagoshima’, *inage* (< \*inaga-i) ‘to the countryside’, *tanbe* (< \*tanbo-i) ‘to the rice field’, *hage* (< \*hago-i) ‘into the box’, *minade* (< \*minado-i) ‘to the port’, *hadage* (< \*hadage-i) ‘to the field’, and *hude* (< \*hude-i) ‘to the brush’, as in (44a). When nouns end in *N*, such as *miN* ‘ear’ **mimi**, *in* ‘dog’ **inu**, and *nezun* ‘mouse’ **nezumi**, *-ni* is used instead, as in *miN-ni*, *in-ni*, and *nezun-ni*, as in (44b).

- (44) a. **hane**    *midj-u*    *kagu-i*  
 flower.DAT    water-ACC    pour-NONPAST  
 ‘(Someone) pours water to a flower.’
- b. *in-ni*    *esa-o*    *ja-i*  
 dog-DAT    food-ACC    give-NONPAST  
 ‘(Someone) gives food to a dog.’

As for nouns ending with *-o* or *-e*, such as *ma<sup>n</sup>o* ‘grandchild’ **magō**, *toRkjo* ‘Tokyo’ **tōkyō**, *nabe* ‘pot’ **nabe**, *oge* ‘tub’ **oke**, it seems that *-N* is preferred to *-i* as in *ma<sup>n</sup>o-N* ‘grandchild-DAT’, *toRkjo-N* ‘Tokyo-DAT’, *nabe-N* ‘pot-DAT’, *oge-N* ‘tub-DAT’, as in (38b), (40), or (54).

### 11.3.4 The basic sentence

#### 11.3.4.1 Declarative

The basic word order of this dialect is SOV, with tense, aspect, and modality markers attached after verbs or adjectives.

Noun predicates are formed with a copula *-zja*, instead of *-da*, the latter of which is widely observed in the eastern dialects including ModJ and Tsuruoka. Thus, *ame-zja* means ‘It is candy’. Adjectival nouns such as *hade* ‘showy, gaudy’, or *mejaQ* ‘bothersome, annoying’ use the same copula *-zja*, as in *hade-zja* ‘It is showy’, *mejaQ-zja* ‘It is bothersome’.<sup>17</sup>

#### 11.3.4.2 Tense and aspect

The non-past form of a verb in Ei is expressed by the suffix *-i*,<sup>18</sup> as in *na-i* ‘become-NONPAST’, *ogi-i* ‘get up-NONPAST’, and *agu-i* ‘open-NONPAST’. When this suffix is attached to consonant-stem verbs, it is deleted in accordance with word-final high-vowel deletion (see (32), (33), and (34)), resulting in forms such as *kaQ* ‘write’ **kak-u** or *ton* ‘fly’ **tob-u**. Attached to /r/-stems, though, the /r/ is changed into *i* as in *nai* ‘become-NONPAST’ **nar-u**, *toi* ‘take-NONPAST’ **tor-u**, and *noi* ‘ride-NONPAST’ **nor-u**.

Past tense is marked by *-da*, allomorph of *-ta*, derived due to the above-mentioned intervocalic voicing as in vowel-stem *ne-da* ‘sleep-PAST’ **ne-ta**, *mi-da* ‘see-PAST’ **mi-ta**, *age-da* ‘open-PAST’ **ake-ta**, and *acjume-da* ‘gather-PAST’ **atsume-ta**, and consonant-stem *ke-da* ‘write-PAST’ **kai-ta**, *me-da* ‘wind-PAST’ **mai-ta**, *ton-da* ‘fly-PAST’ **ton-da**, *non-da* ‘drink-PAST’ **non-da**, and irregular *se-da* ‘do-PAST’ **shi-ta**. The voicing is blocked when the stem ends with *Q* or a devoiced vowel as in *naQ-ta* ‘become-PAST’ **naQ-ta**, *moQ-ta* ‘hold-PAST’ **moQ-ta**, *ogiQ-ta* ‘wake up-PAST’ **oki-ta**, and *ki-ta* ‘come-PAST’ **ki-ta**.

The continuative is expressed by an auxiliary *-zjo-i* ~ *-cjo-i*: e.g. *ne-zjo-i* ‘sleep-CONT’, *mi-zjo-i* ‘see-CONT’, and *ogio-zjo-i* ‘get up-CONT’,<sup>19</sup> *ke-zjo-i* ‘write-CONT’, *ton-zjo-i* ‘fly-CONT’, *naQ-cjo-i* ‘become-CONT’, and *uQ-cjo-i* ‘sell-CONT’. This morpheme developed from *\*-te ori*, which consists of the gerund plus *\*ori* (CJ *wori*) ‘be, exist’. This is a parallel development to that in the Tsuruoka dialect, in which the continuative form is marked by *-oda* or *-deda*, which developed from the gerund + *i-* (CJ *iru*, ModJ *iru*) ‘be, exist’ plus the tense marker *-da* (see 11.2.3.2). This suggests a general tendency in Japanese dialects to derive the continuative aspect by grammaticalization of the lexical verb of existence. OJ/CJ *-tari* developed similarly, from the gerund + *ari* ‘be, exist’.

As in ModJ, progressive and perfect functions are not formally distinguished in Ei; they are both expressed by *-zjoi* ~ *-cjoi*. Thus, ‘(someone) is buying’ and ‘(someone) has bought’ are both expressed by *ko-zjoi*, in contrast to a number of dialects in western parts of the Japanese archipelago, such as the Shikoku or Chūgoku areas, in which progressive vs. perfect aspects are strictly distinguished by different verbal forms.

In Ei, the continuative can also be expressed by *-goQ*:

- (45) *undoRba-zje*      *hasji-goQ-ta*  
 sports.field-LOC    run-CONT-PAST  
 ‘(Someone) was running in the sports field.’

However, this morpheme seems to be used exclusively for actions or events in the past.

#### 11.3.4.3 Negation

Negation is expressed with the suffix *-ran* ~ *-an* for verbs as in *age-ran* ‘open-NEG’ and *ogi-ran* ‘get up-NEG’, *kig-an* ‘hear-NEG’, *nar-an* ‘ring-NEG’, *tob-an* ‘fly-NEG’,

*nom-an* ‘drink-NEG’. Negative past is expressed by attaching *-zjaŋta* after these negative forms, as in *age-ran-zjaŋta* ‘open-NEG-PAST’ and *ogi-ran-zjaŋta* ‘get up-NEG-PAST’, *kig-an-zjaŋta* ‘hear-NEG-PAST’, *nar-an-zjaŋta* ‘ring-NEG-PAST’, *tob-an-zjaŋta* ‘fly-NEG-PAST’, *nom-an-zjaŋta* ‘drink-NEG-PAST’.

For the negative of adjectives, *\*-unaga* is attached to adjective roots, where the suffix-initial underlying vowel *\*/u/* fuses with root-final */a/* and */o/* into *o*. Thus, for example, the negative forms of *aga-* ‘red’, *ama-* ‘sweet’, and *huto-* ‘big’ are *ago-naga* ‘red-NEG’, *amo-naga* ‘sweet-NEG’, and *huto-naga* ‘big-NEG’ respectively. Negative past is expressed by attaching *-ŋta* after these negative forms, as in *ago-naga-ŋta* ‘red-NEG-PAST’, *amo-naga-ŋta* ‘sweet-NEG-PAST’, and *huto-naga-ŋta* ‘big-NEG-PAST’.

The negative of the copula *-zja* is *-zja-naga*, as in *ame-zja-naga* ‘It is not candy’, while the negative past is *-zja-naga-ŋta*, as in *ame-zja-naga-ŋta* ‘It was not candy’. Adjectival nouns such as *hade* ‘showy, gaudy’, or *mejaŋ* ‘bothersome, annoying’ conjugate similarly, as in *hade-zja-naga* ‘It is not showy’, *mejaŋ-zja-naga* ‘It is not bothersome’, *hade-zja-naga-ŋta* ‘It was not showy’, and *mejaŋ-zja-naga-ŋta* ‘It was not bothersome’.

#### 11.3.4.4 Modality

Epistemic modality is expressed by *-zja-ro* as in *juŋ-ŋa fu-i-zja-ro* ‘it may snow’, or *-ro* as in *aŋka-ro* ‘may be red’. The copula *-zja* is inserted after nouns as in *ame-zja-ro* ‘may be rain’. A weaker commitment of the proposition is expressed by *-kamo-sjireN*, or *-kamo-sjitan* as in:

- (46) a. *asjita-wa toRkjo-N iŋ-kamo.sjireN*  
 tomorrow-TOP Tokyo-ALL go-maybe  
 ‘Someone may go to Tokyo tomorrow.’
- b. *asjita-wa ame-zjai -kamo.sjitan*  
 tomorrow-TOP rain-COP.NONPAST maybe  
 ‘It may rain tomorrow.’

Deontic ‘must’ is expressed using the negative forms as a base, suffixed with *-igan* ‘will not go’ or *-naran* (or *-suman* ‘will not do’ for stronger obligation) with an intervening *-nja-*, as in *ogiran-nja-igan* ‘I must get up’, *sen-nja-naran* ‘I must do it’.

- (47) *joga joRhuŋ kite-dji iŋ-an.nja.naran*  
 good clothes wear-GER go-must  
 ‘I must go (there) wearing a good dress.’

The volitional is expressed by *-ro* for vowel-stem verbs and *-o* for consonant-stem verbs, as in *ogi-ro* ‘I will get up’, *ne-ro* ‘I will go to sleep’, and *kag-o* ‘I will write’.

- (48) *asjita-wa hajaga-zje ne-ro*  
 tomorrow-TOP early.NONPAST-because sleep-VOL  
 ‘I will go to bed since I (have to) get up early tomorrow.’

The desiderative is expressed by attaching *-ŋozjaŋ* after the volitional forms: *mi-ro-ŋozjaŋ* ‘I want to see’, *ogi-ro-ŋozjaŋ* ‘I want to get up’, *iŋ-o-ŋozjaŋ* ‘I want to go’, and *kag-o-ŋozjaŋ* ‘I want to write’.

### 11.3.4.5 Potential

The Kyūshū dialects are well known for their retention of two types of ability expressions as mentioned in 11.2.3.5: circumstantial vs. ability potential. In a number of northern Kyūshū dialects, such as those in Fukuoka, Ōita, or Kumamoto prefectures, different predicates exist for these two types; e.g. *jomi-kiru* ‘can read’ means ‘someone has the ability to read’, while *yom-aruru* ‘can read’ means ‘a condition for reading is met’, such as that the room is bright enough to read something (GAJ, vol. 4), although the distribution of these two types varies according to dialect in Kyūshū (Kibe 2008).

However, in Ei, as well as most of the dialects in the southern Kyūshū area, especially in Kagoshima prefecture, such a distinction is no longer made; they are both expressed by non-past + *-ŋa* nominative + *-na-i* ‘become’, as in *ki-i-ŋa-na-i* ‘he can wear’ and *jon-ŋa-na-i* ‘he can read’, whose negative equivalents are *-ŋa-nar-an*, as in *ki-i-ŋa-nar-an* ‘he cannot wear’ and *jon-ŋa-nar-an* ‘he cannot read’.

- (49) a. *kon ko-wa ojoQ-ŋa.na-ran*  
 this child-TOP swim-can-NEG  
 ‘This child is not able to swim.’
- b. *kon un-ŋja annaga-dji ojoQ-ŋa.na-ran*  
 this ocean-TOP dangerous-GER swim-can-NEG  
 ‘You cannot swim in this ocean because it’s dangerous.’

### 11.3.4.6 Non-declarative sentence types

Questions are made by putting *-ka* in informal register (addressed to equals or subordinates), but the more polite form *-kana* is used for formal register; e.g. *oi-ka* ‘Are you there?’ or *oi-kana* ‘Are you there?’. In speaking with superiors, the polite marker *-jai* is required between the verb and *-kana*, as in (50b).

- (50) a. *iQ teŋan-nju kaQ-ka (~ -kana)*  
 when letter-ACC write-Q  
 ‘When are you going to write the letter?’
- b. *iQ teŋan-nju kaQ-jai-kana?*  
 when letter-ACC write-POL-Q  
 ‘When are you going to write the letter?’

Commands are expressed by suffixing *-e* to consonant-stem verbs and, similarly to Tsuruoka dialect, *-re* to vowel-stem verbs, instead of *\*-ro* as in ModJ; thus, the imperative forms for *ogi-* ‘get up’ and *mi-* ‘see’ are *ogi-re* ‘Get up!’ and *mi-re* ‘See!’. However, commands in Ei are often expressed by negative questions such as *ogi-ran-ka* ‘Get up!’.

## 11.3.5 Topic

The topic marker is *-wa*, the allomorphs of which were discussed in 11.3.3.

- (51) a. *minado-wa soqe aQ*  
 port-TOP there.LOC be.NONPAST  
 ‘The port is there.’

- b. *a<sub>N</sub> hito-wa isado<sub>N</sub> zja*  
 that person-TOP doctor COP  
 ‘That person is a doctor.’

Emphatic topic, often conveying contrastive meaning, is expressed by *-nara*, as in:

- (52) *soi-nara sji<sub>Q</sub>-cjo-i*  
 that-TOP know-CONT-NONPAST  
 ‘As for that, I know it.’

### 11.3.6 Passive and causative

The passive is expressed by a suffix *-raru* for vowel-stem verbs and *-aru* for consonant-stem verbs, as in (53b). The whole stem follows the lower bigrade conjugational pattern shown in Table 11.20. For *-rare* for vowel-stem verbs, the suffix-internal *r* is often deleted as in *-rae*, as illustrated in (53c):

- (53) a. *oja-**ŋa** kodon-nju tada<sub>Q</sub>*  
 parent-NOM child-ACC hit.NONPAST  
 ‘The parent hits the child.’  
 b. *kodon-**ŋa** oja-gara tadag-aru-i*  
 child-NOM parent-ABL hit-PASS-NONPAST  
 ‘The child is hit by his parent.’  
 c. *iN-gara ukage-rae-da*  
 dog-ABL chase-PASS-PAST  
 ‘(Someone) was chased by a dog.’

Notice here that the agent is marked by *-gara* ablative, a feature shared with Tsuruoka dialect (23), in contrast with ModJ **-ni**.

The causative is expressed by *-sasu* for vowel-stems, *age-sasu-i* ‘have somebody open something’, and *-asu* for consonant-stems, *kag-asu-i* ‘have someone write something’, or irregular verb *kor-asu-i* ‘have someone come’:

- (54) a. *mano-N to-o age-sasu-i*  
 grandchild-DAT door-ACC open-CAUS-NONPAST  
 ‘(Someone) has a grandchild open the door.’  
 b. *mano-N te<sub>ŋ</sub>an-nju kag-asu-i*  
 grandchild-DAT letter-ACC write-CAUS-NONPAST  
 ‘(Someone) has a grandchild write a letter.’

The *-(s)asu* also conjugates as a lower bigrade as in Table 11.20. The causative passive is marked by *-(s)ase-ra(r)e* as in:

- (55) *aN-ja<sub>Q</sub>-kara kaisja-o jame-sase-rae-da*  
 that-person-ABL company-ACC quit-CAUS-PASS-PAST  
 ‘(Someone) was made to quit the company by that man.’

### 11.3.7 Subordination

Adjectives, genitive phrases, and relative clauses are positioned to the left of the head noun as in *hutoga heja* ‘big room’, *onajo-N heja* ‘woman’s room’, *toRkjo-N i<sub>Q</sub> mae* ‘before going to Tokyo’.

Subordinate clauses are marked by a complementizer *-to* or *-koQ* as in:

- (56) a. *asjita sjiken-ŋa aQ-to sjiQ-cjo-i-ka*  
 tomorrow test-NOM be.NONPAST-COMP know-CONT-NONPAST-Q  
 ‘Do you know that we will have an exam tomorrow?’
- b. *toRkjo-N iQ-kodj-u sjiQ-cjo-i-ka*  
 Tokyo-ALL go.NONPAST-COMP-ACC know-CONT-NONPAST-Q  
 ‘Do you know that (someone) is going to Tokyo?’

Quotations are marked by a complementizer *-tji* as in:

- (57) a. *aN hito-wa kuQ-tji juR-zjoQ-ta*  
 that person-TOP come.NONPAST-QUOT say-CONT-PAST  
 ‘(Someone) was saying that he/she will come.’
- b. *hajo koge keR-tji juR-da*  
 early here.LOC come.IMP-QUOT say-PAST  
 ‘(Someone) told (someone) to come here quickly.’

The conditional/provisional is expressed most commonly by *-reba* for adjectives and vowel-stem verbs or *-eba* for consonant-stem verbs, as in (58):

- (58) a. *SON-nju (sogen) taQka<sup>20</sup>-reba kow-aN*  
 that.much expensive-COND buy-NEG  
 ‘If it is so expensive, I will not buy it.’
- b. *asjita ame-ŋa hur-eba hune-wa de-ran-zjaro*  
 tomorrow rain-NOM fall-COND boat-TOP go.out-NEG-MODAL  
 ‘If it rains tomorrow, the boat will not go out (set sail).’

Finite verb-form + suffix *-ja* ‘when, if’ is also often used for the conditional as in (58c,d):

- (58) c. *soge i-da-ja k<sup>v</sup>ai-ja hiQsun-zjoQ-ta*  
 there.ALL go-PAST-when meeting-TOP finish-PERF-PAST  
 ‘When (someone) went there, the meeting had already finished.’
- d. *midj-i cjike-da-ja aR.sjita iro-ni kawaQ-ta*  
 water-ALL dip.into-PAST-when that.way color-DAT change-PAST  
 ‘When (someone) put it into water, it changed into such a color.’

Thus, instead of (58a), *SON-nju tagaga-ja kow-aN* is also possible.

A hypothetical clause is expressed by non-past + *-nara*, with assimilation of *Q* or *i* to *N*, as in *ogi-N-nara* ‘If (someone) wakes up’, and *kan-nara* ‘If (someone) writes (something)’.

- (59) *kan-nara cjanto kag-e*  
 write.NONPAST-if properly write-IMP  
 ‘If you write it, write it properly.’

Subordinate clauses for reason/cause are marked by *-zje*, as in:

- (60) a. *ame-ŋa huQ-cjoQ-zje in-na*  
 rain-NOM fall-CONT.NONPAST-because go-IMP.NEG  
 ‘Since it is raining, don’t go.’

- b. *kon zji-ja sjiQ-cjoQ-zje kaQ-ŋanai*  
 this letter-TOP know-CONT-because write-POT.NONPAST  
 ‘As I know this letter, I can write it.’

The concessive suffixes include *-baQ* as in:

- (61) a. *hijaga-baQ kibar-o*  
 cold.NONPAST-even.though make.effort-VOL  
 ‘Even though it is cold, I’ll (let’s) do our best.’
- b. *juQ-ŋa huQ-ta-baQ hijaga-naga*  
 snow-NOM fall-PAST-even.though cold-NEG.NONPAST  
 ‘Even though it snowed, it is not cold.’

#### 11.4 CONCLUSION

We have seen that, in spite of the fact that these two dialects (Tsuruoka and Ei) are geographically separated from each other, they share some common features which do not exist in ModJ. To name a few, in both dialects, (1) vowel coalescence (e.g. *ai > ε*, *ai > e*) has occurred, (2) intervocalic voicing of voiceless obstruents has taken place (e.g. *makura > magura*), (3) the intervocalic velar nasal /ŋ/ is retained, (4) old directional particles (*-sa* in Tsuruoka and *-same* in Ei) are preserved, and (5) the agent of the passive is marked by *-gara* ablative as opposed to ModJ *-ni* in both dialects. Such similarity is not only due to shared vestiges of some archaic features of proto-mainland Japanese; some of their common features are thought to have occurred independently.

Especially noteworthy in this respect are their verb conjugation discussed in 11.2.2.3 and 11.3.2.3; due to reanalysis and analogy, both dialects have developed apparently different types of conjugational paradigms from ModJ, but what motivated such developments is similar between these two dialects. That is, there is in both dialects a strong tendency to merge toward the paradigms of certain conjugation types; i.e. the quinquagrade and lower bigrade conjugation patterns are more productive than the others.

Such generalizations about diachronic change in Japanese would not be possible if we only base our study on ModJ data. Thus, we contend that the inclusion of dialects in the study of Japanese will stimulate and enrich the field, by providing significant insight into how language systems change, how they develop their typologically natural features independently, and how seemingly unrelated features of different dialects are motivated by the same cause.

#### NOTES

- 1 The authors are indebted to Prof. Nobuko Kibe of the National Institute of Japanese Language and Linguistics (NINJAL), formerly at Kagoshima University, for giving us useful comments and advice at the initial stage of this research, as well as introducing us to a perfect consultant for the Ei dialect. We also thank Mr. Toshihiro Nishi for painstakingly teaching us the dialect of Ei on a number of occasions. Deepest gratitude is also expressed to the people in Tsuruoka city, who have provided us some excellent data for this chapter. We are also grateful to Nicolas Tranter, the editor of this book, and Prof. J. P. Barron of Tokyo Medical University, who have read and given us valuable comments on the earlier draft of this chapter.
- 2 Recently, Inoue (1994: 52) also reports that these centralized vowels tend to be lost and changed into non-centralized versions, undoubtedly due to the influence of ModJ.

- 3 In Tsuruoka, such fronting of post-coronal /u/ has also shown a strong tendency to disappear in recent years, due to the influence of ModJ, so that the words containing it are now pronounced with [u]: e.g. [tsuzi] ‘soil’ **tsuchi**, [mizu] ‘water’ **mizu**, and [sumi] ‘coal’ **sumi** (Inoue 1994: 68–9, NINJAL 2006b: 39–40). See also consistent *das-u* or *su-ru* verb forms.
- 4 Inoue (1994) reported that in the second and third field trips by NINJAL, this prenasalized quality of /b/, /d/, and /z/ has been totally lost in the Tsuruoka dialect, so that the words listed in (1) are now pronounced as [nabe] ‘pot’, [kabi] ‘mold’, [ūde] ‘arm’, [kado] ‘angle’, [aza] ‘bruise’, and [sadzi] ‘spoon’.
- 5 The palatals /tj/, /sj/, and /zj/ are phonetically [tʃ], [ʃ], and [ʒ] (or [dʒ]), respectively.
- 6 Hereafter, the data cited from other sources are transcribed based on the phonemic system of this chapter; e.g. even though the vowel in *cjoR* ‘butterfly’ in (5a) is described as short in Satō (1994), we will use *R* throughout this section of Tsuruoka to show that, although phonetically subtle, the length distinction still exists in this dialect: *toQ-ta* ‘take-PAST’ vs. *toRQ-ta* ‘pass-PAST’.
- 7 In ModJ, **-koto** is also used in the form of **-no-koto** for the first- or second-person pronouns ‘I’, ‘we’, or ‘you’ or proper names with certain types of predicates, as in **watashi-no-koto dō omoo-te iru** ‘What are you thinking about me?’ or **hanako-no-koto-ga kirai-da** ‘I don’t like Hanako’ (see Tsunoda 1991).
- 8 Inoue (2000: 80) reported that in the field research conducted in 1991, expressions for circumstantial potential and ability potential were distinguished in a village in the suburbs of Tsuruoka; the former marked with *-aeru* as in *ojog-aeru* ‘can swim (the situation allows someone to swim)’, the latter with *-eru* as in *ojog-eru* ‘can swim (someone has an ability to swim)’.
- 9 This change is sometimes observed in the middle of a single word, so that the cognates of ModJ words **kagoshima** ‘Kagoshima (place name)’ and **musume** ‘daughter’ are pronounced *kajoima* and *muime*, respectively, in Ei.
- 10 Although Shibata (1959) reported based on his field trip in the 1950s that /d/ and /b/ also showed a prenasalized quality in this dialect, the present Ei dialect does not seem to retain a nasalized quality for /d/ or /b/, as in *hude* ‘brush’, *mado* ‘window’, or *nabe* ‘pot’.
- 11 As with the Tsuruoka dialect, the palatals /tj/, /sj/, and /zj/ represent [tʃ], [ʃ], and [ʒ], respectively.
- 12 For some unknown reason, the negative suffix *-na-ga* (or any suffixes following the suffix) is not included in the accentual phrase; thus, instead of *\*ago-ná-ga*, *\*siro-na-gá*, *\*ago-na-gáQ-ta*, *\*siro-na-gaQ-tá*, we have *ágo-na-ga*, *siró-na-ga*, *ágo-na-gaQ-ta*, and *siró-na-gaQ-ta*.
- 13 Moreover, in most dialects in the Kinki (Kyoto-Osaka) regions, the old two-pattern accent distinction of adjectives has been totally lost.
- 14 It seems that this dialect has a tendency to contract words including sequences of voiced stops: *taga-ga* ‘expensive’ > *taoka*, *hadage* ‘field’ > *haote*.
- 15 The past tense form of historical upper bigrade verbs *ogi-* ‘get up’ **oki-**, and *iki-* ‘live’ **iki-** contain insert /o/ as in *ogi-o-da* (~ *ogi-o-ta*), or *iki-o-da* (~ *iki-o-ta*). This may also be due to analogy with /r/-stem verbs.
- 16 This is probably due to the resyllabification (and subsequent deletion of the onset consonant of the suffix initial /w/) caused by the suffixation of the topic marker *-wa* to these nouns. Thus, for example, when the suffix *-wa* is attached to *tao* ‘socks’ (derived from underlying /tabi/ by the word-final deletion of /i/), we get /tab-wa/, from which /w/ is deleted because the onset consonant clusters of a syllable are not allowed in this dialect.
- 17 In other western dialects, the corresponding copulae for adjectival nouns are *-ja* in the Kinki area and *-na* in most of the Shikoku and Chūgoku regions (see GAJ, vol. 3).
- 18 This *-i* may have been derived by the same process as the word-final *sj* and *s* are changed to *i*, as illustrated in 11.3.1.1. Due to the word-final deletion of high

vowels /i, u/, -ru, as in *oki-ru* and *ake-ru* is changed into -r, which may then be changed to -i, by the vocalization of the word-final coronal continuant consonants /sj, s, r/.

- 19 The reason why the historical upper bigrade verbs like *ogi-* ‘get up’ contain the insert /o/ as in *ogi-o-zjoi* may be due to analogy with /r/-stem verbs (see note 15).  
 20 Here, *tao-ka* is a contracted form of *taga-ga* (see note 14).

## REFERENCES

- Hidaka, Mizuho (2000) ‘Akita-hōgen no Bunpō’ [Grammar of the Akita dialect], in Akita-ken Kyōiku-iinkai (ed.) *Akita no Kotoba* [The language of Akita], Akita: Mumei-sha: 74–132.
- (2005) ‘Hōgen ni okeru Bunpō-ka’ [The regional differences of the grammaticalization in the Tōhoku dialects], *Nihongo-no Kenkyū*, 1(3): 77–91.
- Inoue, Fumio (1994) ‘Tsuruoka-hōgen no On’in’ [The phonology of the Tsuruoka dialect], in NINJAL (ed.) (1994): 37–79.
- (2000) *Tōhoku-hōgen no Hensen* [The change of the Tōhoku dialect], Tokyo: Akiyama-shoten.
- Kibe, Nobuko (2008) ‘Kyūshū no Kanō-hyōgen no Shosō’ [Various aspects of potential expressions in the Kyushu dialects], *Kokugo Kokubun Satsumaji*, 48: 1–14.
- Kobayashi, Takashi (1997) ‘Shūken-bunpu no Tōzai-sa’ [Variation in the allative case marker -sa between eastern and western dialects], *Kokugo-gaku*, 188: 26–39.
- Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.
- National Institute of Informatics (2008), *Tsuruoka Chōsa Onsei Dēta Bēsu 91* [Speech database for the 1991 Tsuruoka survey], DVD-R.
- National Institute of Japanese Language and Linguistics (Kokuritsu Kokugo Kenkyūjo) (ed.) (1953) *Chiiki-shakai no Gengo-seikatsu: Tsuruoka ni okeru Jittai-chōsa* [Language studies of a local society: field work studies of the Tsuruoka dialect], Tokyo: Shūei-shuppan.
- (ed.) (1959) *Nihon Hōgen no Kijutu-teki Kenkyū* [Descriptive studies on Japanese dialects], Tokyo: Meiji-shoin.
- (1966–1974) *Nihon Gengo Chizu* [Linguistic atlas of Japan (LAJ)], vols 1–6, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (1974) *Chiiki-shakai no Gengo-seikatsu: Tsuruoka ni okeru Nijū-nen mae tono Hikaku* [Language studies of a local society: a comparative study of the Tsuruoka dialect from 20 years ago], Tokyo: Shūei-shuppan.
- (1989) *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 1, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (1991) *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 2, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (1993) *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 3, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (ed.) (1994) *Tsuruoka Hōgen no Kijutsu-teki Kenkyū: Dai-sanji Tsuruoka Chōsa houkoku 1* [A descriptive study of the Tsuruoka dialect: the third language survey in Tsuruoka city, the 1<sup>st</sup> report], Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (1999), *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 4, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (2002), *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 5, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (2006a) *Hōgen Bunpō Zenkoku Chizu* [Grammar atlas of Japanese dialects (GAJ)], vol. 6, Tokyo: Kokuritsu Kokugo Kenkyūjo.
- (ed.) (2006b) *Hōgen Shijō no Bamen-teki Tayō-sei: Tsuruoka-shi niokeru Bamen-sa Chōsa kara* [The variational differences of the dialect use: based on the research on

- the situational differences of (dialect use) in Tsuruoka city], Tokyo: Kokuritsu Koku-go Kenkyūjo.
- Nitta, Tetsuo (1994) 'Tsuruoka-hōgen no Akusento' [Accent in the Tsuruoka dialect], in NINJAL (ed.) (1994): 81–140.
- Onishi, Takuichiro (1994) 'Tsuruoka-shi Ōyama-hōgen no Yōgen no Katsuyō' [Verb conjugation of the Ōyama dialect in Tsuruoka city], in NINJAL (ed.) (1994): 141–235.
- (1998), 'Dōshi-katsuyō no Taiō to Hikaku' [The correspondence and comparison of verb conjugations in Japanese dialects], *Gekkan-gengo*, 27(7): 49–56.
- (1999) 'Hōgen no Dōshi to Nihongo-shi' [Verbs in dialects and the history of Japanese], *Nihongo-gaku*, 18(5): 55–64.
- (2006) 'Hōgen Bunpō Zenkoku Chizu no Igi to Hōgen Bunpu Kenkyū no korekara' [The role of *Grammar Atlas of Japan* and the future studies of dialect distribution], *Gekkan-gengo*, 35(12): 20–7.
- (2008) 'Proto-Japanese and the Distribution of Dialects', in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese: Issues and Prospects*, Amsterdam: John Benjamins, 57–78.
- Satō, Ryōichi (1994) 'Tsuruoka-hōgen ni okeru Joshi "sa" no Yōhō' [The usage of the particle *sa* in Tsuruoka dialect], in NINJAL (ed.) (1994): 267–86.
- Shibata, Takeshi (1959) 'Kagoshima-ken Ibusuki-gun Ei-chō' [Ei in Ibusuki county in Kagoshima prefecture], in NINJAL (ed.) (1959): 315–42.
- Shibuya, Katsumi (1994) 'Tsuruoka-hōgen no Tensu to Asupekuto' [Tense and aspect in the Tsuruoka dialect], in NINJAL (ed.) (1994): 238–66.
- Tsunoda, Tasaku (1991) *Sekai-no Gengo to Nihon-go* [The languages in the world and Japanese], Tokyo: Kuroshio-shuppan.
- Uwano, Zendo (1977) 'Nihongo no Akusento' [Accent of Japanese], in Susumu Ōno and Takeshi Shibata (eds) *Iwanami Kōza Nihongo* [The Iwanami lectures for Japanese], Tokyo: Iwanami-shoten, 281–321.
- Vance, Timothy (1987) *An Introduction to Japanese Phonology*, New York: State University of New York Press.

PART IV

**RYUKYUAN**



# NORTHERN RYUKYUAN

*Michinori Shimoji*

## 12.1 INTRODUCTION

The Ryukyuan languages are in a sister relationship with Japanese, and these two language groups form the Japonic family. The Ryukyuan languages fall into Northern Ryukyuan (Amami and Okinawan) and Southern Ryukyuan (Miyako, Yaeyama, and Yonaguni, see Lawrence's and Izuyama's chapters in this volume). This chapter presents a grammatical sketch of Northern Ryukyuan, with an exclusive focus on the Shuri dialect (Okinawan). Shuri is spoken in an area that was once the capital of the Ryukyu dynasty (AD 1429–1879), and has thus been regarded by the local people as a dominant and prestigious language among all Ryukyuan languages and their sub-varieties. Its written history goes back to the sixteenth century, adapting the Japanese writing system. All Ryukyuan languages including Shuri are, however, in imminent danger of extinction, given that fluent native speakers are in their sixties or older, and that younger generations do not acquire or use Ryukyuan in their daily life, nor do they have a chance to learn it in school. All but two of the eight endangered languages identified by Moseley (2010) are Ryukyuan, including (Shuri) Okinawan as 'definitely endangered'.

The description below is not based on my own fieldwork, but on the vast literature on Shuri, most of which is written in Japanese. However, the analysis and terminology adopted in this chapter are substantially different from those in the past works I refer to. Kokuritsu Kokugo Kenkyūjo (1963), Uemura (1997), Tshako (1997), and Nishioka and Nakahara (2010) were my major sources of information and examples of Shuri. Most examples cited come from previous works like these, and I indicate the source in example sentences.

## 12.2 PHONOLOGY

Shuri phonology is characterized by glottalized nasals and glides. Their phonemic treatment is one tricky area of Shuri phonology (see 12.2.3.3). Most syllables in roots have a cross-linguistically common (C)V(C) structure where the onset and the coda may be absent (see 12.2.3.1 for detail), but there do exist complex structures especially at root-initial position, where a geminate onset (/kkwa/ 'child') and/or a pre-onset glottal stop (as in /?waa/ 'pig' and /?mmee/ 'grandmother') may occur. Shuri has a pitch accent system in which two lexically determined accents are distinguished. Major phonological constraints and processes such as phonotactics, minimality constraint and accent apply to a unit consisting of a morphosyntactic word plus an enclitic or a series of enclitics, and I call this entire phonological unit the phonological word. All phonological words must have at least two morae.

### 12.2.1 Consonants

The consonant phoneme inventory of Shuri is shown in [Table 12.1](#). There are five phonemic places of articulation and six phonemic manners of articulation. Stops and affricates

TABLE 12.1 SHURI CONSONANTS

		Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	voiceless	p	t		k	ʔ
	voiced	b	d		g	
Affricate	voiceless			c [t͡ɕ]		
	voiced			z [d͡ʒ]		
Fricative			s			h
Tap			r [ɾ]			
Nasal		m	n			
Glide		w		j		

except for /ʔ/ contrast in voicing. The phoneme /ʔ/ occurs phonological-word-initially (see 12.2.3.3). The phonemes /c/ and /z/ are palatal rather than alveolar affricates. The phoneme /s/ is pronounced as [ɕ] before /i/ and /e/. The phoneme /n/ is a homorganic nasal, and has allophones of different places of articulation depending on the consonant that follows it. The glide phonemes may combine with other consonants to form complex onsets, as in /hwii/ [ɸi:] ‘day’, /kkwa/ [kk<sup>w</sup>a] ‘child’, /bjooin/ [b<sup>o</sup>:iŋ] ‘hospital’ (Sino-Japanese loanword). The consonants that combine with /w/ are velar stops and the glottal fricative /h/. The consonants that combine with /j/ are any consonants except for /t/, /d/, /ʔ/, and /w/.

### 12.2.2 Vowels

The vowel system of Shuri is straightforward, with the five cross-linguistically common vowels /a/, /i/, /u/, /e/, and /o/, though /e/ and /o/ normally only occur long. Long vowels are treated as a sequence of identical vowel phonemes, as in /kaa/ ‘river’ [ka:]. This analysis is based on the fact that the mora (see 12.2.3.2) is crucial in Shuri phonology (much more so than the syllable) and that there are a number of cases where a single syllable with a long vowel (two morae) behaves like two light syllables (also two morae).

### 12.2.3 Suprasegmentals

#### 12.2.3.1 The syllable

The syllable structure of Shuri varies depending on the position of the syllable within a phonological word. The syllable template is schematized as (#ʔ)((C<sub>i</sub>)C<sub>i</sub>)V<sub>1</sub>(V<sub>2</sub>)(C), where the symbol (#ʔ) is a special pre-onset slot for the glottal stop phoneme, which occurs in phonological-word-initial position. A geminate onset is possible (e.g. /ccu/ [t͡tɕu] ‘person’ and /kkwa/ [kk<sup>w</sup>a] ‘child’). The glottal plus a geminate consonant is also possible, and in this case the geminate consonant must be a nasal: /ʔmmee/ [ʔ<sup>h</sup>mme:] ‘grandmother’ and /ʔnni/ [ʔ<sup>h</sup>nni] ‘rice plant’. Underlying word-initial /i/u/ before nasals typically are realized as /ʔ/, e.g. *uma* ‘there’ > ʔ*mma*. The glide phonemes (/w/ and /j/) do not have a particular position within the syllable. They fill the onset slot that is not filled by other consonants. For example, in /jaa/ [ja:] ‘house’, the glide is analyzed to fill the onset slot, as this slot is not filled by any other consonant. If the onset is already filled, then the glide does not have its own position in the syllable, and it is phonetically realized as labialization or palatalization of the onset.

### 12.2.3.2 *The mora*

As in most other Japonic varieties, the mora is a crucial phonological unit in Shuri. For example, as will be shown in 12.2.3.3, a frequent reference must be made to account for the accentual system. As shown schematically in (1) below, a nucleus ( $V_1$  or  $V_2$ ), a coda ( $C_2$ ), and the left-most onset have one mora each, and all other syllable slots including the pre-onset  $\text{?}$  are mora-less. As in many Japonic varieties, Shuri has a word minimality constraint whereby all phonological words must have at least two morae.

- (1) (#?) (( $C_1$ )  $C_1$ )  $V_1$  ( $V_2$ ) ( $C$ )  
 –      $\mu$     –      $\mu$      $\mu$       $\mu$

A couple of comments are necessary for the glottal pre-onset slot and the onset slots. First, there is no phonological word that has the structure  $\text{?CV}$  (e.g.  $/\text{?wa}/$ ). This fact tells us that the  $\text{?CV}$  has one mora, thus violating the minimality constraint. That is, the glottal pre-onset slot is proven to be mora-less. Second, there are a few phonological words that do have the structure  $C_1C_2V$  (e.g.  $/\text{ccu}/$  [ $\text{t}t\text{eu}$ ] ‘person’ and  $/\text{kkwa}/$  [ $\text{kk}^{\text{w}}\text{a}$ ] ‘child’). This indicates that the initial geminate  $C$  has one mora, whereby the  $\text{CCV}$  structure has two morae in total, thus satisfying the minimality constraint.

### 12.2.3.3 *Phonemic treatment of the glottal stop*

There is a phonemic distinction between  $[\text{j}i]$  and  $[\text{?}i]$  (as in  $/\text{jii}/$  [ $\text{j}i:$ ] ‘picture’ vs.  $/\text{ii}/$  [ $\text{?}i:$ ] ‘stomach’), and between  $[\text{wu}]$  and  $[\text{?u}]$  ( $/\text{wutu}/$  ‘husband’ vs.  $/\text{utu}/$  ‘sound’). As indicated in the phonemic transcriptions here, these distinctions are interpreted as the presence or absence of the glide phonemes  $/\text{j}/$  and  $/\text{w}/$  before the vowel. The other vowels do not have this phonemic contrast; the occurrence of the glottal stop with these vowels is predictable. That is, they carry a clear glottal stop in word-initial position, as in  $/\text{atai}/$  [ $\text{?}atai$ ] ‘kind of garden’ and  $/\text{eesaci}/$  [ $\text{?e:satei}$ ] ‘greeting’; however, in word-medial position this glottal stop disappears, as in  $/\text{hanaatai}/$  [ $\text{hana.atai}$ ] ‘flower garden’ and  $/\text{bueesaci}/$  [ $\text{bu.e:satei}$ ] ‘unfriendly’ (lit. ‘no greeting’).

### 12.2.3.4 *Accent*

Shuri, like other dialects of Ryukyuan, has a pitch accent system in which one mora within a phonological word carries the accent. There is abrupt falling pitch after the accent. Accent is contrastive and lexically determined. Two accentual patterns are identified. One is accentless, with a level tone throughout the phonological word, and I will refer to this pattern as Type A.

Type A: accentless

- (2)  $/\text{hana}/$  [ $\text{hana}$ ] ‘flower’  
 (3)  $/\text{hana}/$  +  $/\text{nu}/$  (nominative case clitic)  $\rightarrow$   $/\text{hananu}/$  [ $\text{hananu}$ ]  
 (4)  $/\text{hwii}/$  [ $\text{?}i:$ ] ‘fire’  
 (5)  $/\text{hwii}/$  +  $/\text{nu}/$  (nominative case clitic)  $\rightarrow$   $/\text{hwiinu}/$

On the other hand, Type B involves falling pitch between the second and the third mora of a phonological word. However, in the case of bimoraic words, the accent falls between the first and the second mora. This accentual pattern applies across different word classes. In the following examples, the accent is marked by the symbol ‘ˆ’ after the accented mora.

Type B: accented

- (6) /ha`na/ [hána] ‘nose’  
 (7) /hana/ + /nu/ (nominative case clitic) → /hana`nu/ [hánanù]  
 (8) /hwi`i/ [φii] ‘day’  
 (9) /hwii/ + /nu/ (nominative case clitic) → /hwii`nu/ [φíinù]

## 12.3 DESCRIPTIVE PRELIMINARIES

### 12.3.1 Shuri morphosyntax: a typological overview

Shuri shares major morphosyntactic features with Japanese: agglutinative and suffixing morphology, verb-final (typically AOV in transitive clauses) and modifier-head constituent order, nominative-accusative case system (in which direct object is unmarked), and a rich inventory of dependent clause markers as verb inflection. Adjectives are a subclass of verbals, having verb-like characters such as inflection and predicative function.

### 12.3.2 Word classes

There are seven word classes in Shuri: nominal, verbal (verbs and adjectives), adnominal, conjunction, interjection, adverb, post-phrasal marker, of which the first two are large classes (Table 12.2). The suggested criteria for word class assignment are: (A) whether the word heads a nominal phrase (NP), (B) whether the word directly fills the modifier slot of an NP, (C) whether the word inflects.

#### 12.3.2.1 Nominals

A nominal is a word that only heads an NP. An NP is independently defined as a constituent that functions as an argument or as a predicate nominal. The nominal word class comprises nouns (e.g. *ccu* ‘person’, *ʔwaa* ‘pig’, *saki* ‘saké’), pronouns (e.g. *wan* ‘1SG’, *kuri* ‘this’, *taa* ‘who’), and numerals (e.g. *miici* ‘three (things)’, *miccai* ‘three people’, *ikuci* ‘how many (things)’).

#### 12.3.2.2 Adnominals

The adnominals form a small and closed set of words. An adnominal functions solely as a modifier of an NP. This class comprises demonstrative adnominals such as *kunu* ‘this’, *unu* ‘that’ (medial), and *anu* ‘that’ (distal), demonstrative approximate adnominals such as *kunna* ‘this kind of’, *unna* ‘that kind of’, and *anna* ‘that kind of’, and interrogative adnominals such as *canu* ‘which’. It may also be possible to include those fossilized adnominal verb forms that do not synchronically function as verbs (i.e. never show conjugation): *aru* ‘one’ (< ‘exist’), *ʔnzaru* ‘leave’ (< ‘leave’), and *cuuru* ‘upcoming’ (< ‘come’).

TABLE 12.2 WORD CLASSES IN SHURI

	(A)	(B)	(C)
Nominal	+	–	–
Adnominal	–	+	–
Verbal (including adjective)	–	–	+
Others (conjunction, interjection, adverb, post-phrasal marker)	–	–	–

### 12.3.2.3 Verbals

Verbals are the only word class in Shuri that shows inflection. They divide into the verb and adjective. Verbals occur in verbal predicate phrases, except for the copula, which is a verb that occurs in a nominal predicate phrase, after a predicate nominal.

Both verbs and adjectives inflect word-finally:

#### (10) Verbs

*tu-ta-ru* (take-PST-ADN) ‘took’ (past adnominal)

*tu-ti* (take-CVB.SEQ) ‘take, and’ (sequential converb)

#### (11) Adjectives

*aka-sa-ta-ru* (red-VLZ-PST-ADN) ‘was red’ (past adnominal)

*aka-sa-i* (red-VLZ-CVB.SEQ) ‘be red, and’ (sequential converb)

### 12.3.2.4 Others

This category subsumes a set of words that do not satisfy any of the criteria (A) to (C). Yet, it is convenient to divide this catch-all category into several divisions according to their syntactic distribution: conjunctions (e.g. *ansi* ‘and’, *jasiga* ‘but’, *unuwii* ‘furthermore’, etc.), interjections (*ii* ‘Yes’, *iiii* ‘No’, *ai* ‘Oh!’, etc.), adverbs (e.g. *dateen* ‘very’, *hwee-ku* ‘quick-ly’, etc.), and post-phrasal markers. Post-phrasal markers are either argument markers (i.e. case markers, limiters such as *=n* ‘too; even’, *=naa* ‘only’, and topic/focus markers such as *=du* (FOC) and *=a* (TOP)), or predicate markers such as bound conjunctions (e.g. *=siga* ‘though’), modal markers (e.g. *=hazi* ‘maybe’), and discourse markers (e.g. *=jaa* ‘isn’t it?’, *=joo* (47), *=sa* ‘I tell you’, *=na* (49), etc.). All of these are clitics that attach phonologically to the last word of the phrase they attach to syntactically.

## 12.4 MORPHOLOGY

Shuri morphology is agglutinative and mostly suffixal. Morpheme boundaries are not always clear, with a very complex system of allomorphy found especially in verbs. Besides affixation, compounding is also common (see 12.4.3). Nouns do not inflect, and case is marked by a case clitic, which is phonologically attached to the last word of a noun phrase it syntactically attaches to. Verbs inflect for tense (past vs. non-past) and mood for finite verbs, or for various adverbial-adsentential relations for non-finite verbs.

### 12.4.1 Nominals

Nominals divide into nouns, pronouns, and numerals.

#### 12.4.1.1 Pronouns

There are personal, location, and direction pronouns in Shuri (Table 12.3), in which only the personal pronouns distinguish number (singular vs. plural). Each pronominal set includes an interrogative form, which is used in questions (e.g. *taa* ‘who’, *nuu* ‘what’, *maa* ‘where’, and *cagata* ‘which direction’). The non-speech-act participant reference (i.e. third person, location, and direction) is intertwined with the demonstrative system, in which the distinction is made between distal, medial, and proximate.

TABLE 12.3 SHURI PRONOUN SYSTEM

	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>			Interrogative		
			proximate	medial	distal			
singular	<i>wan</i>	<i>?jaa</i>	<i>ku-ri</i>	<i>u-ri</i>	<i>a-ri</i>	<i>taa</i>	<i>nuu</i>	personal
plural	<i>wa-ttaa</i>	<i>i-ttaa</i>	<i>ku-ttaa</i>	<i>u-ttaa</i>	<i>a-ttaa</i>	<i>ta-ttaa</i>		
			<i>ku-ma</i>	<i>u-ma</i>	<i>a-ma</i>	<i>maa</i>		location
			<i>ku-gata</i>	<i>u-gata</i>	<i>a-gata</i>	<i>ca-gata</i>		direction

#### 12.4.1.1.1 Plural suffix for personal pronouns

As is shown in the table above, it is possible to identify the plural morpheme *-ttaa* for personal pronouns. The second person form shows irregular root alternation between the singular and plural word forms.

#### 12.4.1.1.2 Social rank and pronominal system

One prominent feature of the personal pronominal system of Shuri (and some other Northern Ryukyuan varieties) is that different forms are used for second and third person referents depending on the relative social rank between the speaker (S) and the referent (2<sup>nd</sup> Prs). Thus, for the second person we have *?jaa/ittaa* (S  $\geq$  2<sup>nd</sup> Prs), *unzu/unzunaa* (S < 2<sup>nd</sup> Prs), and *naa/nattaa* (S > 2<sup>nd</sup> Prs; the referent is older). For the third person, demonstrative location pronouns (rather than demonstrative personal pronouns) are used when the person referred to is higher in social rank (e.g. *kuri* ‘this one’  $\rightarrow$  *kuma* lit. ‘this place’); when the person referred to is lower in social rank and is older than the speaker, then the following forms are used: *kuncu/kuncutaa* (proximate), *uncu/uncutaa* (medial), *ancu/ancutaa* (distal). These last forms developed from an adnominal word (e.g. *kumu* ‘this’) followed by a common noun *ccu* ‘person’.

Besides this elaborate system of pronominal honorification, there is a system of honorification that involves verb morphology. See 12.4.2.1.8 for detail.

#### 12.4.1.2 Nouns

A noun does not inflect, and may thus consist of a root alone, as in *ccu* ‘person’, *taruu* ‘Taruu (personal name)’, *?waa* ‘pig’, etc. Cross-linguistically common inflectional categories for nouns such as number, gender, and case are either absent in Shuri grammar (in the case of gender) or present in other systems than inflection (e.g. case is marked by a clitic attaching to an NP, and plural marking is derivational as shown below).

There are a number of derivational affixes (mostly suffixes). Above all, the diminutive and plural affixes are productive and frequently observed in natural discourse. The diminutive affix is *-gwaa*, which designates tininess of a referent: *tuzi-gwaa* ‘wife’, *taruu-gwaa* ‘Taruu’, *saki-gwaa* ‘(a bit of) saké’, *warabi-gwaa* ‘kid’, etc. The plural affix is *-nucaa*, which basically attaches to human nouns and kinship or social role terms only, as in *?weeka-nucaa* ‘relatives’, *wikiga-nucaa* ‘men’, *warabi-nucaa* ‘children’, etc. Other derivational affixes include the honorific affix *-ganasi* ‘Mr./Ms.’, etc.

### 12.4.1.3 Numerals

A numeral word consists of a numeral root and a classifier suffix. As in other Japonic languages, the numeral root and/or the classifier suffix may differ depending on what is being counted. Some common numeral sets are shown below, where counting is from one to ten, followed by an interrogative form (interrogative root + classifier). In some numerals, as in (12) below, numbers greater than four or five may be expressed by numerals of Sino-Japanese origin.

(12) For humans

*cu-i* ‘one person’, *ta-i* ‘two persons’, *mi-ccai* ‘three persons’, *yu-ttai* ‘four persons’, *icu-tai* ~ *gu-nin* (the latter being Sino-Japanese) ‘five persons’, *ruku-nin* ‘six persons’, *sici-nin* ‘seven persons’, *haci-nin* ‘eight persons’, *ku-nin* ‘nine persons’, *zuu-nin* ‘ten persons’; *iku-tai* ‘how many persons?’

(13) For general non-animate objects

*tii-ci* ‘one thing’, *taa-ci* ‘two things’, *mii-ci* ‘three things’, *juu-ci* ‘four things’, *ici-ci* ‘five things’, *muu-ci* ‘six things’, *nana-ci* ‘seven things’, *jaa-ci* ‘eight things’, *kukunu-ci* ‘nine things’, *tuu* ‘ten (things)’, *iku-ci* ‘how many things?’

## 12.4.2 Verbals (verbs and adjectives)

Verbs and adjectives both inflect for tense and mood (for finite inflection) or for various adverbial-adstantial relations such as sequential, causal, and conditional (for non-finite inflection). As mentioned in 12.3, this morphological homogeneity in verbs and adjectives allows us to conclude that they form a higher-order category *verbal* within the word class system of Shuri. However, a closer look at the two reveals that there are a number of morphological (and syntactic) differences between them, requiring us to subclassify them within the verbal class.

### 12.4.2.1 Verbs

#### 12.4.2.1.1 The structure

The structure of the verb is schematized as Stem (+ Pre-inflection) + Inflection, where pre-inflection (see 12.4.2.1.2) is not obligatory in all verbs. Each component may be internally complex. The verb *kacabirantan* ‘did not write (in polite form)’, for example, consists of a stem *kac-* ‘write’, pre-inflectional suffixes *-abi* (politeness) and *-ran* (negative), and inflectional suffixes *-ta* (tense: past) and *-n* (mood: indicative). By contrast, the verb *kaki* ‘write (imperative)’ only consists of the stem *kak-* and the inflectional suffix *-i* (imperative mood).

#### 12.4.2.1.2 Pre-inflection

Pre-inflection is that part of a verb which cannot be analyzed either as a derivational suffix (i.e. a part of a stem, see 12.4.2.1.10) or as an inflectional affix (tense and/or mood, see 12.4.2.1.5 and 12.4.2.1.6). For a descriptive convenience, therefore, I use the tentative third category ‘pre-inflection’, rather than pushing it into either the stem or inflection, which would only complicate the entire description of verb morphology. Pre-inflectional categories are honorification, polarity, politeness, and aspect (perfective, imperfective, and progressive). On the one hand, a pre-inflectional suffix is not like a

derivational affix since it is not lexically constrained (i.e. it may be attached to various kinds of stems). Also, pre-inflection occurs very frequently in discourse. This is in sharp contrast to verbal derivational affixes like causative and passive (12.4.2.1.10). On the other hand, pre-inflection is still unlike an inflectional affix in that it does not close off the word formation and does not obligatorily occur in all verbs, with a number of co-occurrence conditions. For example, the pre-inflectional affix *-u* (imperfective) may occur in the non-past passive form (e.g. *sugur-arij-u-Ø-n* ‘hit-PASS-IPRF-NPST-IND’), but cannot occur in the past passive form (*sugur-at-ta-n* ‘hit-PASS-PST-IND’, not \**sugur-arij-u-ta-n*).

#### 12.4.2.1.3 Finiteness

A finite verb carries an inflectional affix or affixes, inflecting for tense and/or mood. For example, the verb *icabiin* ‘went (in polite form)’ in (14) below is a finite verb, inflecting for non-past tense and indicative mood. It also carries a pre-inflectional affix *-abii* which indicates politeness. (*-abii* can be interpreted as polite *-abi* + imperfective *-u*, but for simplicity I do not gloss it this way.) A non-finite verb only inflects for adverbial-adsentential relations. For example, the converb *koojiga* ‘to buy’ in (14) below is a non-finite verb, inflecting for the purposive verbal relation with the affix *-iga* (purposive converb).

- (14) *wan=nee macigwaa=nkai kamimun kooj-iga ic-abii-Ø-n*  
 1SG=TOP market=DAT food buy-CVB.PUR go-POL-NPST-IND  
 ‘I am going to buy food at a market.’ [Nishioka and Nakahara 2006: 138]

#### 12.4.2.1.4 Stem alternation

There is a complex system of stem alternation, in which the stem-final segment is changed according to the type of inflection or pre-inflection in which the stem is involved. To describe the stem alternation, three types of stem are identified, Types A, B, and C. Basically, the alternation pattern is predictable from the root-final consonant, and I do not go into detail about exceptional cases.<sup>1</sup>

Table 12.4 shows how the stem alternation system works. Each row shows how the stem-final segment is modified, with the top row in bold indicating the unmodified (i.e. root-final) segment. Roots that end in /r/ and /t/ have all three different stem forms, whereas other roots conflate Types A and B (in the case of roots that end in /b/, /m/, /s/, and /n/) or Types B and C (in the case of roots that end in /k/ and /g/). For example, a root that ends in /t/ (e.g. *mut-* ‘hold’) has the Type A stem form *mut-* as in *mut-a* ‘will hold’, the Type B stem form *muc-* as in *muc-u-ta-n* ‘held’, and the Type C stem form *mucc-* as in *mucc-i* ‘hold, and’. On the other hand, a root that ends in /b/ (e.g. *tub-* ‘fly’) conflates Types A and B, as in *tub-a* ‘will fly’ and *tub-u-ta-n* ‘flew’, but it still has the Type C stem form *tud-* as in *tud-i* ‘fly, and’.

**TABLE 12.4 SHURI STEM FORM AND (PRE-)INFLECTION TYPE**

Root-final C	<b>-b</b>	<b>-m</b>	<b>-s</b>	<b>-n</b>	<b>-r</b>	<b>-t</b>	<b>-k</b>	<b>-g</b>
Type A stem	-b	-m	-s	-n	-r	-t	-k	-g
Type B stem		(~ -n)			-j	-c	-c	-z
Type C stem	-d	-d	-c	-z	-t	-cc		

12.4.2.1.5 Finite inflection

The possible structures of a finite verb are schematized in Table 12.5. Each inflectional pattern is described in the order presented in the table. In what follows, the root *mut*-‘hold’ will be used for illustration, as it distinguishes all three stem types in Table 12.4.

*Finite inflection with Type A stem*

The type A stem may directly be followed by an inflectional affix (mood affix), as shown in Table 12.6, or may be followed by a pre-inflectional affix (negative affix) and then inflectional affixes (tense and mood), as shown in Table 12.7. The (pre-)inflectional affixes denote negation, intention, hortative, and imperative, which are broadly characterized as ‘irrealis modalities’. Even though Shuri verb morphology is fairly agglutinative, there is sometimes tricky allomorphy, such as that found in the combination of the past tense *-ta* and the interrogative mood *-i*, which is *-ti-i* at surface (see *mutantii* in Table 12.7). Note that in the tables the

**TABLE 12.5 SHURI FINITE INFLECTION**

Stem	Pre-inflection	Inflection	Paradigm
Type A		Mood	Table 12.6
	Negative	Tense Mood	Table 12.7
Type B	Politeness	Tense Mood	Table 12.8
	Imperfective	Tense Mood	Table 12.9
Type C	Perfective	Mood	Table 12.10
	Progressive	Tense Mood	Table 12.11
	Resultative	Tense Mood	Table 12.12

**TABLE 12.6 SHURI TYPE A STEM AND INFLECTION**

Root	Word form	stem	inflection (mood)
<i>mut</i> - ‘hold’	<i>muta</i> (intentional) ‘will hold; let’s hold’	<i>mut</i>	<i>-ra</i>
	<i>muti</i> (imperative) ‘you hold’	<i>mut</i>	<i>-ri</i>
	<i>mutee</i> (imperative 2) ‘you hold’ (blunter)	<i>mut</i>	<i>-ree</i>

**TABLE 12.7 SHURI TYPE A STEM, PRE-INFLECTION (NEGATIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection (negative)	inflection	
				tense	mood
<i>mut</i> - ‘hold’	<i>mutan</i> <i>mutani</i> <i>mutanga</i>	<i>mut</i>	<i>-ran</i>	<i>-∅</i> [non-past]	<i>-n</i> [indicative] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative]
	<i>-ta</i> [past]			<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]	

affixes are in their underlying forms rather than surface forms. The underlying /r/ in *-ra*, *-ri*, and *-ree* is deleted when attached to a stem that ends in a consonant.

#### *Finite inflection with Type B stem*

Type B inflections are shown in [Tables 12.8](#) and [12.9](#). Verb forms in [Table 12.8](#) consist of a Type B stem, pre-inflection (politeness), and inflection (tense and mood). Verb forms in [Table 12.9](#) consist of a Type B stem, pre-inflection (imperfective), and inflection (tense and mood). The imperfective affix is often realized as *-i* when attached to a stem that ends in /j/ (i.e. a root that ends in /r/), as in *tur-* ‘take’ → *tuj-u-Ø-n* or *tuj-i-Ø-n* (take-IPRF-NPST-IND).

#### *Finite inflection with Type C stem*

Finite inflections with Type C stem are listed in [Tables 12.10](#), [12.11](#), and [12.12](#). In [Table 12.10](#), the perfective affix only allows past time reference, and never carries the past tense affix.

**TABLE 12.8 SHURI TYPE B STEM, PRE-INFLECTION (POLITENESS), AND INFLECTION**

Root	Word form	stem	pre-inflection	inflection	
				tense	mood
<i>mut-</i> ‘hold’	<i>mucabiin</i> <i>mucabiiru</i> <i>mucabiimi</i> <i>mucabiiga</i> <i>mucabiira</i>	<i>muc</i>	<i>-abii</i> [politeness]	<i>-Ø</i> [non-past]	<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-mi</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]
	<i>-ta</i> [past]			<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]	
	<i>mucabiitan</i> <i>mucabiitaru</i> <i>mucabiitii</i> <i>mucabiitaga</i> <i>mucabiitara</i>				

**TABLE 12.9 SHURI TYPE B STEM, PRE-INFLECTION (IMPERFECTIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection	inflection	
				tense	mood
<i>mut-</i> ‘hold’	<i>mucun</i> <i>mucuru</i> <i>mucumi</i> <i>mucuga</i> <i>mucura</i>	<i>muc</i>	<i>-u</i> [imperfective]	<i>-Ø</i> [non-past]	<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-mi</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]
	<i>-ta</i> [past]			<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]	
	<i>mucutan</i> <i>mucutaru</i> <i>mucutii</i> <i>mucutaga</i> <i>mucutara</i>				

**TABLE 12.10 SHURI TYPE C STEM, PRE-INFLECTION (PERFECTIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection (perfective)	Inflection (mood)
<i>mut-</i> 'hold'	<i>muccan</i> <i>muccaru</i> <i>muccii</i> <i>muccaga</i> <i>muccara</i>	<i>mucc</i>	<i>-a</i>	<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]

**TABLE 12.11 SHURI TYPE C STEM, PRE-INFLECTION (PROGRESSIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection	inflection	
				tense	mood
<i>mut-</i> 'hold'	<i>muccoon</i> <i>muccooru</i> <i>muccoomi</i> <i>muccooga</i> <i>muccoora</i>	<i>mucc</i>	<i>-oo</i> [progressive]	<i>-∅</i> [non-past]	<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-mi</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]
	<i>-ta</i> [past]			<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]	
	<i>muccootan</i> <i>muccootaru</i> <i>muccootii</i> <i>muccootaga</i> <i>muccootara</i>				

**TABLE 12.12 SHURI TYPE C STEM, PRE-INFLECTION (RESULTATIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection	inflection	
				tense	mood
<i>mut-</i> 'hold'	<i>mucceen</i> <i>mucceeru</i> <i>mucceemi</i> <i>mucceega</i> <i>mucceera</i>	<i>mucc</i>	<i>-ee</i> [resultative]	<i>-∅</i> [non-past]	<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-mi</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]
	<i>-ta</i> [past]			<i>-n</i> [indicative] <i>-ru</i> [adnominal] <i>-i</i> [yes-no interrogative] <i>-ga</i> [wh interrogative] <i>-ra</i> [dubitative]	
	<i>mucceetan</i> <i>mucceetaru</i> <i>mucceetii</i> <i>mucceetaga</i> <i>mucceetara</i>				

The semantic difference between the imperfective past (*mucutan*) and the perfective (*mucan*) is explained in terms of evidentiality, and will be further discussed in 12.5.7.

In Table 12.11, the pre-inflection slot is filled by the progressive aspect suffix *-oo*. In Table 12.12, the pre-inflection slot is filled by the resultative aspect suffix *-ee*.

#### 12.4.2.1.6 Relative ordering of pre-inflections

Some pre-inflectional affixes are mutually exclusive, whereas others may not. Specifically, aspectual affixes are mutually exclusive. However, the honorific affix (see 12.4.2.1.8),

a given aspectual affix, the politeness affix, and the negative affix may occur in a row. For example, the honorific suffix *-imisee*, the politeness suffix *-abii*, and the negative suffix *-ran* may co-occur in a row. This is shown in Table 12.13. Here, the inflection that follows is identical to that in Table 12.7. Thus, the inflectional morphology is determined by the final suffix of the successive pre-inflections.

#### 12.4.2.1.7 Non-finite inflection

Non-finite inflection (Table 12.14) involves various adverbial-adsentential relations (or converbial relations), which are marked on the verb that serves as a predicate of a subordinate or coordinate clause. As in the case of finite inflection, the three types of stems carry different kinds of inflectional affixes.

Adverbial-adsentential relations are also encoded by ‘finite converbs’, in which the modal affix of a finite verb is replaced by a conjunctive affix (causal *-kutu*, ‘but’). For example, the ‘but’ relation may be past or non-past with finite inflection. The finite converb *mucutasiga* ‘(x) held, but’, for example, consists of *muc-* (Type B stem), *-u* (imperfective), the past affix *-ta*, and the converb affix *-siga*. Here, the mood affix for the finite inflection of Type B stem (see Table 12.9) is paradigmatically replaced by the conjunctive affix *-siga*. This is a grammaticalized form of a former finite verb (specifically with the adnominal-unmarked mood; see Table 12.9) plus a conjunctive

**TABLE 12.13 SHURI TYPE B STEM, PRE-INFLECTION (POLITENESS + NEGATIVE), AND INFLECTION**

Root	Word form	stem	pre-inflection	inflection	
				tense	mood
<i>mut-</i> ‘hold’	<i>mucimiseebiran</i> <i>mucimiseebirani</i> <i>mucimiseebiraga</i>	<i>muc</i>	<i>-imisee + -abii + -ran</i> [honorific + politeness + negative]	See Table 12.7	
	<i>mucimiseebirantan</i> <i>mucimiseebirantaru</i> <i>mucimiseebirantii</i> <i>mucimiseebirantaga</i> <i>mucimiseebirantara</i>				

**TABLE 12.14 SHURI NON-FINITE INFLECTION**

Root	Word form	stem	Inflection (converbial)
<i>mut-</i> ‘hold’	<i>mutee</i> (anterior-conditional) ‘after (x) hold’ or ‘if (x) hold’	<i>mut</i>	<i>-ree</i>
	<i>mutaa</i> (conditional) ‘if (x) hold’		<i>-raa</i>
	<i>mucaani</i> (sequential) ‘(x) hold, and’	<i>muc</i>	<i>-raani</i>
	<i>muciiga</i> (purposive) ‘in order to hold’		<i>-riiga</i>
	<i>muciinee</i> (anterior-conditional) ‘after (x) hold’ or ‘if (x) hold’		<i>-riinee</i>
	<i>mucagacii</i> (simultaneous) ‘while (x) hold’		<i>-ragacii</i>
	<i>mucci</i> (sequential) ‘(x) hold, and’		<i>mucc</i>

clitic, a structure still attested in other Ryukyuan varieties such as Miyako (*muc-tar=ruga* in Irabu Ryukyuan, which contains =*suga* ‘but’, a cognate form of Shuri *-siga*).

#### 12.4.2.1.8 Honorification

Shuri is characterized by its elaborated system of honorifics, which fall into respectful expressions and humble expressions. In this short grammar sketch I only focus on major honorification processes for respectful honorifics and humble honorifics.

*Respectful honorification.* Respectful honorification indicates the speaker’s respect toward the agent of a sentence. In (15) and (16), for example, the speaker expresses his/her respect toward the teacher by affixing a respectful honorific affix *-imisee* to the verb (15), or by using a special stem for honorification (16).

- (15) *sinsii=ja unu tigami kac-imisee-bi-ta-n*  
 teacher=TOP that letter write-RSP-POL-PST-IND  
 ‘The teacher wrote that letter.’ [Nishioka and Nakahara 2006: 113]
- (16) *sinsii=ga kuma=nkai mensee-Ø-n*  
 teacher=NOM this.place=DAT come.RSP-NPST-IND  
 ‘The teacher is coming here.’ [Nishioka and Nakahara 2006: 110]

As illustrated in the examples above, respectful honorification may be expressed by affixation and stem alternation. The affixation strategy is productive and applied to most verb roots. The respectful honorific affix is *-imisee*, which is a pre-inflectional affix that attaches to a Type B stem. Thus, *kac-imisee-ta-n* (write-RSP-PST-IND) ‘(some respected person) wrote’, for example, consists of the Type B stem form of the root *kak-* ‘write’ (*kak-* → *kac-*), the honorific pre-inflectional suffix *-imisee*, and the finite inflectional affixes *-ta* (past tense) and *-n* (indicative mood). As for stem alternation, there are several roots that are used only for respectful honorification. Three frequently used roots are listed in (17).

- |      |                     |                     |
|------|---------------------|---------------------|
| (17) | Plain root form     | Honorific root form |
|      | <i>ik-</i> ‘go’     | <i>mensee-</i>      |
|      | <i>k-</i> ‘come’    | ”                   |
|      | <i>wur-</i> ‘be’    | ”                   |
|      | <i>kam-</i> ‘eat’   | <i>usagaju-</i>     |
|      | <i>num-</i> ‘drink’ | ”                   |
|      | <i>nzu-</i> ‘look’  | <i>umikakir-</i>    |

*Humble honorification.* Humble honorification indicates the speaker’s respect toward the undergoer of an action. In (18) and (19), for example, the speaker expresses his/her respect toward the teacher by using a special auxiliary verb for humble honorification (18) or by using a special stem for humble honorification (19).

- (18) *taruu=ga sinsii=nkai jud-i usagij-u-Ø-n*  
 taruu=NOM teacher=DAT read-CVB.SEQ HMB-IPRF-NPST-IND  
 ‘Taro read (something) for the teacher.’ [adapted from Nishioka and Nakahara 2006: 121]
- (19) *taruu=ga sinsii=nkai unnukij-u-Ø-n*  
 taruu=NOM teacher=DAT say.HMB-IPRF-NPST-IND  
 ‘Taro said (something) to the teacher.’ [Nishioka and Nakahara 2006: 120]

As illustrated in (18), a productive way of composing a humble honorific expression is to use an auxiliary verb construction in which the auxiliary is a humble auxiliary *usagir-*. Some frequently used roots have suppletive humble honorific forms. For example, the verb ‘say’ has the plain form *i-* and the humble honorific form *unnukir-*, as illustrated in (19).

As mentioned in 12.4.2.1.6, the honorific affix may be followed by the polite affix *-abii*. Honorification is concerned with the relative social rank between the speaker and the subject, whereas politeness is concerned with the relative social rank between the speaker and the hearer. Thus if a speaker needs to, or wants to, express respect toward the subject of a sentence and toward the hearer, both forms are used in a row.

#### 12.4.2.1.9 Existentials and copula

As is common cross-linguistically, existential verbs and copular verbs show certain commonalities among them, and certain morphosyntactic peculiarities as compared with regular verbs. Table 12.15 compares regular verbs, existentials, and the copula in terms of negation strategy and animacy-sensitiveness (i.e. whether the verb in question is restricted to being used with an animate or inanimate subject). The right-most row identifies a feature that distinguishes between regular verbs on the one hand, and existentials and the copula on the other.

The existential verbs comprise *wur-* ‘(animate subject) exist’ and *ar-* ‘(inanimate subject) exist’. The former is negated with the pre-inflectional affix *-ran* just as in the case of regular verbs: *wur-an-ta-n* (exist-NEG-PST-IND) ‘did not exist’. By contrast, the inanimate existential *ar-* cannot carry the negative affix, and a stem alternation strategy is employed to negate this verb, replacing *ar-* by *nee-*: *nee-ran-ta-n* (exist.NEG-NEG-PST-IND) ‘did not exist’. The existentials also show certain irregular inflectional morphology. For example, they do not carry an aspect affix, so the non-past indicative inflection is *a-Ø-n* and *wu-Ø-n* (exist-NPST-IND) ‘exist’ rather than *\*aj-u-Ø-n* and *\*wuj-Ø-n* (cf. *kac-u-Ø-n* write-IPRF-NPST-IND ‘write’).

The copular verb root *jar-* is replaced by *ar-* when negated (*ja-ta-n* COP-PST-IND → *ar-an-ta-n* COP-NEG-PST-IND). Note that this negative stem is different from the existential *ar-*, which cannot carry the negative affix *-ran*. Like existential verbs, the copular verb does not take aspectual affixes.

#### 12.4.2.1.10 The internal structure of the stem

The structure of a verb stem is schematized as Root (+ causative) (+ passive) (+ aspect). The causative suffix is *-as* or *-imir*, the latter being used only when the root ends in /s/. Thus compare: *kak-* ‘write’ → *kak-as-*, whereas *nagas-* ‘drain’ → *nagas-imir-*. The passive suffix is *-rarir*. The aspect affix is *-agir* ‘be about to’ (inceptive), as in *andij-agij-u-Ø-n* (overflow-ICP-IPRF-NPST-IND) ‘is about to overflow’.

TABLE 12.15 SHURI EXISTENTIAL AND COPULAR VERBS

	Negated by <i>-ran</i>	Animacy-sensitive	Aspect marking
Regular verbs (e.g. <i>kak-</i> ‘write’)	+	–	+
Existential <i>wur-</i>	+	+	–
Existential <i>ar-</i>	–	+	–
Copula <i>jar-</i>	–	–	–

As compared with pre-inflections, these affixes are lexically constrained, and occur with much less frequency. Unlike inflectional affixes, they do not usually occur with adjective stems. In (20), the root *nak-* is followed by the causative *-as*, and the entire stem is further followed by the imperfective pre-inflection *-u*, which inflects for tense (non-past) and mood (indicative). In (21), the same verb root is followed by the causative affix *-as* and the passive affix *-rarir* in a row.

(20) *nak-* ‘cry’ → *nak-as-u-Ø-n*  
 cry-CAUS-IPRF-NPST-IND  
 ‘makes (someone) cry’

(21) *nak-* ‘cry’ → *nak-as-arir-u-Ø-n*  
 cry-CAUS-PASS-IPRF-NPST-IND  
 ‘is made to cry’

#### 12.4.2.2 Adjectives

##### 12.4.2.2.1 The structure

As a subclass of verbal, the adjective also inflects, with the structure Stem (+ Pre-inflection) + Inflection. The adjective *maasa-ibii-ta-n* ‘was tasty (polite form)’, for example, consists of the stem *maasa-* ‘tasty’ (which further comprises the root *maa-* ‘tasty’ and the adjective stem formative *-sa*; see below), and the pre-inflectional affix *-ibii* (politeness), and the inflectional affixes *-ta* (past) and *-n* (indicative). The morphological structure of this adjective is thus quite parallel to that of the verb *muc-abii-ta-n* ‘held (polite form)’:

(22) Verb and adjective: structural comparison

	Stem	Pre-inflection (-politeness)	Inflection (-tense -mood)
V: <i>muc-abii-ta-n</i> ‘held (polite form)’	<i>muc</i>	<i>-abii</i>	<i>-ta -n</i>
Adj: <i>maasa-ibii-ta-n</i> ‘was tasty (polite form)’	<i>maa-sa</i>	<i>-ibii</i>	<i>-ta -n</i>

What crucially characterizes the morphology of the adjective is *-sa*, which obligatorily appears stem-finally, and which historically derived from a nominalizer suffix *\*-sa* (cf. Japanese **-sa**) + a state verb *\*ar-*. It is natural, therefore, that adjectives in Shuri inflect just like verbs, given that the former contain what used to function as a state verb. Quite apart from the diachronic development of the affix *-sa*, it is somewhat difficult to pin down the synchronic function of the affix *-sa*. Note that it does not function as a nominalizer in any meaningful sense, since the stem with *-sa* does not carry any nominal affix, nor does it create a stem that functions as the head of a nominal phrase. The affix *-sa* simply functions to derive a verbal stem so that the derived stem may carry verbal affixes (i.e. pre-inflectional and inflectional affixes). For this reason, the gloss VLZ (verbalizer) is given to this affix in this chapter.

Note that the pre-inflectional affix in (22) takes different forms depending on whether the stem is a verb stem (that takes the politeness affix *-abii*) or an adjective stem (that takes *-ibii*). This allomorphy is observed in certain verbs as well, especially in verb roots that end in /r/ (e.g. *tur-* ‘take’ → *tuj-abii-Ø-n* or *tuj-ibii-Ø-n* (take-POL-NPST-IND) ‘take (polite form)’). Remember that an adjective stem historically contained the existential verb *\*ar-*, which triggers the allomorph *-ibii*.

Previous works report that a set of adjective roots which correspond to *-ku* adjective roots in Japanese (see Chapter 7, Table 7.7) such as *taka-* ‘high’, *uma-* ‘tasty’, etc., carry

-*sa*, whereas the other set which correspond to *-shiku* adjective roots in Japanese such as *sabisi-* ‘lonely’, *tanosi-* ‘funny’, etc., carry *-sja*. Thus, whereas we have *takasa-*, *umasa-*, etc., we have *sabisja-* (*sabisi-* ‘lonely’ + *-sa*), *umusja-* (*umusi-* ‘funny’ + *-sa*), etc. However, this distinction seems to be disappearing, with *-sja* being replaced by *-sa*, as in *sabisa-*, *umusa-*, etc. Synchronically, it is still necessary to posit the underlying root form with /si/ (e.g. *sabisi-*, *umusi-*), since this /si/ appears in the infinitive form of an adjective:

- (23) a. *taka-* ‘high’ → *taka-ku*  
 b. *maa-* ‘tasty’ → *maa-ku*  
 c. *huka-* ‘deep’ → *huka-ku*  
 d. *sabisi-* ‘lonely’ → *sabisi-ku*  
 e. *mizirasi-* ‘rare’ → *mizirasi-ku*

#### 12.4.2.2.2 Pre-inflection

The pre-inflectional category for adjectives is politeness (*-ibii*, rather than *-abii* for verbs), but not aspect or polarity. The reason for aspect being absent is obvious, as property concepts always denote a stative meaning. With respect to polarity, the negation of an adjective takes an analytic form, with the infinitive form of an adjective (see 12.4.2.2.5) followed by (the topic marker and) the negative auxiliary *neen*. Thus compare:

- (24) *maa-sa-Ø-n*  
 tasty-VLZ-NPST-IND  
 ‘(It’s) tasty.’
- (25) *maa-ko=o*      *nee-n-Ø-Ø*  
 tasty-INF=TOP    not.exist-NEG-NPST-IND  
 ‘(It’s) not tasty.’

#### 12.4.2.2.3 Finiteness

The inflectional categories for adjectives are tense and mood (for finite inflection) or adverbial-adsentential relation (for non-finite inflection), just as in the case of verbs. However, due to the stative and basically non-volitional nature of adjectives, some inflectional possibilities are absent, such as intentional and imperative (for finite inflection), and simultaneous (for non-finite inflection). Also, the shapes of several inflectional affixes are different from those for verbs, as will be shown below.

The finite and non-finite inflections for adjectives are listed in the sections below. Note that there is no stem alternation as found in verbs.

#### 12.4.2.2.4 Finite inflection

Finite inflection is listed in [Table 12.16](#). The round brackets in the column ‘Word form’ and ‘pre-inflection’ indicate that the bracketed form *-ibii* (politeness) is optional and does not affect inflectional morphology that follows.

#### 12.4.2.2.5 Non-finite inflection

Non-finite inflection is listed in [Table 12.17](#). As mentioned in 12.4.2.2.1, the infinitive form induces the use of the bare root form rather than the root + *-sa*.

TABLE 12.16 SHURI FINITE INFLECTION FOR ADJECTIVES

Root	Word form	stem	pre- inflection	inflection	
				tense	mood
taka- 'high'	<i>takasa(ibii)n</i> <i>takasa(ibii)ru</i> <i>takasa(ibii)mi</i> <i>takasa(ibii)ga</i> <i>takasa(ibii)ra</i>	taka-sa	(-ibii) [politeness]	-∅ [non-past]	-n [indicative] -ru [adnominal] -mi [yes-no interrogative] -ga [wh interrogative] -ra [dubitative]
	-ta [past]			-n [indicative] -ru [adnominal] -i [yes-no interrogative] -ga [wh interrogative] -ra [dubitative]	

TABLE 12.17 SHURI NON-FINITE INFLECTION FOR ADJECTIVES

Root	Word form	stem	inflection
taka- 'high'	<i>takasanu</i> (causal) 'x is high, so'	taka-sa	-nu
	<i>takasai</i> (sequential) 'x is high, and'		-i
	<i>takaku</i> (infinitive) 'high(ly)'	taka-	-ku

The infinitive form occurs in the verbal complement slot which a specific set of verbs takes, as in (26), or as a predicate modifier, as in (27).

(26) *sii=nu*            [*taka-ku*]    *nat-oo-∅-n*  
height=NOM high-INF become-PROG-NPST-IND  
'He has become tall.'

(27) [*taka-ku*]    *tub-i*  
high-INF jump-IMP  
'Jump high!'

### 12.4.3 Word formation

#### 12.4.3.1 Affixation

Shuri morphology is mostly suffixal. However, a small number of prefixes do exist, and most are concerned with honorific expressions: *gu-* (respectful honorification), as in *gu-buree* 'dishonor', *u-* (respectful honorification), as in *u-suba* 'nearby', etc.

#### 12.4.3.2 Compounding

Like other Japonic languages, compounding is very common in Shuri, both in nominal and verbal morphology. *Rendaku* (sequential voicing) is also common (compare other chapters, 8.2.4, 9.2, and 10.2.4). In each of (28a–c), for example, the second stem of a compound is subject to *rendaku*: *kwii* 'voice' → *gwii*, *cin* 'clothes' → *zin*, and *kuri* 'hard' → *guri*.

- (28) a. *[wikiga-gwii]*      b. *[huru-zin]*      c. *[cimu-guri]-sa-Ø-n*  
 male-voice              old-clothes              heart-hard-VLZ-NPST-IND  
 [n-n]<sub>n</sub>                      [a-n]<sub>n</sub>                      [n-a]<sub>a</sub>  
 ‘muscular voice’      ‘old clothes’              ‘is pitiful’

In the examples above, ‘n’ and ‘a’ means a nominal and adjectival stem respectively. The bracketed stem is a compound stem, whose part of speech is indicated in subscript beside the right bracket.

When a verb root participates in compounding, it is turned into a Type B stem, which then carries the infinitive affix *-i* (which is attached to a Type B stem). The infinitive form is nominal in nature, since it may be used as a nominal stem and it cannot carry a verbal inflectional affix.

- (29) a. *[cic-i-kees]-u-Ø-n*  
 ask-INF-return-IPRF-NPST-IND  
 [v-v]<sub>v</sub>  
 ‘asks back’
- b. *[hun-jum-i]*  
 book-read-INF  
 [n-v]<sub>n</sub>  
 ‘book reading’

As indicated in the third layer of the glosses of each example above, it is basically the final stem that determines the part of speech of the entire compound stem (note that the infinitive form of a verb is nominal in nature). In (28c), the compound stem is an adjective stem, which then carries the verbalizer suffix *-sa*, to which the verbal inflectional affixes attach.

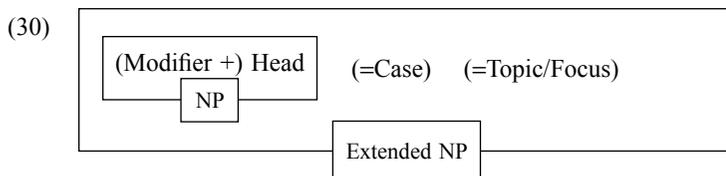
In some cases, it is difficult to distinguish between a compounded stem and an affix. For example, the stem *kuri-* ‘hard’ as illustrated in (28c) is extremely productive, and may be attached to a verb root as well. It is thus difficult to determine whether it is a stem, or a derivational affix that derives an adjective stem from a nominal or verbal root.

## 12.5 SYNTAX

### 12.5.1 The nominal phrase

#### 12.5.1.1 Basic structure

A nominal phrase (NP) consists of the head and optionally a modifier. An argument NP is followed by a case clitic and/or other argument markers (topic marker, focus marker, etc.) and the entire constituent is called an extended NP. Thus the structure of an extended NP is schematized as follows (where ‘=’ indicates a clitic boundary):



In what follows, brief descriptions of modifier and case are made. See 12.4.1 for the description of the head of an NP, i.e. the nominal word class. See 12.5.5 for the description of topic and focus markers, which involves predicate concordance (known broadly as *kakari-musubi*; compare other chapters, 7.1, 7.5.4, 8.5.4, 13.4.4, and 14.5.4).

### 12.5.1.2 Modifier

A modifier may be an adnominal word, as shown in (31), or an adnominal clause, in which the predicate may be either a verb or an adjective, as shown in (32) and (33) respectively. Also, an NP may be recursive, in such a way that the modifier of an NP is an NP itself. In such a case, the embedded NP carries genitive case, as shown in (34).

- (31) [*anu*] *sikuce=e* *acaa=wuti* *subinaj-i-Ø-sa*  
 that work=TOP tomorrow=LOC complete-IPRF-NPST-EMP  
 ‘That work will be completed tomorrow.’ [Tshako 1997: 375]
- (32) [*taruu=ga* *hwic-u-Ø-ru*] *sansin*  
 Taruu=NOM play-IPRF-NPST-ADN guitar  
 ‘the Sansin guitar that Taruu plays’ [Nishioka and Nakahara 2006: 33]
- (33) [*umu-sa-ta-ru*] *hanasi*  
 funny-VLZ-PST-ADN story  
 ‘the story that was funny’
- (34) [*ari=ga*] *jaa*  
 3SG=GEN house  
 ‘his/her house’

First person pronouns (singular *wan* and plural *wattaa*), second person non-honorific pronouns (singular *?jaa* and *ittaa*), and interrogative pronouns (singular *taa* ‘who’ and plural *tattaa* ‘who (plural)’, and *nuu* ‘what’) do not require genitive case marking.

- (35) [*wattaa*] *jaa*  
 1PL house  
 ‘our house’
- (36) [*nuu*] *sikuci* *ja-Ø-ga*  
 what work COP-NPST-Q  
 ‘What work is (it)?’ [Nishioka and Nakahara 2006: 96]

Other pronouns require genitive case marking, as illustrated in (34) above.

### 12.5.1.3 Case markers

Shuri displays a nominative-accusative case system, in which intransitive and transitive subjects (S and A) are marked with the same case marker (=ga or =nu; see 5.1.3.1 below) and transitive objects (O) are left unmarked.

#### 12.5.1.3.1 Nominative and genitive

In Shuri and many other Ryukyuan varieties, the same case form is used for both subject NPs and modifier NPs. This form is either =ga (as in (37) and (38)) or =nu (as in (39) and (40)).

- (37) *ari=ga c-uu-Ø-n*  
 3SG=NOM come-IPRF-NPST-IND  
 ‘S/he is coming.’
- (38) *ari=ga naa*  
 3SG=GEN name  
 ‘his/her name’
- (39) *in-gwaa=nu c-uu-Ø-n*  
 dog-DIM=NOM come-IPRF-NPST-IND  
 ‘A puppy is coming.’
- (40) *in-gwaa=nu naa*  
 dog-DIM=GEN name  
 ‘the puppy’s name’

The =*ga* form is used for a limited set of nominals (proper noun, pronoun, kinship terms, and certain human common nouns such as *ccu* ‘man’), whereas the =*nu* form is used for the other broad range of nominals. That is, the choice is based on the cross-linguistically common semantic-pragmatic referential hierarchy, or animacy hierarchy. This phenomenon where the nominative-genitive case has two variants according to animacy hierarchy is quite common in Ryukyuan.

In this sketch, I use the term nominative case when the form is attached to an NP that serves as subject, and genitive case when the form is attached to an NP that serves as a modifier of a larger NP. However, it is noted that this terminology is only for convenience, and that what actually happens in Shuri is that both subject and NP modifier are encoded by the same case marker, which would be better called ‘nominative-genitive case’.

#### 12.5.1.3.2 Dative

The dative case in Shuri is =*nkai* or =*kai*. These two are not allomorphs, but distinct morphemes. The form =*nkai* may be used for the locus, goal of motion, and agent of a passive sentence, whereas =*kai* may not be used for the agent of a passive sentence. A stem-final *n* changes to *nu* before =*nkai* (similarly before =*n* ‘also’).

- (41) *jaccii=ja jaa=nkai wu-Ø-n*  
 elder.brother=TOP house=DAT exist-NPST-IND  
 ‘My elder brother is in the house.’ [Tshako 1997: 374]
- (42) *naahwa=nkai ic-u-Ø-n*  
 Naha=DAT go-IPRF-NPST-IND  
 ‘(I) go to Naha.’ [Nishioka and Nakahara 2006: 113]
- (43) *ari=nkai sugur-at-ta-n*  
 3SG=DAT hit-PASS-PST-IND  
 ‘(Someone) was hit by him/her.’ [Nishioka and Nakahara 2006: 113]
- (44) *\*ari=kai sugur-at-ta-n*  
 3SG=DAT hit-PASS-PST-IND  
 ‘(Someone) was hit by him/her.’

#### 12.5.1.3.3 Instrumental

The instrumental case is =*ssi*, which developed from the verb *s-* ‘do’. An alternative is =*saani*.

- (45) *taruu=ja irana=ssi daki cicc-oo-Ø-n*  
 Taruu=TOP sickle=INST bamboo cut-PROG-NPST-IND  
 ‘Taruu is cutting bamboo with a sickle.’ [Tzuhako 1997: 375]
- (46) *ucinaa-guci=ssi hanasi s-u-Ø-n*  
 Okinawa-mouth=INST talk do-IPRF-NPST-IND  
 ‘(We) talk in the Okinawan language.’ [Tzuhako 1997: 375]

#### 12.5.1.3.4 Locative

The locative =*wuti* and =*wutooti* developed from the sequential converb form of *wur-* ‘exist’. The semantic difference between the two is unclear; however, given that the latter form developed from a sequential converb with the progressive aspect affix *-oo*, a certain aspectual difference might exist. Some works report that there is another form =*nzi*, which has only recently started to be used as a locative case marker, and which has developed from the verb *ik-* ‘go’.

- (47) *kuma=wuti macc-oo-ri=joo*  
 this.place=LOC wait-PROG-IMP=EMP  
 ‘Wait here, OK?’ [Nishioka and Nakahara 2006: 85]
- (48) *ijo=o naahwa=wutooti ur-att-oo-Ø-n*  
 fish=TOP this.place=LOC sell-PASS-PROG-NPST-IND  
 ‘Fishes are sold in Naha.’ [Nishioka and Nakahara 2006: 85]
- (49) *cuu=ja jaa=nzi sikuci s-a=na*  
 today=TOP house=LOC work do-INT=EMP  
 ‘Today (I) will do the work in (my) house.’ [Tzuhako 1997: 375]

#### 12.5.1.3.5 Ablative

The ablative case is =*kara*. This case denotes a meaning related to the initial point (‘from’), path (‘along’), means (‘by’; only when the case marker co-occurs with motion verbs), and a point in time (‘at’). The first meaning is also found in Japanese, but the others are not. The terminative case is =*madi*.

- (50) *kuma=kara ama=madi*  
 this.place=ABL that.place=TERM  
 ‘from here to there’ [Nishioka and Nakahara 2006: 113]
- (51) *mici=kara acc-u-Ø-n*  
 road=ABL walk-IPRF-NPST-IND  
 ‘(I) walk along the road.’ [Nishioka and Nakahara 2006: 114]
- (52) *basu=kara ic-u-Ø-n*  
 bus=ABL go-IPRF-NPST-IND  
 ‘(I) go by bus.’ [Nishioka and Nakahara 2006: 114]
- (53) *atu=kara hanas-u-Ø-n*  
 later.time=ABL talk-IPRF-NPST-IND  
 ‘(I) will talk at a later time.’ [Nishioka and Nakahara 2006: 114]

### 12.5.2 The predicate phrase

The predicate may be either verbal or nominal.

#### 12.5.2.1 Nominal predicate

A nominal predicate consists of an NP as head of a predicate, which is followed by the copula, which takes over the job of indexing predicate categories (tense, mood, polarity, etc.) for the NP. In the following example, the bracketed phrase is a nominal predicate phrase, consisting of the NP *ari=ga sinsii* ‘his/her teacher’ and the copula verb *jan*, which inflects for tense and mood as a finite verb (see 12.4.2.1.9 for the morphology of the copula).

- (54) *wan=nee [ari=ga sinsii ja-Ø-n]*  
 1SG=TOP 3SG=GEN teacher COP-NPST-IND  
 ‘I am his/her teacher.’

Even though the predicate NP does not carry any case marker, it may be focalized or topicalized with the focus marker or topic marker.

- (55) *wan=nee [ari=ga sinsii=du ja-Ø-ru]*  
 1SG=TOP 3SG=GEN teacher=FOC COP-NPST-ADN  
 ‘I am his/her teacher.’
- (56) *wan=nee [ari=ga sinsii=ja ar-an-Ø-Ø]*  
 1SG=TOP 3SG=GEN teacher=TOP COP-NEG-NPST-IND  
 ‘I am not his/her teacher.’

#### 12.5.2.2 Verbal predicate

A verbal predicate phrase consists of a lexical verbal word (either a verb or adjective) and optionally an auxiliary verb in the case of a complex verbal predicate.

##### 12.5.2.2.1 Complex VP of which the lexical verbal word is a verb

In a complex VP, the lexical verb inflects as a sequential converb, and the auxiliary verb takes over pre-inflection or inflection. In the examples below, the constituent in brackets is a verbal predicate phrase, consisting of the lexical verb (sequential converb) and the auxiliary verb (finite: past indicative). Auxiliary verbs function to denote aspectual, benefactive, and humble honorific meanings.

- (57) *wan=nee ure=e [jud-i neen-ta-n]*  
 1SG=TOP 3SG=TOP read-CVB.SEQ CPL-PST-IND  
 ‘I had read it.’ [completive aspect]
- (58) *sinsii=ga hun [jud-i kwi-ta-n]*  
 teacher=NOM book read-CVB.SEQ BEN-PST-IND  
 ‘The teacher read a book for me.’ [benefactive]
- (59) *taruu=ga sinsii=nkai [jud-i usagij-u-Ø-n]*  
 taruu=NOM teacher=DAT read-CVB.SEQ HMB-IPRF-NPST-IND  
 ‘Taro read (something) for the teacher.’ [Adapted from Nishioka and Nakahara 2006: 121] [humble honorific]

### 12.5.2.2.2 Complex VP of which the lexical verbal word is an adjective

An adjective may fill the lexical verbal word slot in a complex VP. The complex VP is used for negating the adjective. Here, the adjective inflects as an infinitive form and the auxiliary is the negative existential *neen*. In between occurs the topic marker clitic.

- (60) *ari=ga kuruma=a [taka-ko=o nee-n-Ø-Ø]*  
 3SG=GEN car=TOP costly-INF=TOP exist.NEG-NEG-NPST-IND  
 ‘His/her car is not costly.’

### 12.5.3 Sentence type

There are four distinct sentence types that correspond to different speech acts: declarative sentence, interrogative sentence, imperative sentence, and dubitative sentence. These are differentiated according to focus marking and mood marking (Table 12.18).

Focus marking is not obligatory in any sentence type; however, if it does occur in a sentence, =*du* (~ =*ru*) must be used for declaratives and interrogatives, and =*ga* must be used for dubitatives. Imperative sentences do not allow focus marking.

- (61) *ari=ga=du jum-u-Ø-ru*  
 3SG=NOM=FOC read-IPRF-NPST-ADN  
 ‘S/he reads.’ [declarative]
- (62) *ari=ga=du jum-u-Ø-mi*  
 3SG=NOM=FOC read-IPRF-NPST-Q  
 ‘Does s/he read?’ [interrogative (Yes-No)]
- (63) *taa=ga=du jum-u-Ø-ga*  
 who=NOM=FOC read-IPRF-NPST-Q  
 ‘Who reads?’ [interrogative (WH)]
- (64) *ari=ga=ga jum-u-Ø-ra*  
 3SG=NOM=FOC read-IPRF-NPST-DUB  
 ‘I wonder if s/he reads.’ [dubitative]
- (65) *jum-i*  
 read-IMP  
 ‘Read!’

### 12.5.4 Negation

#### 12.5.4.1 Morphological negation

Most action/stative verbs are negated by suffixing the pre-inflection *-ran* to the root. The existential verb root *ar-* (inanimate subject) is negated by replacing it by a special

**TABLE 12.18 SENTENCE TYPES IN SHURI**

Sentence type	Focus marking	Mood marking (see 12.4.2.1.5)
Declarative	= <i>du</i>	Other than the four moods below
Interrogative (Yes-No)	= <i>du</i>	- <i>i</i> [interrogative mood]
Interrogative (WH)	= <i>du</i>	- <i>ga</i> [interrogative mood]
Dubitative	= <i>ga</i>	- <i>ra</i> [dubitative mood] (‘I wonder’)
Imperative	N/A	- <i>ri</i> or - <i>ree</i> [imperative mood]

negative root *nee-* ‘not exist’, which then carries the pre-inflection *-ran*, giving rise to the form *neeran* or *neen*. Likewise, the copula *jar-* is negated by using the negative copula verb root *ar-* ‘be not’ followed by the pre-inflection *-ran*. Unlike the inanimate existential, the animate existential verb *wur-* is regularly negated by the pre-inflection *-ran*.

#### 12.5.4.2 Analytical negation

As mentioned in 12.5.2.2.2, an adjective is negated by using a special auxiliary verb construction, which consists of the adjective (in the infinitive form), the topic marker, and the negative existential verb as an auxiliary.

### 12.5.5 Topic, focus, and emphasis

An argument NP may be marked by a limiter, topic marker, or a focus marker, which follows or replaces a case marker.

#### 12.5.5.1 Limiters

Limiters consist of quantifiers and qualifiers such as *=n* ‘also’ (before which *wan* ‘I’ takes the stem *wanni-*), *=teen* ‘only’, *=naa* ‘only’, *=ndee* ‘and so on’.

#### 12.5.5.2 Topic marker

The topic marker is *=a*, which has a number of allomorphs depending on the final segment of the NP.

- (66) Allomorphy of topic marker
- |           |               |                                  |  |
|-----------|---------------|----------------------------------|--|
| <i>=a</i> | → <i>=ja</i>  | /V <sub>i</sub> V <sub>i</sub> _ | ( <i>cuu</i> ‘today’ + <i>=ja</i> → <i>cuu=j</i> ) |
|           | → <i>=e</i>   | /i_                              | ( <i>tui</i> ‘bird’ + <i>=a</i> → <i>tue=e</i> )   |
|           | → <i>=o</i>   | /u_                              | ( <i>utu</i> ‘sound’ → <i>uto=o</i> )              |
|           | → <i>=oo</i>  | /n_                              | ( <i>in</i> ‘dog’ → <i>in=oo</i> )                 |
|           | → <i>=nee</i> | /wan_ ‘I’                        |  |

#### 12.5.5.3 Focus markers

The focus *=du* requires an indicative predicate verb to inflect for the adnominal-unmarked mood. Thus, whereas in (67) the predicate inflects for the indicative mood, in (68) it inflects for the adnominal-unmarked mood in the presence of the focus marker within the same clause.

- (67) *wan=nee tabaku huc-a-n*  
 1SG=TOP tobacco smoke-PRF-IND  
 ‘I had a cigarette.’ [Adapted from Tshako 1997: 375]
- (68) *wan=nee tabaku=du huc-a-ru. sake=e num-an-ta-n.*  
 1SG=TOP tobacco=FOC smoke-PRF-ADN saké=TOP drink-NEG-PST-IND  
 ‘I had a cigarette; not saké.’ [Tshako 1997: 375]

The focus *=ga* requires the predicate verb to inflect for the dubitative mood.

- (69) *ici=ga naj-i-ra*  
 when=FOC become-IPRF-DUB  
 ‘I wonder when (it) will happen.’ [Nishioka and Nakahara 2006: 104]

### 12.5.6 Tense and aspect

Tense is either past or non-past. Tense is marked by an inflectional affix. Aspect is expressed by a pre-inflectional affix (70), a derivational affix (71), or an auxiliary verb (72). The pre-inflectional aspect marking and derivational aspect marking may be used together. For example, in (71), the inceptive aspect denoted by the derivational affix *-iagir* ‘be about to’ co-occurs with the imperfective aspect affix *-u*.

- |                 |   |  |   |                                  |
|-----------------|---|--|---|----------------------------------|
| (70)            | Non-past  | Past   |   |                                  |
| Imperfective:   | <i>jum-u-Ø-n</i><br>read-IPRF-NPST-IND<br>‘(he) reads’                      | <i>jum-u-ta-n</i><br>read-IPRF-PST-IND<br>‘(he) read’                        |   |                                  |
| Perfective:     |   | <i>jud-a-n</i><br>read-PRF-IND<br>‘(he) read’                                |   |                                  |
| Progressive:    | <i>jud-oo-Ø-n</i><br>read-PROG-NPST-IND<br>‘(he) is reading’                | <i>jud-oo-ta-n</i><br>read-PROG-PST-IND<br>‘(he) was reading’                |   |                                  |
| Resultative:    | <i>jud-ee-Ø-n</i><br>read-RSL-NPST-IND<br>‘(it) has been read’              | <i>jud-ee-ta-n</i><br>read-RSL-PST-IND<br>‘(it) had been read’               |   |                                  |
| (71) Inceptive: | <i>jum-iagij-u-Ø-n</i><br>read-ICP-IPRF-NPST-IND<br>‘(he) is about to read’ | <i>jum-iagij-u-ta-n</i><br>read-ICP-IPRF-PST-IND<br>‘(he) was about to read’ |   |                                  |
| (72) Complete:  | <i>jud-i</i><br>read-CVB.SEQ<br>‘(he) has finished reading’                 | <i>neen-Ø-Ø</i><br>CPL-NPST-IND  | <i>jud-i</i><br>read-CVB.SEQ<br>‘(he) had finished reading’ | <i>nee-n-ta-n</i><br>CPL-PST-IND |

The pre-inflectional aspects are highly productive, and the marking of one of them is obligatory in most main clause predicates (except for those verbs in irrealis mood), and that is why I refer to them as pre-inflectional. This means that these aspects are essential and basic in Shuri semantics. The imperfective aspect expresses habitual or general events/states, or ongoing events. Thus *hun jum-u-Ø-n* (book read-IPRF-NPST-IND) may be used when referring to the action of reading in general (without specifying one particular, actualized action), as in ‘scholars in general read books’, a habitual action of reading, as in ‘I read books every day’, or a particular ongoing action, as in ‘I’m reading books’. The perfective aspect treats the action as a delimited whole. The perfective aspect is intertwined with tense, in such a way that the perfective *-a* simultaneously designates past tense. Thus, whereas the imperfective aspect can co-occur with both non-past and past, as illustrated in (73) and (74), the perfective aspect does not allow an additional tense marking, as in (75).

- (73) *jum-u-Ø-n*  
read-IPRF-NPST-IND  
‘(he) reads (imperfective non-past)’
- (74) *jum-u-ta-n*  
read-IPRF-PST-IND  
‘(he) read (imperfective past)’

- (75) *jud-a-n*  
 read-PRF-IND  
 '(he) read (perfective past)'

An alternative analysis for (75) would be to assume an underlying structure *jum-* + *-ta* + *-n*, which surfaces as *ju-da-n*, with the morphophonological process (/m/ + /t/ → /d/). Note here that there is no aspect marking between the stem and inflection in this analysis. But this analysis has its problem, given that it cannot explain why the corresponding non-past form *jum-* + *-Ø* + *-n* does not exist, even when it should be expected to exist.

### 12.5.7 Modality

As compared with Japanese and certain Southern Ryukyuan varieties such as Miyako, one prominent feature of Shuri and other Northern Ryukyuan is that speech-act distinctions are integrated into the mood system of inflectional morphology. That is, a finite verb must indicate whether it is a statement (indicative mood), a question (interrogative mood), doubt (dubitative), etc. In Modern Japanese, for example, these distinctions are largely carried by sentence-final modal clitics rather than modal inflectional affixes.

Another interesting feature of Shuri modality is evidentiality. Evidentiality marking conspires with the tense-aspect system. As we noted in 12.5.6, a past event may be encoded by either an imperfective form or a perfective form, as in:

- (76) *jum-u-ta-n*  
 read-IPRF-PST-IND  
 'read (imperfective past)'
- (77) *jud-a-n*  
 read-PRF-IND  
 'read (perfective past)'

The difference between the two is not only in aspect, but in evidentiality. The imperfective aspect marker tells us that the speaker saw someone read, whereas the perfective aspect is unmarked in this regard.

### 12.5.8 Passive and causative

The passive marker is a verbal affix *-rarir*. An agent-less passive is more common. If the agent does occur in a passive clause, the agent NP is marked with dative =*nkai* (not =*kai*; see 12.5.1.3.2). The passive morpheme also expresses the potential.

- (78) *ari=nkai sugur-at-ta-n*  
 3SG=DAT hit-PASS-PST-IND  
 '(Someone) was hit by him/her.' [Nishioka and Nakahara 2006: 113]
- (79) *\*ari=kai sugur-at-ta-n*  
 3SG=DAT hit-PASS-PST-IND  
 '(Someone) was hit by him/her.'

The causative marker is a verbal derivational affix *-as* (or *-imir* when the root ends in /s/). The causer is encoded as subject, and the causee agent is marked with dative =*nkai* or =*kai*.

- (80) *siitu=ga hun jum-u-Ø-n*  
 student=NOM book read-IPRF-NPST-IND  
 agent theme  
 ‘The student reads a book.’
- (81) *sinsii=ga siitu=nkai hun jum-as-u-Ø-n*  
 teacher=NOM student=DAT book read-CAUS-IPRF-NPST-IND  
 causer causee agent theme  
 ‘The teacher makes the student read a book.’

An interesting feature of the Shuri causative is that it is used for benefactive as well: the causer is interpreted as a beneficiary, and the causee agent is interpreted as a benefactor.

- (82) *wan=nee sinsii=kai nici hakar-ac-a-Øn*  
 1SG=TOP teacher=DAT temperature take-CAUS-PRF-IND  
 causer causee agent  
 (beneficiary) (benefactor)  
 ‘The teacher kindly took my temperature for me.’ (lit. ‘I made the teacher take my temperature.’)

## 12.5.9 Subordinate clauses

### 12.5.9.1 Complementation

The quotative clause is formed by attaching a clitic =*ndi*. Any kind of clause, finite or non-finite, or any constituent (e.g. an NP), can be embedded with this strategy.

- (83) [*saazaa=nu usiimun cukuj-i-kutu, kuuwu*]=*ndi*  
 heron=GEN soup make-IPRF-CVB.CSL come.IMP=QT  
*i-misee-bi-ta-n*  
 say-RSP-POL-PST-IND  
 ‘“I have cooked heron soup, so come,” said (the king).’ [Nishioka and Nakahara 2006: 144]

The affix *-si* replaces the modal affix and turns the entire clause into a nominalized clause. In each example below, the nominalized clause is in brackets.

- (84) [*kaara=wuti ʔwiiz-oo-Ø-si*]=*nu wu-ta-n*  
 river=LOC swim-PROG-NPST-NLZ=NOM exist-PST-IND  
 ‘There was someone who was swimming in the river.’ [Tshako 1997: 379]
- (85) [*an ʔj-u-Ø-si*]=*ga juta-sa-Ø-n*  
 that.way say-IPRF-NPST-NLZ=NOM good-VLZ-NPST-IND  
 ‘Saying (it) that way is good.’ [Nishioka and Nakahara 2006: 52]
- (86) [*are=e kac-u-Ø-si*]=*ga hwee-sa-Ø-n*  
 3SG=TOP write-IPRF-NPST-NLZ=NOM fast-VLZ-NPST-IND  
 ‘S/he writes fast.’ (lit. ‘She is fast when it comes to writing.’) [Nishioka and Nakahara 2006: 52]

In (84), the affix *-si* denotes an agent who performs the action (attracting the attachment of the nominative =*nu* rather than =*ga*), whereas in (85) and (86) it functions as an

action nominalization (attracting the attachment of the nominative =*ga* rather than =*nu*). The affix *-si* must have been a formal noun meaning ‘man; person; thing’ that headed an adnominal clause, which was then grammaticalized as an affix of the preceding predicate, replacing the adnominal-unmarked mood suffix *-ru*.

### 12.5.9.2 Adnominal clause formation

An adnominal clause precedes a head nominal (12.5.1.2). No relativizer such as a relative pronoun is used. The predicate verbal must be inflected as an adnominal form. In the example below, the adnominal clause is in brackets. Here, the subject of the adnominal clause is relativized.

- (87) [*kurikaa=nkai miikuni cukur-at-ta-ru*] *suba-jaa*  
 around.this.place=DAT newly make-PASS-PST-ADN suba-shop  
*sicc-oo-misee-Ø-mi*  
 know-PROG-RSP-NPST-Q  
 ‘Do you know the suba (Okinawan noodle) shop that was newly established around here?’ [Nishioka and Nakahara 2006: 117]

As is common in Japonic languages in general, an NP of any grammatical relation (core or peripheral argument) can be relativized. Furthermore, an NP that cannot be regarded as an argument of the adnominal clause can establish a modifying semantic relationship with the adnominal clause, where pragmatic inference determines how the adnominal clause narrows down the reference without the head noun playing any role in the adnominal clause.

### 12.5.9.3 Adverbial-adsentential linking

Various adverbial-adsentential relations may be established by making use of non-finite and finite converbs (see 12.4.2.1.7).

- (88) *mazun [kusa kaj-iga] ik-a*  
 together grass gather-CVB.PUR go-INT  
 ‘Let’s go together to get grass.’
- (89) [*saki num-iinee*] *cira=nu aka-ku naj-i-Ø-n*  
 saké drink-CVB.ANT face=NOM red-INF become-IPRF-NPST-IND  
 ‘If (you) drink saké, your face will become red.’
- (90) [*nama=kara jum-u-Ø-kutu*], *kak-i*  
 now=ABL read-IPRF-CVB.CSL write-IMP  
 ‘(I) will read now, so write (it).’
- (91) [*nama=kara jum-u-Ø-siga*], *kak-i*  
 now=ABL read-IPRF-CVB.CNS write-IMP  
 ‘(I) will read now, but do not write (it).’

## 12.6 LEXICON

The Shuri lexicon mainly comprises native roots, Sino-Japanese loans, Japanese loans, Chinese loans, and English loans.

Native roots account for the overwhelming majority of the Shuri lexicon. The ratio of native roots to all kinds of loans in Shuri is much higher than that in Japanese. Most basic lexemes come from native roots, but there do exist culturally important basic lexemes that come from Chinese (Uemura 1997; e.g. *tiida* ‘sun’, *dusi* ‘friend’, and *duu* ‘body; oneself’ come from 天道, 同士, and 胴, respectively).

Sino-Japanese loans came into the Shuri lexicon via Japanese at various periods of history. Especially after the establishment of the Ryukyu dynasty (see 12.1), many Sino-Japanese loans were introduced by the educated ruling class (samurai clan). Sino-Japanese loans of this kind are largely restricted to cultural and sophisticated concepts (Uemura 1997), and are thus commonly found in the written language rather than the vernacular.

Chinese loans are loans that directly came into Shuri, rather than via Japanese (which I call Sino-Japanese here). The existence of direct Chinese loans is owing to the fact that the Ryukyu dynasty was in a frequent and deep contact with China, both politically and commercially. Tsuhako (1997: 388) mentions that there may be 100 or so such loans in Shuri, and lists the following Chinese loans as examples: *sjanpin* (香片) ‘jasmine’, *seehwan* (菜飯) ‘vegetable fried rice (kind of Chinese dish)’, and *cinkunsin* (進貢船) ‘a tribute ship’.

English loans are again loans that directly came into Shuri rather than via Japanese. The influx of English loans was due to the post-Second-World-War occupation of Okinawa by the United States from 1945 to 1972. Tsuhako lists the following English loans: *peedee* ‘payday’, *suubenija* ‘souvenir’, *goobureeki* ‘goldbrick (a lazy person)’, *haanii* ‘honey (meaning a bedfellow of a marine)’, etc. Unlike a vast number of loans of English origin that came into the Shuri lexicon via Japanese, the English loans I mention here are peculiar to Shuri (and other Okinawan dialects), and no Japanese would identify the meaning of *goobureeki* and *haanii*, for example. During the period of the US occupation, no substantial measure was taken in terms of language policy; Okinawans studied Japanese at school as before. This fact, together with the short period of occupation by the US, is reflected in the fact that there is no massive influx of English loans in Shuri and other Ryukyuan varieties.

## NOTE

- 1 Exceptional verbs include the following (forms given are non-past imperfective indicative, negative non-past indicative, and past perfective indicative): *sun*, *san*, *san* ‘do’; *cuun*, *kuun*, *ccan* ‘come’; *icun*, *ikan*, *?nzan* ‘go’; *?jun*, *?jan*, *ican* ‘say’; *nanzun*, *nndan*, *nncan* ‘see’; *-rarijin*, *-raran*, *-rattan* passive/potential.

## REFERENCES AND FURTHER READING

### 1. References

- Kokuritsu Kokugo Kenkyūjo (1963) *Okinawa-go Jiten* [Okinawan dictionary], Tokyo: Printing Bureau, Ministry of Finance.
- Moseley, Christopher (ed.) (2010) *Atlas of the World's Languages in Danger*, 3rd edition, Paris: UNESCO Publishing. (Online version *UNESCO Interactive Atlas of the World's Languages in Danger*, <http://www.unesco.org/culture/languages-atlas/>)
- Nishioka, Satoshi and Jō Nakahara (2000/2006) *Okinawa-go no Nyūmon* [Introduction to Okinawan], Tokyo: Hakusui-sha.
- Tsuhako, Toshiko (1992/1997) ‘Okinawa Chū-Nanbu Hōgen’ [South-central dialect of Okinawa], in Kamei, Takashi, Rokurō Kōno, and Eiichi Chino (eds) *Nippon Rettō no Gengo* [The languages of the Japanese archipelago], Tokyo: Sanseido.

Uemura, Yukio (1992/1997) 'Sōsetsu' [Structure], in Kamei, Takashi, Rokurō Kōno, and Eiichi Chino (eds) *Nippon Rettō no Gengo* [The languages of the Japanese archipelago], Tokyo: Sanseido.

## 2. A suggested list of further reading

- Ashworth, David Edward (1973) 'A generative study of the inflectional morphophonemics of the Shuri dialect of Ryukyuan', PhD thesis, Cornell University.
- Chamberlain, Basil Hall (1964 [1895]) 'Essay in aid of a grammar and dictionary of the Luchuan language', Tokyo: Yushodo Booksellers [Yokohama: Kelly & Walsh].
- Miyara, Shinshō (2000) *Uchināguchi Kōza: Shuri Kotoba no Shikumi* [A lecture on Uchināguchi: the structure of Shuri], Okinawa: Okinawa Times.
- Nishioka, Satoshi (2003) 'Okinawa Shuri Hōgen no Keigo Tsuki Dōshi' [Honorific verbs in the Shuri dialect of Okinawan], *Ryūkyū no Hōgen* 27, Tokyo: Hōsei Daigaku Okinawa Bunka Kenkyūjo.
- (2009) 'Uchinā guchi', in Kaji, Shigeki, Tōru Hayashi, and Yumi Nakajima (eds) *Jiten Sekai no Kotoba 141* [Encyclopedia of the world's languages 141], Tokyo: Sanseido.
- Shibatani, Masayoshi (1990) *The Languages of Japan*, Cambridge: Cambridge University Press.
- Suzuki, Shigeyuki (1960) 'Shuri Hōgen Dōshi no Iikiri no Katachi' [The citation forms of Shuri verbs], *Kokugogaku*, 41: 74–85.
- Takaesu, Yoriko (1999) 'Ryūkyūgo Shuri Hōgen' [The Shuri dialect of Ryukyuan], in Chino, Eiichi and Yoneo Ishida (eds) *Sekai no Kotoba 100go Jiten, Ajia hen* [Encyclopedia of the world's 100 languages: Asia], Tokyo: Sanseido.
- (2007) 'Okinawaken Nahashi Shuri Hōgen no Keiyōshi' [The adjective in the Shuri dialect of Okinawan], in Kudō, Mayumi (ed.) *Nihongo Keiyōshi no Bunpō`Hyōjungo Kenkyū o Koete* [The grammar of Japanese adjectives: beyond the study of standard language], Tokyo: Hitsuji Shobō, 241–62.
- Takahashi, Toshizō (1991) *Omorosōshi no Kokugogakuteki Kenkyū*, Tokyo: Musashinoshoin.
- Uemura, Yukio and Lawrence, Wayne P. (2003) 'The Ryukyuan language', *Endangered Languages of the Pacific Rim* (Series), A4-018, Osaka: ELPR.

# SOUTHERN RYUKYUAN\*

Wayne P. Lawrence

## 13.1 INTRODUCTION

Southern Ryukyuan is the group of languages/dialects spoken on the Sakishima islands, which lie between Okinawa island to the northeast and Taiwan to the southwest. Genealogically the dialect group splits into the Miyako dialects and the Yaeyama dialects. The Miyako dialects, of which there are about 35, are spoken on Miyako island and surrounding islands, and on Tarama island and nearby Minna island; and the Yaeyama dialects, of which there are about 20, are spoken on the large islands of Ishigaki, Iriomote, and Yonaguni, and the surrounding smaller islands. The Yonaguni dialect was the first to diverge from the Yaeyama branch of Southern Ryukyuan, and because it has undergone many major innovations, it will be the focus of Chapter 14 and is not included in this description.

The Southern Ryukyuan dialects are all oral languages, but they have a rich oral literature of songs and plays, and the language of many of these differs from the colloquial language and may be termed literary.

All of the Southern Ryukyuan languages are underdescribed, with the possible exception of the Ishigaki dialect (Yaeyama), which now has a 17,600-word dictionary (Miyagi 2003). A Nakachi dialect (Miyako) dictionary is expected to be published in 2012, and a Hatoma dialect (Yaeyama) dictionary is being compiled.

The following introduction to the Southern Ryukyuan dialect group is based primarily on the Hatoma dialect (Yaeyama), but reference will be made to other Yaeyama and Miyako dialects where appropriate. All forms not specifically identified as being from a particular dialect are to be interpreted as being Hatoma dialect forms.

Southern Ryukyuan dialects are traditionally referred to using a toponym followed by the word ‘dialect’ (Kabira dialect, Tarama dialect, etc.), but hereafter in this description the toponym alone will be used to refer to the dialect.

## 13.2 PHONOLOGY

### 13.2.1 Consonants

The consonant inventory of Hatoma is set out in [Table 13.1](#). *c* is realized as [ts]. *z* ([dʒ]) is functionally the voiced equivalent of both *c* and *s*. These sibilants are palatalized to [tɕ][dʒ][ç] before the front vowels and *j*. *h* is realized as [ç] in this same environment, and as [ɸ] before *u*. *h* is found geminated in only one form (*sahhun* ‘soap’, also attested as *saffun*), and *f* is usually geminated. In cases where it is pronounced short, it is transparently related to a word with a geminate *f*. This suggests that *ff* (and variant *f*) can be considered to be the phonetic realization of geminate *h*. Ikema and closely related Miyako dialects (Nishihara, Sarahama) have a contrast between *f* and *h* (Ikema *funi* ‘boat’, *huni* ‘bone’ [ɸuni]; *fau* ‘eat’, *hau* ‘crawl’), with *h* deriving from historical *p*

TABLE 13.1 CONSONANT INVENTORY OF HATOMA

	Labial		Dental/alveolar		Palatal	Velar		Glottal
Plosive	p	b	t	d		k	g	
Affricate			c	z				
Fricative	(f)		s					h
Nasal		m		n			ŋ	
Tap				r				
Approximant		w			j			

before vowels other than original *u*, and *f* from *k* and *p* before *u*. *ŋ* is pronounced as a moraic nasal with the same place of articulation as a following stop, but in other environments it is a somewhat backed [ŋ]. *ŋ* is found in syllable-coda position, and in word/morpheme-initial position before stops (including affricates and nasals). Geminate nasals are interpreted as a *N* + nasal sequence. The Miyako dialects also have a moraic *m*, which also occurs in word/morpheme-initial position (Nakachi *mnivci* ‘chest’, *cibinum* ‘rectum’), and Ikema and closely related dialects, Nagahama, and Tarama (all Miyako), also have a moraic *r* (retroflex [ɻ]) (Nagahama *par* ‘needle’, Tarama *arki-tar* ‘walk-PAST’).

Geminate voiceless fricatives occur in word-initial position (*ffa* ‘child’, *ssan* ‘louse’) and geminate voiceless consonants occur word-medially (*sippurun* ‘suck on’, *buccu* ‘navel’, *maffa* ‘pillow’). The Miyako dialects and Agarisuji (Kuroshima island, Yaeyama) also have geminate *v* (Ikema, Tarama *vva* ‘you’, Kuroshima *vva* ‘child’). Kuroshima and Aragusuku (Yaeyama) also have singleton *v* (Aragusuku *ava* ‘fat, oil’, Kuroshima *uva* ‘you’), whereas in the Miyako dialects singleton *v* is moraic (Hirara *sivkaï*, Nakachi *sivkam* ‘sour’, Nakachi *jaci-vsã* ‘mugwort’ (cf. *fusa* ‘grass’)) and may be phonological *vu* (or, more abstractly, *gu* or *bu*).

Initial geminate plosives (pronounced with glottal tension) are found in the Miyako dialects in derived environments only (Nakachi *ttã* ‘come-PAST’), and in Taketomi (Yaeyama) they exist as phonetic variants of fricative + high vowel + stop sequences: *hupasa* ‘dark’ [ppãsa ~ ɸɸpãsa], *sikara* ‘strength’ [kkãra ~ ɕkãra]).

In the Yaeyama dialects, vowel- and *N*-initial forms tend to begin with a non-phonemic glottal stop, and this is sometimes preserved word-medially in compounds. In several Miyako dialects (Nakachi, Karimata) and in one form in Sonai (Iriomote island, Yaeyama), the glottal stop is used in reduplicated adverbs. It is always released, often producing creaky phonation: e.g. Nakachi *jamʔ-jam* [jamʔmjam ~ jam:jam] ‘painful’; Karimata *ssõʔ-ssu* [ssoʔossu ~ sso:ssu] ‘white’; Sonai *marõʔ-marõ* [marõʔõmarõ] ‘round’.

Ikema and closely related dialects have voiceless moraic nasals in word-initial position, always homorganic with a following nasal (Ikema, Nishihara *ŋmu* ‘cloud’, *ŋna* ‘rope’), but, at least in Nishihara, compounds such as *ama-gumu* ‘rain cloud’, point to the voiceless nasal being derived (in this case, from morphophonemic *ku* via fricativization and assimilation).

Ogami (Miyako) is unique among the Japonic languages in having no voiced obstruents except for *v*; however, the *ci* and *zi* of neighbouring dialects are distinguished as [kʰũ ~ kʰũ̃] and [kũ] respectively ([kʰũ̃:] ‘breast milk’, [kũ:] ‘letter, character’), suggesting that, at least for velars, the voicing distinction is underlyingly preserved, even if not phonetically realized on the stops.

### 13.2.2 Vowels

Hatoma has six vowels, of which one is marginal.

High	i	(i)	u
Mid	e	o	
Low	a		

All vowels except *i* also occur as contrastively long vowels, although there is a tendency to shorten long vowels in speech at normal speed. The mid vowels are restricted to heavy syllables, with exceptions in recent loanwords and in the sentence-final particle *-gera*. *i* is restricted to light syllables and is always preceded by *s* and followed by *k*, and is thus always devoiced (see below for environment for devoicing). The high vowels are in complementary distribution in this environment, the phonetic realization being determined by the following vowel according to the pattern illustrated below.

<i>sik-</i> ‘to place’	<i>sik-ee</i> ‘place-PERF’	<i>sik-iri</i> ‘place-IMP’
	<i>sik-anu</i> ‘place-NEG’	
	<i>sik-ooru</i> ~ <i>suk-ooru</i> ‘place-HON’	<i>suk-uta</i> (~ <i>sik-uta</i> ) ‘place-PAST’

This *i* is replaced with *i* by younger-generation speakers (under 70 years old). *i* is fully phonemic in other dialects. In the Miyako dialects (except Ōgami) and in Aragusuku (Yaeyama), *i* is commonly pronounced as [z] (or as [s] in devoicing environments: e.g. Nakachi *pītu* ‘person’ [pstu]). The affricate *z* becomes a fricative following *i* (Nakachi *gazimagii* ‘Chinese banyan tree’ [ga<sup>d</sup>zmaqi:], *izu* ‘fish’ [zzu], *izaizi* ‘borrow-INF’ [zzazzi]). In Ōgami (Miyako), *i* is pronounced [ü] and may have a velar fricative off-glide before *a* and a labiovelar off-glide before *u*.

Long *ē* ([ɜ:]) is attested morpheme-internally in Hateruma (Yaeyama) and closely related Shiraho. Ishigaki, Komi (both Yaeyama), Tarama (Miyako) and closely related Minna have long *ē* in morphologically derived environments, deriving from *i-a* (Ishigaki *mukasēē* ‘long ago-TOP’; Tarama *fugēē* ‘nail-TOP’). Tarama also has long *ū* ([y:]) synchronically deriving from *i-u*.

Phonetic nasal vowels are found in Sonai and Taketomi (Yaeyama) (Sonai [pət̚wã] ‘Hatoma’, [a:ɾjã] ‘ant’, [ɸwã:] ‘child’ (cf. [ɸwa:] ‘saddle’)), but in at least Taketomi these alternate with nasal consonants (Taketomi [jū:ndura]~[jumundura] ‘read-FP’, [ʔã:dari]~[ʔamadari] ‘drop (of water)’).

In Hatoma, vowels in light syllables are devoiced between voiceless consonants but vowels in adjacent syllables do not devoice unless separated by a compound boundary ([s̚akae̚iki] ‘wine cup’, [s̚aki-sukka:] ‘wine pourer’, [to:ɸu-p̚aku] ‘tofu box’).

In Nakachi (Miyako), which has vowel devoicing in the same environment as Hatoma, [ʔ] (from *k*) does not count as voiceless (e.g. [s̚aki] ‘bloom-IMP’, [s̚afu] ‘bloom-ATTR’, [s̚aʔaɴ] ‘bloom-NEG’). In a number of dialects, including Hateruma, Kabira, Kohama, Komi, Shiraho, Sonai (Yaeyama) and Karimata (Miyako), vowels in light syllables are devoiced after voiceless consonants, and a following voiced consonant is also devoiced (Karimata [f̚u̚mu] ‘cloud’, [k̚af̚u̚ɾi-ba-du] ‘write-COND-EMPH’).

### 13.2.3 Suprasegmentals

Words in Hatoma have a minimal word length of two moras. Several Yaeyama dialects (Sonai, Taketomi) show a strong tendency towards vowel shortening so have monomoraic

nouns; however, in Taketomi at least, this is synchronic shortening of underlyingly long vowels.

Hatoma is a language where pitch (tone) is phonemic. Most words in Hatoma fall into one of two tonal categories, the unmarked class and the marked class (Lawrence 1997a, 2001). In the marked class, words are high from the syllable containing the second mora (i.e. word-initial light syllables are low) to the end of the word or phrase. The unmarked class differs from the marked class in that the word/phrase-final syllable is also low.

<i>unmarked</i>	<i>marked</i>
pàá 'tooth; blade'	páá 'leaf'
hùcí 'comb; straw sandals'	hùcí 'mouth'
kinài 'quinine'	kinái 'home'
káára 'tile'	káára 'river'
kásánàùN 'carry on back'	kásánáùN 'place one on another'
náNkúkúRù 'by itself'	húNsúkúRú 'pocket'

In word formation processes involving nouns (compounding, affixation), the class of the word-initial component usually becomes the class of the whole noun. In noun compounds with an unmarked first constituent, the sequence of high syllables covers the whole of the first constituent and finishes on the rightmost, non-final, iambic foot head (*itáNdá-ffàì* 'eating for free', *itáNdá-pátàràkì* 'working for free'). In addition to these two tonal classes, there are several peripheral tone classes: a number of nouns are pronounced entirely low (*ràNpù* 'lamp', *zìNzìNjàjàà* 'firefly'), and many adverbs are pronounced with the word-initial iambic foot low (*jààdìN* 'necessarily', *àtààcímá* 'a bit'). There is no requirement that a high tone be present somewhere in a phrase, so it is possible for whole sentences to be on low pitch, e.g. *bàNà kàkìtì kàmài tùrùN* 'set a trap and catch wild boar'.

From the tonal distribution described above, the following details of syllable structure can be deduced (periods mark syllable boundaries): *VuN* is one syllable but *V.iN* is two (*káúN* 'buy', *kù.tàùN ~ kù.tái.rùN* 'reply'; *kù.ìN* 'calendar', *kù.ìN-nù* 'calendar-NOM'); *V:N* is one syllable but *V:i* is two (*ù.kòòN ~ ù.kìN* 'turmeric', *jóóN* 'darkness', *kù.má.rì.kóó.ì* 'hide-and-seek').

Two tonal classes of the type found in Hatoma are found in most Yaeyama dialects, but in Aragusuku the distinction is in the process of being lost. Of the Miyako dialects, the dialects of Irabu island, inland Miyako island, Hirara city centre and environs, and Ōgami no longer have tonal distinctions. The northern dialects of Karimata and Ōura, and the dialects along the southern coast are reported to have two tonal classes (Hirayama 1967), but no further work has been done on their description. A recent brief description of the Tarama tone system (Matsumori 2010) indicates that Tarama (Miyako) has more than two contrasting patterns, and not the two as described in earlier literature.

### 13.2.4 Morphophonology

Hatoma is an agglutinative language, with very little in the way of segmental morphophonology. The only process which is widespread in the language is that by which the sequence *i-a* becomes long *e* when the *i* is in a light syllable, and *ija(a)* when in a heavy syllable: *kakeeN* 'write-PERF' from *kaki-aN*; *kaijaN* 'buy-PERF' from *kai-aN*. In parallel, the sequence *u-a* becomes long *o* when the *u* is in a light syllable, and *uwa(a)* when ending a heavy syllable: *karooN* 'light (adj.)' from *karu-aN*; *kauwan* 'itchy' from

*kau-an* (exception: *aoon* ‘blue’ from *au-an*). However, not all Southern Ryukyuan dialects are as simple. In contrast to Hatoma which possibly has the simplest verb conjugation, in terms of morphophonological processes, of any Japanese or Ryukyuan variety, Ishigaki (Yaeyama) may well have the most complicated of any Japanese/Ryukyuan verb conjugation (Miyara 1995).

In many Miyako dialects the vowel-initial Topic and Object particles (*-a* and *-u* respectively) trigger gemination of an immediately preceding consonant articulation (including *i*) (Nobaru *ammu* ‘net (*am*)-ACC’, *toovvu* ‘tool (*toov*)-ACC’, *fugizu* ‘nail (*fugi*)-ACC’, *kiizu* ‘health (*ki*)-ACC’), coalesce with the vowel of a preceding light syllable (Nobaru *funjuu* ‘boat (*funi*)-ACC’, *naboo* ‘dirt (*naba*)-ACC’), and are pronounced as *-ja* and *-ju* respectively after other heavy syllables (Nobaru *paiju* ‘south (*pai*)-ACC’, *soaju* ‘pole (*soo*)-ACC’). In Tarama (Miyako) the situation is somewhat different, with coalescence after light syllables (including after *i*) (*fugüü* ‘nail (*fugi*)-ACC’) (but *u* does not coalesce with *i* (*funiu* ‘boat (*funi*)-ACC’)), gemination after *r* (*budurru* [-[lu] ‘dance (*budur*)-ACC’) unless a demonstrative (*kuruu* [-ru:] ‘this (*kur*)-ACC’), and *-ja* (topic) and *-ju* (object) after other heavy syllables (*kagamju* ‘mirror (*kagam*)-ACC’, *ciju* ‘breast milk (*ci*)-ACC’, *kiju* ‘voice (*kui*)-ACC’).

In Nakachi (Miyako), velar plosives become [h] between *a* (long or short) when preceded by a voiced consonant, and [ʔ] when not preceded by a voiced consonant (*ma-kaja* [mahaja] ‘cogon grass’, *pak-an* [paʔan] ‘spit out-NEG’, *a-ga* [aʔa] ‘I-NOM’). The *k* of the postposition *-kaa* ‘from’ becomes [h] after *a* regardless of the preceding consonant (*naha-kaa* [nahha:] ‘from Naha’, with the devoiced *a* lost), and the *k* of the adjective-ending *-kam* is [h] after *a* and *o* regardless of the voicedness of the preceding consonant ([jo:ham] ‘weak’, [gabaham] ‘old’, [asaham] ‘shallow’). This sound change was probably once productive after both *a* and *o* (note morpheme-internal *sooha* ‘ginger’ (cf. Ikema *sauka*)), but is now productive only after *a*.

### 13.3 MORPHOLOGY

#### 13.3.1 NOUNS

##### *Pluralization*

Aside from the personal pronouns (see 13.4.2), the plural is usually unmarked, but, if the referent is human, plurality is optionally marked with one of three suffixes: *-taa*, *-nkee*, *-numee*. *-taa* is used with pronouns (*-caa* after the demonstrative series), personal names (*mijazato-san-taa* ‘Mr Miyazato and others’), and a small number of kinship terms. *-nkee* and *-numee* are used after other human referents. Compared with *-nkee*, *-numee* lacks respect and is not used with referents deemed worthy of respect (*sakasa-nkee* / \**sakasa-numee* ‘priestesses’, cf. *kangohu-nkee* / *kangohu-numee* ‘nurses’). Reduplication is commonly used to mark wh-words as plural (*nuu-nuu* ‘what (pl.)’, *nuusi-nuusi* ‘how (= in what ways (pl.))’), and is less commonly used to pluralize inanimate nouns (*jama-jama* ‘places of worship’, *kizaru-kizaru* ‘events’).

##### *Case marking*

The subject is marked with *-nu* or zero-marking, with *nu*-marking being the unmarked option. Zero-marking is used when the subject is a first or second person pronoun (but not third person), the anaphoric pronoun *unaa* (but not *duu*), the interrogative pronouns

'who' and 'where' (but not 'what' or 'when'), given names (but not family names), positions (e.g. 'mayor') referring to a specific individual (but not when used with general reference), plurals in *-taa/-caa* (but not in *-Nkee* or *-numee*), and the kinship terms for mother, father, older sister, aunt, and uncle.

<i>waa</i> 'you-NOM'	<i>kari-nu</i> 's/he-NOM'
<i>taa</i> 'who-NOM'	<i>nuu-nu</i> 'what-NOM'
<i>sinsii</i> 'teacher (you)-NOM'	<i>sinsii-nu</i> 'teachers-NOM'
<i>usitu-taa</i> 'younger.brother-PL-NOM'	<i>buzasa-Nkee-nu</i> 'uncle-PL-NOM'
<i>aaaja</i> 'my.father-NOM'	<i>uja-nu</i> 'parent-NOM'
<i>anmaa</i> 'older.sister-NOM'	<i>siza-nu</i> 'older.brother-NOM'

In Miyako dialects, the subject is marked with *-nu* and *-ga* (cognate with Hatoma's zero-marking), and the distribution is similar to that of Hatoma (Nakama 2000: 177ff).

Overt case marking of the object is optional. If marked, as is common, *-ba* is used, or, in literary style, *-ju*. It is reported (Izuyama 2003: 30) that, in Miyara (Yaeyama), *-ba* may also be used to mark a subject in a situation which is undesirable for the speaker (Miyara: *ami-ba hui, zima-ge har-ar-anu* 'Rain is falling, and (I) cannot go anywhere'; cf. Hatoma: *ami-nu hui, maa-N par-ar-anu*; Yonaguni *-ba*, Chapter 14). In Hatoma this phenomenon is found in one fixed expression: *maa-ba sikibu* 'evil spirit has attached (onto s.o., causing bad behaviour)'. Nakachi (Miyako) marks the object with *-u* (1a), but *-a* is used when marking the object of a conjunctive form (1b) (see also 4b):

- (1) a. *kurjuu asi-ru*                      b. *kurjaa assi-kaa asib-i*  
       this.ACC do-IMP                      this.ACC do.INF-after play-IMP  
       'Do this!'                              'Play after doing this!'

Other case marking particles include *-si*, which marks the instrumental; *-a* ( $\emptyset$  after a heavy syllable, except *pajaa* 'to the south' (*pai* 'south')) denotes movement, usually of the subject, to a place; *-naa* denotes location (in time/place), and also the place to which movement, usually of the complement, takes place; *-naati* marks the place where an action takes place; *-ara* (*-ra* after a heavy syllable) means 'from (time/place)', but is also used to mark the standard of comparison (see 32d) and also means of transportation; *-N* marks the destination of abstract motion, usually to a person (as in speaking, writing, or similarity to), and also the agent of a passive and causee of a causative formed with a transitive verb; *-tu* marks comitative; *-baaki* 'until' and *-baaki-naa* 'by (a point in time/place)'; *-saari* marks a means or cause (e.g. *kazi-saari* 'due to the wind').

### *Diminutives*

Hatoma productively forms diminutives with the suffix *-ama* (*-naama* after *-N*, *-jaama*/*-waama* after heavy syllables ending in *i/u*, and *-ma* after long *a*) (Miyako *-gama*). The diminutive suffix *-naa* is common in Ishigaki and Miyara, and is found fossilized in Hatoma *oo-naa-ma* 'small pig' (cf. *oo* 'pig').

### 13.3.2 Verbs and adjectives

The citation form of Hatoma verbs is the Attributive (stem + *-u*) followed by modal *-N*. The other basic forms are the Negative (stem + *anu*), the Cohortative and Irrealis (stem + *a*), the Infinitive (stem + *i*), and the Imperative (stem + *i*). With the exception of a small

number of irregular verbs, all verb forms are formed as given above. A near exhaustive list of the irregular verbs in Hatoma is: *an* ‘have/be’ (stem *ar-*, citation form irregular), *bun* ‘have/be’ (stem *bur-*, Attributive *buu*, Perfect *bee*, and citation form irregular), *naanu* ‘not have/exist’, *sun* ‘do’ (Passive/Potential *sarin* ~ *sirarin*), *kun* ‘come’ (Cohortative, Irrealis, Imperative all *kuu*, Passive/Potential *kirarin*), *ffaun* ‘eat’ (citation form and Attributive alternative forms based on *ff-*), *ssun* ‘suck (a liquid)’ (stem *ss-* in some forms, *ssu-* in others), *ffurun* ‘give’ (Honorific *ffoorun*), *-arin* passive, potential (stem *-ar-*, citation form irregular), *jan* copula (stem *jar-*). Most verbs with stems ending in *-ir-* have two forms (e.g. *utirun* ~ *utun* ‘fall’). The form with *-ir-* is not used in the Infinitive, or in forms built on the Infinitive (the Perfect and Honorific), but all other forms have two alternative pronunciations (e.g. Neg. *utiranu* ~ *utanu*, Imp. *utiri* ~ *uti*). In Ishigaki (Yaeyama) the alternation is *utirun* ~ *utin* ‘fall’, and most forms are based on the short stem (Neg. *ut(j)unu*, Attr. *uti*). The Perfect (see 13.4.3.2) and Honorific (see 13.4.6) forms themselves conjugate, having Negative, Cohortative, Irrealis, Conjunctive, and Imperative forms.

Negative *-anu* is pronounced as *-anu* in sentence-final position, but becomes *-nu* ‘but’ elsewhere (including before sentence-final particles), except when preceding *-nu* ‘but’. The past negative becomes *-ansen* in sentence-final position (including before sentence-final particles), is *-ansen* or *-anta* when modifying a noun, but is *-anta(n)* elsewhere (see examples 33b, 51). *naanu* ‘not have/exist’ is inflected similarly. Past tense *-utan* becomes *-itan* after verb stems ending in *s*, and is commonly *-itan* with *r*-ending verb stems (*kak-utan* ‘wrote’, *ka-utan* ‘bought’, *panas-itan* ‘spoke’, *tut-tan* ‘took’ (*tur-* ‘take’)).

The citation form of Hatoma adjectives is the Attributive (stem + *-a*) followed by modal *-n*. This Attributive form is homophonous with the Adverbial form, which is used to form the Negative in conjunction with the negative verb *naanu*: *ac-a naanu* ‘not hot’ (*ac-an* ‘hot’), *taka-a naanu* ‘not high’ (*taka-an* ‘high’), *tuu-wa naanu* ‘not far’ (*tuu-wan* ‘far’). The Irrealis form is the stem + (*a*)*ra* (32b). Adjectives may be nominalized by suffixing *-(s)a* to the stem: *ac-a* ‘heat’, *mis-a* ‘goodness’ (*mis-an* ‘good’), *taka-sa* ‘height’, *tuu-sa* ‘distance’.

A very small number of adjectival words do not conjugate, and appear with *-na* in their attributive usage: *deezi-na ujaki-pusu* ‘extremely rich person’, *ifa-na ganbari* ‘odd mischief’, *baa siki-na munu* ‘things I like’. These are probably relatively recent loans from other dialects.

The adjectives *misan* ‘good’, *barasan*, *wassan* ‘bad’, and *maijan* ‘big’ are often replaced by the prefixes *ii-*, *jana-*, and *ubu-* when modifying nouns: *ii-oosiki* ‘good weather’, *ii-pusu* ‘good person’, *ii-tusi* ‘good year’, *jana-husi* ‘bad habit’, *jana-imi* ‘bad dream’, *ubu-kii* ‘large tree’, *ubu-kizaru* ‘big event’.

### 13.3.3 Numerals and classifiers

The basic cardinal numbers in Hatoma are (native series) *pusu-*, *huta-*, *mii-*, *juu-*, *ici-*, *muu-*, *nana-*, *jaa-*, *kunu-*, *tuu-*, interrogative *gjuu-*, and (Sino-Japanese series) *ici-*, *ni-*, *san-*, *jo(n)-/si-*, *go-/gu-*, *roku-*, *nana-/sici-*, *haci-*, *kjuu-/ku-*, *zjuu-*, interrogative *nan-*. In the absence of a classifier or measure word, numbers are given in the native series suffixed with *-ci* (exceptions: *pii-ci* ‘one’, *hutaaci* ‘two’, *kukunuci* ‘nine’, *tuu* ‘ten’).

In Hatoma, native numbers are used for one to ten, and in Ishigaki and Hateruma (both Yaeyama) for one to 19 (Ishigaki *tuu-kukunuci*, Hateruma *tuu-hakonaci* (lit. ‘ten-nine’)). Ishigaki speakers can compose numbers up to 999 using the native system (e.g. 555 *pjaaku-nu icici-tu tuu-nu icici-tu icici* (100-GEN 5-and 10-GEN 5-and 5)), but usually Sino-Japanese numbers are used for 20 or above. In Nakachi (Miyako),

native numbers are used for one to 19; 20 is Sino-Japanese *nizjuu* or native *pataci*; the native numbers for multiples of ten are no longer used, with the exception of *jasici-jaaci* for 88 years old; 100 is Sino-Japanese *pjaaku* or native *mumuci*.

The Southern Ryukyuan dialects have a relatively large number of commonly used classifiers. In Hatoma there are over 50 classifiers used with the native series of numbers (Kajiku 1983; Lawrence 1997b), and over half of these fall into the following categories.

*Classification according to animacy*: People are counted with *-(taa)ru* (*pusuru* ‘1 person’, *hutaaru* ‘2 people’, *micaaru* ‘3 people’, *jutaaru* ‘4 people’), but with Sino-Japanese numbers with the classifier *-niN* being used for numbers over four. Hateruma (Yaeyama) uses *-(ta)rī* for 1–4, and native Japanese numbers with *-pitu* ‘person’ for larger numbers. All animals except shellfish are counted using *-kkara*, and shellfish are counted with *-kku*. In Nakachi (Miyako), *-kuu* (cognate with Hatoma *-kku*) is used for shellfish and crustaceans, but turtles are counted with *-kara* (the shells are counted with *-kuu*).

*Classification according to shape*: *-ira* ‘flat objects’, *-kku* ‘roundish objects (eggs, balls of thread, rice bowls)’ (also root vegetables: e.g. carrots), *-mutu* ‘cylindrical rigid objects (trees, pillars, sticks, needles)’ (also herbaceous plants: e.g. lettuces), *-sizi* ‘long thin non-rigid objects (threads, noodles)’, *-sizi* ‘granular objects’.

*Classifier is an object which counts its function*: *-jaku* ‘strokes (rowing)’ (from *jaku* ‘oar’), *-kabu* ‘fishing trips’ (from *kabu* ‘groundbait’), *-kui* ‘songs’ (from *kui* ‘voice’), *-pan* ‘paces’ (from *pan* ‘leg, foot’); and quantities counted using the word for a container: *-baaki* ‘basketfuls’, *-saban* ‘bowlfuls’, etc.

*Classifier is a deverbal noun*: *-katami* ‘2 bucketfuls (carried on a pole on the shoulder)’ (from *katam(ir)un* ‘to carry on the shoulder’), *-kisi* ‘slices’, *-maaru* ‘cycles’, *-maraki* ‘30 bundles of rice’ (from *marakun* ‘to bind’), *-tabaru* ~ *-taburu* ‘2 handfuls of rice stalks bundled together’ (from *tabarun* ‘to bundle’), etc.

*Classifier is a part which counts the whole*: *-kkubi* ‘shirts, kimonos’ (related to *hubi* ‘neck, collar’), *-kiburu* ‘houses’ (related to *kiboosi* ‘smoke’).

*Units of length*: [spatial] *-busi* ‘inches; internode stems of bamboo or sugar cane’, *-(h)iru* ‘fathoms’; [temporal] *-juu* ‘evenings’, *-ka* ‘days’, *-nanka* ‘weeks after death’, *-siki* ‘months’, *-tusi* ‘years’, etc.

Sino-Japanese and Western loanwords function primarily as measure words (*-hun* ‘minute’, *-kin* ‘600 g’, *-kiro* ‘kilogram’, *-doru* ‘dollar’) rather than as classifiers, and they are used with the Sino-Japanese series of numbers.

Approximate amounts are indicated with the particle *bukara*: *ici-zikan-bukara* ‘about one hour’, *tu-kkara-bukara* ‘about ten (animals)’.

Ordinal numbers are formed with the suffix *-mii* or *-me(e)* after the classifier.

- (2) *mii-musi-mii-na izu hoos-i-N gee-ta pinma-a*  
3-time-ORD-at fish catch-INF-PURP went (suppletive) time-TOP

*izu tu-kkara hoos-ee-tan*

fish 10-CL catch-PERF-PAST

‘The third time (I) went fishing, (I) caught ten fish’

Fractions are expressed according to the model *miici-nu piici* ‘one third’ (i.e. three-GEN one), but the quantity ‘half’ is usually expressed using *panbun*.

### 13.4 SYNTAX

#### 13.4.1 Noun phrase structure

Nominal expansions of the demonstrative stems are given in Table 13.2. Although classified in Table 13.2 as proximal-medial-distal, the distinction in usage between proximal and medial forms is not clear-cut.

##### *Noun modification*

When used in an attributive relationship with a following noun, the location, time, and amount forms above are followed by the Genitive particle *-nu* (*un-nu maara* ‘around that time’, *ubi-nu cigai* ‘just that much difference’). The “thing”-series suffixes *-nu* directly onto the stem (*u-nu panasi* ‘that story’). “Kind of” is denoted by *kai-bu(u)/kai-buca*, *ai-bu(u)/a(i)-buca*, and indefinite *nuusi-buu/nuu-cca* (*kai-bu kazi* ‘this kind of wind’).

Possession is marked by the genitive particle *-nu* ~  $\emptyset$ , the distribution of the two allomorphs being identical to that of the subject marker (see 13.3.1). Note the distinction between *bee sinta* ‘behind us’ and *bee-nu sinta* ‘behind our house’. Modification of a noun by a postpositional phrase is mediated by the Genitive particle (3a), and clausal modification is achieved by embedding the clause-final conjugated element in the Attributive form before the modified noun (3b).

- (3) a. *isankeera-nu situ*  
Ishigaki.from-GEN souvenir  
‘souvenir from Ishigaki’
- b. *noon sii-s-an pusu*  
anything do.INF-POT-NEG person  
‘person who cannot do anything’

Internally headed relative clauses are formed using the nominalizer *-mu* (4a). Nakachi (Miyako) (4b) and Hateruma (Yaeyama) use cognate *munu* ‘thing’.

- (4) a. *basannaru-nu namandauumi s-ii-buu-moo hucisuboo-nu*  
banana-NOM half.ripe do-INF-be.ATTR-NMR.TOP tart-so  
*ffa-ar-anu*  
eat-POT-NEG  
‘Bananas which are half ripe are tart and are thus inedible’

TABLE 13.2 HATOMA DEMONSTRATIVE PRONOMINALS

	thing	location	point in time	small amount
proximal	<i>ku-ri</i>	<i>ku-ma</i>		<i>ku-bi</i>
medial	<i>u-ri</i>	<i>u-ma</i>	<i>u-N</i>	<i>u-bi</i>
distal	<i>ka-ri</i>	<i>ka-ma</i>	<i>u-N</i>	<i>u-bi</i>
interrogative	<i>ziri</i> ‘which’ <i>nuu</i> ‘what’	<i>ma-</i>		

- b. *basanaai-nu namaumma s-ii bui-munoo fa-aiN*  
 banana-NOM half.ripe-ACC do-INF be-NMR.TOP eat-POT.NEG  
 ‘Bananas which are half ripe are inedible’

Ishigaki, Miyara (Yaeyama) and Hirara (Miyako) use the nominalizer *-su*, but this is not attested in Hatoma, and in Nakachi (Miyako) cognate *si* is very restricted in the structures it is used in (e.g. *ikan-si-du masi* ‘(It) is better not to go’).

#### *Noun phrase coordination*

Noun phrases are coordinated using *-tu* (also the comitative marker) (*kabi-tu enpicu-nu an* ‘There are paper and pencils’). A partial listing is given using the particles *-nujaa* or *-tiba* (*kabi-tiba enpicu-tiba, nuun-kuiN micun-ken an* ‘There are all sorts of things, paper and pencils and the like’). *-N* ‘also’ (*-nun* or *-jun* after *N*, and *-jun* after other heavy syllables) is used to indicate that all marked nouns are involved (*saki-N saa-jun numi buu* ‘I drink both rice wine and tea’, *isa-N jabu-N ooranu* ‘There are neither doctors nor herbalists’).

#### *Quantification*

Positive indefinite quantifiers are formed from wh-words by suffixing *-nnaa-ka*, and are formed from quantity words using *-kajaa*: *noo-nnaa-ka-nu an* ‘There is something’, *taa-nnaa-ka-ba mireeru* ‘(I) saw someone’, *gjutaaru-kajaa oottan* ‘Several people came’.

Wh-words followed by *-nnaa-N* and quantity words with *-N* give universal quantifiers (‘everything’ is also expressed as *nuu-N-kui(-N)*, and ‘every time’ is *iccin* ‘always’).

- (5) a. *mukasee maa-nnaa-N oo-nu makee ari-b-uta*  
 long.ago.TOP where-nnaa-N pig-GEN farm.TOP have-be-PAST  
 ‘In the past there used to be pig-sties everywhere’
- b. *husee noo-nnaa-N an*  
 fault.TOP what-nnaa-N have  
 ‘Everything has its faults’
- c. *pakoo gjuu-ci-N ari-b-uta-nu . . .*  
 box.TOP how.many-CL-N have-be-PAST.ATTR-but  
 ‘(There) were many boxes, but . . .’

Negative indefinite quantification is expressed by a wh-word + *-N* followed by a negative predicate, and a wh-word + *-N* followed by non-negative predicate gives a free choice quantifier.

- (6) a. *taa-N sikai-oor-anseN*                      b. *noo-N ffa-ar-anu*  
 who-N use-HON-NEG.PAST                      what-N eat-POT-NEG  
 ‘Nobody used (it)’                                      ‘(I) cannot eat anything’
- c. *aibu munoo taa-N tur-arin-doo*  
 that.kind thing.TOP who-N take-POT-FP  
 ‘Anybody can take that kind of thing’

The anti-quantifier (forcing distributivity) *naa* appears as a prefix with reduplicated forms but is a suffix with a quantity: *naa-jaajaa* ‘everyone’s respective houses’,

*naa-meemee* ‘each person respectively’, *naa-taraasitaraasi* ‘supplementing each other’s deficiencies’, *hutaaru-naa* ‘two people each; in groups of two people’, *kubi-naa* ‘this small amount each’, *Nmeema-naa* ‘a little each; little by little’.

### 13.4.2 Pronouns and anaphora

Hatoma personal pronouns are set out in Table 13.3, along with the parallel “chez” (house of) series.

The interrogative pronouns are *nuu* ‘what’, *nuusi* ‘how’, *nunti* ‘why’, *ma-* ‘where’, *ici* ‘when’.

Hatoma has two anaphoric pronouns: *duu* and *unaa*. *duu* is coreferent with the nearest c-commanding subject.

- (7) a. *baa<sub>i</sub> usitu<sub>j</sub>-nu duu<sub>\*i<sub>j</sub></sub>-nu jaa peer-u-muu*  
 I.TOP younger.brother-NOM self-GEN house enter-ATTR-NMR  
*mit-taN*  
 see-PAST  
 ‘I saw my younger brother enter (his) own house’
- b. *baa<sub>i</sub> usitu<sub>j</sub>-N duu<sub>i<sub>j</sub>/n<sub>j</sub></sub>-nu zi<sub>N</sub> batas-ita<sub>N</sub>*  
 I.TOP younger.brother-to self-GEN money pass-PAST  
 ‘I passed my younger brother (my) own money’
- c. *taroo<sub>i</sub>-ja ziroo<sub>j</sub> mutikee-ta duu<sub>i<sub>j</sub>/n<sub>j</sub></sub>-nu bentoo-ba*  
 Taro-TOP Jiro.NOM bring-PAST self-GEN packed.lunch-ACC  
*ffii<sub>t</sub>-ta*  
 give-PAST  
 ‘Taro gave me (his) own lunch which Jiro had brought’

In (7a,c) *duu-nu* may be replaced by *unaa* with no change in meaning. However, replacing *duu-nu* with *unaa* in (7b) results in a change of the referent to the younger brother. This is because *unaa* always has third person reference.

- (8) *hanako/\*baa/\*waa-ja unaa-duu-si duu-nu sigutu misikee-taN*  
 Hanako/\*I/\*you-TOP self-self-by self-GEN work discover-PAST  
 ‘Hanako/\*I/\*you found her/\*my/\*your job by her/\*my/\*your-self’

In Nakachi (Miyako), the anaphoric pronoun is *una*, but the form *duu* is used in two expressions: *duu-si* ‘by oneself’ (*\*una-si*) (e.g. *duu-si asi-ru* ‘Do (it) yourself!’), and *duu-katti* ‘selfish’.

TABLE 13.3 HATOMA PERSONAL PRONOUNS (SEE ALSO 13.4.6)

	1p	2p	3p			interrog.
			prox.	med.	dist.	
sg.	<i>baa</i>	<i>waa</i>	<i>kuri</i>	<i>uri</i>	<i>kari</i>	<i>taa/taru</i>
pl.	<i>bee</i> (incl.) <i>bantaa</i> (excl.)	<i>wa(a)taa</i>	<i>kuccaa</i>	<i>uccaa</i>	<i>kaccaa</i>	<i>taataa</i>
chez	<i>bante, bee</i>	<i>watte</i>	<i>kunne</i>	<i>unne</i>	<i>kanne</i>	<i>tatte</i>

In Hatoma, *duu*, with no overt antecedent, may also be used to refer to the speaker.

- (9) *duu-ja kai-ru umu-i-buu-juu*  
 self-TOP this.way-EMPH think-INF-be.ATTR-FP  
 ‘I am thinking along these lines’

### 13.4.3 The basic sentence

#### 13.4.3.1 Declarative

The Southern Ryukyuan dialects are tense/aspect-final, with tense/aspect marked on verbs (including copula) and adjectives. The copula, however, is used only when required by morphological considerations, as in (10b) where the emphatic particle requires a following Attributive form (see also 36b).

- (10) a. *kuree sumuci*      b. *kuree sumuci-du jaru*  
 this.TOP book            this.TOP book-EMPH COP.ATTR  
 ‘This is a book’            ‘This is a book’

The negative of the copula *jan* is *aranu*, optionally preceded by the topic marker *-a*: *sumuci aranu ~ sumucee aranu* ‘(It) is not a book’. (Note that the negative of the verb *ar-* ‘have/exist’ is suppletive *naanu*.) The copula is also used in isolation as a complete utterance: *jan* ‘That’s right’; *aranu* ‘That’s wrong’.

Existence is expressed using the verbs *an* and *bun*. *bun* is used of entities which are perceived to be capable of self-controlled change of location, and *an* is used of all other subjects.

- (11) a. *bee taa-ja paita-naa an*  
 we.GEN paddy.field-TOP southern.place-at be  
 ‘My rice fields are on Iriomote island’  
 b. *kazee iramuti-nu pajaaki-naa-ru b-ee*  
 wind.TOP Iriomote-GEN southerly.direction-at-EMPH be-PERF.ATTR  
 ‘The typhoon is to the south of Iriomote island’

Possession may be expressed using the existential construction with the possessor as the topic.

- (12) *karee usitu-nu bun*  
 s/he.TOP younger.brother-NOM be  
 ‘S/he has a younger brother’

In Hatoma the possessor is not given dative-marking, but in Nakachi (Miyako) the dative may also be used to indicate possession (13).

- (13) *abaa utu-nu-du bui* or *anna utu-nu-du bui*  
 I.TOP younger.brother-NOM-EMPH be I.DAT.TOP  
 ‘I have a younger brother’

#### 13.4.3.2 Tense and aspect

The citation forms of verbs are perfective, with stative verbs being non-past, and dynamic verbs (e.g. *numun* ‘drink’) future.

Non-past Habitual (14a) is marked by the Infinitive form followed by *buu* ‘be’ (e.g. *numibuu* ‘drink’, Neg. *numiburaanu*), and this form is also used to mark a single event with the predictive sense of ‘of course’ (14b).

- (14) a. *buusoo-ja bikidumoo ffa-i-ja nar-anu-ti*  
 rice.offering-TOP man.TOP eat-INF-TOP become-NEG-QUOT  
*az-ar-i-b-uu*  
 say-PASS-INF-be-AFF  
 ‘(It) is said that men must not eat rice offerings’ [Kajiku 1991: 90]
- b. *acaa baa kair-i k-uu-baaki-naa sigutoo*  
 tomorrow I.NOM return-INF come-ATTR-until-at work.TOP  
*katazik-i-b-uu-naa-ree*  
 clear.away-INF-be-ATTR-FP-FP  
 ‘By the time I return tomorrow, you will have finished the work, won’t you!’

The Continuative is formed from the Infinitive form followed by the Perfect of *buu* (e.g. *numibee* ‘is drinking’, Neg. *numanbee*) and is used with both an action in progress (15a), or when the resulting state continues (15b,c).

- (15) a. *kari-nu k-uu-mu-ba mat-i-b-ee*  
 s/he-NOM come-ATTR-NMR-ACC wait-INF-be-PERF  
 ‘(I) am waiting for him/her to come’
- b. *jana-ssa-nu pana-nu sak-i-b-ee*  
 bad-grass-GEN flower-NOM bloom-INF-be-PERF  
 ‘The weeds are in flower’
- c. *ku-naa sik-ar-i-b-ee-moo nuu-ja*  
 here-at place-PASS-INF-be-PERF-NMR.TOP what-Q  
 ‘What is this thing put here?’

The Perfect is formed by fusing the conjunctive form with *an* ‘have’ (e.g. *numeen* ‘has drunk’, Neg. *numeenaanu*).

- (16) a. *waa ffa-a gju-taaru nas-ee-wa*  
 you child-TOP how.many-CL give.birth-PERF-Q  
 ‘How many children have you given birth to?’
- b. *hunee kisa nZ-i-par-eeN*  
 boat.TOP already leave-INF-go-PERF  
 ‘(The) boat has already left’

Past tense forms are Perfective *numutan*, Habitual/predictive *numibutan* (Neg. *numiburaanSEN*), Continuative *numibeetan* (Neg. *numanbeetan*), and Perfect *numeetan* (Neg. *numeenaanSEN*).

Aspectual forms formed from other verbs include the following:

Habitual	<i>num-i-aak-un</i>	(drink-INF-walk-AFF)
Continuative	<i>num-ee-ti aak-un</i>	(drink-PERF-QUOT walk-AFF)
Conative	<i>num-un-ti s-un</i>	(drink-AFF-QUOT do-AFF)
Prospective	<i>num-un-ti b-un,</i> <i>num-un-ti aak-un</i>	

Irreversible	<i>num-i-naanu</i>	( <i>naanu</i> ‘not exist’)
Preparative	<i>num-i-suk-un</i>	( <i>sik-</i> ‘place’)
Experiential	<i>num-i-mit-tan</i> <i>num-uta kutoo an</i>	( <i>mittan</i> ‘saw’) (drink-PAST thing.TOP have.AFF)
Completive	<i>num-i-kis-un</i>	( <i>kis-</i> ‘cut’)
Delimitative	<i>num-i-mir-un</i>	( <i>mir-</i> ‘see’)
Inchoative	<i>num-un-joo-ni nar-un</i>	(lit. ‘become as to drink’)
Inceptive	<i>num-i-pazimir-un</i>	( <i>pazimir-</i> ‘begin (intr.)’)

The Irreversible aspect formative *-naanu* is the negative of *an* ‘exist’ (and thus the past is *-naansen*). This usage of negative existence as the Irreversible aspect marker pervades the Ryukyu dialect area (cf. Miyako *-niin* (Nakachi), *-njaan* (Nishihara)). The Experiential ‘have done before’ is formed with the past tense form of *mirun* ‘see’, but the Negative is formed with the non-past, i.e. *numi-miranu*. This last expression is ambiguous between the Experiential ‘have not drunk (it) before’ and the Delimitative ‘will not have a taste’.

### 13.4.3.3 Modality

Conjecture with little epistemic commitment to the proposition is expressed using an embedded question (this structure also means ‘I do not know whether . . .’, the interpretation being determined according to the context).

- (17) (*hjottu s-uu-kaa*)                      *par-un-juu ss-anu*  
 perhaps do-ATTR-COND go-AFF-Q know-NEG  
 ‘(He) might go’

Conjecture with a higher degree of commitment to the truth of the assertion is expressed using *pazi*, with the degree of certainty of the outcome optionally expressed by modal adverbs: (*noosin/jaadin*) *parun pazi* ‘(He) will probably/definitely go’. Speaker’s commitment to the truth of the assertion is also expressed by verb/adjective-final *-N*: *jum-i-bee-ta* ‘(He) was reading’, *jum-i-bee-ta-N* ‘(I declare he) was reading’; *kisar-i-bee* ‘(It) has been cut’, *kisar-i-bee-N* ‘(Clearly it) has been cut’. A past event/situation which has not been witnessed first-hand but which the speaker is certain of from the available evidence (including hearsay) is expressed using the verb-ending *-eeru*: *jadu-nu ak-i-bur-eeru* ‘(No doubt) the window was open’, *umussa ar-eeru* ‘(No doubt it) was interesting’.

Evidentiality is expressed by suffixes. Hearsay is marked by the reportative evidentials *-ti* (the quotative particle) or *-coo* (on a low pitch): *parun-ti doo* and *parun-coo* ‘I hear that (he) will go’. Appearance (and also conclusions based on other evidence, including hearsay) is marked with the adjective *-ngisan*: *pari-ngisan* ‘It looks/seems like (he) will go’. This suffix is commonly used with the Adverbial of adjectives which describe feelings and so, without a marker such as *-ngisan*, can only be used to describe the speaker, or the addressee if interrogative: *jaasa-ngisan* ‘(He) is hungry’, *pijja-ngisan* ‘(He) is cold’, *umussa-ngisan* ‘(He) is happy’. Also used is the Adverbial form of adjectives followed by *SUN* ‘do’: *sabisa sun* ‘(He) is (acting) lonely’.

The deontic modality of obligation [must] is expressed as a conditional.

- (18) *s-a-ba-ru*                                      *nar-u*                                      (also less commonly *subaru naru*)  
 do-IRR-COND-EMPH become-ATTR  
 ‘(You) must do (it).’ (lit. ‘It will do only if you do it.’)

This is also expressed using a conditional linking two negatives (i.e. ‘if not done, it will not do’), as in *s-an-kaa nar-anu* and, with the added implication that there will be an undesirable consequence if not done, *s-an-too nar-anu*. Also commonly used is the structure Infinitive + *-sariru* (e.g. *num-i-sariru* ‘(you) must drink’). Prohibition [must not] is expressed by making the prohibited action the topic of *nar-anu* ‘become-NEG’: *s-ee* (do-INF.TOP) *nar-anu* ‘(You) must not do (it)’ (lit. ‘doing (it) will not do’).

Permission [may] and dispensation [need not] are expressed by *misan* ‘good’ preceded by positive and negative Infinitive forms respectively: *sii misan* ‘(It) is all right to do (it)’, *san misan* ‘(It) is all right not to do (it)’.

Responsibility [should] is expressed using the noun *biki* after the Infinitive form of the verb, and this structure does not convey imperative force.

- (19) a. *kuree waa-ru kak-i-biki*  
 this.TOP you-EMPH write-INF-should  
 ‘You should write this’
- b. *kak-i-biki-nu kutu*  
 write-INF-should-GEN thing  
 ‘what (one) should write’
- c. *kunu gwansoo waa mut-i-bikee ar-anu*  
 this funeral.tablet.TOP you hold-INF-should.TOP COP-NEG  
 ‘You should not inherit this funerary tablet’

Volition in Hatoma is expressed using the adjective *-pusan* (Hateruma *-boh-*, Ishigaki *-pus-* ~ *-pis-*, Miyara *-pis-*): *par-i-pusan* ‘want to go’. In some other Yaeyama dialects, a different form (Ishigaki, Taketomi *-tta-*, Miyara *-t’a-*) is also used to express volition. The cognate Hatoma form, *-cca-*, is used only with verbs expressing involuntary physical actions (*sibaru s-i-cca-an* ‘want to (= be about to) urinate’, *ssu mar-i-cca-an* ‘want to defecate’, *pak-i-cca-an* ‘want to vomit’, *pana pus-i-cca-an* ‘want to sneeze’). This same distribution is found in Nakachi (Miyako) (*jum-busi-munu* ‘want to read’, but *sibai s-i-ta-munu* ‘want to urinate’, *niv-ta-munu* ‘want to sleep’).

Intention is conveyed by the nouns *kangai* (from *kanga(ir)UN* ‘consider’) and, less commonly, *sanmin* ‘calculation’.

- (20) a. *baa acaa par-u kangai / sanmin*  
 I tomorrow go-ATTR consideration/calculation  
 ‘I intend to go tomorrow’
- b. *enma-a ukinaa par-u sanmin jar-i-ngisaa*  
 next.year-TOP Okinawa.to go-ATTR calculation COP-INF-seems  
 ‘(It) seems (he) plans to go to Okinawa next year’

The potential modality (possibility) is expressed by a number of forms, with the distinction between ability potential (*-juus-*; cognate *-bus-* in Ishigaki, Komi, and Miyara (Yaeyama)) and circumstantial potential (*-ar-*) important. Note the following examples (potential morphemes are underlined): *baa ui-juusanu* ‘I cannot [ability] swim’; *uijaa narunudu*, *pan buriti oor(ar)anu* ‘(I) can [ability] swim, but (I) broke (my) leg and cannot [circumstantial] swim’; *kjuu-ja pancasanu*, *uin pararanu* ‘(I) am busy today, so (I) cannot [circumstantial] go swimming’; *jarabee duu-si kin kisi-juusanu* ‘The child cannot [ability] put on clothes by (him)self’; *kumu kinmaa gumaanu*, *kis(ar)anu* ‘These clothes are small, so (I) cannot [circumstantial] put (them) on’.

Ability potential can also be expressed using *-seen* (*ui-seen* ‘can swim’, *ui-sanu* ‘cannot swim’). This is apparently a grammaticalization of the Perfect form of the verb *ssun* ‘to know’, a pan-Yaeyama development not found in Miyako (Uchima 1997). *narun* (*uijaa narun* ‘can swim’) is used with both ability and circumstantial meanings. *-kantii* (*ui-kantii sun* ‘cannot swim’) is a circumstantial potential with no positive equivalent.

In Nishihara (Miyako), the form *-(r)ai-* (*uugai-du si* ‘can swim’, *uugain* ‘cannot swim’), cognate with the Hatoma circumstantial potential, is used for both ability and circumstantial potentials, whereas *-juus-* (*uuzi-juusi-du si* ‘can swim’, *uuzi-juuhan* ‘cannot swim’), cognate with the Hatoma ability potential, is used with acquired ability (but not innate ability) and also where agent-internal circumstances (i.e. the agent’s physical or mental condition) determine potentiality, but not agent-external circumstances (e.g. the weather or rules) (Nakama 2000: 225–33). Unlike Nishihara, Hatoma does not use a potential construction to denote innate ability (‘fish can swim’, ‘birds can fly’).

#### 13.4.3.4 Non-declarative sentence types

##### Questions

Non-past wh-questions are marked with sentence-final Attributive + *-wa*: *waa nuu sibe-wa* ‘What are you doing?’, *ujaa taa-ru miru-wa* ‘Who will look after the parents?’, *nunti kakan-wa* ‘Why won’t (you) write?’. However, if the sentence ends in a noun phrase which includes an interrogative pronoun, the particle *-ja* is used.

- (21) a. *paaree-nu tumu-jakoo taa-ja*  
 dragon.boat-GEN stern-oar.TOP who-Q  
 ‘Who is the steerer of the dragon boat?’
- b. *taa jui-ja* c. *kjuu-ja nuu-nu pii-ja*  
 who.GEN cause-Q today-TOP what-GEN day-Q  
 ‘Whose fault is it?’ ‘What day is it today?’

Yes/no (nexus) questions and past wh-questions are typically marked by intonation alone.

- (22) a. *waa aca-N k-ii ffiir-un*  
 you tomorrow-also come-INF give.me-AFF  
 ‘Will you come tomorrow too?’
- b. *nuusi nat-taa*  
 how become-PAST  
 ‘What happened?’

Questions (both wh- and yes/no) are often made somewhat less direct by the use of clause-final *-kajaa* (*-kjaa* in fast speech), which is formally a self-directed expression of doubt, but which is often used to ask for information: *taa-ru buu-kajaa* ‘(I wonder) who is (here)?’, *taa-kajaa* ‘(I wonder) who (is it)?’.

In Nakachi (Miyako), wh-questions are marked with *-ga* on the wh-word (*ta-?a ttaa* ‘Who came?’, *karjaa NZa-N-ga bui* ‘Where is he?’). Yes/no questions are marked with *-ru* on the non-final word being questioned (*jaa cinuu-ru ttaa* ‘Did you come yesterday?’) or with *-ro* (*-mo* after *m*, *-no* after *N*) sentence-finally (*kjuu-ja ami-ro* ‘Is it rain today?’, *kafutam-mo* ‘Did (you) write?’).

Alternative questions are formed by coordinating the alternatives, each with the relevant intonation.

- (23) a. *kuree turu kaburee*      b. *kuree turu-kajaa kaburee-kajaa*  
 this.TOP bird bat                      this.TOP bird-Q bat-Q  
 ‘Is this a bird or a bat?’                      ‘(I wonder) is this a bird or a bat?’

The interrogative particle *-ka* is used rhetorically, when the speaker is in doubt and is trying to build up confidence to make a decision: *nibun-ka* ‘Shall I sleep? (Yes, I think I will)’, *pararin-ka* ‘I wonder if I can go?’, *kari-ka* ‘Is it him? (It seems it may be)’. Negative + *-noo* also forms rhetorical questions, expressing the speaker’s judgment: *waa jarabee aran-noo* ‘You’re (just) a child, aren’t you!’

Questions which function as negative assertions are formed with *-wa* following the *N*-form of verbs and adjectives: *kaijan-wa* ‘Is it pretty? (No it is not!)’, *narun-wa* ‘Can (he) do it? (No, he cannot!)’. A similar effect (with slight differences in nuance) is gained by the endings *-tuu*, *-noo*, *-noo-coo*, and *-ree*, also after *N*-forms. All of these constructions express the speaker’s opinion, so the subject must be third person. With a first-person subject, because the decision is in the hands of the speaker, ‘Will I go? (No, I won’t!)’ is rendered as *paran-tuu-hjaa* (with a derogatory attitude towards the addressee) or simply as *paran-doo* ‘(I) won’t go!’.

*-i* after the Cohortative form with a rising intonation marks a statement accompanied by a request for agreement: *baa sakinari para-i* ‘I will go on ahead, OK?’.

Negative leading questions are most naturally answered with *aai* ‘no’ followed by the appropriate form of the verb or adjective.

- (24) Q. *basannaroo naanu* ‘Aren’t there any bananas?’  
 A. *aai, naanu* ‘No, there aren’t.’  
 A. *aai, an-daa* ‘Yes, there are.’

#### *Commands, requests and proposals*

*-mii* (on a high pitch), grammaticalized from *mirun* ‘to see’, is used to mark a weak (low coercion) imperative (*kaki-mii* ‘write!’). *-nooree* follows a negative to mark a weak imperative or invitation (*san-nooree* ‘Won’t you do (it)?’). The Imperative form of verbs (*jaa pari* ‘Go home!’) can be followed by *-ba* to make it more forceful (*pari-ba* ‘Go!’) and further by *-joo* for still more force (*pari-ba-joo* ‘Go!’). *kurin kaka-naa* (*-naa* on low pitch) ‘Write this too!’ carries the implication that there will be an undesirable consequence if not done. The prohibitive (negative imperative) is formed with *-na* (*paru-na* / *pan-na* ‘Don’t go!’).

Requests are typically formulated using the verb *ffirun* ‘give me’ or an equivalent at the appropriate speech level (see 13.4.6).

- (25) a. *waa pimaa jar-u-kaa teenai s-ii ffiir-i*  
 you free COP-ATTR-COND assistance do-INF give.me-IMP  
 ‘If you are free, please help me’  
 b. *tanam-ar-i ffiir-ar-anu*  
 request-PASS-INF give.me-POT-NEG  
 ‘Could you do me a favour?’ (lit. ‘Couldn’t you give me (the favour of) being asked (to do something)?’)

Proposals are made using the Cohortative form, usually followed by a sentence-final particle (*para-dii* (*-dii* on falling pitch) ‘Let’s go’; *kurin kaka-naa* (*-naa* on high pitch) ‘Shall (I/we) write this too’ (asking for permission)), or by using a conditional.

- (26) a. *kak-u-kaa nuusi-jaa*  
 write-ATTR-COND how-Q  
 ‘How about writing?’
- b. *kak-i-ba-ru masi*  
 write-INF-COND-EMPH better  
 ‘It is better to write’

### Exclamations

Exclamations are formed by the suffixation of sentence-final particles to the attributive form of adjectives: *atarasa-wa-ree* ‘How cute!’, *ffaipusa-wa-ree* ‘How I want to eat!’, *kazi-nu pijja-joo* ‘How cold the wind is!’.

### 13.4.4 Topic, focus, and emphasis

The topic is marked with the particle *-a* (*-ma* after *N*, *-wa* after a heavy syllable ending in *u*, and *-ja* after all other heavy syllables). This particle is also used to mark contrastive focus. The topic is marked as zero after *baa* ‘I’, *waa* ‘you’, except for the purpose of contrast.

- (27) a. *taroo-ja isanke-naati kaimunu s-un*  
 Taro-TOP Ishigaki-at shopping do-AFF  
 ‘Taro will do the shopping in Ishigaki’
- b. *uree imee wakar-an-ban*  
 that.TOP meaning.TOP understand-NEG-FP  
 ‘I don’t understand the meaning of that’

A clause may be topicalized by putting the verb into the Infinitive form, and placing tense/aspect/polarity markers on the light verb *SUN* ‘do’ (28a). In (28b) this process has been repeated.

- (28) a. *uree baa-ja wakar-ee s-anu*  
 that.TOP I-TOP understand-INF.TOP do-NEG  
 ‘I don’t understand that’
- b. *ai-ja nar-ee s-ee s-anu* [Kajiku 1973: 48]  
 like.that-TOP able-INF.TOP do-INF.TOP do-NEG  
 ‘(I) am not able to do (it) like that’

In the same way that topicalized verbs are supported by *SUN* ‘do’, topicalized adjectives are supported by the verb *AN* ‘have’.

- (29) *kaija-a ar-u-nu . . .*  
 pretty-TOP have-ATTR-but . . .  
 ‘It is pretty, but . . .’

The Topic particle is not used with wh-words (*\*nuu-ja masi-ja* ‘What is better?’). This sentence would be *nuu-ru masi-ja*, using the emphatic focus particle *-ru* (also pronounced as *-du*, especially after *N*). This particle requires an Attributive form in clause-final position (see 10), and is incompatible with the modal suffix *-N* (cf. 30 with 27a).

- (30) a. *taroo-ru isanke-naati kaimunu s-uu*  
 Taro-EMPH Ishigaki-at shopping do-ATTR  
 'Taro will do the shopping in Ishigaki'
- b. *taroo isanke-naati-ru kaimunu s-uu*
- c. *taroo isanke-naati kaimunu-ru s-uu*

*-ru* is unnatural in relative clauses (*baa(??-ru) par-u pinma . . .* 'when I go . . .'), but is possible in embedded interrogatives. It is limited to one occurrence per clause, in both declarative and interrogative clauses.

- (31) *taa-ru nuu-ba(\*-ru) ici(\*-ru) ka-i k-uuta*  
 who-EMPH what-ACC(-EMPH) when(-EMPH) buy-INF come-PAST.ATTR  
 'Who bought what when?'

As a particle marking exclusive focus, *-ru* carries the strong implication that there are no other relevant candidates (32a,b), and it is also used in comparative constructions (32c,d).

- (32) a. *pakoo hutaaci-ru at-ta*  
 box.TOP 2-EMPH have-PAST.ATTR  
 'There were only two boxes'
- b. *karee taka-ra-ba-ru ka-u*  
 s/he.TOP high-IRR-COND-EMPH buy-ATTR  
 'S/he only buys expensive (things)' (lit. 'S/he buys only if expensive')
- c. *maija-kaa maija-mutee misa-ru ar-u*  
 big.ATTR-COND big.extent-TOP good-EMPH have-ATTR  
 'The bigger, the better'
- d. *tii-ra-N pan-du jam-u*  
 hand-from-even foot-EMPH hurt-ATTR  
 '(My) feet hurt more than (my) hands'

For increased emphasis, these focus particles are commonly doubled, but only when the pronunciations differ (i.e. *\*-ja-ja* and *\*-ru-ru*).

- (33) a. *kuree-ja baa-ru kak-u*  
 this.TOP-TOP I-EMPH write-ATTR  
 'As for this, I will write it'
- b. *cisa-a-ja nanzoo sukur-oor-anse-anaa*  
 lettuce-TOP-TOP so.much make-HON-NEG.PAST-FP  
 '(They) didn't grow lettuce very much'
- c. *sakusi-ru-du jaa-ja sig-u*  
 oldest.son-EMPH-EMPH house-TOP inherit-ATTR  
 'The oldest son inherits the house'

This particle doubling is also found in Taketomi (Yaeyama) (several examples in Kajiku 1998), but is not permitted in Nakachi (Miyako).

Other emphatic particles are *-baaki* 'even' (lit. 'until'), the exemplifier *-(a)ncan* (*-(ma)ncan* after *N*, *-wanCAN* ~ *-janCAN* after heavy syllables ending in *u*, and *-janCAN* after other heavy syllables), *-saagi* 'even' and *-N* 'even' (lit. 'also').

- (34) *rakkjoo-nu-saagi ar-u-kaa katimunoo naan-tantiN*  
 shallot-NOM-EMPH have-ATTR-COND main.dish.TOP not.have-CONC  
*misaN*  
 good  
 'If (I) only have shallots, I don't need any other dishes (with my rice)'

All of the above focus particles appear following case particles with no co-occurrence restrictions. The sequence of subject- and topic-marking particles expresses a strong feeling of unexpectedness, and is usually followed by a negative predicate, but given sufficient context may also be used without a following negative.

- (35) a. *uri-noo s-ii-juus-anu*  
 s/he-NOM.TOP do-INF-POT-NEG  
 '(Amazingly) s/he cannot do it'
- b. *uri-noo par-i s-u*  
 s/he-NOM.TOP go-INF do-ATTR  
 '(Completely unexpectedly) s/he will go'

Focus is also expressed by clefting. The cleft structure is formed using the nominalizer *-mu* (a truncated form of *munu* 'thing').

- (36) a. *baa uja-tu par-u-moo isanakee-ru ja-ru*  
 I parent-with go-ATTR-NMR.TOP Ishigaki.to-EMPH COP-ATTR  
 'It is to Ishigaki that I will go with (my) parent(s)'
- b. *uja-tu isanakee pat-ta-moo baa(-ru ja-ru)*  
 parent-with Ishigaki.to go-PAST.ATTR-NMR.TOP I(-EMPH COP-ATTR)  
 'It is I who went to Ishigaki with (my) parent(s)'

Hatoma is rich in sentence-final particles, and many of these are used to express emphasis. For example, '(I) don't understand!' can be expressed as *wakaran-saa*, *wakaran-daa*, or *wakaran-doo*, in order of increasing emphasis. *-baN*, *-gera* (*-gja* in fast speech) and *-curee* all express light exclamatory force (but stronger than *-saa*), but *-baN* additionally conveys unexpectedness and *-gera* a sense of confidence on the part of the speaker (similar to English 'of course'). *-coo* (on a high pitch) and *-joo* are both forceful indicators of conviction, but *-joo* may also simultaneously convey a request for confirmation.

### 13.4.5 Passive and causative

The Passive is formed with the morpheme *-ar-*, and with the logical subject (agent) marked with the particle *-N*.

- (37) *pabu-N hooreera deezi-daa (hooreera = hu- + ar-i- + -ara)*  
 snake-by bite.PASS.after big.problem-FP  
 'Having been bitten by a snake, it's a big matter'

Hatoma uses the same verb form as an 'adversative passive', where the logical object, if there is one, remains the object, the logical subject is marked with *-N*, and the subject is the affected party.

- (38) a. *agajaa kari-N koosi ffa-ar-inaanu*  
 oh s/he-by sweet eat-PASS-IRREV  
 'Oh! S/he has eaten my cake!' (lit. '(I) was eaten the cake by him/her')

- b. *uri-N paN hun (si)k-ar-i-ti jam-i nar-anu*  
 s/he-by foot stand.on-PASS-INF-CONJ hurt-INF become-NEG  
 ‘S/he stood on (my) foot, and it hurts to distraction’
- c. *ami-N hoor(-ar)-i-ti sigutu-N nar-anu*  
 rain-by fall.PASS(-PASS)-INF-CONJ work-also become-NEG  
 ‘It is raining, so (I) cannot even do any work’

The Passive form is also used to express ‘spontaneous’ (Shibatani 1990: 332–3).

- (39) a. *karee sidas-itan-ti umoor-i*  
 s/he.TOP make.up-PAST-QUOT think.PASS-AFF  
 ‘It occurs to me that she put on make up’
- b. *unu panasi suk-u-kaa nak-i-ru sir-ar-i-coo*  
 that story hear-ATTR-COND cry-INF-EMPH do-PASS-ATTR-FP  
 ‘When I hear that story, I cry’

The Causative has two forms, *-as-* and *-asim(ir)-* (*-im(ir)-* after *s-* ending verb stems): *kakasUN ~ kakasim(ir)UN* ‘cause to write’ (*kak-* ‘write’), *umoosUN ~ umoosim(ir)UN* ‘cause to think’ (*umu-* ‘think’), *ibasUN ~ ibisim(ir)UN ~ ibasim(ir)UN* ‘cause to plant’ (*ib(ir)-* ‘plant’), *panasasUN ~ panasim(ir)UN* ‘cause to speak’ (*panas-* ‘speak’), *sasUN ~ sim(ir)UN* ‘cause to do’ (*s-* ‘do’).

With the Causative of a transitive verb, the causee is marked with *-N* (40a), and the causee of an intransitive verb is marked as the direct object (40b,c).

- (40) a. *baa kari-N-du tigamee kak-as-i-pusa-coo*  
 I s/he-N-EMPH letter.TOP write-CAUS-INF-want.ATTR-FP  
 ‘I want to make her write a letter’
- b. *karee par-i-pusa naan-ti s-ita-N-du*  
 s/he.TOP go-INF-want NEG-QUOT say-PAST.ATTR-but-EMPH
- baa macaa-ba iramutee par-as-ita*  
 I Matsu-ACC Iriomote.to go-CAUS-PAST.ATTR  
 ‘He said he didn’t want to go, but I made Matsu go to Iriomote island’
- c. *karee par-i-pusa s-ii bur-ee-Nda-ru*  
 s/he.TOP go-INF-want do-INF be-PERF-so-EMPH
- baa macaa-ba iramutee par-as-ita*  
 I Matsu-ACC Iriomote.to go-CAUS-PAST.ATTR  
 ‘He was wanting to go, so I let Matsu go to Iriomote island’

The Passive of the Causative is formed as in (41), with the causee of the Causative verb becoming the subject of the Passive-Causative.

- (41) *baa unza-N taku ffa-as-ar-itan*  
 I s/he-by octopus eat-CAUS-PASS-PAST  
 ‘I was made to eat (the) octopus by him/her’

In addition to being a Causative suffix, *-as-* is also a common formative in lexical transitive/intransitive pairings, the main patterns of which are given below.

intransitive	transitive
-ir-	-as-
niirUN (ni-ir-) 'boil'	neesUN (ni-as-) 'boil'
NZirUN (NZ-ir-) 'go out'	NZasUN (NZ-as-) 'put out'
-ar-	-ir-
sumarUN (sum-ar-) 'be dyed'	sumirUN (sum-ir-) 'dye'
tasikarUN (tasik-ar-) 'help'	tasikirUN (tasik-ir-) 'help'
-r-	-s-
birUN (bi-r-) 'sit'	bisUN (bi-s-) 'place'
nuurUN (nuu-r-) 'get on'	nuusun (nuu-s-) 'place on'
Ø	-as-
sinkUN (sink-) 'sink'	sinkasUN (sink-as-) 'sink'
taraUN (tara-) 'be sufficient'	taraasUN (tara-as-) 'supplement'
-ir-	Ø
barirUN (bar-ir-) 'split'	barUN (bar-) 'split'
tarirUN (tar-ir-) 'drip'	tarUN (tar-) 'make (sugar, alcohol)'

### 13.4.6 Speech levels and respect

Respect language is a very important part of the Hatoma dialect. Age and social position are the primary factors which determine the level of respect language used, and of these age is the most important. When a person older than the speaker is the subject, subject honorification is used, and a speaker will use subject-honorific forms of a group which includes an elder, even if the group also includes the speaker.

- (42) *nika-a*                      *bugarinoosi*                      *s-oor-a*  
 this.evening-TOP    after.work.gathering    do-HON-COH  
 'Let's have a party this evening'

Other examples include *oora* 'Let's go' and *nkoora* 'Let's eat'. Also, an older person speaking to a younger person usually uses subject-honorific forms when speaking of himself.

Subject-honorifics are also used when the grammatical subject is an important natural entity (*siki-nu aar-oor-un* 'The moon rises'), or rice (*mai-nu uum-oor-an-ken* . . . 'Before the rice ripens . . .'), an important crop which is not grown on Hatoma and which, until 1962, the people of Hatoma had to sail to their fields on Iriomote to tend.

The subject-honorific form of most verbs is formed with the suffix *-oor-* (*i-oor-* after vowel-ending verb stems: *kaioorUN* from *kaUN* 'buy', *kangaioorUN* from *kanga(ir)UN* 'consider'; adjectives use the Adverbial form followed by the Honorific form of *an* 'have': *suuwa ar-oor-un* 'strong'), but some verbs have suppletive subject-honorific forms: *oorUN* 'go, come, be', *nkoorUN* 'eat, drink (alcohol)', *maarasUN* 'die'. This last form may be made more honorific by suffixing *-oor-* (*maarasoorUN*). Another word for 'die', the euphemistic *juusigirUN* (*juu* 'world' + *sigirUN* 'pass by'), is also subject-honorific but the degree of honorificity is not as high as *maarasUN*, and it too may be suffixed with *-oor-* (*juusigiroorUN*). *ffoorUN* 'give me', the subject-honorific form of *ffirUN*, is used when the subject is a superior (elder), and *taboorUN* is used of superiors of high social position (school principal, mayor) and of people the speaker feels particularly indebted to. *tabooroorUN* is still more honorific. *oosUN* 'give' and *ssarUN* 'say' are forms used when the speaker (who is not necessarily the subject) expresses respect towards the indirect object of the verb. These verbs may be converted to subject-honorific by suffixing *-oor-*. An object-honorific verb is *sikaasUN* 'accompany'.

TABLE 13.4 HATOMA PRONOUNS AND SPEECH LEVELS

	superiors	equals and subordinates	derogatory
2p sg.	<i>waa</i>	<i>waa</i>	<i>wANZA</i>
pl.	<i>wa(a)taa</i>	<i>wa(a)taa</i>	<i>wANZA-Nmee</i>
3p prox. sg.	<i>kunu-pusu</i>	<i>kuri</i>	<i>KUNZA</i>
prox. pl.	<i>kunu-pusu-Nkee</i>	<i>kuccaa</i>	<i>KUNZA-Nmee</i>
med. sg.	<i>unu-pusu</i>	<i>uri</i>	<i>UNZA</i>
med. pl.	<i>unu-pusu-Nkee</i>	<i>uccaa</i>	<i>UNZA-Nmee</i>
dist. sg.	<i>kanu-pusu</i>	<i>kari</i>	<i>KANZA</i>
dist. pl.	<i>kanu-pusu-Nkee</i>	<i>kaccaa</i>	<i>KANZA-Nmee</i>
anaphoric	<i>unaa</i>	<i>unaa</i>	<i>UNANZA</i>
indef. sg.	<i>taa</i>	<i>taa / taru</i>	<i>TANZA</i>
pl.	<i>taataa</i>	<i>taataa / taccaa</i>	<i>TANZA-Nmee</i>

*-juu* is a sentence-final particle, used in positive declarative sentences, which expresses respect towards the addressee by depreciating the actions or belongings of the speaker: *parun-juu* ‘I will go’, *kuree baa-mu-juu* ‘This is mine’, also (9). *-neera* is a subject-honoric interrogative particle used in sentence-final position in both wh- and yes–no questions.

In addition to verbs, personal pronouns have different forms to indicate respect, as shown in Table 13.4. (Note: *pusu* ‘person’.)

The forms used for ‘yes’ also depend on the addressee, with *oo* used towards superiors, and *NN* towards equals and subordinates. ‘No’ is *aai* to all addressees.

The Causative Imperative has a reduced pronunciation (*-aai*) which is used only towards subordinates (cf. ‘Please pass me (it)’ *jar-as-oor-i* [to superiors], *jar-as-i* [to equals or subordinates], *jar-aai* [to subordinates]).

### 13.4.7 Adverbs

The adverbial form of adjectives (and negative verbs) is homophonous with the attributive form: *kataa makun* ‘sow (seeds) thickly’ (*kataan* ‘thick, dense’); *majja narun* ‘become big’ (*majjan* ‘big’); *sabisa narun* ‘become lonely’ (*sabisan* ‘lonely’); *tatan narun* ‘become unable to stand’ (*tatun* ‘stand’).

As for adverbial suffixes, *-ni*, *-si*, *-ti*, and *-(t)tu* are common.

*atta-ni* ‘suddenly’, *sama-ni* ‘soberly’, *takaa-ni* ‘much’, *z(j)un-ni* ‘really’  
*jaata-si* ‘soon’, *jaarama-si* ‘slowly and carefully’, *zeezee-si* ‘in a drunken state’  
*biccu-ti* ‘identical’, *dakka-ti* ‘soaking’, *gjan-ti* ‘jam-packed’, *kisaa-ti* ‘already’  
*sikai-(t)tu* ‘for certain’, *musi-(t)tu* ‘the least bit’, *pasi-ttu* ‘refreshingly’

Adverbs ending in *-ku* are rare and are likely loans (*cjuraa-ku* ‘completely’, *paa-ku* ‘quickly’). Also relatively rare are adverbs ending in *-sa* (*pai-sa* ‘quickly’, *gjuu-sa* ‘how much’, *ikaa-sa* ‘how much’).

Reduplication, often followed by *-si*, is very common with an intensifying function: *ffoo-ffoo* ‘jet black’ (*ffoon* ‘black’), *naa-naa* ‘for a long time’ (*naan* ‘long’), *pisoo-pisoo-si* ‘widely’ (*pisoon* ‘wide’), *turii-turi-si* ‘quietly’ (*tur(ir)un* ‘become calm’).

The demonstrative adverbs are *kai* ‘in this way’ and *ai* ‘in that way’. Conjunctive adverbs include *andaa* ‘and then’, *andeekaa* ‘if that is the case’, *asi-ti* ‘and’, *asu-nu* ‘however’, *bee-ti* ‘thus’.

Ryukyuan adverbs often display a characteristic suprasegmental structure. In Hatoma, adverbs often have the word-initial iambic foot on a low pitch (e.g. the adverbs ending in *-ti* above), or the word-final mora alone is on a high pitch (*kànnààzì* ‘by all means’, *tàkààtàkà* ‘high up’), and reduplicated adverbs in the Miyako dialects of Karimata and Nakachi are characterized by the presence of a glottal stop, which is not part of the underlying phonemic inventory of these dialects.

### 13.4.8 Clause linking

Verb-ending predicates and clauses may be coordinated using the Infinitive form of verbs (43a), but adjective-ending predicates and clauses are coordinated by converting the adjectives into the Adverbial form and using the verb *an* ‘have’ in the Infinitive form (43b) (cf. 29, 32c). A similar structure is also possible with verbs, using the light verb *sun* (43c) (cf. 28).

- (43) a. *dansi mai ssa-i pamai sikoor-i*  
promptly rice polish-INF table.rice prepare-IMP  
‘Quickly polish the rice and prepare table rice’
- b. *jassa-N ar-i nmaa-N an*  
cheap-also have-INF delicious-also have  
‘(It) is both cheap and delicious’
- c. *waa-ja kak-i baa-ja jum-i s-aa*  
you-TOP write-INF I-TOP read-INF do-COH  
‘You write and I’ll read’

The Infinitive form may be extended by the conjunctive particle *-ti*. However, *-ti* after adjectives and stative forms of verbs expresses reason/cause (*isikaa-ti* ‘is short, so . . .’; *kak-ee-ti* ‘have written, so . . .’). *-nu* after adjectives expresses the same but from a more subjective (less logical) viewpoint. *-nu* after verbs, including stative forms and the negative *naanu* ‘not have/exist’, functions as a contrastive conjunctive particle.

- (44) a. *kirama-a mir-ar-i-nu mii-nu macee mir-ar-anu*  
Kerama-TOP see-POT-ATTR-but eye-GEN eyelash.TOP see-POT-NEG  
‘(You) can see the Kerama islands, but (you) cannot see (your) eyelashes’
- b. *gee-ta-nu oor-anSEN*  
go-PAST-but be.HON-NEG.PAST  
‘(I) went, but (you) weren’t (there)’

The contrastive conjunction of an adjective is achieved by introducing the verb *an* ‘have’.

- (45) *kaija-a ar-u-nu azee naanu*  
pretty-TOP have-ATTR-but flavour.TOP not.have  
‘(It) looks good, but (it) lacks flavour’

Reason/cause with verbs is expressed using *-ki* after the Infinitive form (46a), or by using the bare past tense form (46b).

- (46) a. *garazi-nu jam-i-ki-ru ukir-anSEN*  
head-NOM hurt-INF-because-EMPH get.up-NEG.PAST.ATTR  
‘(I) had a head-ache, so I didn’t get up’

- b. *ami-nu hu-uta(a) simikkeer-i-ti muir-anu*  
 rain-NOM fall-PAST damp-INF-CONJ burn-NEG  
 ‘It has rained, so (it) is damp and won’t burn’

The conjunctive particle *-ba* after the Irrealis *-a* (cf. *kak-a* ‘let’s write’) introduces a conditional. (Infinitive – actually a frozen realis, perhaps borrowed from Okinawan – is also found, cf. (26b).)

- (47) *baa kak-a-ba waa sik-i*  
 I write-IRR-COND you listen-IMP  
 ‘When/if I write, you listen’

The same *-ba* after the non-past attributive form of a verb is used when the speaker has made a firm decision: *baa kak-u-ba waa siki* ‘I will write, so you listen’ (also *baa kak-u-nda waa siki*); *baa kak-an-ba waa kaki* ‘I won’t write, so you write’.

The Conditional particle is *-kaa*, which follows the Attributive form (see also 25a, 32c, 34).

- (48) *tana-nu naan-kaa sikai-gurusAN*  
 shelf-NOM not.be.ATTR-COND use-difficult  
 ‘If/when there are no shelves, (it) is difficult to use’

There are two hypothetical concessive structures, one based on the irrealis conjunctive form *-aba*, and the other on the past tense form *-utan*.

- (49) a. *baa kak-abaN waa kak-una*  
 I write-CONCESS you write-PROH  
 ‘Even if I write, don’t you write’  
 b. *baa kak-utantiN waa kak-ee nar-an-daa*  
 I write-CONCESS you write-INF.TOP become-NEG-FP  
 ‘Even if I write, you must not write’

In (49a), the speaker has no intention of writing at the time of speaking. In contrast, in (49b) the speaker is signalling that it is quite likely that he will write. Also concessive is the particle *-munu*, which is also used at the end of counterfactual statements.

- (50) a. *ai-ni sanija s-iibee-munu gakkoo-naato par-un-ree*  
 that.much-ADV happy do-CONT-CONCESS school-such.like go-AFF-FP  
 ‘(With him) so happy, will he go to school or the like? (Of course he won’t!)’  
 b. *kak-ee misa-munu*  
 write-INF.TOP good.ATTR-CONCESS  
 ‘It would be good if you wrote’  
 c. *kak-ee-ba misa-ta-munu*  
 write-PERF.ATTR-COND good-PAST.ATTR-CONCESS  
 ‘(You) should have written’

*kutoo* (grammaticalized *kutu* ‘thing.TOP’) after the past tense form is used to introduce an immediate and unexpected occurrence in the following clause.

- (51) *acica hum-ee-ta kutoo buu kis-inaansen*  
 clog wear-PERF-PAST.ATTR thing.TOP cord cut-IRREV.PAST  
 ‘When (I) put on (my) wooden sandals, the thong broke’

Purposive, with a following verb of motion, is marked by *-N* (*-na* in most other Yaeyama dialects and cognate *-ga* in the Miyako dialects):

- (52) *manama mir-i-N par-ibee-ti u-naa mat-ibeer-i*  
 now see-INF-PURP go-CONT-so there-at wait-CONT-IMP  
 '(I) will go and meet you now, so wait there'

Other purposive clauses use *-joo-ni*. *-joo-ni* and *-kataci-ni* both create manner adverbial phrases, but whereas *-joo-ni* is restricted to being used after verbs and adjectives, *-kataci-ni* is used predominantly after the Genitive of nouns. It is also used after verbs, but is much less frequent than *-joo-ni* in this usage.

- (53) a. *waa kak-u joo-nee kak-i-juus-anu*  
 you write-ATTR like-ADV.TOP write-INF-POT-NEG  
 '(I) cannot write like you do'
- b. *pusoo bara-an joo-ni s-ii*  
 person.TOP laugh-NEG.ATTR like-ADV do-IMP  
 'Act so as not to laugh at people'
- c. *waa kataci-ni kak-i-juus-anu*  
 you.GEN like-ADV write-INF-POT-NEG  
 '(I) cannot write like you'
- d. *isi-nu kataci-ni koor-ibee*  
 stone-GEN like-ADV harden-CONT  
 '(It) has hardened like stone'

Embedded questions, even if sentence-final through right dislocation, are marked with clause-final *-juu* or *-saa*.

- (54) *par-un-juu par-an-juu ss-an-ban*  
 go-AFF-Q go-NEG-Q know-NEG-FP  
 'I don't know whether (I) will go or not'

In Nakachi (Miyako), an embedded question has normal interrogative marking and is followed by *ga(a)ra*.

- (55) *abaa taroo-ga noo-ju-ga asi-taa-?aara ss-an*  
 I.TOP Taro-NOM what-ACC-Q do-PAST-Q know-NEG  
 'I don't know what Taro did'

The negative connectives *-amuti* and *-andoosi* are used as in (56), with *-amuti* used when the following clause is an undesirable situation in the circumstances, and *-andoosi* when it is a desirable situation (often an imperative).

- (56) a. *sigutoo s-amuti saki-kaani num-ibee*  
 work.TOP do-NEGCONJ alcohol-only drink-CONT  
 '(He) isn't working, but is only drinking'
- b. *sigutoo s-andoosi uja-nu kanbjoo s-uuta*  
 work.TOP do-NEGCONJ parent-GEN caring do-PAST  
 '(He) didn't work, but looked after (his ill) parent(s)'

Temporal relationships between clauses may be expressed using *-pin* 'time', *-ken* 'period', *-ara* 'after' (lit. 'from'), *-mai* 'before' (lit. 'front') and *-nteena* 'while' (two actions performed by the same subject).

- (57) a. *tii sim-u piN-maa sahnUN sika-i-ba*  
hand wash-ATTR time-TOP soap use-IMP-FP  
'When you wash your hands, use soap'
- b. *tii sim-ibeen-ken maja-nu peer-i k-ee-tan*  
hand wash-CONT.ATTR-period cat-NOM enter-INF come-PERF-PAST  
'While I was washing my hands, a cat came in'
- c. *ami-nu hoon-ken-naa jaa par-a*  
rain-NOM fall.NEG.ATTR-period-at house.to go-COH  
'Before it starts to rain, let's return home'
- d. *tii sim-i-teera ii ffa-a-dii*  
hand wash-INF-CONJ.after food eat-COH-FP  
'After washing (our) hands, let's eat'
- e. *ff-uu mai-naa tii sim-i*  
eat-ATTR front-at hand wash-IMP  
'Wash your hands before eating'
- f. *ii ffa-i-nteena sinbUN jum-ibee*  
food eat-INF-while newspaper read-CONT  
'While eating, (he) is reading the newspaper'

A non-exhaustive listing of predicates may be expressed using the conjunctive form *-tari*, based on the past tense + *-ri*, or the Attributive form with *aaru*, both followed by *SUN* 'do'.

- (58) a. *uta iz-i-mit-tari buduru s-ii-mit-tari s-UN*  
song say-INF-EXPER-*tari* dance do-INF-EXPER-*tari* do-AFF  
'(I) will (do such things as) sing and dance'
- b. *uta iz-u-aaru buduru s-uu-aaru s-ii ikkena*  
song say-ATTR-*aaru* dance do-ATTR-*aaru* do-INF very  
*umussa s-uutan*  
interesting do-PAST  
'(I) did such things as sing and dance and thoroughly enjoyed (myself)'

Two actions by the same subject which take place repeatedly one after the other are expressed using an A-B-B-type reduplication followed by *SUN* 'do'.

- (59) a. *tat-i-tee bir-i-bir-i s-UN*  
stand-INF-CONJ.TOP sit-INF-sit-INF do-AFF  
'repeatedly stand up and sit down'
- b. *ffa-i-tee nib-i-nib-i s-UN*  
eat-INF-CONJ.TOP sleep-INF-sleep-INF do-AFF  
'repeatedly eat and sleep'

A variation on this common pattern is *ffaiti nibee sii sii SUN* 'repeatedly eat and sleep'.

### 13.5 LEXICON

Southern Ryukyuan dialects are straightforwardly identifiable by having an initial *i* in the word for 'sea', due to a change in proto-Southern-Ryukyuan from *\*umi* to *\*im*

((Miyako) Hisamatsu, Ikema, Tarama *im*; (Yaeyama) Hateruma *inaga*, Hatoma *in* (literary), Ishigaki *in-*, Sonai *in*, Taketomi *inna*). Although there are some words of unknown provenance, the bulk of Hatoma vocabulary is inherited from proto-Ryukyuan.

Chinese loanwords are comparatively uncommon in Hatoma. A count of a 6,450 word glossary gives around 9 per cent to be of Chinese origin, or containing morphemes which are of Chinese origin. Most of the loans are Sino-Japanese, probably entering Southern Ryukyuan via Okinawa. Some Chinese loanwords have entered the Ryukyus directly from southern China (e.g. Hatoma *koosa* ‘clenched fist’ (cf. Fuzhou *k’au<sup>213</sup>tsɔ<sup>53</sup>*)), but these too entered Southern Ryukyuan from Okinawa. A possible direct loan from a Chinese dialect into Southern Ryukyuan is Hatoma *daipaa* ‘mortar’ (Chinese *lěi bō*), with Miyako *naipa* suggesting a southern Chinese origin (cf. Fuzhou *lai<sup>55</sup>pua<sup>24</sup> ~ nai<sup>55</sup>pua<sup>24</sup>*).

A number of Sino-Japanese loanwords are so assimilated that native speakers are unaware of their relationship to the Japanese forms: *sinnin* ‘expressly’ (Sino-Japanese *sen* ‘devote oneself to’), *sippai* ‘to the best of one’s ability’ (Sino-Japanese *sei-iopai*)

Loanwords of Western origin constitute less than 1 per cent of the Hatoma vocabulary. A number of boating terms have been borrowed from English, probably around the period of the Russo-Japanese war (1904–1905) when Hatoma men served in the imperial navy which had been modelled on the British navy: *anaka* ‘anchor’, *burizi* ‘bridge’, *goohee* ‘go ahead’, *goositan* ‘go astern’. (These forms are also found in the dialects of a number of mainland Japanese fishing ports.) The word *kinai* ‘quinine’ (for treating malaria, previously a major health problem on Iriomote island) was probably borrowed directly from English rather than from Standard Japanese, which has *kinine* (from Dutch).

Loans from Northern Ryukyuan dialects include *haizoo* ‘stern sprit’ (the inherited Hatoma form would be *\*parizau*) and *haragoo* ‘skipjack tuna belly’. These particular words would have been introduced in the early twentieth century by fishermen from Itoman, Kudaka, Ou-jima, and Motobu (Okinawa).

Hatoma nouns are steadily being replaced by Standard Japanese forms, even in the speech of the oldest speakers: e.g. *akaru* ‘paper sliding door’ by *soozi* (Sino-Japanese *shōji*), *bunaru-bikiru* ‘siblings’ (lit. ‘female siblings–male siblings’) by *kjoodai* (Sino-Japanese *kyōdai*), *ikusajuu-nu atoo* ‘post-war period’ by Sino-Japanese *sjuusengo* (*shūsengo*), *maikari* ‘rice harvest’ by Sino-Japanese *sjukaku* (*shūkaku*), *pairu* ‘vinegar’ by *sii* (Native Japanese *su*), *pii* ‘guttering’ by *tui* (Native Japanese *toi*), and *sahhun* ‘soap’ by Sino-Japanese *sekken*. Other generational changes include *azun* ‘say’ being pronounced as *asun* and *ffun* ‘eat’ as *fuun* in the speech of the younger generation. Older-generation *barasan* ‘bad’ has become *wassan* in the younger generation, probably a loan from a Northern Ryukyuan dialect.

Words of mixed ancestry are rare but include *kakuri-zuu* ‘hiding place’ (native or borrowed *kakuri-* + *-zoo* (Sino-Japanese *-jo* via Okinawa)), *kwajaa-izu* ‘fish which are biting’ (borrowed Okinawan *kwajaa* + native *izu* ‘fish’), *nerukin* ‘flannel clothing’ (loan *neru* + native *kin* ‘clothing’) and *tamagarasu* ‘glass’ (native *tama-* ‘glass’ + loan *garasu*).

Mimetic vocabulary is predominantly adverbial and frequently involves reduplication. Typical examples are: *doon-toon-si* ‘knocking (at a door)’, *goon-koon-si* ‘gulping (down a drink)’, *ngoo-ngoo-si* ‘babbling’, *pacara-pacara-si* ‘twinkling’.

The use of *agamunu ~ agamunooma* (lit. ‘red-thing(-DIM)’) for *pii* ‘fire’ and *naamunu* (lit. ‘long-thing’) for *pabu* ‘poisonous snake’ in circumstances with strong religious significance is an example of linguistic taboo.

### 13.6 DIALECT ATTRITION

The youngest competent speakers of Hatoma are now in their 60s. The same is the case with Nakachi (Miyako). These are endangered languages, as are most Southern Ryukyuan dialects, yet until the mid-1950s these dialects were thriving. The following account of the decline of the Hatoma dialect draws heavily on Kajiku (2002).

The first school was established on Hatoma in 1896. Prior to this time, the inhabitants of Hatoma would have had no exposure to mainland Japanese language, or indeed to any written language, tied as they were to their island by the crippling poll-tax which was levied upon them from 1637 until its abolition in 1903. With very few exceptions, the furthest inhabitants would have ventured was to their paddy fields on Iriomote island seven kilometres away. Records of rice and cloth production used as payment of the poll-tax were kept using symbols denoting family groupings, types of produce and quantities (Uesedo 1976: 331–5). Some of this information was also recorded using knotted straw (*barazan*).

In 1898 Japan introduced conscription into Okinawa (although it was only in 1899 that the inhabitants of Okinawa gained Japanese citizenship). A number of men from Hatoma took part in the Russo-Japanese war, and Japan's victory in this war played a large rôle in moulding a sense of Japanese identity in the people of Okinawa. However, although conscripts and school children were exposed to Japanese, even in the late 1940s people on Hatoma who were proficient in Standard Japanese were the exception rather than the rule. There were only two radios on the island (one at the school, and one in the post office), so dialect was the normal language of all aspects of life outside of school.

In the late 1950s and early 1960s, mainland Japan was experiencing intense economic growth, and one result of this was that junior high school graduates from Hatoma would go to the Japanese mainland to seek employment. This led to rapid depopulation of the island. The 1955 population of 567 had more than halved to 220 by 1965. With the realization that the economic success of their children lay in mainland Japan, parents and grandparents began to make a conscious effort to use Standard Japanese towards the children in their families.

In 1967 OHK (Okinawa Public Broadcasting Corporation), which merged with NHK (Japan Broadcasting Corporation) in 1972 when Okinawa reverted to Japan, started television broadcasts into Miyako and Yaeyama, so from this year children became exposed to Standard Japanese from birth. The following year Hatoma's dried skipjack tuna factory closed due to falling catches. The skipjack tuna fishing and processing industry had been the main source of income for the island for over 50 years, and its collapse resulted in increased migration, so that by 1970 the population had fallen to 69. As of December 2011 it is 52 (36 households).

With increased travel to Okinawa island and mainland Japan, marriages between people of different dialect areas became common, and, because of the mutual unintelligibility of dialects, in such households the language used was usually, of necessity, Standard Japanese.

Although the youngest dialect-speaking generation is in its 60s, there are many in this age group who do not speak the dialect. Most confident speakers are now in their 70s or older. There are some rare cases of speakers of Southern Ryukyuan dialects who are in their 20s or 30s, but these are people raised in linguistically exceptional circumstances, such as those who were brought up by their grandparents who used dialect.

## NOTE

- \* I would like to express my deepest appreciation to Professor Kajiku Shin'ichi, who tirelessly responded to my questions on all aspects of the Hatoma dialect, and to Mr Tomihama Sadakichi, who provided all of the Nakachi (Miyako) data.

## REFERENCES AND FURTHER READING

- Bentley, John R. (2008) *A Reconstruction of the Proto-Language of the Southern Ryukyus*, Folkestone, UK: Global Oriental.
- Hirayama Teruo (ed.) (1967) *Ryūkyū Sakishima-Hōgen-no Sōgōteki-Kenkyū* [Comprehensive research into the Sakishima dialects of the Ryukyus], Tokyo: Meiji shoin.
- (ed.) (1983) *Ryūkyū Miyako-Shotō-Hōgen-Kiso-Goi-no Sōgōteki-Kenkyū* [Comprehensive research into the basic vocabulary of the dialects of the Miyako islands of the Ryukyus], Tokyo: Ōfūsha.
- (ed.) (1988) *Minami-Ryūkyū-no Hōgen-Kiso-Goi* [The basic dialect vocabulary of the Southern Ryukyuan dialects], Tokyo: Ōfūsha.
- Iitoyo Ki'ichi (ed.) (1984) *Kōza Hōgen-gaku 10: Okinawa / Amami-Chihō-no Hōgen* [Dialectology 10: The dialects of the Okinawa and Amami regions], Tokyo: Kokusho Kankōkai.
- Izuyama Atsuko (2003) 'The grammar of Ishigaki Miyara dialect in Luchuan', in Izuyama Atsuko (ed.) *Studies on Luchuan Grammar*, ELPR (Endangered Languages of the Pacific Rim) Publications Series A4-024. [Original Japanese version published as ELPR Publications Series A4-004 (2002).]
- Kajiku Shin'ichi (ed.) (1973) *Okinawa-ken Yaeyama Hatoma-jima-Hōgen* [Hatoma island dialect of Yaeyama, Okinawa prefecture], (Dialect recording material series 15), National Language Research Institute.
- (1983) 'Okinawa-ken Taketomi-chō-hōgen [The dialect of Taketomi district, Okinawa prefecture]', in Hirayama Teruo (ed.) *Zenkoku Hōgen Jiten* [Japanwide dialect dictionary] 1, Tokyo: Kadokawa, 297–301.
- (1984) 'Yaeyama-hōgen-gaisetsu' [Overview of the Yaeyama dialects], in Iitoyo Ki'ichi (ed.) *Kōza Hōgen-gaku 10: Okinawa/Amami-Chihō-no Hōgen* [Dialectology 10: The dialects of the Okinawa and Amami regions], Tokyo: Kokusho Kankōkai, 289–361.
- (1991) 'Hatoma-hōgen-no saishi-kankei-goi (1)' [The ritual-related vocabulary of the Hatoma dialect (1)], *Ryūkyū-no Hōgen*, 16: 56–104.
- (1998) 'Taketomi-hōgen-no kiso-goi: Bun'ya 3 Shokubutsu' [The basic vocabulary of the Taketomi dialect: Field 3 Vegetation], *Ryūkyū-no Hōgen*, 23: 95–107.
- (2002) 'Yaeyama Taketomi-chō-no kiki-gengo-no hyakunen' [100 years of the endangered languages of Taketomi district, Yaeyama], <http://www.let.osaka-u.ac.jp/~sbj/genchijouhou.htm>.
- Kamei Takashi, Kōno Rokurō and Chino Ei'ichi (eds) (1992) *Gengogaku Dai-jiten Dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido.
- Karimata Shigehisa (1992a) 'Ryūkyū-rettō-no gengo (Miyako-hōgen)' [The languages of the Ryukyu archipelago (Miyako dialects)], in Kamei Takashi, Kōno Rokurō and Chino Ei'ichi (eds) *Gengogaku Dai-jiten Dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido, 388–403.
- (1992b) 'Ryūkyū-rettō-no gengo (Yaeyama-hōgen)' [The languages of the Ryukyu archipelago (Yaeyama dialects)], in Kamei Takashi, Kōno Rokurō and Chino Ei'ichi (eds) *Gengogaku Dai-jiten Dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido, 403–13.
- Lawrence, Wayne P. (1997a) 'Hatoma-hōgen-no akusento: Meishi' [Hatoma dialect accentuation: Nouns], *Okinawa Bunka*, 85: 1–26.

- (1997b) ‘Hatoma-hōgen-no akusento: Sūshi/josūshi’ [Hatoma dialect accentuation: Numbers/classifiers], *Okinawa Geijutsu-no Kagaku*, 9: 1–21.
- (2001) ‘Tone and/or accent in Ryukyuan dialects,’ in Kaji Shigeki (ed.) *Proceedings of the Symposium Cross-linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, 195–204.
- Matsumori Akiko (2010) ‘Tarama-jima-no 3-kei-akusento-to “keiretsu-betsu goi”’ [The 3-class accentual system of Tarama island and a ‘classified vocabulary’], Uwano Zendō (ed.) *Nihongo-Kenkyū-no 12-shō* [12 Chapters of Japanese language research], Tokyo: Meiji shoin, 490–503.
- Miyagi Shin’yū (2003) *Ishigaki-Hōgen-Jiten* [Ishigaki dialect dictionary], Naha: Okinawa Times.
- Miyanaga Masamori (1982) [1931] *Yaeyama-Goi Kō-hen* [Yaeyama vocabulary Part 1], Tokyo: Dai’ichi Shobō.
- Miyara Shinshō (1995) *Minami-Ryūkyū Yaeyama Ishigaki-Hōgen-no Bunpō* [A grammar of the Ishigaki dialect of Yaeyama, Southern Ryukyu], Tokyo: Kurosio.
- Nakama Mitsunari (1991) ‘Miyako Nishihara-hōgen-no asupekuto’ [Aspect in the Nishihara dialect of Miyako], *Kokubungaku Kaishaku-to Kanshō*, 56: 1.86–92.
- (2000) *Ryūkyū-Hōgen-no Imiron* [The semantics of the Ryukyuan dialects], Tokyo: Rukku.
- Nippon Hōsō Kyōkai (ed.) (1972) *Zenkoku Hōgen Shiryō Dai-11-kan Ryūkyū-hen II* [Japanwide dialect materials, vol. 11, Ryukyu II], Tokyo: Nippon Hōsō Shuppan Kyōkai. [CD-ROM version published in 1999.]
- Ōno Makio (1989) ‘Ryūkyū Hateruma-hōgen-no josūshi: Sono keitai-to imi-kōzō’ [Hateruma dialect classifiers: Their morphological and semantic structure], *Ryūkyū-no Hōgen*, 14: 107–17.
- Shibatani Masayoshi (1990) *The Languages of Japan*, Cambridge: Cambridge University Press.
- Suzuki Shigeyuki *et al.* (2001) *Ryūkyū Yaeyama-Hōgen-no Dōshi-no Kenkyū: Ishigaki-Hōgen-no Dōshi-no Asupekuto-to Tensu* [Research into the verbs of the Yaeyama dialects: Aspect and tense in Ishigaki dialect verbs], Ministry of Education Research Fund (C)(2) research report. Project no. 11610438.
- Uchima Chokujin (1984) ‘Miyako-shotō-no hōgen’ [The dialects of the Miyako islands], in Iitoyo Kī’ichi (ed.) *Kōza Hōgen-gaku 10: Okinawa / Amami-Chihō-no Hōgen* [Dialectology 10: The dialects of the Okinawa and Amami regions], Tokyo: Kokusho Kankōkai, 251–87.
- (1997) ‘Ryūkyū-hōgen-ni okeru kanō-hyōgen’ [Potential expressions in the Ryukyu dialects], *Okinawagaku* [Okinawan Studies], 1: 59–76.
- Uesedo Tooru (1976) *Taketomi-Tōshi: Minwa / Minzoku-hen* [Chronicles of Taketomi island: Folklore edition], Tokyo: Hosei University Press.

## YONAGUNI\*

*Atsuko Izuyama*

## 14.1 INTRODUCTION

Yonaguni, in the Yaeyama islands in Okinawa prefecture, is a small island (28.88 sq. km) at the westernmost point in Japan. Located at 24° 27' N, 123° 00' E, Yonaguni is 111 km from Taiwan and 509 km from Okinawa island. The total population of the island is 1,682. The main population centers are: Sonai (*ubu-mura*), where the town office is located, on the north coast; Kubura (*kubura*), a fishing port on the west coast; and Higawa/Hinai (*ndi*) in the south. The differences in dialect among these three villages appears small. Almost all data in this chapter was obtained from Sonai dialect speakers.

Population as of April 2007, from the Yonaguni-chō homepage ([www.town.yonaguni.okinawa.jp/](http://www.town.yonaguni.okinawa.jp/)):

<i>By village</i>		<i>By age</i>	
Sonai	1,040	Age 0–14	19.5%
Kubura	509	Age 15–24	6.7%
Higawa	133	Age 25–64	52.6%
		Age ≥65	21.2%

There are about 340 residents aged 65 or over. However, it is said that a number of these people, particularly in Kubura, have moved to the island from other places, so it is probable that less than half of these people can speak the dialect. The majority (57.9 percent) of working islanders are in tertiary industries, with 14.1 percent in primary industry and 28 percent in secondary industry.

Due to the lack of a high school on Yonaguni, young boys and girls leave the island to go to high school and later to college. For them it is difficult to come back to their home town, so the number of resident dialect speakers is continually decreasing.

While Yonaguni belongs to Yaeyama-gun (district) governmentally, people on the main Yaeyama islands (e.g. Ishigaki) and people on Yonaguni cannot mutually understand each other's local languages. Yonagunians are somewhat familiar with the Okinawan (Shuri, Naha) dialect because the Okinawa local theater used to come to Yonaguni for their entertainment.

The history of Yonaguni is not well known. In 1477, fishermen from the Korean peninsula drifted to an island, which was probably Yonaguni, and described it. A local legend tells of a woman ruler named *sanaï-isoba*, who lived around 1500. In historical documents, it is described that Nakasone Tomigusuku came to Yonaguni from Miyako to defeat the ruler 'Onitora'. A famous folk tale about the beautiful girl that was taken away to Miyako is still talked about on both Yonaguni and Miyako.

Before 1945, large passenger ships sailed regularly between Japan and Taiwan via Yonaguni; during that time Yonaguni people felt closer to both Taiwan and mainland Japan than Yaeyama people did.

Nowadays television has a great influence on local life. Yonaguni people use mostly Standard Japanese (ModJ). Since the introduction of Japanese, people have had personal names such as *Kazu*, *Toshi*, *Miyo*, *Nae*, etc. which contain sounds that the traditional Yonaguni dialect does not contain. People over about 65 years old, however, still have a *cima na* (lit. ‘island name = local name’). Now we can easily go to Yonaguni by airplane and by ship, from Naha in mainland Okinawa or from Ishigaki in Yaeyama. Today Yonaguni is better known for sightseeing and diving than for its peculiar dialect. Yonaguni dialect has been recognized as a ‘severely endangered’ language by UNESCO (Moseley 2010).

## 14.2 PHONOLOGY

Yonaguni phonology is important comparatively for three things: (1.) Its three-vowel system. (2.) Where Japanese has /w/ and /j/, Yonaguni has *b* and *d* (e.g. *dama* : OJ *yama* ‘mountain’, *bur-* : OJ *wor-* ‘break’). (3.) Its glottalized consonants, largely derived from clusters, in turn deriving from the loss of a vowel that in other varieties is devoiced, e.g. *\*sita* > *\*sta* > *t’a* ‘tongue’ (= OJ *sita*). (cf. the similar process of Middle Korean clusters becoming Modern Korean emphatics, Chapters 4 and 5.)

### 14.2.1 Vowels

Yonaguni is famous for its three-vowel system, that is, /a, i, u/. /i/ is realized as [i ~ e.] and /u/ as [u ~ o.]. The vowel system forms a classic triangle:

i	u
a	

Long vowels are observed phonetically, but they do not contrast phonemically with short vowels. In some cases, we shall use double notation, e.g. *aa*, for long vowels where appropriate. Vowel length seems to depend on emphasis or accentuation. Shibata (1972: 54) observes that, “All words end in long vowels except when the final sound is /N/, or /i/ as the final element in a diphthong. In two-syllable words, the vowel in the first syllable is likewise long, with the same exceptions as above, as well as cases where it is followed by a glottalized consonant.”

[e] and [o] appear exceptionally in some sentence-final particles, e.g. [do:] [je:] [jo:]. These are always long, and are generally emphasized.

/i/ and /u/ have several allophones. After [s], labial rounding of /u/ is scarcely observed, e.g. *usubi* [usubi] ‘Lie down!’ vs. *ubu munu* [ubumunu] ‘a large thing’. /i/ in word-final (especially sentence-final) position tends to open and approach [e]. /u/ in word-final position, e.g. *ita hiruu* ‘Let’s go’, has the same tendency and approaches [o].

/i/ and /j/ palatalize a preceding dorsal consonant, just as in ModJ: /si/ → [ei], /sja/ → [ea], /sju/ → [eu], /ci/ → [tʰi], but /ca/ → [tsa], /cu/ → [tsu].

Semivowels are /j/ and /w/. /j/ is followed by the vowels [a, u, ee, oo], e.g. [ja, ju, jee, joo]. /w/ can occur between /p, b, t<sup>h</sup>, d, m, n, k<sup>h</sup>, kʰ, g, h, s/ and /a/.

### 14.2.2 Consonants

Yonaguni consonants are presented in [Table 14.1](#). The dialect has a phonological contrast between glottalized (non-aspirated) voiceless stops and nonglottalized (aspirated) ones word-initially (1b). Word-medially this contrast is not observed, and almost all voiceless

**TABLE 14.1 YONAGUNI CONSONANT INVENTORY**

	Labial	Dental/ Alveolar	Palatal	Velar	Glottal
Plosive	b	d		g	
Glottalized	pʷ	tʷ		kʷ	
Aspirated		tʰ		kʰ	
Affricate					
Glottalized		tʃʷ	[tɕʷ]		
Fricative	ɸ	s (z)	[ɕ] ([z])		h
Nasal	m	n		ŋ	
Tap		r			
Approximant	w		j		
Moraic Consonant:					N

stops and affricates are glottalized with some exceptions (compound words). The voiceless labial stops are always glottalized (1a).

- (1) a. [pʷaŋai]  
hoe
- b. [tʰa] [tʰa] [tʰi] [tʷi] [tʰu] [tʷu]  
who tongue hand moon ten person
- [tʰa] [tʰa] [tʰi] [tʷi] [tʰu] [tʷu]  
grass blood wear
- [kʰai] [kʷaŋ] [kʰi] [kʷi] [kʰu] [kʷu]  
buy-IMP use-PERF tree pickle-IND this lungs

We do not need to assign Q phonemic status. Immediately before a glottalized consonant we hear a short moment that retains the duration of a stop, accompanied by glottalization. However, we often notice gemination at the beginning of the second syllable, which might be related to accentuation.<sup>1</sup> As Shibata (1959) points out, in Yonaguni /N/ is syllabic but [Q] is not syllabic. Whereas in ModJ /N/ generally nasalizes following vowels, in Yonaguni the vowels retain their orality. We sometimes could hear a light [ŋ] as a glide before vowels. It seems that speakers keep uvular closure and pronounce [m] before labial stops and [n] before dental stops. It is difficult to tell the difference between -Nŋ and -ŋ, especially in the second syllable, e.g. [pʷaŋai] (Shibata 1972) or [pʷaŋŋai] (N. Ikema 1998). This chapter will choose the former. This phenomenon is probably related to the case of [Q] above.

The voiceless affricate does not seem to have a contrast in glottalization. There is slight aspiration in the [s]-release, but it is normally glottalized. The symbol /c/ represents the allophones [tɕʷ] before /j, i/ and [tɕʷ] elsewhere.

[i] and [j] palatalize preceding consonants, so /si/, /sja/ and /zja/ represent [ei], [ea] and [za].

Some scholars, for example Hirayama and Nakamoto (1964, 20–3), and Kajiku (2002, 268), describe speakers of Yonaguni dialect as having [çi] as realization of /hi/. As far as I can tell, however, speakers do not have [ç]. /t/ (= [ɾ]) does not occur word-initially and sounds similar to [d] intervocally. /g/ in word-initial position is very rare. /z/ is found in a very limited number of words, as noted by Shibata (1959). Two such words are *zjahi* [zahi] (literally ‘snake skin’), referring to the shamisen, a stringed musical

instrument, and *zjaku* [zaku], a small fish eaten by bonitos. In former times, there was a Yonaguni shamisen made with special paper, but later they imported the more expensive shamisen made with snake skin from Okinawa.

### 14.2.3 Accent

According to Uemura (2003), there are four contour patterns for word accent. (Shibata 1959 earlier identified just three.) The following is adapted from Uemura (2003, 34–41):

(2) Low level	Falling	Rise from a mid tone	High level
<i>hana</i> ‘flower’	<i>hatu</i> ‘bird’	<i>kadi</i> ‘wind’	<i>ki</i> ‘hair’
<i>ami</i> ‘rain’	<i>tidan</i> ‘sun’	<i>huci</i> ‘star’	<i>ci</i> ‘blood’
<i>cima</i> ‘island’	<i>nni</i> ‘boat’	<i>unti</i> ‘sweet potato’	<i>t’u</i> ‘person’
<i>tii</i> ‘hand’	<i>cii</i> ‘blood’	<i>nda</i> ‘you’	
<i>anu</i> ‘I’	<i>tun</i> ‘wife’	<i>utu</i> ‘sound’	
<i>t’uci</i> ‘one’	<i>nai</i> ‘now’	<i>t’aci</i> ‘two’	
	<i>maci</i> ‘pine tree’	<i>miN</i> ‘water’	

Those classified as ‘high level’ above begin with voiceless consonants. This might be influencing the word accent contour.

### 14.2.4 Transcription

We use phonemic transcription; however, for practicality we have simplified it in some ways, while complicating it in others. Firstly, word-initially, i.e. where voiceless stops have a glottalized-nonglottalized contrast, we mark the glottalized ones. Word-medially, where they do not contrast (all of them are phonetically glottalized), we do not mark glottalization. For example, an unmarked *k* indicates a different sound depending on its position: [k<sup>h</sup>] word-initially and [k] word-medially. Secondly, though vowel length is not contrastive, in some cases (for discourse purposes, etc.) long vowels are indicated with double signs, e.g. *aa*.

(3) Examples:

<i>a</i> ‘millet’	/p/	<i>pa</i> ‘sprout’	/b/	<i>ba</i> ‘particle’	/t/	<i>t’a</i> ‘tongue’	/t <sup>h</sup> /	<i>ta</i> ‘who’	/d/	<i>da</i> ‘parlour’
<i>i</i> ‘cooked rice’		<i>pirun</i> ‘pushed’		<i>bi</i> ‘plant’		<i>t’i</i> ‘mouth’		<i>ti</i> ‘hand’		<i>di</i> ‘letter’
<i>u</i> ‘it’		<i>pui</i> ‘suck’		<i>bu</i> ‘string’		<i>t’u</i> ‘person’		<i>tu</i> ‘ten’		<i>du</i> ‘body’
<i>k’a</i> ‘priest’	/k <sup>h</sup> /	<i>ka</i> ‘well’	/g/	<i>gansu</i> ‘mortuary tablet’	/s/	<i>sa</i> ‘tea’	/h/	<i>ha</i> ‘leaf’		
<i>k’i</i> ‘pickle’		<i>ki</i> ‘tree’		<i>gici</i> ‘pickles’		<i>si</i> ‘themselves’		<i>hi</i> ‘wind’		
<i>k’u</i> ‘lungs’		<i>ku</i> ‘this’		<i>gui</i> ‘stake’		<i>su</i> ‘tide’		<i>hu</i> ‘luck’		
<i>ca</i> ‘grass’	/ts/	<i>ma</i> ‘already’	/m/	<i>na</i> ‘name’	/n/	<i>ηasa</i> <sup>2</sup> ‘deep’	/ŋ/	<i>karan</i> ‘light’	/r/	
<i>ci</i> ‘blood’		<i>mi</i> ‘eye’		<i>ni</i> ‘root’		<i>anji</i> ‘lift’		<i>kari</i> ‘he’		
<i>cu</i> ‘white’		<i>mui</i> ‘hill’		<i>nu</i> ‘what’		<i>anun</i> ‘lift’		<i>turun</i> ‘take’		
<i>an</i> ‘be’	/N/	<i>nai</i> ‘rice plant’		<i>nba</i> ‘urine’		<i>nkadi</i> ‘centipede’		<i>nta</i> ‘soil’		
<i>inna</i> ‘don’t do’		<i>nma</i> ‘horse’		<i>nda</i> ‘you’		<i>ngi</i> ‘beard’		<i>nn</i> ‘yes’		
<i>bun</i> ‘be’		<i>npa</i> ‘no’		<i>nnu</i> ‘yesterday’		<i>niŋ-</i> no examples				

/Cwa/ (verb examples are cited in their perfect forms)

*pwa* *bwan* *twan* *dwai* *k'wan* *kwan* *gwan* *swan* *hwan* *mwan* *nwan*  
 'sour' 'break' 'take' 'festivity' 'make' 'cross' 'send' 'bitter' 'rain' 'burn' 'ride'

/j/ *ja* PART *kati ja?* 'Did you write?'

*ju* PART *kati juu* 'I wrote.'

*joo* PART *kagi joo* 'Write!'

*jee* PART *jee, isi nai?* 'Oh! Is that so?'

/bj/ *bjAN* Perfect form of *bi-* 'plant'

*bjuka* 'nephew'

/pj/ *pjaakiN* 'leprosy'

/tʃj/ *tʃAN* Perfect form of 'hear'

(/tʰj/) *tjuu* *anu-nki tjuu ~ anu-nki ti uu* 'Give it to me.'

/dj/ *djadi* 'serious matter'

/sj/ *sjana-* 'happy'

(/zj/) *zjahi* 'musical instrument')

/kʲj/ *kʲAN* Perfect form of *kʲi-* 'pickle'

/kʰj/ *kʰAN* Perfect form of *ki-* 'do'

/gj/ *gjagisi* 'obstruct'

/mj/ *mjaaku* 'rubbish'

/nj/ *njan* Perfect form of *ni-* 'cook'

*njun* Perfect form of *ni-* 'cook'

/hj/ *hjaagu* '100'

*hjUN* Perfect form of *hi-* 'go'

### 14.3 TRADITIONAL WRITING

The Yaeyama islands have never had a full-featured writing system – few attempts have even been made at recording the dialects with written characters. Even today, their transmission is almost exclusively oral.

A simple method of recording family names, items, and numerals was devised by the islanders in order to record the burdensome taxes imposed on them by mainland authorities. This system, called *kaida* writing, was used until the introduction of Japanese-language education near the end of the nineteenth century, when it began to decline.

*Kaida* writing focused on personal names (expressed with arbitrary, non-semantic symbols called *dahan* 'house markings'), physical items such as animals and foodstuffs, and counters for units of volume and weight, which eventually evolved into numerals. No written characters ever evolved to express the rich system of verbs found in Yonaguni – verbs are only implied in *kaida* writing, and are limited to simple actions such as 'send', 'contribute', etc. Its non-phonographic nature meant that essentially the same system could be intelligible on both Yonaguni and Taketomi despite their significant phonetic differences.

Attempts were made to extend this writing system to more complicated expressions, but ultimately the superior expressibility of the newly mandatory Japanese language left this simpler script at a disadvantage. As of (2007) only a few older residents could remember its active use, and fewer still were capable of writing with it.

See the Appendix for a list of *kaida* characters known to have been in use. See also N. Ikema (1998, 2003), Rosa (2006), Shibata (1959, 1972), Uemura (2003).

## 14.4 MORPHOLOGY

### 14.4.1 Nouns

Noun suffixes are diminutive *-ti* and plural *-nta* (4a). The plural means ‘and the like’, and can be attached to persons, animals, or inanimates (4b).

- (4) a. *agami agami-ti agami-ti-nta*  
 ‘child’ ‘small child’ ‘small children’
- b. *ija-nta abuta-nta asa-nta duci-nta*  
 ‘fathers’ ‘mothers’ ‘grandfathers’ ‘friends’  
*maju-nta inu-nta hatu-nta*  
 ‘cats’ ‘dogs’ ‘birds’  
*acuko-nta*  
 ‘Atsuko and others’  
*unti-nta* ‘potatoes and the like’ *ubuni-nta* ‘Japanese *daikon* radishes’  
*dagu-nta* ‘(many kinds of) insects’ *iju-nta* ‘(many kinds of) fish’  
*da-nta* ‘houses’ *icibugu-nta* ‘stones’  
*dama-nta* ‘mountains’ *nnani-nta* ‘traditional kimono and the like’

#### 14.4.1.1 Case

Subjects are marked by either zero or *-ŋa*. It seems that intransitive verbs are acceptable without a particle for the nominative case when the speaker judges the action (5, 6a, 7, 8, 9b). In reporting purely objective facts, however, it appears necessary (6b, 9a). In general, if the subject is modified in any sense, *-ŋa* is likely to be added to that noun (10, 11). Where the object is topicalized, *-ŋa* appears obligatory for the subject as a rule (12a), even if the topic is ellipted (12b). In a subordinate clause, *-ŋa* is obligatory (13). As any case particle, *-ŋa* may co-occur with contrastive *-ja* or additive *-N* (14, 15).

- (5) *madagaŋti ami-Ø hur-un do*  
 soon rain rain-CONC PART  
 ‘It will rain soon.’
- (6) a. *su-ja kadi-Ø k'-un di do*  
 today-TOP wind blow-CONC QUOT PART  
 ‘It is said that it will be windy today.’
- cf. b. *suu-ja nan-ŋa taga-N sujaa*  
 today-TOP wave-NOM high PART  
 ‘The waves look high today.’
- (7) Waiting for the birth of a baby:  
*agami-Ø mar-un!*  
 child born-PERF  
 ‘A child has been born!’

- (8) People waiting to see the night-blooming cereus (which open rapidly enough to be visible) blooming late at night:  
*hana-Ø sag-un!*  
 flower bloom-CONC  
 ‘A flower is blooming!’
- (9) a. *su magu-ŋa s-utan* (Grandma is reporting this fact to someone.)  
 today grandchild-NOM come-PAST  
 ‘My grandchild came today.’  
 b. *aa magu-Ø ku-n!* (Grandma sees her grandchild coming.)  
 ah! grandchild come-CONC  
 ‘Ah! My grandchild is coming!’
- (10) *sidasa-ru kadi-ŋa k'-un suja*  
 cool-CONC wind-NOM blow-CONC PART  
 ‘There’s a cool wind blowing.’
- (11) *ndja-nu agami-ŋa tinamai k-i*  
 your.house-GEN child-NOM mischief do-IND  
 ‘Your child did mischief!’
- (12) a. *u-ja t'u-ŋa k-an do*  
 it-TOP person-NOM buy-PERF PART  
 ‘That, another person has (already) bought.’  
 b. They are looking at something.  
*(u-ja) abu-ŋa araagu atara-ki*  
 (that-TOP) grandma-NOM very.much treasure-IND  
 ‘(That is [something]) grandmother treasures/has treasured.’
- (13) a. *maisa-ru icibugu-ŋa a-i-bi h-iraninu-tan*  
 big-CONC stone-NOM be-IND-due.to go-POT.NEG-PAST  
 ‘There was a very big stone (on the road), so I could not go.’  
 b. *ndi-ŋa kidita-ba agami ug-i*  
 you (pl.)-NOM make.noise-because child get.up-IND  
 ‘You made a noise, so the child woke up.’
- (14) a. Q. *duda bur-un?* A. *anu-ja bur-anun*  
 branch break-CONC I-TOP break-NEG  
 ‘Was the branch broken?’ ‘I didn’t break it.’  
 b. *\*aŋa-ja bur-anun*  
 I.NOM-TOP break-NEG
- vs. c. *kari-ŋa-ja b-utan-tin a-ŋa-ja bur-anun*  
 he-NOM-TOP break-PAST-even.if I-NOM-TOP break-NEG  
 ‘Even if he had broken it, I would not break it.’
- (15) Two people are talking on the banks of the stream.  
*a-ŋa-n tub-jar-ja nda-n tub-i*  
 I-NOM-even jump-PERF-ja you-also jump-IMP  
 ‘Even I myself have jumped, so you jump too!’

Yonaguni dialect does not have any marker for the accusative case. However, as any other case, objects can be accompanied by topic *-ja*, selective *-ba*, or focus *-du*. The genitive may be marked variously:

- (16) a. zero: Names (*-nu* can be added instead), forms of address, 2nd and 1st plural pronouns:  
*abu-Ø ~ asa-Ø ~ ija-Ø ~ sinsi-Ø ~ nda-Ø ~ baa-Ø munu*  
 Grandma's ~ Grandpa's ~ father's ~ teacher's ~ your ~ our thing
- b. *-ja*: Other pronouns:  
*kari-ja ~ u-ja ~ ku-ja munu*  
 that person's ~ his ~ this person's thing
- c. *-nu*: Other nouns:  
*unu agamiti-nu ~ unu t'u-nu ~ acuko-nu munu*  
 that child's ~ that person's ~ Atsuko's thing

The dative is marked with *-nki*, which is also used for direction, the complement of 'become', the causee of causative structures, and the agent in passive sentences.

- (17) *nma-nki h-juN ga*  
 where-DAT go-PERF Q  
 'Where has he (/she/they) gone?'
- (18) *agami-nki kag-am-iruN*  
 child-DAT write-CAUS-CONC  
 'I let the child write.'
- (19) *kama-nu agami-ja sinsi-nki n-an tuna*  
 there-GEN child-TOP teacher-DAT become-PERF QUOT  
 'It is said that the child of that house has become a teacher.'
- (20) Speaking to a sick person, referring to having a nurse help a patient stand up.  
*(nda) kari-nki tat-am-iru naa*  
 (you) that. person-DAT stand-CAUS-CONC Q  
 'I will get him to help you stand up, OK?'
- (21) *kari-nki ut-ar-i*  
 he-DAT hit-PASS-IND  
 'I was hit by him.'
- (22) a. An artificial causative sentence:  
*acuko-ja kazu-nki tosiko aig-amj-an*  
 Atsuko-TOP Kazu-DAT Toshiko walk-CAUS-PERF  
 'Atsuko ordered Kazu to let Toshiko walk.'
- b. Yonaguni people will say in practice:  
*acuko-ja kazu-nki tosiko aiga-si tura-i ndi*  
 Atsuko-TOP Kazu-DAT Toshiko let.walk-IND give-IMP QUOT  
*nd-jaN*  
 say-PERF  
 'Atsuko said to Kazu, "Let Toshiko walk."'

Other case particles are instrumental *-si* (23–27), ablative *-gara* (148), terminative *-kuta* (113), and locative *-ni* (28–30). Demonstratives of place have alternative forms for the locative, *kuma* ‘here’ > *kuma-ni* ~ *kumi*, etc.

- (23) Q. *nuu-si* *baga-N* *ga* A. *dagan-si* *baga-si* *tura-i*  
 what-INST boil-CONC Q kettle-INST boil-IND give-IMP  
 ‘What do you boil [water] with?’ ‘Please boil it with a kettle.’
- (24) *kumu* *duda-ja* *kadi-si-du* *bur-i* *juu*  
 this branch-TOP wind-INST-FOC break (intr.)-IND PART  
 ‘This branch has been broken by the wind.’
- (25) *nda* *t’ui-si* *ii* *ha-i* *joo*  
 you alone-INST meal eat-IMP PART  
 ‘Have a meal by yourself.’
- (26) *nda-si/ŋa* *aig-am-iri* \**nda-nki* does not occur.  
 you-INST/NOM walk-CAUS-IMP  
 ‘You, please let me walk.’
- (27) *kazu-nki/si* *nd-am-irun*  
 Kazu-DAT/INST say-CAUS-CONC  
 ‘I will let Kazu say it.’
- (28) a. *acuko-ja* *nmi* *bu-N* *ga*  
 Atsuko-TOP where.LOC be-CONC Q  
 b. *acuko-ja* *nma-ni* *bu-N* *ga*  
 Atsuko-TOP where-LOC be-CONC Q  
 ‘Where is Atsuko?’
- (29) Looking for Atsuko:  
*nmi-ba* *bu-N* *ga*  
 where.LOC-SEL be-CONC Q  
 ‘Where is Atsuko?’
- (30) *duci-nu* *da-ni-du* *bu-ru*  
 friend-GEN house-LOC-FOC be-CONC  
 ‘She is in her friend’s house.’

#### 14.4.1.2 *-ja, -ba*

The topic marker is *-ja*, and cannot be placed after an interrogative. *-ja* carries a nuance of selection or specification rather than simply marking the topic. In negative sentences, the actor of the verb is usually marked by *-ja*.

- (31) a. *anu-ja* *acuko* (*do*) b. *anu-ja* *kag-anun*  
 I-TOP Atsuko PART I-TOP write-NEG  
 ‘I am Atsuko.’ ‘I don’t write.’
- (32) It seems that there was some traffic accident:  
*nuu* *k-i* *ndi* *ta-ba*  
 what do-IND QUOT who-SEL  
 ‘What happened? Who was it?’

- (33) *tagu tuba t'u-ja taa ja*  
kite fly.CONC person-TOP who Q  
'Who is the one flying a kite?'

*-ba* is used usually with interrogative words, giving a nuance of selection or contrast and specification. However, it can also occur with nouns in the nominative case, and give some modal connotations (14.5.4). Yonaguni does not allow the sequences with focal *-du* \**ici-ba-du* 'when' or \**nma-nki-ba-du* 'where to', even though it does allow *ta-ba-du* 'who' and *nu-ba-du* 'what' (36–37).

- (34) Q. *nda nuu ki-ba muja-N ga*  
you what wood-SEL burn-CONC Q  
'Which wood do you burn?'

- (35) A. *kanu ki-du muja*  
that wood-FOC burn.CONC  
'I burn that wood.'

- (36) Q. *ta-ba-du tata-N ga*  
who-SEL-FOC make.stand-CONC Q  
'Whom do you make stand up?'

- A. *acuko-du tata*  
Atsuko-FOC make.stand.CONC  
'I make Atsuko stand.'

- (37) Q. *nu-ba-du ut-iru kaja*  
what-SEL-FOC fall-CONC DOUBT  
'I wonder which one will fall down.'

- A. *kari-ŋa-du ut-iru*  
that-NOM-FOC fall-CONC  
'That one will fall down.'

- (38) *dubi-ja nmi-ba cii a-tan ga*  
last.night-TOP where.LOC-SEL fire be-PAST Q  
'Where was the fire last night?'

- (39) *nda ici-ba s-un ga*  
you when-SEL come-PERF Q  
'When did you come?'

- (40) *ici-du i-run ga* (about the fixed date of a meeting or party, etc.)  
when-FOC do-CONC Q  
'When are you going to do it?' (*ici irun ga* is also acceptable)

## 14.4.2 Verbs and adjectives

### 14.4.2.1 Verbs

Japanologists specializing in Ryukyuan languages normally present verb conjugation using almost the same categories as in traditional Japanese grammar. In Yonaguni, the 'volitional' form's function has a slight, but important, difference from its operation in Japanese. In Japanese, a volitional such as **shiyō**, **kakō** can express a single person's desire to perform some action, whereas in Yonaguni the form is used only to express an invitation to others in the speaker's group (*banta* 'we') to perform the action together.

It therefore functions more like English ‘let’s’ than its Japanese cognate. The last vowel of the volitional form is accented and always long. Its phonetic value is a little bit different from the last vowel of the conclusive form: [ʷo.]. Some speakers do not use the volitional form, but instead use an interrogative form, that is, *kiru na?*, *ugiru na?*, *kagu na?*, etc., or a form in *-Ndangi*, e.g. *hi-Ndangi* ‘let’s go’.

Table 14.2 is a different arrangement proposed by the author. The conjugational type of *i-r-* is irregular, *ki-r-* and *ug-i-r-* correspond to ModJ monograde, and *kagu-* corresponds to ModJ quinquagrade. The fifth, *ugu-*, corresponds to ModJ *s*-stem quinquagrade; the *s* is lost in most forms. The last one (*sama-*) does not directly correspond to Japanese, but it is also an important regular type of conjugation. There are many other irregular types.

Some verbs that correspond to ModJ quinquagrade have two stems (*-g-/t-*, *-ŋ-/d-*), the second used for the indefinite form and derivatives.

Finite and attributive forms end in *-i*, *-a*, *-u* with or without *-N*. Table 14.2 treats volitional as the same form as the non-past without *-N*.

Stative stems cannot be used independently, but combine with negative (*-nu-*), causative (*-miru-*), and passive/potential (*-riru-*) morphemes.

The verb has two negative forms: *kag-anuN* (stative stem + *-nuN*) and *kat-i minuN* (indefinite + *minuN*). The former indicates the speaker’s subjective statement, and the latter is the speaker’s objective statement, e.g. *kaganuN* ‘I will not write’, *madi kati minuN* ‘I have not yet written’. *minuN* is an independent word meaning ‘there is not’, and *kati minuN* has two meanings: one is the literal ‘have not yet done’, while the other expresses the speaker’s disappointment or regret at something having been

TABLE 14.2 YONAGUNI VERB MORPHOLOGY

	Irreg.	<i>ir-u</i> -Type	<i>u</i> -Type	Derived <i>u</i> -Type	<i>a</i> -Stem	
	‘do’ (intr.)	‘do’ (tr.)	‘get up’	‘write’	‘wake’ (tr.)	‘put [fire] out’
Non-Past	<i>irun</i> <i>iru</i>	<i>kirun</i> <i>kiru</i>	<i>ugirun</i> <i>ugiru</i>	<i>kagun</i> <i>kagu</i>	<i>ugun</i> <i>ugu</i>	<i>saman</i> <i>sama</i>
Imperative	<i>iri</i>	<i>kiri</i>	<i>ugiri</i>	<i>kagi</i>	<i>ugui</i>	<i>samai</i>
Prohibitive	<i>inna</i>	<i>kinna</i>	<i>uginna</i>	<i>kagunna</i>	<i>uginna</i>	<i>samanna</i>
Stative stem	<i>ira-</i>	<i>kira-</i>	<i>ugira-</i>	<i>kaga-</i>	<i>uga-</i>	<i>sama-</i>
Indefinite	<i>isi</i>	<i>ki</i>	<i>ugi</i>	<i>kati</i>	<i>ugusi</i>	<i>samasi</i>
Conjunctive	<i>isiti</i>	<i>kiti</i>	<i>ugiti</i>	<i>katiti</i>	<i>ugusiti</i>	<i>samasiti</i>
Perfect	<i>isjan</i> <i>isjaru</i>	<i>kjan</i> <i>kjaru</i>	<i>ugun</i> <i>uguru</i>	<i>katjan</i> <i>katjaru</i>	<i>ugusjan</i> <i>ugusjaru</i>	<i>samasjan</i> <i>samasjaru</i>
Past: Conclusive	<i>itan</i> <i>itaru</i>	<i>kitan</i> <i>kitaru</i>	<i>ugitan</i> <i>ugitaru</i>	<i>katitan</i> <i>katitaru</i>	<i>ugusitan</i> <i>ugusitaru</i>	<i>samasitan</i> <i>samasitaru</i>
Perfect	<i>isjatan</i> <i>isjataru</i>	<i>kjatan</i> <i>kjataru</i>	* <i>ugutan</i> * <i>ugutaru</i>	<i>katjatan</i> <i>katjataru</i>	<i>ugusjatan</i> <i>ugusjataru</i>	<i>samasjatan</i> <i>samasjataru</i>
<i>Subordinate:</i>						
CONC + <i>-ba</i>	<i>iruba</i>	<i>kiruba</i>	<i>ugiruba</i>	<i>kaguba</i>	<i>ugu(i)ba</i>	<i>sama(i)ba</i>
CONC + <i>-ja</i>	<i>irja</i>	<i>kirja</i>	<i>ugirja</i>	<i>kaja</i>	–	–
Past + <i>-ja</i>	<i>itaja</i>	<i>kitaja</i>	<i>ugitaja</i>	<i>katitaja</i>	<i>ugutaja</i>	<i>samataja</i>
Perfect + <i>-ja</i>	<i>isjarja</i>	<i>kjarja</i>	<i>ugurja</i>	<i>katjarja</i>	<i>ugusjarja</i>	<i>samasjarja</i>
Perfect + <i>-ba</i>	<i>isjaruba</i>	<i>kjaruba</i>	–	<i>katjaruba</i>	<i>ugusjaruba</i>	<i>samasjaruba</i>
Perfect-Past + <i>-ja</i>	<i>isjataja</i>	<i>kjataja</i>	–	<i>katjataja</i>	<i>ugusjataja</i>	<i>samasjataja</i>

unexpectedly done, and carries the connotation '[I/he] shouldn't have done that!' The negative form of the potential of the verb (*kag-arirun*) is the stative stem + *-ninun*.

The indefinite form is formally cognate with the ModJ *ren'yōkei* (termed 'infinitive' form in other chapters). It has both dependent and independent uses in sentences. Speakers often treat this form as the citation form, as reflected in Ikema's (1998) dictionary.

The subordinate forms seem to contrast between *-ba* and *-ja*. *a*-stem and derived *u*-stem verbs do not have the conditional B form. Rather, this is expressed using the conditional A form or the past form + *-ja* (*sama-ta-ja*, *ugu-ta-ja*). The conditional B of derived *u*-stem verbs has two forms dependent on who is the actor. When the actor is the speaker, *-ba* is added to the non-past. Otherwise, *-ba* is added to the non-past + *i*. *a*-stem verbs have the same tendency but it happens less:

- (41) a. *a-ŋa ugu-ba* . . . 'When I awaken [someone] . . .'  
 b. *kari-ŋa ugu-i-ba* . . . 'When he awakens [someone] . . .'
- (42) a. *a-ŋa sama-ba* . . . 'When I put it out . . .'  
 b. *kari-ŋa sama-i-ba* . . . 'When he put it out . . .'

A considerable number of regular type verbs have two forms for the perfect. Most such verbs correspond to the CJ bigrade, and most pairs contrast between *-a* and *-u*.

(43)	'open'	'raise'	'get over'	'burn'	'snap'	'cut'
(CJ bigrade:	<i>ake-</i>	<i>age-</i>	<i>koye-</i>	<i>moye-</i>	<i>ore-</i>	
Non-past	<i>ag-irun</i>	<i>aŋ-irun</i>	<i>ku-irun</i>	<i>mu-irun</i>	<i>bur-irun</i>	<i>cic-irun</i>
Perfect	<i>ag-jaN</i>	<i>aŋ-jaN</i>	<i>kw-aN</i>	<i>mw-aN</i>	<i>bur-un</i>	<i>cic-jaN</i>
	<i>ag-un</i>	<i>aŋ-un</i>	<i>k-un</i>	<i>m-un</i>	<i>b-un</i>	<i>cic-un</i>
	'ride'	'cook'	'die' (honorific)			
(CJ:	<i>nor-</i>	<i>ni-</i>	monograde)			
Non-past	<i>nur-un</i>	<i>n-irun</i>	<i>ma-i-run</i>			
Perfect	<i>nw-aN</i>	<i>n-jaN</i>	<i>ma-i-sjaN</i>			
	<i>n-un</i>	<i>n-jun</i>	<i>ma-i-sun</i>			

The form in *-a* is a conventional 'perfect' form and expresses a situation after the action is completed. If the actor is the first person singular (*aŋa/anu*) this perfect form is the most natural. The form in *-u* expresses the speaker's judgment on the action as completed at that moment. It is not used in statements with an agent + *-ŋa*, but used mainly when the speaker discovers a given situation. The same restriction applies to (mostly intransitive) verbs which only have the perfect in *-u*:

- (44) a. *nai utun* 'A fruit has fallen.' \**nai-ŋa utun*  
 b. *acuko ugun* 'Atsuko has got up.' \**acuko-ŋa ugun*

The distinctive feature of this *u*-perfect is that the speaker recognizes the action at that moment. It focuses on the end of the whole action in a moment. The speaker sees the action itself and judges that it has finished.

The (*j*)*a*-perfect, on the other hand, cannot generally occur with intransitive verbs (especially human acts) such as \**tatjan* 'stand up', \**natjan* 'cry', \**aitjan* 'walk'. Unlike with transitive verbs, such acts are performed by human beings without leaving traces of what they have done, so it is difficult to imagine such a situation. For example, when a child stands or walks for the first time, non-perfect forms '*tatun!*', '*aigun!*' can be uttered. More examples:

- (45) a. Despite the door being difficult to open, A was able to open it.  
 A: *dadu ag-jan do* B: *ag-un?* C: *ag-u na?*  
 door open-A.PERF PART open-U.PERF open-U.PERF Q  
 ‘I have opened the door.’ ‘The door opened?’ ‘Did it open?’
- b. Looking at A, persons B and C can also comment:  
 B: *ag-un!* C: *ag-jan!*  
 open-U.PERF open-A.PERF  
 ‘(He) opened it. ~ It opened!’ ‘He has opened it!’
- (46) A mother and her child Toshi were going to get on a ship. Missing him on the wharf, the mother cried: Toshi! Toshi! Her friend, who saw Toshi getting on the ship, tells her:  
 Mother’s friend: *tosi nmi-nki n-un*  
 Toshi ship-DAT board-U.PERF  
 ‘Toshi has boarded the ship.’  
 Toshi (on deck): *anu nw-an \*anu n-un*  
 I board-A.PERF  
 ‘I’m on board.’

#### 14.4.2.2 Copula

Most Ryukyuan varieties have a distinct copula to the existential verb ‘exist, be (located); have’, but in Yonaguni the verb *a-* ‘be’ fills both these roles. It has an irregular conjugational pattern (Table 14.3).

It appears as a predicate copula only when focus is being placed on the complement (47b vs. 47a). *aranun*, the negative form of *an*, is not used to mean ‘does not exist’ – the word *minun* is used for that – rather, it is used as a negative copula.

- (47) a. *u-ja sagi Ø*  
 it-TOP saké  
 ‘It is saké.’
- b. *u-ja sa a-ranun sagi-du a-ru*  
 it-TOP tea be-NEG saké-FOC be-CONC  
 ‘It’s not tea; it’s saké.’
- (48) Q. *u bagin sagi Ø na*  
 it also saké Q  
 ‘Is that saké, too?’
- A. *a-ranun*  
 be-NEG  
 ‘(No,) it isn’t.’

TABLE 14.3 YONAGUNI COPULA

Concl.	Stative stem	Inf.	Conj.	Past	CONC + <i>-ba</i>	CONC + <i>-ja</i>	PAST + <i>-ja/-ba</i>
<i>a-N</i> <i>a-ru</i>	<i>a-ra-</i>	<i>a-i</i>	<i>a-i-ti</i>	<i>a-tan</i>	<i>a-ru-ba</i>	<i>a-r-ja</i>	<i>a-ta-ja</i>

(49) Q. *u-ja nu Ø ja*  
 it-TOP what Q  
 ‘What is this?’

A. *saa Ø hadi*  
 tea INFER  
 ‘It’s probably tea.’

(49 A.) expresses a higher degree of probability than English *probably*; *hadi* is used because rules of evidentiality prevent the speaker from making an outright statement of fact about something not seen in person (see 14.5.3.2).

(50) a. *sa-du a-ru hadi*  
 tea-FOC be-CONC INFER  
 ‘It must be tea.’ (The level of certainty is even higher than (49 A.))

b. *saa a-ranu kaja*  
 tea be-NEG DOUBT  
 ‘I wonder if this is tea.’ (The speaker presumes that it is tea.)

c. After opening the package and seeing that it was indeed tea:  
*saa-du a-N sai*  
 tea-FOC be-CONC PART  
 ‘I found that it was tea.’ (Unlike Japanese, in Yonaguni this is in the present.)

(51) *nuu a-ru-ba-N minu na*  
 what be-CONC-*ba*-even not.have Q  
 ‘Do you have anything (to eat)?’

(52) *kari-ja saa munu a-i-bi-du mut-i h-juru*  
 that.man-TOP his.own thing be-IND-because-FOC bring-IND go-PERF  
 ‘He took (it) because it is his.’

(53) Mother told her daughter to bring a book:  
 Daughter: *mik-ita-ŋa minUN suja*  
 find-PAST-but not.exist PART  
 ‘I tried to find it but it is not there.’

Mother: *a-i-du bu-ru nai*  
 be-IND-FOC be-CONC PART  
 ‘I am sure that it is there.’

(54) *i-si kaja a-ranu kaja*  
 do-IND DOUBT be-NEG DOUBT  
 ‘I wonder if it is so or not.’

*isi* is a conjugal form of the verb *iruN* ‘do’ (intransitive). That is to say, the opposite of ‘It is not so (*aranuN*)’ is literally expressed as ‘It is *doing* so’ using *iruN*.

#### 14.4.2.3 Adjectives

Adjectives are divided into two classes: the descriptive adjective and the emotive adjective (Table 14.4). In its base form the latter can normally only be used to describe the speaker in statements, or the hearer in questions. When a speaker wishes to describe another person using an emotive adjective, the stem + *kir-* ‘do’ is used, by which the

TABLE 14.4 YONAGUNI ADJECTIVE MORPHOLOGY

Stem		Conclusive	Judgment	Negative	Adverbial	Conditional
Descriptive: <i>taga-</i> 'high'		<i>taga-N/ru</i>	<i>taga-nu</i>	<i>taga-minuN</i>	<i>taga-gu</i>	<i>taga-rja</i> <i>taga-ruba</i>
Emotive: <i>sjana-</i> 'happy'	Speaker:	<i>sjana-N/ru</i>	<i>sjana-nu</i>	<i>sjana-minuN</i>	* <i>sjana-gu</i>	<i>sjana-rja</i> <i>sjana-ruba</i>
	Else:	<i>sjana kiru(-N)</i>		<i>sjana kiranuN</i>	<i>sjana ki</i>	<i>sjana kirja</i> <i>sjana kiruba</i>

speaker expresses his recognition of a situation that he himself cannot personally experience (*kir-* can also be used of the speaker for particular nuances). The adjective has only one negative form, stem + *-minuN*.

(55) a. *nda sjana na*      b. *anu sjana-N do*      c. *anu-ja sjana-minuN*  
 you happy Q      I happy PART      I-TOP happy-NEG  
 'Are you happy?'      'I am happy.'  
 'I am not happy.'

(56) *kari (/anu)-ja sjana k-iruN*  
 he (/I)-TOP happy do-CONC  
 'He is (/I am) happy.'

(57) a. *banu-N sjana-N do*  
 we-also happy PART  
 'We are happy, too.'

b. *kantati-ja sjana ki-buja banta bagin sjana-N sa*  
 they-TOP happy do.IND-but we also happy-CONC PART  
 'They are happy and we [speaker's group] are, too.'

Past = stem + *-ta-*: *taga-ta-N* '(It) was high. (I saw it.)', *sjana-ta-N* '(I) was happy. (I felt so.)'. Negative = stem + *-minu-N*: *taga-minu-N* '(It) is not high. (I saw it.)', *sjana-minu-N* 'I am not happy'. Negative past = negative stem + *-ta-N*: *taga-minu-ta-N* 'It was not high', *sjana-minu-ta-N* 'I was not happy'.

The stems of descriptive adjectives function as modifiers but those of emotive adjectives do not.

(58) *taga-munu*      \**sjana-agami*  
 'expensive thing'      'happy child'

Descriptive adjectives can directly modify nouns but the emotive adjectives cannot because we cannot directly feel what another person is feeling, but infer it from seeing his attitude.

(59) a. *taga-ru*      *munu*      b. \**sjana-ru*      *agami*  
 expensive-CONC      thing      happy-CONC      child  
 c. *nigura-ru*      *t'u*  
 fearful-CONC      person

- (60) a. *taga-ru hadi* high-CONC must  
'It must be high.'  
b. *kari-ja sjana-ru hadi* he-TOP happy-CONC must  
'He must be happy.'
- (61) a. *kanu ki-ŋa-du taga-ru* that tree-NOM-FOC high-CONC  
'That tree is tall.'  
b. *\*kari-ŋa-du sjana-ru*

Both *sjanaN* and *niguraN* indicate the speaker's feelings, but *nigura-* indicates the subject's feelings toward someone or something, so it is possible for this word to modify the noun corresponding to the latter. Another difference between those two adjectives is that *nigura-gu na-i* 'become afraid' is acceptable but *sjana-gu na-i* 'become happy' is not. The reason might be that the feeling of delight springs out in an instant. Adjectives such as *nigura-* seem to be in between the two classes.

- (62) I talked about my experience when I was walking alone at midnight.  
*da-nki h-iru basu nigura-gu na-i*  
house-to go-CONC when fearful-ADV become-IND  
'I became afraid while going to the house.'

Emotive adjectives take verbal forms, stem + *kir-* 'do' to describe others' feelings. A descriptive adjective stem + *kir-* is also possible, indicating someone other than what the adjective describes insists it is the case:

- (63) a. *unu ki-ja taga-N*  
this tree-TOP high-CONC  
'This tree is high.'  
b. *\*unu ki-ja taga k-irun*  
this tree-TOP high do-CONC  
c. *kari<sub>i</sub>-ja taga k-irun*  
'He<sub>i</sub> insists that it<sub>i</sub> is high.'

This collocation functions more freely than the ModJ **-garu** construction.

- (64) a. Atsuko is talking about her flowers:  
*baja-nu hana sat-i-ti araagu abja-N do*  
our.house-GEN flower blossom-IND-CONJ very beautiful-CONC PART  
'The flowers in my house are in bloom and are very beautiful.'  
b. Another person, having heard this, tells a friend:  
*acuko-ja sija-nu hana araagu abja k-i bu-N*  
Atsuko-TOP own.house-GEN flower very beautiful do-IND be-CONC  
'Atsuko thinks (~ repeats) that the flowers in her house are very beautiful.'
- (65) a. At the table, Atsuko said to Kazu:  
*ku-ja ma-N suja*  
this-TOP delicious-CONC PART  
'This is delicious.'  
b. Later Kazu said to someone:  
*acuko-ja maa k-i ha-tan do*  
Atsuko-TOP delicious do-IND eat-PAST PART  
'Atsuko ate [it] saying that it was delicious.'

The difference between the conclusive and judgment forms is as follows. In (a), the child judges his mother to be noisy, and in (b) the child uses his own feelings to conclude that it is noisy, roughly opinion vs. perceived fact.

- (66) Mother: *ug-iri ug-iri*  
get.up-IMP get.up-IMP  
'Get up! Get up!'
- Child: a. *kacimasa-nu ug-iruN*  
noisy-JUDG get.up-CONC  
'(I judge) you are noisy. I will get up.'
- b. *kacimasa-N ug-iruN*  
noisy-CONC get.up-CONC  
'(I feel) it's noisy. I will get up.'

As with verbs, the adjective has the form *-ta + -ja* to indicate a hypothesis.

The focus-construction, *sjana-du aru, taga-du aru*, indicates the speaker's subjective opinion, and so is frequent with emotive adjectives. With descriptive adjectives, however, it is more limited:

- (67) a. *kanu ki-ja taga-N*  
that tree-TOP high-CONC  
'That tree is high.'
- b. \**kanu ki-ja taga-du a-ru*  
that tree-TOP high-FOC be-ATTR
- (68) *ku-ja taga-du a-ru. k-anuN*  
this-TOP high-FOC be-ATTR buy-NEG  
'This [price] is high (for me). I will not buy (it).'

Examples of usages of other forms:

- (69) *taga-gu cim-i*  
high-ADV pile-IMP  
'Pile them up high.'
- (70) *kuma taga-rja nu-i-gurisa-N*  
here high-COND climb-IND-difficult-CONC  
'Since this place is high up, it is difficult to climb to.'
- (71) *da@ca-ru-ba k-uu munu taga-bi k-anuN*  
inexpensive-CONC-*ba* buy-CONC but high-because buy-NEG  
'If it were inexpensive, I would buy it, but since it's expensive, I won't buy it.'
- (72) *da@ca-ta-ja k-u-ŋa taga-ta-ja k-anuN*  
inexpensive-PAST-*ja* buy-CONC-but expensive-PAST-*ja* buy-NEG  
'If it is inexpensive, I would buy it, but if it is expensive, I won't buy it.'

Yonaguni has a number of inchoative suffixes, which combine with *bun* 'be' predicatively. If *minuN* replaces *bun*, it does not form a negative statement (for which *buranuN* is used) but rather indicates that the result is unexpected.

- (73) *-tai*  
a. *iri tidan aga-ntai bu-N suja*  
setting sun red-INCH be-CONC PART  
'The setting sun has become red.'

- b. *aa aga-ntai minun*  
 ah! red-INCH not.CONC  
 ‘Ah! It has become (too) red (unexpectedly).’ (In the case of dyeing.)
- c. *aga-nnani* ‘red clothes’ \**aga-ntai nnani*  
*aga-mi buranun* ‘It isn’t going red.’
- (74) *-dari*
- a. *cu-dari-nu munu* cf. *u-ja cuu-munu*  
 white-INCH-GEN thing it-TOP white-thing  
 ‘something which became white’ ‘It is a white thing.’ (Originally white.)
- b. *ara-ta-ba cu-dari bu-N*  
 wash-PAST-*ba* white-INCH be-CONC  
 ‘When I washed it, it became white.’
- c. *cu-dari minun*  
 white-INCH not.CONC  
 ‘It has become white unexpectedly (because of washing it.)’
- (75) *-cici*
- huru-cici bu-N*  
 black-INCH be-CONC  
 ‘It has become black.’
- (76) *-tati*
- a. *hici-tati bu-N*  
 thin-INCH be-CONC  
 ‘It has become thin.’
- b. *hici-tati-nu nnani*  
 thin-INCH-GEN clothes  
 ‘clothes which have worn thin’
- c. *hici-tati c-ii*  
 thin-INCH cut-IMP  
 ‘Slice it thinly!’

#### 14.4.3 Numerals and classifiers

The traditional way of counting is as follows: *t’uci*, *t’aci*, *miici*, *duuci*, *icici*, *muuci*, *nanaci*, *daaci*, *kugunuci*, *tuu* ‘10’, *tuututi* ‘11’, *tuutati* ‘12’, *tuumiti* ‘13’, etc. This system is related to the ‘native’ Japanese counting system seen in ModJ as **hitotsu**, **futatsu**, etc. While ModJ switches to Sino-Japanese roots beyond 10 (**jū ichi** ‘11’, **jū ni** ‘12’, etc.), Yonaguni only does this from 20, creating *nindu* ‘20’, *ninduiti* ‘21’, *ninduni* ‘22’, . . . *ninduhati* ‘28’, *nindugu* ‘29’, *sandu* ‘30’, etc., continuing onward using these just as ModJ does.

There are also special forms used in recitation: In a sacred place or in front of the family altar, people recite together: *t’ii*, *t’aa*, *mii*, *juu*, *ici*, *muu*, *nana*, *jaa*, *kukunu*, *tuu* (only up to ten). These lack the typical Japanese-Yonaguni [j] vs. [d] correspondence, and ‘9’ has /k/ in intervocalic position, pointing to them being Japanese loans. Some people used these words to count weaving strokes.

Examples of dates are *t’iitati* ‘the first day of the month’, *duju-niti* ‘the fifteenth night (of a lunar month); a night with a full moon’. Yonaguni has calendar-based

memorial services: *t'u-nanka* 'seventh day' (lit. 'first seventh day'), *t'a-nanka* 'fourteenth day' (lit. 'second seventh day'), *mi-nanka*, *du-nanka*, *ici-nanka*, *mu-nanka*, . . . *cindukuniti* 'forty-ninth day of memorial' (Sino-Japanese roots). See also *hjakaniti* 'hundredth day memorial' (Sino-Japanese roots), *duniduni* 'first anniversary of someone's death'.

Examples of Yonaguni classifiers are: (a) Straw rice bags *-tara*: *t'utara*, *t'atara*, *mitara*, *datara*, *icitara*, *mutara*, *nanatara*, *daatara*, *kugumutara*, *tuutara*. These bags held 2½ to 3 **to** (斗) = about 45 litres. (b) Clothes *-ha*: *t'uha*, *t'aha*, *miha*, *duha*, *iciha*, *muha*, *nanaha*, *daha*, *kugunuha*, *tuuha*. (c) Dishes and papers *-ira*: *t'uira*, *t'aira*, *mira*, *duira*, *icira*, *muira*, *nanaira*, *daira*, *kugunuiira*, *tuiira*. There are many other classifiers, such as: Trees, chopsticks *-mutu*, e.g. *t'u-mutu*, *t'a-mutu*, etc. Steps *-mata*, e.g. *t'u-mata*, *t'a-mata*, etc. *Acida* sandals (Japanese **geta**) *-sa*, e.g. *t'u-sa*, *t'a-sa*, etc. Animals *-gara*, e.g. *t'u-gara*, *t'a-gara*, etc. Bowls *-magai*, e.g. *t'u-magai*, *t'a-magai*, etc. Houses *-bugui*, e.g. *t'u-bugui*, *t'a-bugui*, etc. Ships *-su*, e.g. *iqsu* ~ *iisu*, *nisu*, etc. This last example uses Sino-Japanese roots from the beginning, rather than Yonagunian ones.

The *kaida* writing system mentioned above was frequently used to count objects and animals such as these, but with the classifier unwritten and implied.

## 14.5 SYNTAX

### 14.5.1 Noun phrase structure

Nouns are modified by demonstratives, quantity expressions, genitives, relative clauses, adjectives, and postpositional phrases. The modifier precedes the noun. The *-N*-less form of verbs and adjectives indicates general recognition by the participants in the conversation, and is the form used in modification, e.g. *kagu t'u* 'a person who writes'; *tagu tuba t'u* 'a person who flies a kite'. Whereas the personal recognition *-N* is always subjective and concrete, the general recognition form can be comparatively objective and abstract – in other words, something commonly recognized. A sentence can be transformed into a relative clause by using this form.

- (77) a. *munu* 'thing'  
 b. *a-ŋa munu* 'mine'  
 c. *duci-nu munu* 'a friend's thing'  
 d. *ba sinsi-nu munu* 'our teacher's thing'  
 e. *ba gaoku-nu sinsi-nu munu* 'our school teacher's thing'
- (78) a. *baja-nu hana* 'the flowers in our house'  
 b. *baja-ni sag-u hana* 'the flowers which bloom in our house'  
 c. *abja-ru hana* 'the beautiful flowers'  
 d. *baja-ni sat-i b-uru hana* 'the flowers which are in bloom in our house'  
 e. *baja-nu abu-ŋa atara k-i b-utaru hana* 'the flowers which our grandma was treasuring'
- (79) a. *baja-nu agami-ŋa kat-jaru dii* 'the characters which our child has written'  
 b. *baja-nu agami-ŋa nnu kat-itaru dii* 'the characters which our child wrote yesterday'

### 14.5.2 Pronouns and anaphora

ModJ does not have personal pronouns in the strict definition of the word, but has several first person forms depending on the age, sex, etc. of the speaker (Table 14.5). In Yonaguni, however, as well as in other Ryukyuan varieties, there is only one word for the first person singular. Though the honorific system always distinguishes sex, and the kin terms differ depending on it, Yonaguni pronouns do not. In addition, unlike Shuri and other varieties which have honorific equivalents for personal pronouns, Yonaguni does not have any honorific form for the second person.

It has long been observed (earliest in Iwakura, 1941) that a number of Ryukyuan varieties have two kinds of first person plural, namely Amami (Kikai island), the northern part of the main Okinawan island, its southern part, Kume island, the Miyako islands, and the Yaeyama islands.

Both the singular and the plural of the first person pronoun change their forms: independent form (*anu*, *banu*), nominative form (*anya*, *banja*), and genitive form (*anya*, *ba*). Other particles follow the independent form.

There is no honorific form for the second person. The plural form *ndi-nta* sounds somewhat disagreeable.

- (80) *ndi-nta nu-ba k-i b-un ga jee*  
 you-PL what-SEL do-IND be-CONC Q PART  
 (With a connotation of annoyance) ‘What on earth are you doing?’

The genitive case is expressed by the particle *-ja* for persons and *-nu* for others. However, some forms can be used without *-ja* for the genitive case.

- (81) Q. *kumi a-ru munu-ja taa munu ja*  
 here.LOC exist-CONC thing-TOP whose thing Q  
 ‘This thing here, whose is it?’

**TABLE 14.5 YONAGUNI PERSONAL (AND REFLEXIVE), DEMONSTRATIVE, AND INTERROGATIVE PRONOUNS**

	1st		2nd		3rd		
	‘I’		‘we’ (incl.)		‘you’	‘he/she/it’	‘-self’
		Nom.	Gen.				
Sg.	<i>a-nu</i>	<i>a-ja</i>	<i>a-ja</i>	–	<i>nda</i> <sup>+</sup>	<i>u ~ u-ja</i>	<i>sa</i> <sup>+</sup> ~ <i>-ja</i>
Pl.	<i>ba-nu</i>	<i>ba-ja</i>	<i>ba</i>	<i>ba-nta</i>	<i>ndi/ndi-nta</i>	<i>untati</i>	<i>si</i> <sup>+</sup> ~ <i>-ja</i>
	Demonstrative			Interrogative			
	‘this’		‘that’		‘what’	‘who’	
Sg.	<i>ku ~ -ja</i>		<i>kari ~ -ja</i>		<i>nu</i>	<i>ta</i> <sup>+</sup> ~ <i>-ja</i>	
Pl.	<i>kuntati</i>		<i>kantati</i>		<i>nuu-nu</i>	<i>taa-ta</i> (~ <i>ta-nta</i> ~ <i>taa-ta-nta</i> )	
Place	<i>kuma</i>		<i>kama</i>		<i>nma</i>		
Adj.	<i>kunu</i>		<i>kanu</i>		<i>munu</i>		
Adv.	<i>kunni</i>		–		<i>munni</i>		
					‘why’	<i>mundi</i> , <i>nuu ki bi</i>	
					‘when’	<i>ici</i>	

+ : *nda*, *sa*, *si*, and *ta* can also be used as genitive forms.

- A. *nda/a-ŋa/baa/u-ŋa ~ ku-ŋa/kari-ŋa/kantati/banta/*  
 ‘your/my/our/his ~ this person’s/that person’s/those persons’/our (S.G.)/  
*kanu t’u-nu/untati/abuta/ija/sinsi/acuko(-nu)/unu agamiti-nu munu*  
 that person’s/their/mother’s/father’s/teacher’s/Atsuko’s/that child’s thing’

The third person pronoun can refer to a person (or thing) contextually indicated who is not present with the speaker and the hearer, as well as a person (or thing) that is there with them. It is desirable to refer the person who is before their eyes with *kanu t’u* ‘that person’, not *ku* or *kari*, especially when they have respect for him/her.

- (82) Just after Kazu went out, a friend came to see her. Kazu’s mother says:

*u-ja tund-i hj-un aa u-ŋa munu a-N suja*  
 she-TOP hurry-IND go-PERF ah! she-GEN thing be-CONC PART  
 ‘She has just gone out. Oh! She left her thing here.’

Yonaguni does not have any form corresponding to ‘which’. Instead, *ta* ‘who’ or *nu* ‘what’ are used. The selective particle *-ba* is used with these pronouns to indicate choice or contrast.

- (83) a. The speaker, seeing many women, asks the child:

*nda abuta-ja ta ja*  
 your mother-TOP who Q  
 ‘Which is your mother?’

- b. Seeing a picture of the mothers, the speaker could say:

*ta-ba ndja-nu abuta ja*  
 who-SEL your.house/family-GEN mother Q  
 ‘Which is your mother?’

- (84) Q. *nu-ba nda munu ja* A. *ku-du a-ŋa munu*  
 what-SEL your thing Q this-FOC I-GEN thing  
 ‘Which is yours?’ ‘This is mine.’

*ku-*, *u-*, and *ka-* are often compared with ModJ **ko-**, **so-**, and **a-**. They are, however, different from the Japanese words in usage: *u-* is contextually based, as both hearer and speaker have to be looking at that same thing. Demonstrative pronouns have plural forms to refer to persons, and for things in some special situations. *kuntati* and *kantati* for persons have a disagreeable connotation.

- (85) *u-ja nda i-ri ku-ja a-ŋa i-run*  
 it-TOP you do-IMP this-TOP I-NOM do-CONC  
 ‘That, you do. This, I’ll do.’

- (86) *untati-nta-ja kantati-n-ja buru ab-iru na*  
 they-PL-TOP those.people-also-TOP all invite-CONC Q  
 ‘These people and those people, do you invite all of them?’

- (87) *kuntati-nta-N abiri ndi ja*  
 these.people-PL-also invite-IMP QUOT Q  
 ‘Do you say: “Invite these people, also”?’ (I would not invite them.)

- (88) Various things are put all together in a place:

*kantati(-nta) ~ kuntati(-nta) ~ untati(-nta)-ja baa munu do*  
 those(-PL) these(-PL) they(-PL)-TOP our thing PART  
 ‘These are ours.’

The demonstrative adverbs *kunni* ‘like this’ and *unni* ‘like that’ are used, but the logically possible *\*kanni* ‘like that’ does not occur.

Yonaguni, like some other Yaeyama dialects, has special words to denote the family community. The members are not necessarily living altogether: *baja* ‘our home’, *ndja* ‘your home’, *sija* ‘their own home’. *\*a-ŋa-ja* ‘my home’ is not permissible even for a single unmarried person. Note that ‘house’ in Yonaguni otherwise is *da*, which corresponds to Japanese *ya*.

(89) *uma a-ŋa da* ‘Here is my [a single unmarried person’s] house.’

Only the third person uses the reflexive pronoun. Other persons use different devices. The reflexive indicates persons who are not in the presence of those having the conversation, or do not participate in the conversation.

(90) a. I saw someone take a newspaper from our house.  
*kari-ja sinbun mut-i h-jun do*  
 he-TOP newspaper take-IND go-PERF PART  
 ‘He took a newspaper and went off.’

b. My mother answered:  
*kari-ja sa munu a-i-bi-du mut-i h-juru*  
 that-TOP REFL thing be-IND-because-FOC take-IND go-PERF  
 ‘That is his own thing, so he took it and went off.’

c. She continued:  
*ku-N kari-ŋa munu \*sa munu*  
 this-also he-GEN thing  
 ‘This is his, too.’

(91) A: *kanu sandu-ja kanu da-nki ha-i h-jun do*  
 that Sandu-TOP that house-to enter-IND go-PERF PART  
 ‘That Sandu has gone in that house.’

B: *kama-ja si da*  
 there-TOP REFL.PL house  
 ‘That is his (lit. their) own house.’

(92) Looking at her friends going out.  
*anu ab-iranun k-i si bagai h-irun suja*  
 I call-NEG do-IND REFL.PL only go-CONC PART  
 ‘Without calling me, they are going out by themselves.’

(93) a. *sa-ŋa i-si-ti c-anu katarai k-ii* (To a single hearer.)  
 b. *si-ŋa i-si-ti c-anu katarai k-ii* (To multiple hearers.)  
 REFL-NOM do-IND-CONJ know-NEG manner do-IND  
 ‘You yourself [yourselves] did (that) and you acted nonchalantly.’

(94) a. *du-si i-ri*  
 self-by do-IMP  
 ‘Do it yourself.’

b. *a-ŋa du-nu du-si i-run*  
 I-NOM body-GEN self-by do-CONC  
 ‘I’ll do it myself.’

- c. *banta du-nu du-si i-ru na*  
 we (S.G.) body-GEN self-by do-CONC PART  
 ‘Let’s each do it in our own way.’
- d. *banu muuru-si i-run*  
 we together-by do-CONC  
 ‘We’ll all do it together.’
- (95) *sa t’ui-si mu-i bu-N suja*  
 itself alone-by grow-IND be-CONC PART  
 ‘Plants are growing by themselves [spontaneously, without anyone having sown seeds].’

Some contemporary scholars classify the first person plural pronouns into ‘exclusive’ and ‘inclusive’. This opposition is strictly ‘we, the opposite of second person’ vs. ‘we, the opposite of third person’. The so-called inclusive Yonaguni form, however, means ‘we, the speaker’s group’, that is, the speaker and anyone participating in the speaker’s action, or otherwise associated with the speaker. We will use the term ‘speaker’s group (S.G.)’, as introduced in Izuyama (1992). *ba-* (first person exclusive) indicates the speaker + other(s), who are contrasted with the hearer (second person *nda ~ ndi*). *banta* (first person inclusive) indicates the speaker + hearer(s), who are contrasted with others (third person *kari ~ untati*). There are co-occurrence restrictions with verb forms, e.g. the *-N* conclusive form cannot be used by itself with *banta*; *-N* has to be followed by other particles as the need arises, e.g. *sa*:

- (96) a. *banu-N i-run*  
 we-also do-CONC  
 ‘We are going to do (it), too.’
- b. *banta-N i-ru na*  
 we (S.G.)-also do-CONC Q  
 ‘Shall we do it, too?’
- (97) *banu-N sjana-N (do)*  
 we-also happy-CONC (PART)  
 ‘We are happy, too.’
- (98) *kantati-ja sjana k-i-buja banta bagin sjana-N sa*  
 they-TOP happy do-IND-but we (S.G.) also happy-CONC PART  
 ‘They are happy and so are we [speaker’s group]. (I know.)’
- (99) a. One group of people was overtaken by another, whereupon one of its members said:  
*untati-ja haja-aci-si h-iru-ŋa banta-ja durii.duri hindagi*  
 they-TOP fast-foot-by go-CONC-but we (S.G.)-TOP slowly let’s go  
 ‘They are walking quickly, but we will go slowly.’
- b. He then shouted to the other group:  
*ndi-ja haja-gu h-iri joo banu-ja duri.duri h-iru jungara*  
 you-TOP fast-ADV go-IMP PART we-TOP slowly go-CONC since  
 ‘You go fast! We will go slowly.’
- (100) A mother-in-law shows her son’s wife the family tomb:  
*uma-du banta haga do*  
 here-FOC our (S.G.) tomb PART  
 ‘This place is our tomb.’

cf. *uma-ja ba haga* (This would be said to the people of another family.)  
 here-TOP our tomb  
 ‘This place is our tomb.’

(101) A grandmother is speaking to her grandson who just arrived at Yonaguni airport for the first time. The boy is not living in the same house, but *baja* cannot be used, nor can *banta da* or *a-ŋa da*. The grandmother can refer to herself as *abu* ‘grandma’.  
*ita! abu da-nki di!*  
 INTERJ grandma’s house-DAT INTERJ  
 ‘Now! To Grandma’s house! Come on!’

(102) A woman is explaining to her elderly mother-in-law, who is becoming senile, that their house is in a different direction:  
*banta da-ja kama do \*baja*  
 our house-TOP that.way PART  
 ‘Our house is that way!’

*baja* cannot be used for ‘our house’: it implies that the house is the speaker’s (plus others’), not the hearer’s. Here, a stranger can say *ndja* (‘your house’) *-ja kama do*.

### 14.5.3 The basic sentence

#### 14.5.3.1 Declarative

Sentences can take two intonation patterns: assertion and question. Some particles that emphasize the speaker’s assertion cannot take the question intonation, e.g. *do*, *ju*, etc. In assertive sentences which have neither verbs nor adjectives, the speaker must be seeing (or feeling) the objects. They are not so-called ellipsis phenomena.

As the grammatical category of evidentiality is fundamental in Yonaguni, it is not strictly accurate simply to characterize word order as SOV. Though SOV is generally the case, the order Topic-(Actor/Object-)Verb is arguably more important.

(103) a.	Topic-(Actor/Object-)Verb T O V	(104) a.	NonTopic-(Actor/Object-)Verb S O V
	<i>anu dii kag-un</i>		<i>a-ŋa dii kag-un</i>
	I character write		I-NOM character write
	‘I will write characters.’		‘I will write characters.’

The above two sentences are different. If we add the topic particle *-ja* to the actor of both sentences, the difference becomes clear.

(103) b.	T O V	(104) b.	(A=)T O V
	<i>anu-ja dii kag-un</i>		<i>a-ŋa-ja dii kag-un</i>
	I-TOP character write		I-NOM-TOP character write
	‘I will write characters.’		‘I myself will write characters.’ (Others, on the other hand, will not.)
	c.		c. (O=)T A V
	<i>*dii-ja anu kag-un</i>		<i>dii-ja a-ŋa kag-un</i>
			‘As for the character, I will write it.’

(104b) is not permissible at the beginning of a conversation. It would be as a response to *kari-ja kag-anun do* ‘He will not write’.

14.5.3.2 *Tense, aspect, and evidentiality*

Yonaguni has distinctions of tense (non-past *kagUN* ‘writes; will write’ vs. past *katitan* ‘wrote’) and aspect (unmarked; perfect *katjan* ‘has written’; progressive *kati bun* ‘is writing’). Other forms include indefinite + *nmUN* ‘see’ to express an experimental aspect, e.g. *kati nni* ‘Try writing!’.

Ryukyuan varieties have evidentiality as a basic grammatical category. We find a three-term system (Aikhenvald 2004, 42) in Yonaguni: direct, inferred, and reported.

*Evidentiality and tense*

The word-final *-N* of verbs and adjectives (traditionally called *shūshikei*) expresses information directly acquired by the speaker. In the present tense, *-N* by itself expresses directly acquired information (105), whereas reported information requires the addition of a quotative marker (106).

(105) People are having a dance party in the town hall:

1st person:

- a. *anu su-ja budi k-iru-N (do)*  
 I today-TOP dance do-CONC-*N* PART  
 ‘Today I am going to the dance.’

2nd person:

- b. *Nda budi k-iru-N*  
 you dance do-CONC-*N*  
 ‘You are going to the dance?’
- c. *\*Nda budi k-iru-N (do)*  
 you dance do-CONC-*N* PART  
 ‘You are going to the dance.’

3rd person:

- d. *\*acuko su-ja budi k-iru-N (do)*  
 Atsuko today-TOP dance do-CONC-*N* PART  
 ‘Atsuko is going to the dance today.’

(106) Atsuko told her friend about the dance party, and her friend told her family:

- a. *acuko su-ja budi k-iru-N tuna*  
 Atsuko today-TOP dance do-CONC-*N* QUOT
- b. *acuko su-ja budi k-iru-N di*  
 Atsuko today-TOP dance do-CONC-*N* QUOT  
 ‘Today Atsuko is going to the dance(, I heard).’

The assertive sentence (105c) is not permissible. The only way it would be is if all the other people, except the woman in question, have decided to involve her in their dance group, with her being almost unable to deny their request. This would constitute a kind of strong order. In (105d), only if Atsuko is the speaker’s daughter, and the speaker knows that she practices dancing every day and has made a dress for her, etc., would it be acceptable. If the speaker is not personally familiar with the situation, the grammatical noun *hadi*, expressing a kind of inferred statement, will be used (107).

- (107) *acuko su-ja budi k-iru hadi do*  
 Atsuko today-TOP dance do-CONC INFER PART  
 ‘Atsuko is expected to dance today./Atsuko is probably going to dance today.’

In this example, the speaker knows that other women of the same age as Atsuko also usually dance and that Atsuko belongs to the women’s society which is the host organization for the dance party, etc. But the speaker is not Atsuko, so he cannot directly say *acuko su-ja budi kiru<sub>N</sub>*, and must insert *hadi*.

For the past tense, the morpheme *-ta-N* means that the speaker experienced the action directly in his past (108a,b; 110b). If he did not witness it directly, an explicit quotative marker (110a,c) is again obligatory.

- (108) a. Atsuko’s mother saw Atsuko dance in the town hall. Later, she told her friend:  
*su acuko-ja budi ki-ta-N do*  
 today Atsuko-TOP dance do-PAST-*N* PART  
 ‘Atsuko danced today.’

- b. That night, Atsuko told her brother, who could not see her dance:  
*anu su budi ki-ta-N do*  
 I today dance do-PAST-*N* PART  
 ‘I danced today.’

- (109) A friend of her brother who was with him and heard the above conversation later told his own family:

- a. *su-ja acuko-ja budi k-ita-N tuna*  
 today-TOP Atsuko-TOP dance do-PAST-*N* QUOT
- b. \**su-ja acuko-ja budi k-ita-N (do)*  
 ‘I heard that Atsuko danced today.’

- (110) a. In the town hall, two girls are talking:

Girl A: *nda budi k-ja na*  
 you dance do-PERF Q  
 ‘Have you danced?’

- b. Atsuko: *NN budi k-ja-N do*  
 yes dance do-PERF-*N* PART  
 ‘Yes, I have danced.’

- (111) Later A’s friend asked her whether Atsuko danced or not. A did not see her dance, but she did hear Atsuko and the audience talking about Atsuko’s dancing, so replied:

- a. *anu-ja nn-anu-ta-ŋa budi k-ja-N tuna*  
 I-TOP see-NEG-PAST-but dance do-PERF-*N* QUOT
- b. \**anu-ja nn-anu-ta-ŋa budi k-ja-N (do)*  
 ‘I did not see her dance but apparently she did.’

We do not have any memory of the day when we were born, but our mothers of course remember it, hence the evidential difference between (112a) and the mother’s words (112b). Because we cannot witness them ourselves, we can only use reported structures to discuss them, for example the 1771 Meiwa tsunami (113).

- (112) a. *a-ŋa mari-taru ci-ja ubu-ami hu-ta-N tuna*  
 I-NOM be.born-PAST day-TOP heavy-rain fall-PAST-*N* is.said  
 ‘(It is said that) it rained heavily on the day when I was born.’

- b. *nda-ŋa mari-taru ci-ja ubu-ami hu-ta-N do*  
 you-NOM be.born-PAST day-TOP heavy-rain fall-PAST-N PART  
 ‘On the day you were born, it rained heavily.’
- (113) *nkaci-ja caba-nu dasiki-kuta nan-ŋa s-uta-N*  
 long.ago-TOP Tsaba-GEN residence-TERM wave-NOM come-PAST-N  
*di do (~ tuna)*  
 is.said QUOT  
 ‘(It is said that) long ago, a tsunami came all the way up to the Tsaba residence  
 [situated at the highest place in Yonaguni].’

The question sentence is (*sutan*) *di?* and the answer is *di do*. *do* is necessary in the answer.

- (114) The words used in reported speech are the same as those used in transmitting usual messages.  
*nka-i-ndi ku-N di do / tuna*  
 meet-IND-PURP come-CONC QUOT PART / QUOT  
 ‘It is said (~ I heard) that someone is coming to meet you.’

Inferred situations are illustrated in (115–116):

- (115) Atsuko’s friend came to the town hall and saw Atsuko take off her dress and her make-up. She said to herself (or to someone at her side).  
*acuko-ja budi k-jaN suja \*budi k-jaN*  
 Atsuko-TOP dance do-PERF PART  
 ‘Atsuko must have danced, I guess./It seems that Atsuko has danced.’
- (116) When I visited my friend’s house, it was very clean:  
 a. Speaking to someone or to myself:  
*sudi k-jaN suja*  
 sweeping do-PERF PART  
 ‘They (including my friend) must have done sweeping.’  
 b. My friend:  
*NN k-jaN do kazu-ŋa-du sudi k-i ju*  
 yes do-PERF PART Kazu-NOM-FOC sweeping do-IND PART  
 ‘Yes, we did. It is Kazu who did it.’

In this sentence, it is directly implied that my friend did see Kazu sweep the house, but I did not. I guessed it seeing the clean house.

### *Evidentiality and aspect*

The indefinite (*ren’yōkei*) form indicates an action which has come into existence. We have to see it in order to take it into our consciousness. At the point when the action began, we had not yet noticed it. We have to see it and then know that it has taken place. That is why this form is used in questions, asking back, and mentioning the present situation. It is not used at the moment when the action takes place or ends. It has relation neither to the beginning nor to the end of the action. That is the reason why the indefinite form is fit to use as the dictionary form.

This form can be accompanied with the question particle *na* and has such connotation as you did not need to do so. Sometimes, depending on the speaker, it can occur with

*ja* and *ju* to express the action which occurred in the past (and may still continue). The former is for questioning and the latter is for assertion.

- (117) My glass is empty, so I asked my boy  
*num-i na* Answer: *num-i-minun*  
 drink-IND Q drink-IND-NEG  
 'Have you drunk it?' 'I have drunk it.' (with regret)
- (118) Seeing a notebook, Ms. B said that the writing is good. Ms. A answered proudly:  
 A: *kunu di-ja kazu-ŋa-du kat-i ju*  
 this character-TOP Kazu-NOM-FOC write-IND PART  
 'This character, Kazu wrote it.'  
 B: *maamuni kazu-ŋa-du kat-i ja*  
 really Kazu-NOM-FOC write-IND PART  
 'Is it true that Kazu wrote it?'
- (119) *ndi-ŋa kid-ita-ba agami ug-i*  
 you (pl.)-NOM make.a.noise-PAST-*ba* baby get.up-IND  
 'Since you made a noise, the baby has woken up.' (Mother said to her children.)
- (120) *daatusi ug-i aig-ar-irun suja*  
 with.difficulty get.up-IND walk-POT-CONC PART  
 '(He) can get up and walk with difficulty.'
- (121) The baby fell asleep at last, and mother went to the kitchen to wash dishes:  
 Grandma: *agami ug-un*  
 baby wake.up-PERF  
 'The baby has woken up.'  
 Mother: *mata ug-i na*  
 again wake.up-IND Q  
 'He got up again?'
- (122) A friend of Kazu wants to see Kazu dance in the town hall. There, she asks her mother whether Kazu danced already, and her mother, who had seen Kazu dance in person, answered:  
 Kazu's friend: *kazu budi k-i*  
 Kazu dance do-IND  
 'Did Kazu dance?'  
 Mother: *nai k-jaŋ \*k-i*  
 now do-PERF  
 'She did just now.' (Mother saw her dance.)  
 cf. *nai-gara-du k-iru*  
 now-ABL-FOC do-CONC  
 'She'll do it (starting) now.'
- (123) Mother found the bed of her little boy wet:  
 Mother (to herself): *nbai k-i . . .*  
 pee do-IND  
 'He wet his bed.'

- (124) Kazu's mother describes how much her father loves Kazu.  
*asa-ŋa araagu atara k-i (ju)*  
 Grandpa-NOM much cherish do-IND (PART)  
 'Grandpa has always cherished her very much (and still does now).'

Not only the result but also recognition of the action is necessary for using the perfect form. In the case of not seeing with one's own eyes, the perfect form is not possible. The perfect form would be said only by the person who performed the action, or by someone who saw him/her perform it.

- (125) Mr. A came to take the paper which he asked Atsuko to write.  
 Mr. A: *tarum-jaru munu kat-jaN ~ kat-ja-na*  
 ask-PERF thing write-PERF write-PERF-Q  
 'Have you written what I asked you to write?'  
 Atsuko: *kat-jaN*  
 write-PERF  
 'I have written it.'
- (126) When Mr. A came, Atsuko was away from home, but her daughter who came back from Tokyo a while ago was home.  
 Mr. A: *acuko-ja kat-jaru kaja \*kat-ja-na, \*kat-jaN*  
 Atsuko-TOP write-PERF DOUBT  
 'I wonder if Atsuko has written it.'
- (127) The daughter found the paper which she *thinks* was written by her mother, but she could not say *katjan*, because she did not actually see Atsuko write it.  
*kat-i a-N do \*kat-jaN*  
 write-IND be-CONC PART  
 'It is written (by someone).'
- (128) If the daughter saw Atsuko write and her mother explained the details about the paper to her, she could say *katjan*.  
*nnu kat-ja-ta-ŋa tu-i-ndi h-ir-ja minun*  
 yesterday write-PERF-PAST-but take-IND-PURP go-CONC-ja be.not  
*suja*  
 PART  
 'She wrote it yesterday, but even though I went to get it, it wasn't there.'

### *Evidentiality and the adjective*

The emotive adjective in its base form can only be used to describe the speaker as a general rule. It occurs with the second person only in question sentences.

When a speaker wishes to describe another person using an emotive adjective, the stem will be followed by the verb *kir-* (to do), with which the speaker expresses his recognition of conditions which he himself cannot personally experience.

We know our own feelings or internal situations; however, we cannot directly experience (see, touch, etc.) those of others, and must infer them by watching their attitudes. This difference forms the basis for the two different adjective types.

The emotive adjective cannot occur as a predicate with the 3rd person, so *acuko-ja sjanaN* is not permissible. To describe another person's inner world, an objective expression using verb collocation should be adopted. Otherwise *tuna* ('it is said' or 'I heard') should be added.

(129) Atsuko said to Kazu, “*anu-ja sjana-N* (I am glad),” and so Kazu reported this to her friend:

- a. *suu-ja acuko-ja sjana-N tuna \*acuko-ja sjana-N*  
 today-TOP Atsuko-TOP happy is.said  
 ‘Apparently Atsuko is happy today.’

That friend later told her family members.

- b. *acuko-ja sjana-tan tuna*  
 Atsuko-TOP glad-PAST be.said  
 ‘Apparently Atsuko was happy.’

cf. The friends who saw Atsuko’s attitude could say:

- c. *acuko-ja sjana k-i bu-tan*  
 Atsuko-TOP happy do-IND be-PAST  
 ‘Atsuko was happy (in our opinion).’

The speaker and the subject of emotive adjective + *kir-* can be different. We have to notice also the time difference of the occurrence and the speaker’s recognition.

(130) Seeing that Atsuko looks happy:

- acuko-ja sjana k-i bu-N suja \*sjanaN*  
 Atsuko-TOP happy do-IND be-CONC PART  
 ‘Atsuko seems happy.’ (Atsuko is in a happy situation. I am looking at her.)

(131) My friend gave me some cake for my child Kazu, and so I gave it to her. I report this to my friend:

- a. *kazu-ja kaci tur-a-ta-ba sjana k-i ju*  
 Kazu-TOP cake take-CAUS-PAST-*ba* happy do-IND PART  
 ‘[I/someone] passed the cake to Kazu and she looked happy.’
- b. *kazu-ja kaci tur-a-ta-ba sjana k-i bu-tan*  
 Kazu-TOP cake take-CAUS-PAST-*ba* happy do-IND be-PAST  
 ‘[I/someone] passed the cake to Kazu and saw that she was happy. [And she is presumably still happy now.]’

(132) Looking at the cake that she is receiving, she can also say the following, despite it being grammatically in the past tense:

- anu araagu sjana-tan*  
 I very happy-PAST  
 ‘I was very glad. (I felt very much happiness.)’

(133) a. The same use of the past tense immediately upon receiving the gift can be used in answering a question:

- nda sjana na* Answer: *hugarasa araagu sjana-tan*  
 ‘Are you happy?’ ‘Thank you, I’m very happy.’

b. General question and answer:

- nda sjana na* Answer: *araagu sjanaN*  
 ‘Are you happy?’ ‘I am very happy.’

#### 14.5.3.3 Modality

Conditional forms + *nsaru munu* (lit. ‘[it’s] a good thing’) are used for desiderative or advice (‘I wish . . .’; ‘you should . . .’):

- (134) a. *kagja / kaguba / katitaja nsaru munu* 'You should write it.'  
 b. *kirja / kiruba / kitaja nsaru munu* 'You should do it.'  
 c. *nnirja / nniruba / nnitaja nsaru munu* 'I wish I would die.'  
 d. *ugutaja / uguba nsaru munu* 'I wish you'd woken me up.'  
 cf. *ugurja nsan do* 'As he's awake, it's OK [you don't have to go and wake him].'  
 e. *irja / iruba / itaja nsaru munu* 'You should do it.'

Conditional forms are also used for the obligative ('must, should, ought to', etc.), in construction with focal *-du + naru* 'become'. Some verbs have two forms for this function, but others do not. The rules governing this are unclear.

- (135) a. *nnirja-du naru / nniruba-du naru* 'It must die.'  
 b. *aigja-du naru / aiguba-du naru* 'She has to walk.'  
 c. *irja-du naru / iruba-du naru* 'I must do it.'  
 d. *ugirja-du naru* 'I must get up.' \**ugiruba-du naru*  
 e. *uguba (~ uguiba)-du naru* 'I (~ she) must wake her.' \**ugutaja-du naru*  
 f. *haiba-du naru* 'You have to eat it.' \**huba-du naru*  
 g. *nurja-du naru* 'I have to get on board.' \**nuruba-du naru*

The potential is formed from stative stem + *-r-irun* (the passive), but in the negative *-ninun* is used: *kag-arirun* 'I am able to write', 'It will be written'; *kag-aninun* 'I am unable to write'. Another potential construction is indefinite + *cu-N*, e.g. *kat-i-cu-N* 'I am able to write'. Other constructions are indefinite + *-daca-N* 'it is easy to' and *-gurisa-N* 'it is hard to', e.g. *kat-i-daca-N* 'it is easy to write'.

Desiderative meanings are expressed through the indefinite + *-busa-N* < *husa-N* 'be desirable', e.g. *kat-i-busa-N* 'I am eager to write'.

#### 14.5.3.4 Non-declarative sentence types

##### Questions

Information questions use the interrogative pronouns listed in 14.5.2. Unlike in English, the speaker can indicate that more than one person or thing is questioned about (136b, 139).

- (136) a. *ta-ŋa bu-N ga*  
 who-NOM be-CONC Q  
 'Who's there?' (The speaker knows someone is there.)  
 b. *taa-ta-ŋa bu-N ga*  
 who-who-NOM be-CONC Q  
 'Who's there?' (The speaker hears some sounds in the next room.)
- (137) a. *nu ja*  
 what Q  
 'What?' (Answering someone or asking to a stranger coming into the room, etc.)  
 b. *nu-ba*  
 what-SEL  
 'What [did you say]?' (In response to not having heard something.)
- (138) Q. *nu-ŋa (/ba) maci ja*  
 which-NOM(/SEL) better Q  
 'Which is better?'

- A. *ku-du maci*  
 this-FOC better  
 ‘This is better.’
- (139) *nuu-nu-ba a-N ga*  
 what-what-SEL be-CONC Q  
 ‘What’s there?’
- (140) *nma-nki(-ba) h-irun ga*  
 where-to(-SEL) go-CONC Q  
 ‘Where are you going?’
- (141) Casual greeting on the road:  
 Q. *nda nma-nki ja \*nma-nki-ba*  
 you where-to Q  
 ‘Where are you going?’ (In greeting)
- A. *kama-nki / uma-nki*  
 over.there-to / there-to  
 ‘Over there.’ / ‘Just over there.’
- (142) *nma jaa nma jaa*  
 where Q where Q  
 ‘Where, where?’ (In the case that people hear ‘Fire! Fire!’)

### Commands

Commands are expressed by the imperative (affirmative) and prohibitive (negative): *h-iri!* ‘Go out!’ *h-inna!* ‘Do not go!’ Addition of *joo* to a command softens it (25). Polite commands/requests are formed from indefinite + *turai* ‘give!’: *kat-i turai* ‘Please write’.

### Sentence particles

Yonaguni has many sentence-final particles, such as: *na* and (expressing greater speaker certainty) *nai* (both after *-N*-less forms or nominals) = yes/no question; *ga* (after verbals in *-N* form), *ja(a)* (after nominals) = question-word question; *kaja(a)* = question with doubt (‘I wonder . . .’); *sa* = tag; *jee* = emphasis with suspicion or disappointment; *suja(a)* = affirmation with direct evidence; *do(o)* = assertion; *sai* = attraction of attention to the statement; indefinite + *juu* = politeness; imperative + *joo* = softened command. Note that some of these particles show exceptions to the phonological system: the sentence-final particles *doo* and *jee*, like the interjection *hee*, contain [o] and [e] sounds, which are generally not found in this dialect. These particles can be seen in examples throughout this chapter.

### 14.5.4 Topic, focus, and emphasis

In addition to the topic particle *-ja*, Yonaguni has a focus particle *-du* and a selective particle *-ba*.

The focus particle *-du* can co-occur with *-ba* but not with *-ja*. Ryukyuan is believed to mirror the *kakari-musubi* phenomenon found in Old and Classical Japanese; that is, in a sentence containing a *-du*, the sentence-final verb or adjective has the *N*-less form, which

expresses general recognition, instead of the *-N* form, which expresses the speaker's recognition. That is to say, all participants in the conversation know the situation expressed by the sentence except for the information marked by *-du*. *-du* can be used with an imperative.

*-ba* indicates a selection from several choices, and it is often used with *-du* in question sentences. It also has another usage described below.

- (143) Q. *nu-ba nda munu ja*  
 what-SEL your thing Q  
 'Which is yours?'  
 A. *ku-du a-ŋa munu*  
 this-FOC I-GEN thing  
 'This is mine.'
- (144) Q. *u-ja nmi-ba a-N ga*  
 it-TOP where.LOC-SEL be-CONC Q  
 'Where is it?'  
 A. *sakibara-nu misija-ni-du a-ru*  
 Sakibara-GEN shop-LOC-FOC be-CONC  
 'It is at Sakibara's shop.'
- (145) Q. *umi a-ta-ru kaci-ja ta-ŋa h-a-ru*  
 there.LOC be-PAST-CONC cake-TOP who-NOM eat-PERF-CONC  
*kajaa*  
 DOUBT  
 'Who has eaten the cakes which were there?'  
 A. *a-ŋa-du h-arū*  
 I-NOM-FOC eat-PERF.CONC  
 'It is I who has eaten them.'
- (146) *unni-ba-du i-si jaa*  
 to.that.extent-SEL-FOC do-IND Q  
 'You could only do that much?'
- (147) Being asked to help a child stand up:  
 Q. *ta-ba-du tata-N ga*  
 who-SEL-FOC make.stand-CONC Q  
 'Who (shall) I make stand up?'  
 A. *acuko-du tata / acuko-du tata-i*  
 Atsuko-FOC make.stand.CONC / Atsuko FOC make.stand-IMP  
 'Make Atsuko stand up.'
- (148) Q. *nda budi k-i*  
 you dance do-IND  
 'Have you danced?'  
 A. *nai-gara-du k-iru*  
 now-ABL-FOC do-CONC  
 'I will do it (starting) now.'

Only *nai-gara* is new information. The action *budi* is understood by the speaker and the hearer in the response.

- (149) *nai-du s-u-ru* (Said in front of the door when visiting someone's house.)  
 now-FOC come-PERF-CONC  
 'Now I have come.' (People had been waiting for him.)

The particle *-ba* often occurs with a subject noun or complement noun to mark it as somewhat undesirable.

- (150) a. *ubu-ami-ba hu-i-bi ubuni maga-ninu-ta<sup>N</sup>*  
 terrible-rain-SEL fall-IND-due.to radish sow-POT.NEG-PAST  
 'It was raining heavily, so we could not plant radishes.' (The speaker did not want the heavy rain.)
- b. *ami hu-i-bi ubuni mag-u<sup>N</sup> dii \*ami-ba*  
 rain fall-IND-because radish sow-CONC INTERJ  
 'It is raining, so we can plant radishes. Let's go!'
- (151) *ui t'u-ba na-i sikama k-iraninun*  
 old person-SEL become-IND work do-POT.NEG  
 'I am becoming an old man, so I cannot work.' (The speaker is lamenting this.)
- (152) *ca-ba pa<sup>N</sup>-i bu-<sup>N</sup> sujaa*  
 weeds-SEL grow-IND be-CONC PART  
 '(A lot of) weeds are growing.' (The speaker dislikes having to remove them.)

### 14.5.5 Passive and causative

#### 14.5.5.1 Transitive/intransitive pairs

Many, but not all, verbs have a pair of intransitive and transitive/causative forms. Those which correspond to ModJ quinquagrade (conclusive *-u<sup>N</sup>*) mostly change to *-a-<sup>N</sup>*, and those which correspond to ModJ monograde have curiously the same shapes as their perfect forms, as in *ug-u<sup>N</sup>* (Tables 14.6 and 14.7). Note that in ModJ, the transitive verb **akeru** 'open' has its corresponding intransitive verb **aku**. In Yonaguni, *ag-iru<sup>N</sup>* 'open' does not have a corresponding intransitive. Instead, people use the passive form *ag-irar<sup>iN</sup>* 'be opened'; Yonaguni people always say 'How can a door open by itself?'. However, it is important to note that one of the perfect forms, *ag-u<sup>N</sup>*, corresponds phonetically to Japanese intransitive **aku**. Note that the negative forms of 'walk' and its causative 'make [someone] walk' are the same: *aigam<sup>N</sup>*. In causative sentences, another negative morpheme *-min<sup>uN</sup>* seems to be preferred.

#### 14.5.5.2 Causative and passive

Causative verbs are formed by adding *-m-iru<sup>N</sup>* to the stative stem, and passive verbs are formed by adding *-r-iru<sup>N</sup>* to the stative stem.

**TABLE 14.6 YONAGUNI INTRANSITIVE/TRANSITIVE VERBS: INFLECTION**

	conclusive	stative	imperative	indefinite	perfect
'walk'	<i>aig-u<sup>N</sup></i>	<i>aig-a-</i>	<i>aig-i</i>	<i>ait-i</i>	<i>ait-ja<sup>N</sup></i>
Transitive:	<i>aig-a-<sup>N</sup></i>	<i>aig-a-</i>	<i>aig-a-i</i>	<i>aig-a-si</i>	<i>aig-a-sja<sup>N</sup></i>
'get up'	<i>ug-iru<sup>N</sup></i>	<i>ug-ira-</i>	<i>ug-iri</i>	<i>ug-i</i>	<i>ug-u<sup>N</sup></i>
Transitive:	<i>ug-u-<sup>N</sup></i>	<i>ug-a-</i>	<i>ug-u-i</i>	<i>ug-u-si</i>	<i>ug-u-sja<sup>N</sup></i>

**TABLE 14.7** YONAGUNI INTRANSITIVE/TRANSITIVE VERBS: DERIVATIONi. Corresponding to ModJ *quinquagrade*, OJ *quadrigrade*

	‘walk’	‘fly’	‘cry’	‘stand’	‘enter’	‘take’
Intransitive	<i>aig-un</i>	<i>tub-un</i>	<i>nag-un</i>	<i>tat-un</i>	<i>haj-un</i>	<i>tur-un</i>
	‘make O walk’	‘fly O’	‘make O cry’	‘make O stand’	‘put O in’	‘give’
Transitive	<i>aig-a-N</i>	<i>tub-a-N</i>	<i>nag-a-N</i>	<i>tat-a-N</i>	<i>haj-a-N</i>	<i>tur-a-N</i>

ii. Corresponding to ModJ *monograde*, OJ *bigrade*

	‘get up’	‘fall’	‘get down’	‘peel off’	
Intransitive	<i>ug-irUN</i>	<i>ut-irUN</i>	<i>ur-irUN</i>	<i>hag-irUN</i>	
Perfect	<i>ug-un</i>	<i>ut-un</i>	<i>ur-un</i>	<i>hag-un</i>	
	‘awaken O’	‘drop O’	‘put O down’	‘peel O off’	
Transitive	<i>ug-un</i>	<i>ut-un</i>	<i>ur-un</i>	<i>hag-un</i>	
Perfect	<i>ug-u-sjaN</i>	<i>ut-u-sjaN</i>	<i>ur-u-sjaN</i>	<i>had-jaN</i>	
	‘burn’	‘go out’	‘break’	‘snap’	‘cut’
Intransitive	<i>mu-irUN</i>	<i>nd-irUN</i>	<i>bar-irUN</i>	<i>bur-irUN</i>	<i>cic-irUN</i>
Perfect	<i>mu-N/mw-a-N</i>	<i>nd-un</i>	<i>bar-un</i>	<i>bur-un/b-un</i>	<i>cic-un/cic-jaN</i>
	‘burn O’	‘take O out’	‘break O’	‘snap O’	‘cut O’
Transitive	<i>mu-ja-N</i>	<i>nd-a-N</i>	<i>bar-un</i>	<i>bur-un</i>	<i>c-un</i>
Perfect	<i>mu-ja-sjaN</i>	<i>nd-a-sjaN</i>	<i>b-a-N</i>	<i>bw-a-N</i>	<i>c-jaN</i>

iii. *Mixed type*

	‘[fire] be put out’	‘die’	‘stop’	‘hang down’
Intransitive	<i>sam-ar-un</i>	<i>nn-irUN</i>	<i>tum-ar-un</i>	<i>saŋ-ar-un</i>
Perfect	<i>sam-a-N</i>	<i>nn-un</i>	<i>tum-a-N</i>	<i>saŋ-u-N</i>
	‘put [fire] out’	‘kill O’	‘stop O’	‘hang O down’
Transitive	<i>sam-a-N</i>	<i>nn-a-N</i>	<i>tum-irUN</i>	<i>saŋ-irUN</i>
Perfect	<i>sam-a-sjaN</i>	<i>nn-a-sjaN</i>	<i>tum-jaN</i>	<i>saŋ-jaN</i>

- (153) a. *kag-am-irUN* (Causative) ‘I make [someone] write.’  
 b. *ut-ar-irUN* (Passive) ‘I am [someone is] hit.’  
 c. *ug-ar-irUN* (Passive) ‘I am [someone is] gotten up.’

Yonaguni speakers prefer to choose passive forms instead of intransitive verbs, because the latter might be misunderstood as perfect forms. Another construction is illustrated by *kat-i a-N* ‘It was written’, where the speaker sees the result of the action and the identity of the agent is unimportant.

#### 14.5.6 Speech levels and respect

Yonaguni has no polite forms of verbs, but indefinite + *juu* adds politeness to statements (24). Some verbs do have honorific forms, e.g. *b-un* ‘be’, *hir-un* ‘go’, *k-un* ‘come’ > *war-un*; *h-un* ‘eat’ > *uja-N*; *nnir-un* ‘die’ > *mais-un*. *war-un* is used as an honorific auxiliary.

- (154) a. *suu-ja* *kama-nu* *asa-ja* *war-anUN* *sujaa*  
 today-TOP over.there-GEN grandfather-TOP come.HON-NEG PART  
 ‘Today the grandfather of that house is not (has not come) here.’

- b. *uja-si*                      *war-i*  
 eat.HON-IND            be.HON-IMP  
 ‘Please eat it.’

### 14.5.7 Clause linking

There are conditional *-ba* and *-ja* pairs for each verb form at several different levels, and we observe interplay of modality and evidentiality that distinguish them.

- (A) Conclusive + *-ba/-iba* and conclusive minus *-u* + *-ja*  
 (B) Past + *-ba* and past + *-ja*  
 (C) Perfect stem + *-ru-ba*, + *-r-ja* and + *-ta-ja*  
 (+ *-ta-ba* is rare, and only one example occurs in our materials.)  
 (D) Negative form: *-nu-ta-ja*

(A) The conclusive minus *-u* + *-ja* is used to indicate situations where the action is going to take place or has begun and is still continuing. With the conclusive + *-ba/-iba*, the speaker assumes a relatively desirable situation. However, if the actor of the *ba*-form is the speaker, the whole action is almost fixed; that is to say, it is more an expression of the speaker’s will than an assumption as to what will happen. In the *ba*-form, the aspectual phase is also notable.

- (155) a. *ami hur-ja nma-nki ja* (It is raining now.)  
 rain fall-*ja* where-to Q  
 ‘Where are you going in spite of the rain?’
- b. *ami hur-ja unti b-irar-irun* (It is raining now.)  
 rain fall-*ja* potato plant-POT-CONC  
 ‘It is raining, so potatoes can be planted.’
- c. *ami hu-ru-ba unti b-irar-irun* (A rainy day is desirable.)  
 rain fall-CONC-*ba* potato plant-POT-CONC  
 ‘If it rains, we can plant potatoes.’
- d. An excursion will be held tomorrow: (A fine day is desirable.)  
*ami hu-ta-ja h-iraninun \*hu-ru-ba*  
 rain fall-PAST-*ja* go-POT.NEG  
 ‘If it rains, we cannot go (to the excursion).’
- (156) a. *cinmai nabi a-ta-ja ir-am-iri \*a-ru-ba*  
*chinmai* pot be-PAST-*ja* borrow-CAUS-IMP  
 ‘If you have a *chinmai* pot, let me borrow it.’
- Answer: b. *baja-ni a-r-ja k’a-i*  
 our.house-in be-CONC-*ja* use-IMP  
 ‘It’s in our house, so use it.’
- c. *baja-ni a-ru-ba ir-am-iru munu*  
 our.house-in be-CONC-*ba* borrow-CAUS-CONC thing  
 ‘If it were in our house, we would let you use it.’
- (157) a. Father and his son are sowing flower seeds. Father says to his son:  
*hana-ŋa sag-u-ba bu-i kuu joo \*sag-ja*  
 flower-NOM bloom-CONC-*ba* pick-IND come.IMP PART

‘When [you see the] flowers bloom, pick them and bring them.’ (The speaker knows and desires that flowers eventually bloom.)

- b. *hana-ŋa sag-ja muc-i-ŋa niguraruta s-ii juu*  
 flower-NOM bloom.CONC-*ja* insects-NOM frightfully come-IND PART  
 ‘Since the flowers bloomed, frightfully many insects have come.’
- (158) *hana-ŋa sat-i bu-r-ja bu-i kuu*  
 flower-NOM bloom-IND be-CONC-*ja* pick-IND come.IMP  
 ‘Since the flowers are in bloom, pick them and come back.’ (The speaker saw the flowers in bloom himself.)
- (159) Elder sister said to her younger sister in the morning:  
*anu ug-ir-ja nda-N ug-iri*  
 I get.up-CONC-*ja* you-also get.up-IMP  
 ‘I have gotten up, so you get up too.’
- (160) *a-ŋa tigami kag-u-ba nda mut-i h-iri joo \*kag-ja*  
 I-NOM letter write-CONC-*ba* you bring-IND go-IMP PART  
 ‘I will write a letter, then you bring it back with you.’
- (161) *nda-ŋa dum-ja imi bagar-u na*  
 you-NOM read.CONC-*ja* meaning understand-CONC Q  
 ‘If you read it, do you understand the meaning?’
- (162) *acuko ug-u-ba maduN su-i h-iri*  
 Atsuko awaken-CONC-*ba* together bring-IND go-IMP  
 ‘I will wake Atsuko up, so bring her with you.’
- (163) *kari-ŋa acuko ug-u-i-ba maduN su-i h-iri*  
 he-NOM Atsuko awaken-CONC-*i-ba* together bring-IND go-IMP  
 ‘He will wake Atsuko up, so bring her with you.’
- (164) *tura-i-ba tura-nuN tuna*  
 give.CONC-*i-ba* take-NEG QUOT  
 ‘(Someone) tried to give it to him, but he said “I’m not taking it.”’

(B) The past + *-ba* indicates a fixed fact which the speaker recognizes, whereas the past + *-ja* indicates a hypothetical situation.

- (165) *nnu ami hu-ta-ba nuci mu-i juu*  
 yesterday rain fall-PAST-*ba* life sprout.up-IND PART  
 ‘Since it rained yesterday, everything has been brought to life.’
- (166) a. *tumuti ug-ita-ba ta-N bu-ranu-taN*  
 morning get.up-PAST-*ba* who-even be-NEG-PAST  
 ‘When I got up in the morning, there was nobody in my house.’
- b. *tumuti ug-ita-ja nda t’ui-si ii ha-i joo*  
 morning get.up-PAST-*ja* you alone-with rice eat-IMP PART  
 ‘When you get up (tomorrow) morning, eat breakfast by yourself.’
- (167) a. *acuko ugu-ta-ba nur-i juu*  
 Atsuko awake-PAST-*ba* grumble-IND PART  
 ‘(I) woke Atsuko up, and then she grumbled.’

- b. *acuko ugu-ta-ja ii h-am-iri*  
 Atsuko awake-PAST-*ja* rice eat-CAUS-IMP  
 ‘When you have woken Atsuko up, let her eat breakfast.’
- (168) a. *muja-ta-ba mu-itan*  
 burn-PAST-*ba* burn-PAST  
 ‘(I) tried to burn it, and it burned.’
- b. *muja-ta-ja mu-iranu kaja*  
 burn-PAST-*ja* burn-NEG DOUBT  
 ‘I wonder whether it will burn if I try to burn it.’
- (169) *budi k-ita-ja ma-ru munu tura-N*  
 dance do-PAST-*ja* delicious-CONC thing give-CONC  
 ‘If you dance, I will give you delicious food.’
- (170) *sa-ŋa a-ta-ja husa-N sujaa*  
 tea-NOM be-PAST-*ja* desirable-CONC PART  
 ‘If there is tea, I would like to have it.’
- (171) *kat-ita-ba tu-i h-jun*  
 write-PAST-*ba* take-IND go-PERF  
 ‘I wrote it, and (someone) went out with it.’
- (172) *kari-ŋa tat-ita-ja nda-N tat-i joo*  
 he-NOM stand-PAST-*ja* you-also stand-IMP PART  
 ‘When he stands up, stand up, you also.’
- (173) *ug-uta-ja haja-gu ugu-si-du k-iru*  
 awake-PAST-*ja* early-ADV awake-IND-FOC do-CONC  
 ‘You woke me up, but it would be OK if you had awoken me earlier.’ (The speaker is thinking ‘You woke me up too late.’)

(C) Similarly, with the perfect.

- (174) *budi k-jaru-ba ma-ru munu tura-N*  
 dance do-PERF-*ba* delicious-CONC thing give-CONC  
 ‘If you finish dancing, I’ll give you a reward.’
- (175) *budi k-jar-ja ma-ru munu tura-i*  
 dance do-PERF-*ja* delicious-CONC thing give-IMP  
 ‘She finished dancing, so give her a reward.’
- (176) a. *kari-ŋa kat-jaru-ba nsa-ta-ŋa jee*  
 he-NOM write-PERF-*ba* good-PAST-*ŋa* PART  
 ‘I wish he had written it.’
- b. *a-ŋa kat-jar-ja nda-N kag-u na*  
 I-NOM write-PERF-*ja* you-also write-CONC Q  
 ‘I have written (a letter). Do you add a word?’
- (177) *anu ug-ur-ja madun su-i h-iri*  
 I get.up-PERF-*ja* together bring-IND go-IMP  
 ‘I just got up so bring me together with you.’

- (178) a. *acuko ugu-sjaru-ba su-i h-irar-ita-ŋa*  
 Atsuko awake-PERF-*ba* bring-IND go-POT-PAST-*ŋa*  
 ‘If (I/someone) had awakened Atsuko, I could have brought her with me.’
- b. *acuko ugu-sjar-ja su-i h-iri*  
 Atsuko awake-PERF-*ja* bring-IND go-IMP  
 ‘As I awoke Atsuko, bring her with you.’
- (179) *hana sat-ur-ja nn-i-ndi kuu*  
 flower bloom-PERF-*ja* see-IND-PURP come.IMP  
 ‘Flowers are in bloom so come to see them.’
- (180) a. *sama-sjaru-ba nsa-ta-ŋa jee*  
 extinguish-PERF-*ba* good-PAST-but PART  
 ‘I wish you had put the fire out.’
- b. *cii sama-sja-ta-ja (/ sama-sjar-ja) h-ii nsa-N*  
 fire extinguish-PERF-(PAST)-*ja* go-IND good-CONC  
 ‘If you have put out the fire, you may go out.’
- c. (*cii*) *samar-ja maa h-ii nsa-N*  
 (fire) go.out.CONC-*ja* now go-IND good-CONC  
 ‘In the case that the fire is out, you may go out.’ (CONC-*ja* and PERF-*ja*  
 of the intransitive *samar-un*, PERF *sam-an*, have the same forms.)
- (181) *dadu ag-ur-ja nn-i-ti kuu*  
 door open-PERF-*ja* see-IND-CONJ come.IMP  
 ‘The door opened, so go and see what happened.’ (The speaker felt the wind.)

(D) Negative form *-anu-ta-ja*:

- (182) *hur-anu-ta-ja h-irar-irun*  
 rain-NEG-PAST-*ja* go-POT-CONC  
 ‘If it does not rain, we can go.’

Like the verbs, the adjectives have conjunctive forms.

- (183) *kuma taga-r-ja nu-i-gurisa-nu*  
 here high-CONC-*ja* climb-IND-difficult-JUDG  
 ‘Since this place is high up, it is difficult to climb to.’
- (184) *taga-bi k-anun / k-anu-tan*  
 high-because buy-NEG / buy-NEG-PAST  
 ‘It is expensive, so I won’t buy it.’/‘It was expensive, so I didn’t buy it.’
- (185) a. *daŋca-ru-ba k-uu munu . . .*  
 inexpensive-CONC-*ba* buy-CONC thing  
 ‘If it were inexpensive, I would buy it.’
- b. *daŋca-ta-ja k-u-ŋa taga-ta-ja k-anun*  
 cheap-PAST-*ja* buy-CONC-but high-PAST-*ja* buy-NEG  
 ‘If it’s cheap, I’ll buy it, but if it’s expensive, I won’t.’
- c. *ma ibitati hudu-ŋa taga-(ta-)ru-ba nsa-ru munu*  
 some more body-NOM high-(PAST)-CONC-*ba* good-CONC thing  
 ‘I wish I were a little taller.’

- d. *nda sjana-ta-ja budi k-iri*  
 you happy-PAST-*ja* dance do-IMP  
 ‘Dance, if you are happy.’
- e. *kazu-ja sjana-ta-ja budi k-irun sa*  
 Kazu-TOP happy-PAST-*ja* dance do-CONC PART  
 ‘If Kazu feels happy, she will dance. I know.’

Conditional forms + *-N* ‘also; even’ create concessive forms:

- (186) *kag-ja-N kag-aninun*  
 write.CONC-*ja*-even write-POT.NEG  
 ‘I tried to write it, but I couldn’t.’
- (187) *i-r-ja-N i-raninun*  
 do.CONC-*ja*-even do-POT.NEG  
 ‘I couldn’t do it even though I tried.’

Other clause-linking devices include: indefinite + *-ti* ‘and (then)’, negative *N*-form + *-ki* ‘not . . . and; without . . . ing’; indefinite + *-ndi* ‘in order to’; indefinite (or adj. root) + *-bi*, conclusive (*N*-less) + *jungara*, conclusive (in *-N*) + *gara* ‘because’; conclusive (*N*-less) + *ɲa* ‘and then, but’.

#### 14.6 LEXICON: CASE STUDY OF KINSHIP TERMS

Yonaguni dialect is very rich in kinship terms. Most are used as address terms, even to persons who are not related by birth. ○ indicates that any speaker can use a term as an address form.

<i>Term</i>	<i>Kin use</i>	<i>Non-Kin use</i>
○ <i>umi asa</i>	great-grandfather	To address a very elderly man.
○ <i>umi abu</i>	great-grandmother	To address a very elderly woman.
○ <i>asaa</i>	grandfather	To address an elderly man.
○ <i>abuu</i>	grandmother	To address an elderly woman.
<i>butu</i>	husband	Does not function as an address term.
<i>tun</i>	wife	Does not function as an address term.
<i>huga tun</i>	mistress (lit. ‘another wife’)	
<i>mugu</i>	son-in-law	
	e.g. <i>baja-nu mugu</i> : ‘our son-in-law’	
<i>dumi</i>	daughter-in-law	
	e.g. <i>baja-nu dumi</i> : ‘our daughter-in-law’	
○ <i>ija</i>	father	To address a middle-aged man.
○ <i>abuta</i>	mother	To address a middle-aged woman.
<i>atabuta</i>	relative’s wife (lit. ‘the following mother’)	Addressed by relatives.
○ <i>ami</i>	a (middle-aged) lady	Addressed by anybody.
<b>Uncles</b>		
○ <i>ubu-ija</i>	○ <i>ubu-abuta</i>	
○ <i>naga-ija</i>	○ <i>naga-abuta</i>	
○ <i>naga-ijati</i>	○ <i>naga-abutati</i>	
○ <i>ijati</i>	○ <i>abutati</i>	
○ name + <i>ijati</i>	○ name + <i>abutati</i>	

*ijati* and *abutati* refer to uncles and aunts younger than the speaker's father or mother, and name + *ijati* and name + *abutati* mean uncles and aunts who are younger than *ijati* or *abutati*. The other three terms refer to uncles (or older sisters' spouses) and aunts (or older brothers' spouses) older than the speaker's father or mother in order of their birth.

- *ubai(ti)* aunt who is younger than the speaker's parents, but this word is not much used
- *ami* middle-aged woman who is not related by birth

### Nephews and nieces

- buiha* nephew/niece  
*buihati* very young nephew or niece  
*bjuka* nephew Nephew who is older than the speaker.  
*bjukati* nephew Nephew who is younger than a *bjuka*, but still older than the speaker.  
*bunti* nephew Nephew who is younger than the speaker.  
*buntiti* nephew Nephew who is younger than a *bunti*.  
*mimani* niece Niece who is older than the speaker. (Ikema 1998)

### Brothers and sisters

Address forms used by brothers and sisters vary according to the sex of the speaker. (Examples in the case of seven children.)

- (A) Words for older sisters used by both boys and girls:  
 1st-born 2nd-born 3rd-born 4th-born 5th-born 6th-born 7th-born  
*ub-ani nag-ani nag-ani-ti ubu-ati naga-ati guma-ati ati*
- (B) Words for older brothers used only by boys:  
*ub-uda nag-uda nag-uda-ti ubu-sunati naga-sunati guma-sunati sunati*
- (C) Words for older brothers used only by girls:  
*u-bja naga-bja naga-bja-ti ubu-bja-ti naga-bja-ti guma-bja-ti bjaati*
- (D) Older people call younger people by their personal names.

Examples of kinship terms that cannot be used as address terms:

- (188) Q. *nma-nu agamiti ja*  
 where-GEN child Q  
 'Whose child is he?'  
 A. *sakibara-nu buihati ~ baja-nu buihati*  
 Sakibara-GEN nephew our.home-GEN nephew  
 'He is the Sakibaras' nephew.' ~ 'He is our nephew.'
- (189) Q. *kanu t'u-nta-ja nma-nu t'u kaja*  
 those person-PL-TOP where-GEN person Q  
 'Those people, where are they from?'  
 A. *kama-nu ija.abuta do*  
 over.there-GEN couple PART  
 'They are the couple from over there.'

## NOTES

- \* In completing this chapter regarding Yonaguni dialect in Luchuan, I would like to thank University of Tokyo graduate student Mark Rosa, who translated the Yonaguni examples into English, formatted the accompanying gloss, and edited and proofread the descriptive text. He also prepared Section 14.3 and the Appendix and its chart of kaida characters which came from his Master's thesis. Also, I would like to thank the speakers from Yonaguni named below.

Sonai:	Ikema, Nae (b. 1919)	Kikuyama, Miyo (b. 1926)
	Sakihara, Toshi (b. 1924)	Sakihara, Kinu
	Sakihara, Hatsuko	Shinjou, Yoshimi (b. 1945)
	Miyara, Hozen (b. 1918)	Miyara, Setsu (b. 1924)
Higawa:	Maecawakura, Onari (b. 1906)	Tomari, Natsu (b. 1916)

I would also like to thank the family of Ms. Toshi Sakihara, from whom most of the examples were drawn. She was born in Sonai, and she moved to mainland Japan when she was 16 years old. In 1945, she came back to Sonai and lived there, but in the 1980s she moved to the suburbs of Tokyo where she lives to this day. Her husband, who was also from Sonai, died several years ago. Her daughter understands Yonaguni dialect, but only speaks it on rare occasions.

- 1 Forms with gemination in this chapter reflect ModJ influence, either loans (*gaoku* < *gaokō*, *daokyu* < *raokyo*) or younger speakers' reinterpretation of medial glottalized consonants (*ciqka* < *cika*).
- 2 *ŋasa* is in Ikema (1998), but was not known to my informants.

## REFERENCES AND FURTHER READING

- Aikhenvald, Alexandra Y. (2003) 'Evidentiality in typological perspective', in Alexandra Y. Aikhenvald and R. M. W. Dixon (eds) *Studies in Evidentiality*, Amsterdam: John Benjamins.
- (2004) *Evidentiality*, Oxford: Oxford University Press.
- Chamberlain, Basil Hall (1895) *Essay in Aid of a Grammar and Dictionary of the Luchuan Language*, Transactions of the Asiatic Society of Japan, vol. XXIII, Tokyo: Z. P. Maruya & Co.
- Hattori Shirō and Kinjō Chōei (1955) 'Ryūkyū-go' [Ryūkyū language], in Ichikawa Sanki and Hattori Shirō (eds) *Sekai-gengo Gaisetsu* [An introduction to the languages of the world], vol. 2, Tokyo: Kenkyūsha, 307–56.
- Hirayama Teruo and Nakamoto Masachie (1964) *Yonaguni Hōgen no Kenkyū* [Research on Yonaguni dialect of Luchuan], Tokyo: Tōkyō-dō.
- Hirayama Teruo, Oshima Ichiro, and Nakamoto Masachie (1967) *Ryūkyū Sakishima Hōgen no Sōgōteki Kenkyū* [Research on *Sakishima* dialects of Luchuan], Tokyo: Meiji Shoin.
- Ikema Eizō (1959) *Yonaguni-no Rekishi* [History of Yonaguni], Yonaguni: Privately printed.
- Ikema Nae (1998) *Yonaguni Kotoba Jiten* [Yonaguni language dictionary; Yonaguni-Japanese], Yonaguni: Privately printed.
- (2003) *Yonaguni-go Jiten* [Yonaguni language dictionary; Japanese-Yonaguni], Yonaguni: Privately printed.
- Iwakura Ichirō (1941) *Kikai-jima Hōgen-shu* [Word book of Kikai-jima], Tokyo: Chuō Kōron.
- Izuyama Atsuko (1992) 'Ryūkyū-hōgen no 1-ninshō daimeishi' [First person pronoun in Luchuan language], *Kokugogaku*, 171: 106–24.

- (1997) ‘Ryūkyū hōgen keiyōshi seiritsu no shiteki-kenkyū’ [A historical study on the adjectives in Luchuan dialects], *Journal of Asian and African Studies*, 54: 92–3.
- (2002a) ‘Ryūkyū Yaeyama (Yonaguni) hōgen no bunpō kiso-kenkyū’ [A study on the grammar of Yaeyama-Yonaguni dialect], in Sanada Shinji (ed.) *Grammatical Aspect of Endangered Dialects in Japan (2)*, Kyoto: ELPR Publication A4-012.
- (2002b) ‘Evidentiality on Yaeyama Dialect in Luchuan’, *Mathesis Universalis*, 4(1): 43–55.
- (2005) ‘Grammaticalization of \*i- “to do” (intransitive) in Luchuan’, *Nihongo no Kenkyū* [Research of Japanese], Tokyo: Musashino Shoin, 155–70.
- (2006a) ‘Shōko-yōtai Yonaguni-hōgen no ba’ ai’ [Evidentiality – the case of Yonaguni dialect of Luchuan], *Mathesis Universalis*, 7(2).
- (2006b) ‘Evidentiality – Ryūkyū-go no ba’ ai’ [Evidentiality – in the case of Ryukyuan], *Research on the Process and Varieties of the Grammatical Forms*, National Institute for Japanese Language, Publication series (B).
- Jōdai-go Jiten Henshu Iinkai [Dictionary of ancient Japanese] (eds) (1967) *Jidai betsu kokugo dai-jiten* [Dictionary of ancient Japanese], Tokyo: Sanseido.
- Kajiku Shin-ichi (2002) ‘On Yonaguni Dialect’, in Karimata Shigehisa, Tshako Toshiko, Kajiku Shinichi, and Takahashi Toshizo (eds) *Preliminary Research on Endangered Ryūkyūan Language*, Tokyo: Ministry of Education, 263–280.
- Kamei Takashi, Kōno Rokurō and Chino Ei’ichi (eds) (1992) *Gengo-gaku dai-jiten dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido.
- Kawamura Tadao (1939) *Nantō-Bunka no Tankyū* [Research on the culture of the southern islands], Okinawa: Okinawa Bunkyo Shuppan.
- Miyara Tosō [Miyayama Masamori] (1930) *Yaeyama-go’i* [Yaeyama vocabulary with general remarks], Tokyo: Tōyō-Bunko.
- Moseley, Christopher (ed.) (2010) *Atlas of the World’s Languages in Danger*, 3rd ed., Paris: UNESCO Publishing. (Online version *UNESCO Interactive Atlas of the World’s Languages in Danger*, <http://www.unesco.org/culture/languages-atlas/>)
- Nakamatsu Takeo (1987) *Ryūkyū Hōgen Jiten* [Dictionary of Luchuan dialect], Naha: Naha Shuppan.
- National Language Research Institute (1963, ninth printing 2001) *Okinawa-go Jiten* [Dictionary of the Okinawan language], NLRI materials 5, Tokyo: Ministry of Finance Printers.
- Rosa, Mark (2006) ‘Kaida Writing in the Yaeyama Islands’, Master’s thesis, Linguistics Department, University of Tokyo.
- Shibata Takeshi (1959) ‘Ryūkyū Yonaguni-jima hōgen no on’in ni tsuite’ [Phonology of Yonaguni dialect of Luchuan], *Kotoba no Kenkyū-1* [Language research-1], Tokyo: National Institute for Japanese Language, 103–20.
- (1972) *Zenkoku Hōgen Shiryō* [Materials of Japanese dialects] vol. 11, Luchuan-2, NHK.
- Takahashi Toshizō (1986–1987) ‘Yaeyama and Yonaguni-jima’ [Dialects of Luchuan 11–12], Tokyo: Okinawa Bunka Kenkyūjo, Hōsei University.
- (1992) ‘Yonaguni Hōgen’ [Yonaguni dialect], in Kamei, Takashi, Kōno, Rokurō, and Chino, Ei’ichi (eds) (1992) *Gengo-gaku Dai-Jiten Dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido, 873–82.
- Uemura Yukio (1992) ‘Ryūkyū rettō no gengo’ [Languages of the Ryukyus], in Kamei, Takashi, Kōno, Rokurō, and Chino, Ei’ichi (eds) (1992) *Gengo-gaku Dai-Jiten Dai-4-kan* [Sanseido encyclopedia of linguistics, vol. 4], Tokyo: Sanseido, 771–814.
- (2003) *The Ryūkyūan Language*, Endangered Languages of the Pacific Rim A4-018 (Translated by Wayne P. Lawrence), Tokyo: Ministry of Education.

## APPENDIX: CHART OF KAIDA CHARACTERS (PREPARED AND HAND-WRITTEN BY MARK ROSA)

The written symbols in Table 14.8, called *kaida-di* (kaida characters), were used until the beginning of the twentieth century to record tax payments, transactions, and various other items (see 14.3). Here, they are presented in approximately the order in which they were first recorded by previous researchers, such as Eizō Ikema.

The character *inokin* was recorded by Gisuke Sasamori in 1893; its meaning is unknown, as is *koroji* (which does not fit into Yonaguni phonology with that pronunciation, but may be related to ModJ **kōrogi** ‘cricket’). *Funa* is a kind of fish. One *jō* is a length equal to approximately six English feet. A *furumami* is a kind of bean.

The geometric characters at the end of the list, called *sucuma*, were used to indicate volumes of liquid or dry goods, and were either repeated or followed by kanji-like numerals to indicate more than one of a given measure. Kanji-like numerals could also be appended to characters for animals to indicate numbers of livestock. The distinction between male animals (heavier bodies) and female (thin bodies) is particularly noteworthy. For actual examples of use, see Rosa (2006), from which this chart is adapted.

TABLE 14.8 YONAGUNI KAIDA CHARACTERS

Kaida	English	Yonaguni
	‘rice’	<i>mai</i>
	‘millet’	<i>aa</i>
	‘bean’	<i>mami</i>
	‘saké’	<i>sagi</i>
	‘barley’	<i>mun</i>
	‘fish’	<i>iju</i>
	‘shrimp’	<i>tarunja</i>
	‘funa’	<i>funa</i>
	‘firewood’	<i>timunu</i>
	‘board’	<i>ita</i>
	‘jō’	<i>nanji</i>
	‘bamboo’	<i>tagi</i>
	‘ <i>inokin</i> ’	(unknown)
	‘ <i>hechima</i> gourd’	<i>nabira</i>
	‘ <i>tōgan</i> gourd’	<i>cibui</i>
	‘thatch’	<i>kaja</i>
	‘cricket?’ (‘ <i>koroji</i> ’)	(unknown)
	‘canoe’	<i>itanni</i>
	‘yanbaru boat’	<i>maranni</i>
	‘ <i>furomame</i> ’	<i>furumami</i>
	‘burdock’	<i>gunbu</i>
	‘vegetable’	<i>nanuha</i>

TABLE 14.8 (cont'd)

Kaida	English	Yonaguni
	'garlic'	<i>hiru</i>
	'pot'	<i>cibu</i>
	'handled pot'	<i>mincibu</i>
	'person'	<i>t'u</i>
	'mare'	<i>miinma</i>
	'(young) horse'	<i>biginma</i>
	'goat'	<i>hibida</i>
	'bird/rooster'	<i>bigimita</i>
	'egg'	<i>kaiju</i>
	'cow'	<i>miuci</i>
	'pig'	<i>wa</i>
	'cat'	<i>maju</i>
	'dog'	<i>inu</i>
	'carrot'	<i>kindaguni</i>
	'scallion'	<i>daoku</i>
	'pumpkin'	<i>nankuu</i>
	' <i>daikon</i> radish'	<i>ubuni</i>
	'sweet potato'	<i>tumaikuru</i>
	'peanut'	<i>dimami</i>
	'oil'	<i>anda</i>
	'ox'	<i>bigiuci</i>
	'hen'	<i>miimita</i>
	'mountain'	<i>dama</i>
	'straw rope'	<i>baranna</i>
	'water'	<i>min</i>
	'potato'	<i>unti</i>
	'rice plant'	<i>nii</i>
	'grain'	<i>mainuhi</i>
	'red bean'	<i>agamami</i>
	'soy sauce'	<i>suju</i>
	'vinegar'	<i>hairi</i>
	'miso'	<i>nsu</i>
	'house'	<i>da</i>
	'grave'	<i>haga</i>
	'salt'	<i>masu</i>
	'watermelon'	<i>cioka</i>
	'Japanese onion'	<i>cinda</i>
	'black rope'	<i>anquinna</i>

TABLE 14.8 (cont'd)

Kaida	English	Yonaguni
	'octopus'	<i>tagu</i>
	'fabric'	<i>nunu</i>
	'scale'	<i>hagai</i>
	'paper'	<i>kabi</i>
	'straw rice bag'	<i>tara</i> (cf. 14.4.3)
	'kudzu yam'	<i>sungaru</i>
	'konbu/kelp'	<i>kubu</i>
	'noodles'	<i>sumin</i>
	'hyō' = 45 l	<i>tara</i> (cf. 14.4.3)
	'to' = 18 l	<i>tu</i>
	'shō' = 1.8 l	<i>cu, bagaci</i>
	'gō' = 180 ml	<i>gu</i>
	'shaku' = 18 ml	<i>sagu</i>
	'sai' = 1.8 ml	(rarely used)



PART V

**AINU**



# SOUTHERN HOKKAIDO AINU\*

*Anna Bugaeva*

## 15.1 INTRODUCTION

### 15.1.1 Genetic, typological, and dialectal profile of Ainu

The genetic affiliation of the Ainu language is unknown. Ainu is agglutinating, polysynthetic, and incorporating, with SOV constituent order and a mixed (but basically tripartite) alignment. It is predominantly head-marking and prefixing.

In the past, the Ainu (their self-name meaning ‘person’), traditionally hunter-gatherers, occupied not only Hokkaido but also a considerable part of the island of Honshu until the middle of the eighteenth century, the Kurile Islands until the beginning of the twentieth century, the southern part of Sakhalin until the middle of the twentieth century, and presumably the southern part of Kamchatka. The three primary divisions are geographically based, and distinguish between the dialects once spoken on Hokkaido, Sakhalin, and the Kurile Islands. Sakhalin and the Kuriles form part of the Russian Federation today, with Southern Hokkaido being the last autochthonous location of a couple of native speakers/rememberers of the language.

The Kurile dialects are different from both Sakhalin and Hokkaido dialects and there are almost no data on them, except for a word list in Torii (1903) and a few other sources. The Sakhalin dialects can be roughly divided into east and west coast groups (SEC and SWC), and Hokkaido dialects into northeastern (HNE: northern, eastern, and central) and southwestern (HSW: southern and southwestern) groups, which are further subdivided into local sub-dialectal forms, see Hattori (1964: 18) and Asai (1974). The Taraika dialect once spoken in the northeast of Sakhalin is different from all other Sakhalin dialects but it shares some common features with southern Hokkaido dialects. Murasaki (2009: 1) notes that there is no mutual intelligibility between any Sakhalin and Hokkaido dialects of Ainu. According to Hattori (1964: 19), the number of cognates between Biratori (Southern Hokkaido Ainu) and Raichishka (Western Sakhalin Ainu) is 73.8 percent, which means that the respective dialects are not only closer than Tokyo and Miyako (59 percent) or Tokyo and Shuri (66 percent) varieties of Japonic, but also closer than Miyako and Shuri Ryukyuan (72 percent) (Hattori 1959: 228).

At present, Ainu should be best classified as a language-isolate which is presumably a remnant of some very old language family. There were attempts to relate Ainu to such language families as Indo-European (Batchelor 1889; Naert 1958; Lindquist 1960),<sup>1</sup> Austronesian (Gjerdman 1926; Murayama 1992, 1993), Altaic, as particularly belonging to the Japanese-Korean subgroup being closer to Korean (Patrie 1982), or to individual languages such as Japanese (Hattori 1959), Nivkh (Austerlitz 1976), Eskimo, and Basque, but none of them presented convincing evidence arrived at by a consistent application of the comparative method. The biggest problem is that for the lack of old records, in

most cases, the comparison is made between modern languages. Vovin (1993) is the only serious attempt to reconstruct proto-Ainu, i.e. a language spoken in the last centuries of the first millennium AD. This work is based on the data of all three dialect groups, including those from early Japanese xylographic and manuscript dictionaries of Ainu, such as *Moshiogusa* [Seaweeds] (1792) and others.<sup>2</sup> Furthermore, Vovin (1993) ventures a hypothesis that proto-Ainu may be related to proto-Austroasiatic.<sup>3</sup> Leaving apart the discussion of the validity of this particular hypothesis, I would like to emphasize that Vovin's (1993: 158) general suggestion "to look for the origins of the Ainu language in the southern direction," which is based on his reconstruction of proto-Ainu phonology, sounds plausible. It might not be accidental that Ainu also shows certain typological characteristics of southern languages, such as an extensive use of prefixation, which is fairly uncommon in the languages of the North Pacific Rim (Nakagawa 2009). From the point of view of vocabulary, Ainu definitely represents a kind of mixed language, with several unrelated strata inside it, including those which are due to neighboring, northern languages, such as Japanese and Nivkh (Vovin 1993: 158).

Archaeological finds suggest that the Ainu were the creators of the Jōmon culture in Japan which started between 6000 and 10 000 BC (and terminated around 200 BC in Honshu and 500 AD in Hokkaido) and thus have spoken a Jōmon language. However, it still remains unclear whether a predecessor of Japanese has also at some point been spoken as a Jōmon language related or unrelated to the Ainu stock (ICHL20 International Symposium). It is widely recognized that the comparative method, which is the only reliable tool in historical linguistics, may hardly be applied when we are dealing with a time depth of more than 6,000 years.

### 15.1.2 Sociolinguistic situation

Ainu was a spoken language until the 1950s; now all the Ainu speak Japanese in everyday life.

Traditionally, the Ainu settled along rivers toward the inland, engaged in fishing, hunting, and gathering. They lived in harmony with nature and believed every natural phenomenon to be a *kamuy* 'god/spirit',<sup>4</sup> which is typical for an animistic religion.

Intensive contacts of the Ainu and Japanese started in 1604 when the Japanese warrior clan Matsumae received Hokkaido (then called Ezo) from shogun Ieyasu as a fief and monopolized trade with the Ainu. In those days, the Ainu were not allowed to learn Japanese and Japanese–Ainu communication was carried out by specially trained Japanese translators. Thus, the Ainu remained monolingual in Ainu till the end of the eighteenth century.

For Russia, the eighteenth century was a golden age of science and sailing, which among other achievements resulted in Russia's exploration of the Far East and Kurile Islands, eventually threatening Hokkaido. As a response to that, in 1799, the central government of Japan assumed direct control over eastern Hokkaido (and later over the whole island) and started an assimilation policy for the Ainu, encouraging them to learn Japanese. This period is regarded as the beginning of language shift (Refsing 1986: 63).

In 1899, the Japanese parliament enacted the *Hokkaido Former Aborigines Protection Act*. Originally, this law was designed to help the Ainu to adapt within modernizing Japan and to merge them with the ethnic Japanese. In practice, the law significantly contributed to the progress of language shift and collapse of traditional Ainu lifestyle

due to the introduction of obligatory elementary schooling and other measures, but failed to achieve the full assimilation of the Ainu, because the mere fact of singling the Ainu out as former aborigines became the source of prejudice and discrimination against them within Japanese society. All these led to the rapid abandonment of language and its eventual loss by succeeding generations: in the early part of the twentieth century language shift completed (Refsing 1986: 63).

The exact number of ethnic Ainu is unknown because questions about ethnicity are not included in Japanese censuses. According to the survey of actual living conditions of the Ainu held by Hokkaido Government, Department of Health and Welfare in 2006, the number of people in Hokkaido who identified themselves as Ainu was 23,782, which is probably only half of the real number in Hokkaido. There is also a considerable Ainu population on Honshu, concentrated mainly in the Kanto area (about 10,000); thus the overall number of ethnic Ainu in present-day Japan is likely to reach 100,000.

Since the Ainu experienced severe ethnic and linguistic repression from the state in the beginning of the twentieth century, for years many of them have tried to conceal their Ainu origins even from their own children. However, the attitude of many Ainu toward their native culture and language has changed to positive after the official adoption of the *Law for the Promotion of the Ainu Culture and for the Dissemination and Advocacy of the Traditions of the Ainu and the Ainu Culture* (1997) and the official recognition of the Ainu as the indigenous population of Hokkaido (2008).

Established in accordance with this law was the Foundation for Research and Promotion of Ainu Culture which is, among other promotion activities, running 14 Ainu language schools across Hokkaido in all regions with a high concentration of Ainu population and one school in Tokyo at the Ainu Culture Center. Recently, more and more Ainu are experiencing a new sense of self-identity and becoming aware of the importance of revitalizing their language. There is a strong demand for community-oriented textbooks for the practical learning of Ainu, which are still few,<sup>5</sup> viz. Hokkaidō Utari Kyōkai ed. (1994), Nakagawa and Nakamoto (1997), and Nakagawa (2007); see also an online talking dictionary of Ainu which is a first attempt of this kind aimed to support language revitalization efforts (Bugaeva and Endō 2010).

It is worth mentioning that only ethnic Ainu and their family members are allowed to enroll in the above-mentioned Ainu language schools. There is little chance for ethnic Japanese or other nationals to learn Ainu, unless one gets enrolled as a student in one of the few Japanese universities which have courses on Ainu in their curriculum.<sup>6</sup> For a general audience the only chance to learn practical Ainu is a 15-minute weekly radio programme “Ainugo Rajio Kōza” presented by STV Radio in Hokkaido.<sup>7</sup>

### 15.1.3 Previous description and research of Ainu

Ainu was not a written language, but the first attempts to transcribe it with the Latin alphabet or Japanese katakana go back to the seventeenth century. The earliest word list (54 words) was compiled by a Jesuit missionary, Jeloramo de Angelis, who attended Hokkaido in 1618 and 1621 and sent a report with a word list to the pope (*Relatione del Regno di Iezo, Relatione di alcune cose*, Milan, 1625). At about the same time the earliest Japanese written record of Ainu, *Matsumae no Kotoba* [Words of Matsumae] (117 words), was compiled, which presumably belongs to one of the eastern Hokkaido dialects (Satō 2008: 6); the author, year and place of compilation are unknown (published in Meiji-mae nihonkagakushi kankōkai ed. 1971).

Since then several dozens of word lists have been compiled by Japanese and Europeans<sup>8</sup> but it was not till the beginning of the twentieth century that any sizable bulk of texts had been written down. Transcribing Ainu texts is associated with the next stage of Ainu research, i.e. linguistic research of Ainu, which was pioneered by the Japanese researcher Prof. Kyōsuke Kindaichi, who published the first academic grammar of Ainu (1931) based on his collected texts.

Now more than a century has passed since linguistic research of Ainu began, and this research has produced a few comprehensive dictionaries: Batchelor (1889 [1938, 1995]), M. Chiri (1975, 1976 [1953, 1954, 1962]), Hattori (1964), Nakagawa (1995), Tamura (1996), and Kayano (1996), several Ainu grammars/grammar outlines of Sakhalin (Murasaki 1976: 3–9; M. Chiri 1942) and Hokkaido dialects of Saru (Kindaichi 1931, Tamura 1988), Saru vs. Horobetsu (M. Chiri 1936), Horobetsu (Kirikae 2003), Ishikari (Asai 1969), Shizunai (Refsing 1986), and Chitose (Bugaeva 2004; Satō 2008), and numerous articles on various grammatical phenomena.<sup>9</sup> Yet, none of those grammars is complete, as we are still at a rather early stage of Ainu research.

Ainu folklore is extremely rich: there are at least three major genres, viz. *yukar* ‘heroic epics’, *kamuy yukar* ‘songs of gods’, and *uwepeker* ‘folktales’ (for a more elaborate classification see Nakagawa 1997). Many folklore texts of Hokkaido and Sakhalin Ainu were recorded by Japanese and European researchers, viz. Kannari and Kindaichi (1993 [1959–1975]; Horobetsu HSW), M. Chiri (1937; Horobetsu HSW), Kubodera (1977; Saru HSW), Tamura† (1984–2000; Saru HSW), Nevskij (1972; Saru HSW) – available in Japanese in Uoi (1991), Kayano† (2005 [1974], 1998; Saru HSW), Satō (1995–1998; Saru HSW), Nakagawa (1998–2010; Chitose HSW), Bugaeva (2004; Chitose HSW), Bugaeva (2007; Mukawa HSW), Kirikae (1996; Tokachi HNE), Okuda (1991–1995; Shizunai HNE), Shitaku Yae no Denshō Kankōkai ed.† (2007, 2011; Shiranuka HNE), Pilsudski (1912; SEC), Murasaki† (1976: 3–9, 2001; Raichishka SWC), etc. There are also a few examples of texts written down by the Ainu introspectively, viz. a wonderful collection of *kamuy yukar* by Y. Chiri (1923; Horobetsu HSW) – available in English in Katayama (2003) and Strong (2011) – and an autobiography by Sunazawa (1983; Ishikari HNE).

In fact, Ainu is a relatively well-documented language compared with other endangered languages. However, the problem is that most Ainu documentation focused on recording folklore texts and there are only very few conversational texts left, viz. Tamura (T3) and Murasaki (1976: 3–9), and unfortunately nothing can now be done about it.

Another problem is that Ainu texts are rarely accompanied by audio files (texts with audio are marked with a †). Therefore, generally, there is a very strong demand for Ainu texts of a ‘new generation’, i.e. fully glossed and annotated texts with both Latin and *katakana* transcriptions of Ainu, Japanese, and English translations, and attached audio/video files. A good example of this is Nakagawa and Bugaeva (forthcoming; Saru HSW). It should be noted that there are still a lot of unpublished old audio recordings of Ainu folklore in private collections in Japan which are awaiting talented transcribers.

### 15.1.4 The present description of Southern Hokkaido Ainu

This chapter is an overview of phonology and morphosyntax of Southern Hokkaido Ainu (southwestern group), viz. Saru and Chitose dialects. I have chosen this particular group of dialects because they are much better described and documented than any other Ainu dialects, which may be explained by the availability of native speakers in the second half of the twentieth century and the beginning of the twenty-first. It should

be noted that these dialects show a few grammatical features which are hardly representative of “Ainu” as a whole;<sup>10</sup> I will specially comment on these features. For Chitose, I will use my own field data (OI; Bugaeva 2004) and those published by Nakagawa (1995, 2000–2011) and Satō (2008). For Saru, I will use the data from Nakagawa and Bugaeva (forthcoming), Tamura (1984–2000, 1988/2000, 1996), Kubodera (1977), and other sources. Data from specialist literature are also used (see Sources). The difference between Chitose and Saru dialects is not great, but I will provide a note where it is of any significance to the description. Reference to other Ainu dialects will be made where appropriate.

## 15.2 PHONOLOGY

### 15.2.1 Consonants

Just like all other Ainu dialects, Southern Hokkaido Ainu has 12 consonant phonemes /p/, /t/, /k/, /c/, /s/, /r/, /m/, /n/, /w/, /y/, /h/, /ʔ/ (Table 15.1). All of them, except the glottal stop /ʔ/, are used in the commonly accepted phonemic transcription of Ainu such as that of Tamura (1988). The glottal stop /ʔ/ (also represented as /ʔ/) is conventionally omitted in writing because its occurrence is predictable from the environment: it occurs syllable-initially if there is no other initial consonant; there is no full agreement on the phonological status of the glottal stop.

/p/, /t/, /k/ are realized as regular stops in syllable initial (onset) position, but as unreleased [p̚], [t̚], [k̚] in syllable-final (coda) position, e.g. [kap̚] ‘skin/bark’, [kut̚] ‘girdle’, [sik̚] ‘eye’, and often (but not necessarily) as voiced [b], [d], [g] in intervocalic position, and after nasals /m, n/ or the tap /r/, e.g. *sampe* [sambe] ‘heart’ and *kirpu* [kirbu] ‘fat’. As is clear, there is no voiceless–voiced distinction in modern Ainu dialects but it seems to have existed in proto-Ainu (Vovin 1993: 175). In Sakahlin Ainu (except for the Taraika dialect), the syllable final stops /p/, /t/, /k/ and also the tap /r/ have neutralized into /h/, which should probably be regarded as a later development, e.g. *kah* ‘skin’, *kuh* ‘girdle’, *sih* ‘eye’ (cf. the above Hokkaido forms), and *utah* ‘people’ (cf. *utar* of Hokkaido Ainu).

/s/ is realized either as [s] or as palatalized [ʃ]. Most Japanese authors, for example (Satō 2008: 11), describe /s/ as palatalized [ʃ], but Tamura (2000 [1988]: 19) notes for Saru that there is individual variation in the degree of palatalization. As a native speaker of Russian, which has a non-palatalized/palatalized distinction, I find that in most environments /s/ of Southern Hokkaido Ainu is much less palatalized than /s/ of Tokyo Japanese, cf. my description of Chitose: “As a rule, it is not palatalized, unless it is followed or preceded by /i/. Cf. [mus] ‘a fly’, [urayusnai] ‘the name of a village in folktales’, [pas] ‘to run’, [os] ‘after’, but [ʃine] ‘one’, [ʃinutapka] ‘the name of a village in folktales’, [piʃ] ‘beach’, [niʃpa] ‘rich man’” (Bugaeva 2004: 11).

TABLE 15.1 CONSONANT INVENTORY OF SOUTHERN HOKKAIDO AINU

	Bilabial	Alveolar	Palatal	Velar	Glottal
Plosive(=stop)	p	t		k	ʔ
Fricative		s			h
Affricate		c			
Tap		r			
Nasal	m	n			
Glide	w		y		

/c/ is realized as a voiceless alveolar affricate [tʃ] or in the speech of some speakers as [ts], e.g. *cup* ‘moon/sun’ [tʃup] or less common [tsup]. After nasals /m, n/, /c/ is realized as a voiced affricate [dʒ], e.g. *konci* [kondʒi] ‘hood’.

/r/ is a voiced alveolar tap sound [ɾ], similar to that of Japanese, but in the speech of some speakers it is occasionally realized as alveolar trill [r]. Frequently, after a syllable final /r/ the preceding vowel is copied, which is, however, not phonological: *pirka* [pirika] ‘good’, *kor* [koro] ‘have’, *utar* [utara] ‘people’. This copied vowel is often not clearly articulated and sounds more like the schwa sound [ə]: *kor* [korə] ‘have’, *punkar* [punkarə] ‘vine’ (Bugaeva 2004: 12). This “copying” phenomenon might be due to contact with Japanese, which is a rather syllabic language; “copying” is even more prominent in the speech of the last generation of Ainu speakers.

/m/ and /n/ are realized as velar [ŋ] before the stop /k/ and as bilabial [m] and alveolar [n] in all other environments.

/w/ and /y/ are glides [w] and [j], respectively.

/h/ is a glottal fricative [h] which is hardly distinguishable in intervocalic position: *uwekohopi* “separately from each other” (Bugaeva 2004: 12), cf. Tamura (2000 [1988]: 20) who notes that “it is frequently weakened intervocalically and undergoes voicing,” i.e. turning into voiced glottal fricative [ɦ]. Before /u/, /h/ is often realized as a voiceless bilabial fricative [ɸ], as in Japanese.

As mentioned, /ʔ/ is a glottal plosive (glottal stop), which occurs syllable-initially if there is no other initial consonant: *’arpa* ‘go (SG)’. It is conventionally omitted in writing, but I prefer to keep it after syllables ending in a consonant, i.e. in cases when it may potentially influence syllabification and accentuation processes: *wén.ʔe.kot* ‘die a hard death’ (one would expect *we.né.kot* if there were no glottal stop), see 15.2.3. The glottal stop has a tendency to weaken in intervocalic positions, but not in accented syllables such as *’o’ár* ‘completely’ (Bugaeva 2004: 12). As reported by Tamura (1988: 13), the glottal stop is often omitted in the syllable immediately following an accented syllable: *ukáomare* ‘put sth on top of one another in one pile’ (instead of *uká’omare*). Generally, the last speakers of Ainu are not very consistent with the use of the glottal stop as it often falls out in their speech under the influence of Japanese.

### 15.2.2 Vowels

Ainu has five vowels, which are more or less the same as the respective vowels of Japanese.

i                      u  
e                      o  
a

Unlike the Japanese vowels, both /u/ and /o/ are slightly rounded. In addition, /u/ is pronounced as a more back vowel than the respective sound in Japanese and may easily be mistaken for /o/, especially in the unaccented position. Distinguishing /u/ from /o/ in Ainu seems to be no easier for native speakers of English or Russian as is obvious from the title ‘Aino fairy tales’ by Chamberlain (1887) and multiple mistakes in Batchelor (1889) or Dobrotvorskij (1875).

Kindaichi (1993 (1931): 148–9) and M. Chiri (1974 (1936): 7) suggest that there are diphthongs in Ainu,<sup>11</sup> but most contemporary scholars of Ainu prefer to regard the second member of sequences in question as glides (/w/ or /j/).

There is no contrast in vowel length in any Hokkaido dialects, but it is retained in Sakhalin dialects (Murasaki 1976: 3–9), where it is commonly regarded as a vestige of proto-Ainu (Tamura 2000 [1988]: 22).<sup>12</sup>

### 15.2.3 Syllable structure

The syllable structure in Ainu is CV(C), which is either CV or CVC, e.g. *re* ‘three’, *cip* ‘boat’, and *ʔa* ‘sit (SG)’. Basically, there are no vowel-initial syllables (\*V or \*VC), but they sometimes occur in the speech of the last speakers of Ainu due to the omission of the glottal stop. Any consonant can occur as a syllable initial C (onset). The following consonants are not allowed as a syllable final C (coda) in Southern Hokkaido Ainu: /c/, /h/, and /ʔ/. There are a number of individual restrictions on syllable formation, thus, in Southern Hokkaido Ainu, syllables *ti*, *wi*, *ʔuw*, and *ʔiy* are ruled out.

### 15.2.4 Suprasegmentals

Southern Hokkaido Ainu has a clearly distinguishable pitch accent,<sup>13</sup> unlike some eastern Hokkaido dialects, e.g. Shizunai, Samani, etc. (Satō 2008: 14; cf. Refsing 1986: 73). It is a so-called rising kernel accent since the rise from low to high is distinctive, which is the opposite of the accent of Tokyo Japanese, which has a falling kernel accent with the fall from high to low being distinctive.

The accent in southern Hokkaido is said to be distinctive, cf. *niná* ‘crush sth’ vs. *nina* ‘collect firewood’, *nisáp* ‘shin’ vs. *nisap* ‘suddenly’, but the number of minimal pairs distinguished by accent only is very small, as in most cases the accent is assigned automatically in accordance with existing accentuation rules. The rules are as follows: the accent falls on the first syllable if it is a closed syllable, and on the second one if the first syllable is open: *’áp.to* (CVC.CV) ‘rain’ vs. *ko.tán* (CV.CVC) ‘village’ and *sa.pá.ha* (CV.CV.CV) ‘his/her head’.

There are many exceptions from these accentuation rules, e.g. *nú.pe.he* (CV.CV.CV) ‘the tears of’, *tére* ‘wait’ (CV.CV). Historically, these irregular accent patterns of Hokkaido Ainu correspond to the long vowels in Sakhalin Ainu (and presumably of proto-Ainu, cf. footnote 12): *mína* ‘laugh’ (HA) – *miina* (SA). The irregular accent is also observed in some Japanese loanwords, e.g. *káne* ‘metal, money’.

Another reason for irregular accent is the impact of additional morphology on accentuation. For example, the allomorphic causative *-re/-e/-te* suffix does not affect accentuation<sup>14</sup>: *kú* ‘drink’ – *kú-re* ‘make/let drink’, but not *\*ku-ré* as we would expect in accordance with the rules. The same is true of personal affixes: some of them (those which are closer to clitics) do not affect accentuation, but some do (genuine prefixes); for details, see Tamura (1970).

### 15.2.5 Morphophonology

There are certain restrictions on the distribution of phonemes, hence phonological alternations (sandhi phenomena) occurring at syllable boundaries to avoid unfavorable sequences of phonemes. Some of the restrictions apply only to the collocation of phonemes of certain morphemes and therefore trigger morphophonological alternations. All these alternations are subject to dialectal and individual variation.

Here is a list of phonological alternations reported for Saru (Tamura 1979, 2000). I have classified them into different assimilation and dissimilation types.

- regressive assimilation (examples from Tamura 1979: 1):

- (1) /t+i-/ > /ɕi/: *mat-ikor* > /maɕikor/ ‘female treasures’  
 woman-treasure
- (2) /r+t/ > /tt/: *ku=kor tasiro* > /ku=kot̚ tasiro/ ‘my sword’  
 1SG.A=have sword
- (3) /r+c/ > /t̚c/: *ku=kor cise* > /ku=kot̚ cise/ ‘my house’  
 1SG.A=have house
- (4) /r+n/ > /nn/: *ku=kor nonno* > /ku=konn̚ nonno/ ‘my flower’  
 1SG.A=have flower
- (5) /n+s/ > /ys/: *pon sisam* > /poys̚ sisam/ ‘little Japanese’  
 little Japanese

- progressive assimilation (Tamura 2000 [1988]: 24):

- (6) /m+w/ > /mm̚/: *isam wa* > /isamm̚ ma/ ‘(he/she) disappeared and’  
 not.exist and

- mutual assimilation (Tamura 2000 [1988]: 24):

- (7) /n+w/ > /mm̚/: *an wa* > /amm̚ ma/ ‘(he/she) was and’  
 exist.SG and

- regressive dissimilation (Tamura 1983b: 1):

- (8) /r+r/ > /nr̚/: *ku=kor rusuy* > /ku=konn̚ rusuy/ ‘I want to have (it).’  
 1SG.A=have DESID

There are also glide insertion and glide formation phenomena in Southern Hokkaido Ainu and other dialects. Traditionally, both of them were regarded as purely phonological processes for avoiding vowel sequences (hiatus), similar to the above-mentioned ones (Kindaichi 1993 [1931]: 5; M. Chiri 1974 [1936]: 13–14; Tamura 1996: 820, 833). The insertion of the glide /j/ (transcribed as “y”) occurs between two vowels if the first vowel is /i/ and the second vowel is /a, e, o, u/ (9a), while the insertion of the glide /w/ occurs between two vowels if the first vowel is /u/ and the second vowel is /a, e, o/ (9b); in both cases the glottal stop before the second vowel is dropped. Glide formation, i.e. alternation types /i/ > /j/ (10a) and /u/ > /w/ (10b), occurs when the underlying vowels /i/, /u/ are in an intervocalic position.

- (9) a. *si-y-oka*  
 REFL-EP-behind  
 ‘behind oneself’ (AB 16)
- b. *u-w-ekap=an*  
 REC-EP-salute=IND.S  
 ‘we saluted each other’ (AB 16)
- (10) a. *ku=y-ómap (<ku=i-omap)*  
 1SG.S=APASS-love  
 ‘I love children.’ (Satō 2003: 19)
- b. *ko-w-é-peker (<ko=ú-e-peker)*  
 to.APPL-REC-with.APPL-become.bright  
 ‘tell stories to sb’ (AB 17)

However, as shown in Satō (2003: 12) for Chitose, both glide insertion and glide formation “are related to morphological factors as well,” i.e. they are morphophonological

processes in their nature. According to Satō (2003: 18–19), glide insertion occurs only after certain prefixes, e.g. *i*= <IND.O>, *i*- <APASS>; *si*- <REFL>; *u*- <REC>, etc., while a number of other prefixes, e.g. *eci*= <2PL.A/S/O> (11), etc. and incorporated noun stems do not trigger glide insertion under the same phonological conditions.

- (11) *eci*=*y*<sup>h</sup>*opitta*  
 2PL.A=all  
 ‘all of you’ (AB 17); \**eci*=*y*-*opitta*

Similarly, glide formation occurs only when /i/ and /u/ are preceded by a particular set of prefixes, i.e. (C)V type prefixes which have a property of assigning accent on the next syllable (10), e.g. *ku*= <1SG.A/S>, *e*= <2SG.A/S/O>, *i*= <IND.O>, *i*- <APASS>, *e*- <APPL>, *ko*- <APPL>, *o*- <APPL>, but not *a*= <IND.A> or *eci*= <1PL.A/S/O>.

There are also a few morphophonological alternations which apply to the personal prefixes *ku*= <1SG.A/S> and *ci*= <1PL.EXC.A>, as in (12) and (13); both of them occur only in southern Hokkaido.

- (12) *ku*= + *a / e / o / u* → *k*= *a / e / o / u* (/u/ of *ku*= falls out)  
*ku*= + *apkas* → *k*=*ápkas* ‘I walk’
- (13) *ci*= + *a / e / o / u* → *c*= *a / e / o / u* (/i/ of *ci*= falls out)  
*ci*= + *eramuan* → *c*=*éramuan* ‘we understand’

### 15.3 MORPHOLOGY: WORD CLASSES

Tamura (2000 [1988]: ii) sets up the following six word classes.<sup>15</sup> I would suggest just a minor re-classification within the subclasses and the use of a more widely accepted terminology (parts in brackets).

- (1) Verbs: complete [=zero-argument] verbs, intransitive verbs, transitive verbs, and copula;
- (2) Pronouns [personal/interrogative];
- (3) Nouns: common nouns, locative [=relational] nouns, dependent [=bound] nouns, (nominalizing words);
- (4) Adnominals [=Determiners]: adnominal numerals [=numeral determiners], spatial demonstratives [=nominal demonstratives], and conceptual demonstratives [=anaphoric demonstratives].<sup>16</sup>
- (5) Adverbs: normal [=common] adverbs, postpositional adverbs;
- (6) Conjunctions;
- (7) Particles: auxiliary verbs, nominalizing particles, case particles, adverbial particles, (conjunctive particles and conjunctions) [=conjunctions], sentence-final particles;
- (8) Interjections.

We may further tentatively classify word classes into open and closed ones as follows.

- (a) Open lexical classes: all verbs, common nouns, common adverbs;
- (b) Closed lexical classes: locative nouns, bound nouns, numeral determiners, postpositional adverbs, interjections;
- (c) Closed classes of shifters: all pronouns, nominal and anaphoric demonstratives;
- (d) Closed grammatical systems: all conjunctions, all particles.

The categorization of word classes/subclasses is not without problems, but the scrutiny of this issue is a matter of future research.

Word formation in Ainu includes affixation, reduplication, and compounding (for details see Tamura 2001 [1972b, 1973], 2000 [1988]: 193–224; Satō 2008: 265–79). In subsequent sections, I will give an overview of the two major word classes, i.e. nouns and verbs, and also determiners as they are of special interest for linguistic typology.

### 15.3.1 Nouns

According to their morphosyntactic properties, nouns are classified into three major subclasses: (i) common nouns, (ii) relational nouns, (iii) bound nouns.

Common nouns may have “conceptual” forms, which are free and unmarked, and “possessive” forms, which are bound and marked with possessive suffixes and pronominal prefixes of the A (= transitive subject) series to cross-reference the possessor, e.g. *sik* ‘an eye’ – *ku=sik-ih* ‘my eye’. Not all common nouns have possessive forms in Southern Hokkaido Ainu (see Tamura 2001 [1964, 1966]), unlike those of Sakhalin Ainu.

Relational nouns express spatial and temporal relations which are usually expressed by adverbs or prepositions in European languages: *ka* ‘on/over’, *corpok* ‘under’, *etok* ‘before’, etc. (see Tamura 2001 [1982, 1993], 1983b; Nakagawa 1984). Just like common nouns, they have conceptual and possessive forms but the distribution of the two is not the same as with common nouns (see the examples below). However, a major morphosyntactic difference is in that relational nouns are marked for the person and number of the possessor with the prefixes of the O (= object) series: *en=sam* ‘near me’, cf. *menoko sam-a* ‘near the woman’, *cikuni sam* ‘near the tree’ (there are no pronominal prefixes for the third person; possessive suffixes occur only with third person animate possessors such as ‘woman’, but not ‘tree’).

Bound nouns cannot occur on their own as they must be modified with a determiner, noun or relative clause, e.g. *tan kur* ‘this man/person’, but not *\*kur* ‘a man/person’. They are few: *kur* ‘person’, *utar* ‘people’, *uske(he)* ‘(the) place/time (of)’, *pe/p* ‘thing/person’, *hike* ‘the one/side’ (similar to J *hō*), and *hi/i* ‘place/time’. All of them are in the process of grammaticalization into suffixes; the latter three have already developed additional usages as derivational suffixes (15.3.3) and subordinating conjunctions (15.4.6).

There is a group of “nominalizing words” with evidential meanings which are also subsumed under nouns in Tamura (2000 [1988]: 92). Despite their undoubtedly nominal origin, manifested in the transparency of their original meanings and presence of possessive suffixes, I prefer to regard them as particles since their primary present-day function is marking evidentiality, viz. inferential *ruw-e* (< ‘the trace of’), reportative *haw-e* (< ‘the voice of’), non-visual (= semblative) *hum-i* (< ‘the sound of’), visual *sir-i* (< ‘the sight of’); for details, see 15.4.4.2.1.

### 15.3.2 Verbs

The opposition of nouns and verbs is clear-cut (although not encoded morphologically) but many intransitive verbs can also be used as nouns. Verbs are classified into four major subclasses: (i) zero-argument verbs, (ii) intransitive verbs, (iii) transitive verbs, and (iv) the copula. There is no morphologically distinct class of adjectives. The content expressed by adjectives in other languages is expressed by intransitive verbs in Ainu, cf. (14b).

The class of zero-argument (= valent) verbs consists of meteorological predicates which, most commonly, contain an incorporated subject *sir* ‘appearance/land’, as in *sir-pirka* ‘it is good weather’. In Ainu studies, these verbs are traditionally referred to as

**TABLE 15.2 PERSON/NUMBER MARKING IN THE SARU AND CHITOSE DIALECTS OF AINU**

	A/S/O pronouns	A markers	S markers	O markers
1SG	<i>kani</i> ‘I’	<i>ku=</i>	<i>ku=</i>	<i>en=</i>
1PL.EXC	<i>cóka(y)</i> ‘we (I and (s)he/them)’	<i>ci=</i>	<i>as=</i>	<i>un=</i>
2SG	<i>eani</i> ‘you.SG’	<i>e=</i>	<i>e=</i>	<i>e=</i>
2PL	<i>ecioká</i> ‘you.PL’	<i>eci=</i>	<i>eci=</i>	<i>eci=</i>
3SG	<i>sinuma</i> ‘he/she’	$\emptyset$	$\emptyset$	$\emptyset$
3PL	<i>oka</i> ‘they’	$\emptyset$	$\emptyset$	$\emptyset$
Indefinite (IND)	<i>(aoka)</i>	<i>a=</i>	<i>an=</i>	<i>i=</i>

“complete verbs” because they cannot take an external subject; even a dummy subject is not possible.

All other types of verbs are obligatorily cross-referenced for the person and number of their arguments with pronominal verbal affixes. Intransitive (monovalent) and transitive (bivalent) verbs employ different pronominal verbal marking, S vs. A and O (Table 15.2).<sup>17</sup>

There are also trivalent transitive verbs which employ the same A and O series of pronominal affixes as bivalent transitive verbs and are differentiated from them only by the number of unmarked objects. All trivalent verbs are derived by applicative and/or causative derivations, see 15.4.4.1.1 and 15.4.4.1.2.

The equative copula *ne* ‘be/become’ is separate from the above-mentioned subclasses, as it takes A series cross-referencing markers (for the subject), like transitive verbs, but unlike them takes no O markers (for the complement), cf. (16a) and (16b).<sup>18</sup>

Some Ainu verbs employ different stems (= suppletion) or suffixes for singular and plural. In the case of intransitive verbs, plurality refers to the number of S referents while, in the case of transitive verbs, plurality refers to the number of O or patientive A referents. There are a number of aspectual, modal, and evidential markers encoded by particles but no pure tense markers.

### 15.3.3 Determiners

Following Japanese philological tradition, Tamura (2000 [1988]: 36) terms this class “adnominals” (*rentaishi*), which describes a position in a syntactic structure rather than a word class. In European tradition such words are usually referred to as determiners: “words that fill the determiner slot function to specify, identify, or quantify the following noun phrase” (Payne 2006: 124).

There are three subclasses of determiners in Ainu: (i) numeral determiners, (ii) nominal demonstratives, and (iii) anaphoric demonstratives.

All the determiners may occur only in prenominal position and cannot make up a complete NP on their own, which seems to be not so common cross-linguistically (Dixon 2010: 225), e.g. *tan cise* ‘this house’, cf. *\*tan* ‘this’; to express the “demonstrative pronoun” function a bound noun *pe/p* ‘thing’ should be used: *tan pe* ‘this’, see 15.3.1. Similarly, numeral determiners: *tu cise* ‘two houses’, but *tu-p* ‘two (things)’ (not *\*tu*); note that in this case, the bound noun *pe/p* is reanalyzed as a noun-deriving suffix with an additional meaning of “non-human” classifier, cf. *tu-n* ‘two (people)’ or *iwan-iw* ‘six people’ with the “human” classifier suffix *-n* (after V)/*-iw* (after C).

Southern Hokkaido Ainu has an interesting three-term deictic system, with two proximate and one distal nominal demonstratives (Tamura 2000 [1988]: 261).

- *tan*, emphatic form *tapan*. Expresses something that is present, visible, is the new topic of conversation, or that is located where the conversation is taking place: ‘this, here’.
- *taan*. Something that is in the immediate vicinity of the speaker: ‘this, here’.
- *toan*. Separated from oneself: ‘that, there’.

Another cross-linguistically unusual feature of these nominal demonstratives is that they are not used for an anaphoric/cataphoric function (Dixon 2010: 250). There is a separate set of demonstratives used just for anaphora which may tentatively be classified into (a) most recent anaphora, (b) recent anaphora, (c) distant anaphora, (d) most distant anaphora; the latter three forms are distinguished for number.

- *ne* ‘this’, lit. ‘the one that is’ (< copula *ne*). Comes before the noun referring to the person or thing that is being discussed now.
- *ne wa an* (SG), *ne wa oka* (PL) ‘just exactly that’ (< copula *ne* + *wa* ‘and’ + *an/oka* ‘exist.SG/PL’).
- *néa* (SG), *nérok* (PL) ‘that’ (< copula *ne* + perfect marker *a/rok* (SG/PL)). Reminds the listener of someone or something that came up in the conversation earlier.
- *ikia* (SG), *ikirok* (PL) ‘that’ (< *iki* ‘do’ + perfect marker *a/rok* (SG/PL)). Refers to a person or animal that did something previously (Tamura 2000 [1988]: 263–4).

## 15.4 SYNTAX

The basic constituent order in Ainu is SV/AOV. Modifiers are prepositive, and subordinate clauses come before main clauses. All particles (15.3) are postpositional. Topical NPs tend to come first.

For reasons of space, I focus on person marking (15.4.1), postpositions (15.4.2), NP structure (15.4.3), and VP structure (15.4.4). Topics not dealt with are illustrated in various examples, e.g. clause-linking conjunctions and complementizers, see numbers of examples in 15.4.6.

### 15.4.1 Marking of basic syntactic functions

There is no flagging (= case marking) on arguments (S/A/O), either on pronouns or on nouns. Obliques are marked by postpositions (15.4.2).

Personal pronouns (Table 15.2) are uninflected (S/A/O) and are often omitted in subject (A/S) and object (O) position;<sup>19</sup> Ainu is a so-called pro-drop language but indexing (= verbal cross-reference marking) is obligatory. With respect to indexing, Ainu shows mixed alignment: There is neutral alignment in second and third person indexing, viz. A, S, and O are marked by *e*= in 2SG and by *eci*= in 2PL, and the third person is always zero-marked; there is nominative/accusative alignment in the first person singular: 1SG *ku*= marks A/S, and *en*= O; and one finds a tripartite alignment (distinct marking for S, A, and O) in the first person plural exclusive and in the so-called “indefinite person”.

Person markers have varying morphological status: some exhibit properties of words, others are closer to clitics or affixes. For instance, *=an*, *=as* may be separated from the stem by other words and are autonomous prosodic units, so can be classified as words. *a*=, *eci*= cannot be separated and have no accent of their own, characteristic of bound morphemes. On the other hand, they do not change the position of the accent kernel in

the stem, which distinguishes them from regular prefixes. Thus *a=* and *eci=* should rather be classified as clitics. *ku=*, *ci=*, *e=*, *en=*, *un=*, *i=* are full-fledged prefixes since they may not be separated and *do* affect the position of the accent kernel in accordance with the general accentuation rule (Tamura 1970: 305). Their varied morphological status is indicative of their being at different stages of grammaticalization from independent pronouns to obligatory cross-referencing affixes, a commonly attested grammaticalization scenario. Moreover, it has been pointed out that “existing personal pronouns can be replaced by a set of new personal pronouns but survive as verbal clitics or affixes” (Heine and Kuteva 2007: 95). This is exactly what has happened in Ainu: sets of “new” personal pronouns in southern Hokkaido (Table 15.2: A/S/O pronouns) exhibit clear features of being the result of secondary development. Each contains a respective “old pronoun” + existential/locative verb *an* (SG)/*oka* (PL) + nominalizer *-i/-y* (< *hi* ‘thing/place/time’), e.g. *eani* ‘you (SG)’ < *e=an-i* (2SG.A/S/O=exist.SG-thing.NMR), lit. ‘the place where you (sg) exist’ (Bugaeva 2011a: 524).

The basic intransitive and transitive constructions are shown in (14) and (15). Intransitive predicates are indexed for S, and transitive predicates for A and O.

- (14) a. *k=unu-hu*                       $\emptyset$ =*omke*    *a*             $\emptyset$ =*omke*    *a*  
 1SG.A=mother-POSS    3.S=cough    ITR    3.S=cough    ITR  
 ‘My mother coughed and coughed.’ (based on N 130)
- b. *réra*     $\emptyset$ =*ruy*  
 wind    3.S=be.violent  
 ‘The wind is strong.’ (KS #0762)
- c. (*káni*) *ku=mina*  
 1SG    1SG.S =laugh  
 ‘I laughed.’ (OI)
- (15) a. (*káni*) *cikap ku= $\emptyset$ =nukar*  
 1SG    bird    1SG.A=3.O=see  
 ‘I saw a bird.’ (T1 15)
- b. *toan hekaci*     $\emptyset$ =*en=koyki*  
 that    boy    3.A=1SG.O=bully  
 ‘That boy bullied me.’ (T1 30)
- c. *eci=en=hotuyekar yak pirka p*  
 2PL.A=1SG.O=call    if    be.good    but  
 ‘You(PL) may have called out to me.’ (T1 36)
- d. *eci=nukar rusuy kusu te ta eci=hunara*  
 1SG.A+2SG.O=see    DESID    because    here    LOC    1SG.A+2SG.O=search  
 ‘I was looking for you here because I wanted to meet you.’ (KS #0009)

The copula clause is a special type since its predicate is indexed for A (= subject), but never for O (= complement) (16a).

- (16) a. [*cóka*]<sub>CS</sub>    [*tu-n*]<sub>CC</sub>                      *ci=ne*                                      *na, tu-p*  
 1PL.EXC    two-person.CL    1PL.(EXC).A=COP    FIN    two-thing.CL
- en=kor-e*                                      *yan*  
 1SG.O=have-CAUS    IMP.POL  
 ‘There are two of us [lit. ‘we are two’], so give us two pieces.’ (KS #2961)

- b. [cóká]<sub>CC</sub> Ø=*ne*  
 1PL.EXC 3.A=COP  
 ‘It is us (me and (s)he).’ (T2 53)

Theoretically, A markers precede O markers, but in practice one can find few examples where both markers are overtly expressed as in (15c); recall that there is no overt marking for the third person. Moreover, regarding the interaction of first and second person participants, person marking on verbs is not always analyzable as two prefixes (15d), and it is, indeed, subject to considerable dialectal variation. For instance, in Southern Hokkaido Ainu, the combination of first person singular/plural A and second person singular/plural O is never marked as \**ku=e=*, \**ku=eci=*, \**ci=e=*, and \**ci=eci=*, but invariably with the portmanteau *eci=* (1SG/PL.A+2SG/PL.O=) (Table 15.3), which has a key function of second person plural marker for A, S, or O.

The difference between intransitive and transitive verbs is clear-cut because they employ different indexing for S and A (and also O) in the case in 1PL and IND (17).

- (17) a. *nenó é=iki yak a=e=kóyki na*  
 like.this 2SG.S=do if IND.A=2SG.O=scold FIN  
 ‘If you do that, you will be scolded [lit. **one/they** will scold you].’ (T2 71)
- b. *te ta rok=an ciki pirka?*  
 here LOC sit.PL=IND.S if be.good  
 ‘May **one** sit here?’ (KS #1660)

The indefinite person is a cover term for a number of functions expressed by the same marking. Indefinite proper,<sup>20</sup> referring to the indefinite speaker or addressee, does not allow the use of the respective pronoun *aoka* and, at least originally, could be used with plural forms of verbs only (17b) regardless of the actual number of (indefinite) referents. Other important functions are first person plural inclusive (18), second person singular/plural honorific (19), and logophoric, referring to the author of the discourse (20).<sup>21</sup> The latter is regarded as ‘first person in quotes’ in Tamura (2000 [1988]: 74), which is hardly intelligible for non-Ainologists, and may be easily mistaken for a kind of “genuine” first person; I have suggested “logophoric” instead, similar to that of West African languages (Bugaeva 2008a, 2008b). In Ainu, logophorics occur in the embedded clause (= the quote) when the subject of the main clause is third/second person singular (20b) or plural (20a); they show coreference of main clause subject with the S/A/O or possessor in the embedded clause (Bugaeva 2008a: 43–4).

**TABLE 15.3 SUBJECT/OBJECT INDEXING OF TRANSITIVE VERBS IN THE SARU AND CHITOSE DIALECT OF AINU (ADAPTED FROM TAMURA (2000 [1988]: 59) WITH SLIGHT MODIFICATION)**

A/O	1SG	1PL	2SG	2PL	3SG/3PL	IND
1SG	--	--	<i>eci=</i>		<i>ku=Ø=</i>	<i>ku=i=</i> (/kuy/)
1PL	--	--	<i>eci=</i>		<i>ci=Ø=</i>	<i>a=i=</i>
2SG	<i>en=</i>	<i>un=</i>	--	--	<i>e=Ø=</i>	<i>e=i=</i> (/ey/)
2PL	<i>eci=en=</i>	<i>eci=un=</i>	--	--	<i>eci=Ø=</i>	<i>eci=i=</i>
3SG/PL	<i>Ø=en=</i>	<i>Ø=un=</i>	<i>Ø=e=</i>	<i>Ø=eci=</i>	<i>Ø=Ø=</i>	<i>Ø=i=</i>
IND	<i>a=e=</i>	<i>a=un=</i>	<i>a=e=</i>	<i>a=eci=</i> (Saru) cf. <i>aci=</i> (Chitose)	<i>a=Ø=</i>	<i>a=i=</i>

- (18) a. “minimal inclusive” – (1+2) ‘I and you’  
*suy u-nukar=**an** ro*  
 again REC-see=IND.S HOR  
 ‘Let’s (me and you) meet again.’ (AB 94)
- b. “augmented inclusive” – (1+2+3) ‘I and you and s/he/they’  
***aoka** anak, kamuy renkayne, ri uske ta oka=**an***  
 IND TOP god thanks.to high place LOC exist.PL=IND.S  
*kusu i-sitoma ka somo a=Ø=ki*  
 because APASS-fear even NEG IND.A=3.O=do  
 ‘Thank god(s), we (me and you and s/he/they) lived at the high place, so we (me and you and s/he/they) did not fear the danger (of the flood).’ (T3 50)
- (19) a. ***aoka** yaykata a=Ø=kar ruwe?*  
 IND oneself IND.A=3.O=make INF.EV  
 ‘Have **you** made it yourself?’ (lit. ‘Have **one** made it himself?’) (T4 381)
- b. *ku=**i**=nukar rusuy korka*  
 1SG.A=IND.O=see DESID but  
 ‘But I wanted to meet with **you**.’ (lit. ‘But I wanted to meet with **one**.’) (T4 381)
- (20) a. “[***aoka***] *oya-pa suy arki=**an** kusu ne na!”*  
 IND.PL next-year again come.PL=IND.S intention COP FIN  
*sekor Ø=haweoka kor Ø=paye wa orano*  
 QUOT 3.S=say.PL while 3.S=go.PL and then  
*k=Ø=okaramotte-pa*  
 1SG.A=3.O=feel.reluctant.to.part-PL  
 ‘They left while saying, “We [lit. selves] will come again next year,” and I felt reluctant to part with them.’ (T2 75)
- b. “[***asinuma*** *arpa=**an** kusu ne” sekor Ø=hawean*  
 IND.SG go.SG=IND.S intention COP QUOT 3.S=say.SG  
 ‘(She) said, “I [lit. self] will go.”’ (T2 74) (the glossing is mine – A.B.)

Southern Hokkaido dialects differ from others in having developed a “dedicated” singular logophoric pronoun *asinuma*<sup>22</sup> ‘one, someone’ (Tamura 1988: 22–4), cross-referenced on verbs by the same originally non-singular indefinite affixes *a*= A, *=an* S, and *i*= O. According to Nakagawa (1988: 246), *asinuma* has appeared as a response to the reanalysis of verbs with non-singular affixation as singular forms (i.e. in the case of singular referent, the use of *arpa=an* (go.SG) instead of original *paye=an* (go.PL)), clear from other Hokkaido dialect data.

The logophoric function (20) is important because it also commonly occurs in folktales (21) to refer to the protagonist in the quote, since folktales have the structure of reported discourse with the whole story being a quote.

- (21) “. . . *cis=**an**, cis=**an** kor patek an=**an** ayne, . . .*  
 cry=IND.S cry=IND.S and only exist.SG=IND.S finally  
*ray=an ma isam ruwe ne” sekor; sine menoko Ø=itak*  
 die.SG=IND.S and not.exist INF.EV COP QUOT one woman 3.S=tell

“... I [lit. self] cried and cried. I [self] was always crying. Finally, ... I [self] died”. One woman told.’ (AB 200)

In early studies (Kindaichi 1993 [1931]; M. Chiri 1974 [1936]; Kubodera 1977), the regular occurrence of the indefinite in folklore, especially *yukar* ‘heroic epic’, was attributed to the genre’s elegant style, and all Ainu literature was presented as “first person literature,” i.e. direct style narration. It was Tamura who first documented the logophoric use of colloquial Ainu in Southern Hokkaido (Saru), rejecting the stylistic interpretation in folklore. Until recently, due to the lack of relevant descriptive data, logophoric marking in colloquial Ainu has not been attested in other dialects. However, Bugaeva (2008a: 51) has revealed its employment in non-folklore texts of one of the central Hokkaido dialects (Asahikawa; HNE), which proves that, as in southern Hokkaido, its common use in folklore is triggered by logophoric contexts, since folktales are also a kind of reported speech. The logophoric is consistently glossed IND here to emphasize its diachronic relatedness to other functions of the indefinite person, but is translated as ‘I’ for convenience.

#### 15.4.2 Marking of peripheral syntactic functions

As mentioned, case markers are not employed for arguments; their use is restricted to peripheral syntactic functions, i.e. adjuncts. Adjuncts are marked by the following post-positions: locative *ta*, allative *un* (inanimate Goal), dative *e-un* (< head-ALL) (animate Goal), ablative *wa*, instrumental *ani*, comitative *tura/tura-no*, traversals *peka* (‘over’), and mutative *ne* (‘as’). Most of these clearly originate in verbs and *ani* and *tura/tura-no* are still in a very early stage of grammaticalization as they can often occur without nouns. In (22), the *itanki* before *ani* ‘with’ (< *ani* ‘hold sth’) can easily be omitted and the respective sentence would trigger a zero-anaphora interpretation: ‘holding (that bowl)’. However, if *ani* were still a verb here, we would need a coordinating conjunction after *ani*, which is not the case, meaning that it really has intermediate status between verb and postposition.

- (22) *itanki huraye hine (itanki) ani Ø=i=ko-i-puni*<sup>23</sup>  
 bowl wash and bowl INST 3.A=IND.O=to.APPL-APASS-raise  
 ‘She washed a bowl and with (that bowl) she served me food.’ (OI)

I provide examples of locative, allative, dative, and ablative. Static location is marked with the locative (23).

- (23) *apa sam ta a=an*  
 door near LOC sit.SG=IND.S  
 ‘I sat down near the front door.’ (K7803232UP.035)

Goal is expressed by the allative (24a), but also with the locative (24b). In the latter, achievement seems to be implied. *e-un* (25) is reserved for marking animate Goal, hence the label “dative”.

- (24) a. *sine-an-ta . . . kuca or un arpa=an*  
 one-be.SG-LOC . . . hunting.cabin place ALL go.SG=IND.S  
*kun-ak a=Ø=ramu kusu,*  
 going.to-COMP IND.A=3.O=think because  
 ‘One day, I thought I’d go to a hunting cabin.’ (K8109171UP.010)

b. *kuca or ta arpa=an ine*  
 hunting.cabin place LOC go.SG=IND.S and

*inaw-roski=an*

prayer.stick-stand.PL=IND.S

‘I went to the hunting cabin, and put up prayer sticks.’ (K8106233UP.007)

(25) *ne wen-kur okkay-po eun anak-ne iteki arpa=an*  
 this/that poor-person man-DIM DAT TOP-COP PROH go.SG=IND.S

*yak pirka*

if be.good

‘You better never go to that poor man.’ (K7803232UP.099)

Source is marked with the ablative (26a). It is also used for Actor demotion in the case of indefinite proper (= the impersonal passive) (26b); for details see Bugaeva (2011a: 527–8). Cf. (17a), which is the same construction without the oblique Actor phrase.

(26) a. *nisat-ta tunas-no tan kotan wa k=arpa kusu*  
 dawn-at be.early-ADV this village ABL 1SG.S=go.SG intention

*ne na*

COP FIN

‘I will leave this village early tomorrow.’ (KS #0053)

b. *hapo or-o wa a=en=koyki*  
 mother place-POSS ABL IND.A=1SG.O=scold

‘I was scolded by mother.’ (T2 72) (lit. ‘Someone scolded me by mother.’)

Most nouns cannot be used with locative and spatial postpositions directly: a special relational noun (see 15.3.1) – neutral *or(-o)* ‘place (of)’ (24), (26b) or specific *sam(-a)* ‘near (of)’ (23), etc. – is added between noun and postposition.<sup>24</sup> Only a few common nouns (e.g. *kotan* ‘village’ (26a), *kim* ‘mountains’, *pis* ‘beach’, *rep* ‘open sea’, *ya* ‘shore’, *put* ‘mouth of a river’) and geographical names do not require a relational noun; in Tamura (1984: 39–40), the former are described as inherently “thing” nouns and the latter as inherently “place” nouns.

### 15.4.3 Noun phrase structure

A noun can be head of an NP with prepositional modifiers: a determiner (numeral determiner or nominal/anaphoric demonstrative), an adverb, a noun, a possessor NP, and a relative clause. (See 15.3.3 for “determiner + head noun.”)

The “Adv + head noun” type is not very common (27).

(27) *teeta huci*

long.ago elderly.woman

‘a grandmother from long ago’ i.e. ‘great-grandmothers and previous generations’

(T2 188)

“N + head noun” is an attributive construction; dependent and head nouns are juxtaposed (28).

(28) a. *sisam uwepeker*

Japanese old.tale

‘an old Japanese folk story’

b. *kamuy rus*

bear fur

‘bear skin’ (T2 187)



“case-retention” strategies facilitating the recoverability of respective case relations, which is different from Japanese which allows the gap strategy for some obliques. All positions on Keenan and Comrie’s hierarchy (1977) except object of comparison can be relativized: SU > DO > IO > OBL > GEN > OCOMP.

*SU: relativization of S.* Ainu shows no distinction in the treatment of agentive (32a) and non-agentive (32b) subjects.

- (32) a. [*cise soy pak-no Ø=arki*]                    *utar anak-ne*  
 house outside till-ADV 3.S=come.PL men TOP-COP  
*a=Ø=ahun-ke*    *yak pirka wa*  
 IND.A=3.O=enter.SG-CAUS if be.good FIN  
 ‘As to the **men** who came up to the doorway, I may let them in.’ (AB 95)
- b. [*Ø=pirka*]    *menoko a=Ø=etun*  
 3.S=good/beautiful woman IND.A=3.O=get.as.a.bride  
 ‘I got a beautiful woman as a bride.’ (AB 94)

*SU: relativization of A.* This is rare (33).

- (33) [[*anrur Ø=Ø=un*]                                    *kamuy Ø=Ø=ray-ke*]                                    *kur*  
 Anrur 3.A=3.O=belong.to god/bear 3.A=3.O=die.SG-CAUS man  
 ‘a **man** who killed a bear from Anrur’ (KI 292)

*DO: relativization of O* (Patient/Theme of a bivalent transitive).

- (34) [*toan-i un ku=Ø=nukar*]                    *cise hemanta Ø=an?*  
 that-place ALL 1SG.A=3.O=see house what 3.S=exist.SG  
 ‘What is that **house** I see over there?’ (KS #3347)

*IO: relativization of O* (Goal of a trivalent transitive; there is no formal distinction in Ainu between direct and indirect objects).

- (35) [*kampi a=Ø=e-pakasnu*]    *hekattar*  
 paper/letter IND.A=3.O=about.APPL-teach.to children  
 ‘**children** who are taught to read and write’ (KS #1237)

*OBL: relativization of obliques.* Oblique relations, which are typically expressed by oblique NPs in other languages, often appear as applicative direct objects in Ainu (see 15.4.4.1.1) and are relativized with the gap strategy. On the syntactic level, they are technically DO.

- (36) *katkemat* [*a=Ø=e-hotke*]    *usi Ø=Ø=kar wa*  
 housewife IND.A=3.O=at.APPL-lie.down place 3.A=3.O=make and  
*Ø=i=kor-e*  
 3.A=IND.O=have-CAUS  
 ‘The housewife prepared a sleeping place [lit. ‘**place** to lie at’] for me.’ (AB 400)

Relativization of true obliques is possible only with the use of special “case-retention” strategies to disambiguate the case role of the relativized NP: (a.) Retention of the instrumental case postposition *ani* in the relative clause (37). (b.) Retention of the comitative case postposition *tura(-no)* in the relative clause (38).

- (37) [*ani ku=yupo kamuy Ø=Ø=tukan*] *teppo*  
 INST 1SG.A=elder.brother.POSS bear/god 3.A=3.O=shoot gun  
 ‘the **gun** with which my elder brother shot the bear’ (AB 95)

cf. base construction:

*teppo ani ku=yupo kamuy Ø=Ø=tukan*  
 gun INST 1SG.A=elder.brother.POSS bear/god 3.A=3.O=shoot  
 ‘My elder brother killed a bear with a gun.’

- (38) [*tura-no ku=yupo kamuy Ø=Ø=cotca*] *acapo*  
 COM-ADV 1SG.A=elder.brother.POSS bear/god 3.A=3.O=shoot uncle  
 ‘the **uncle** with (whom) my elder brother shot the bear’ (AB 96)

cf. *acapo tura-no ku=yupo kamuy Ø=Ø=cotca*  
 uncle COM-ADV 1SG.A=elder.brother.POSS bear/god 3.A=3.O=shoot  
 ‘My elder brother shot a bear together with the uncle.’

(c.) Retention of the respective relational nouns with locative postpositions within relative clauses. As mentioned in 15.4.2, most nouns in Ainu cannot appear in locative phrases without *or(-o)* ‘the place (of)’, etc., so *or-o* is retained in its possessive form within the relative clause together with locative postposition (39). There is some similarity with demonstratives as “the mirror images of resumptive pronouns” in Japanese adnominal clauses (Tsunoda, forthcoming).

- (39) [*or-o ta cep poro-n-no Ø=hemesu*] *pet*  
 place-POSS LOC fish many-EP-ADV 3.S=go.upstream river  
 ‘the **river** where a lot of fish go upstream’ (AB 96)

cf. *pet or ta cep poro-n-no Ø=hemesu*  
 river place LOC fish many-EP-ADV 3.S=go.upstream  
 ‘At the river, a lot of fish go upstream.’

*GEN*: relativization of the possessor (only inalienable). Possessee is retained in possessive form, highlighting the possessor’s presence.

- (40) [*sara-ha Ø=tanne seta*]  
 tail-POSS 3.S=long dog  
 ‘a **dog** with a long tail’ (lit. ‘a **dog** (whose) tail is long’) (OI)

cf. *seta sara-ha Ø=tanne*  
 dog tail-POSS 3.S=long  
 ‘The dog’s tail is long.’

Bound nouns (15.3.1) are the most common heads of relative clauses – see (32), (33), and (36).

One other interesting type of noun-modifying construction<sup>25</sup> occurs, which cannot be subsumed under the traditional relative clause because there is no possible non-relative paraphrase in which the head noun can appear as an argument or adjunct (41). A similar type is attested in Japanese (the famous *soto no kankei* type of Teramura (1992: 192–205) and its in-depth research in Matsumoto 1997); its existence in Ainu has been recognized only recently (Bugaeva and Matsumoto 2009).<sup>26</sup>

- (41) [*e=munin*] *Ø=hura*  
 2SG.S=rot 3.A=smell.POSS  
 ‘your rotten smell’ (lit. ‘the smell of you rotting’) (N 338)

## 15.4.4 Verb phrase structure

## 15.4.4.1 Morphological structure of the verb: changing valency

According to Fukuda (Tamura) (1955: 55), there are six slots in verb structure (Table 15.4): (I) applicative *e-/ko-/o-*; (II) reciprocal *u-*, reflexive *yay-/si-*, generalized object *i-*; (III) applicative *e-/ko-/o-*, which can be doubled; (IV) suffixes of singularity/plurality; (V) suffixes of transitivity/intransitivity; (VI) causative *-re/-e/-te* and *-yar* (*-ar*).

A personal stem may consist of a base alone or with one or more affixes. All slots except IV encode valency-changing alternations (or voice in a broad sense), hence they will be dealt with here. I would add a few major modifications to Table 15.4, however, as it lacks anticausatives and contains only two causatives. I suggest the following reinterpretation based on forms' syntax and function: Slot II *si-* 'oneself' → *anticausative/reflexive si-*; Slot V suffixes of transitivity/intransitivity → *anticausative ke-* and *direct causative -V, -ke, -ka, -re/-e/-te*; Slot VI causative *-re/-e/-te* → *indirect causative -re/-e/-te*. It is also important to account for the possibility of noun incorporation: right before the base (0) and before each applicative slot (I/III).

To summarize, there are two devices for increasing valency (applicative *e-, ko-, o-* and causative<sup>27</sup> *-e/-te/-re, -ka, -ke, -V*) and five for decreasing valency (reflexive *yay-, si-*, reciprocal *u-*, antipassive *i-*, anticausative *si-, -ke*, and noun incorporation). Although no verb has been found to contain all six kinds of affix and noun incorporation, there are many possibilities for multiple derivational processes on a single verb (42).

- (42) a. *ruska* 'be angry because of sth' (vt)  
 b. *ko-ruska* <APPL-be.angry.because.of> 'be angry with sb because of sth' (OI) (vd)  
 c. *yay-ko-ruska* <REFL-APPL-be.angry.because.of> 'be angry with oneself because of sth' (OI) (vt)  
 d. *i-ruska* <APASS-be.angry.because.of> 'be angry' (OI) (vi)  
 e. *ko-i-ruska* <APPL-APASS-be.angry.because.of> 'be angry with sb' (OI) (vt)

TABLE 15.4 VERB STRUCTURE IN THE SARU DIALECT OF AINU

personal stem					
I	II	III	0	IV	VI
' <i>e-</i> [APPL] 'with[/about/at]'	<i>i-</i> [APASS] 'something'	' <i>e-</i> [APPL] 'with[/about/at]'	base	singular suffix or plural suffix	<i>-re/-e/-te</i> [CAUS] 'make/let sb'
	<i>yay-</i> [REFL] 'oneself'			V	
<i>ko-</i> [APPL] 'to/with'	<i>u-</i> [REC] 'each other'	<i>ko-</i> [APPL] 'to/with'		intransitive suffix or transitive suffix	[indefinite] <i>-yar</i> [- <i>ar</i> ] [CAUS] 'let someone do'
( <i>o-</i> [APPL] 'to/at some place')	<i>si-</i> [ACAUS] 'oneself'	( <i>o-</i> [APPL] 'to/at some place')			

- f. *si-ko-i-ruska* <ACAUS-APPL-APASS-be.angry.because.of>  
'get angry', cf. 'get angry with another' (B 453) (vi)
- g. *i-ruska-re* <APASS-be.angry.because.of-CAUS>  
'make sb angry' (OI) (vt)
- h. *ko-i-ruska-re* <APPL-APASS-be.angry.because.of-CAUS>  
'make sb angry with sb' (OI) (vd)
- i. *si-ko-i-ruska-re* <ACAUS-APPL-APASS-be.angry.because.of-CAUS>  
'make sb angry with oneself' (KII 243) (vt)
- j. *u-ko-i-ruska* <REC-APPL-APASS-be.angry.because.of>  
'be angry with one another' (OI) (vi)
- k. *u-ko-i-ruska-re* <REC-APPL-APASS-be.angry.because.of-CAUS>  
'make sb angry with each other' (OI) (vt)
- l. *\*u-ruska* <REC-be.angry.because.of>  
(ungrammatical because (a) has no personal O) (OI) (vi)

#### 15.4.4.1.1 Applicatives

The applicative is typically described as a syntactic construction signaled by overt verbal morphology which allows the coding of a semantically peripheral argument or adjunct as a core object argument (Peterson 2007: 1). In most languages, including Ainu, applicatives have an important discourse function, i.e. "the entity the construction refers to has a greater discourse salience (foregrounded information) or topic continuity than would otherwise be expected of it" (Peterson 1999: 61, with reference to Givón 1983).

Ainu applicatives are derived with *e-/ko-/o-*. These prefixes increase valency, but their specific semantic function may vary. The range of the constructions stretches over several semantic roles and the exact role is attributed to interaction between the semantics of the prefix and verb; the following is based on a thorough sample and corpus-based study (Bugaeva 2010a: 760): *e-* Content, Location, Instrument, Theme, Cause/Purpose, Other (Comitative – Co-agent, Manner, Beneficiary, Path); *ko-* Addressee, Goal, Recipient/Beneficiary, Comitative – Co-patient, Malefactive Source, Other (Cause/Purpose); *o-* Goal, Location.

In Bugaeva (2010a), the three applicatives are defined as instrumental *e-*, dative *ko-*, and locative *o-* because most other roles can be explained in terms of their relatedness to the proto-role.

The prefixes turn intransitive verbs into bivalent transitives. The S-affix *=an* of the base construction (43a) is replaced by its A-counterpart *a=* in (43b), and the *ta*-marked adjunct is promoted to object.

- (43) a. *casi upsor ta ahun=an ruwe ne*  
castle inside LOC enter.SG=IND.S INF.EV COP  
'I entered the castle.' (OI)
- b. *casi upsor a=Ø=o-ahun ruwe ne*  
castle inside IND.A=3.O=to.APPL-enter.SG INF.EV COP  
'I entered the castle.' (OI)

Applicativized bivalent transitives turn into trivalent transitives (44b), which have two unmarked objects that are rather similar in terms of encoding and behavioral properties.<sup>28</sup> In principle, either object can be cross-referenced on the verb (one at a time) if first/second person. In practice, it is usually Recipient objects that are cross-referenced

because they are typically animate and may have first/second person referents, while Theme/Instrument objects are typically inanimate.

- (44) a. [*e=matak-i*]<sub>p</sub> . . . [nea *mat-ikor*]<sub>INST</sub> *ani* *a=Ø=resu*  
 2SG.A=younger.sister-POSS that woman-treasure INST 3.A=3.O=raise  
 ‘I raised your younger sister with those treasures.’ (modified K8109193UP.088)
- b. [*matkaci anak-ne*]<sub>p</sub> [*nea iskar emko Ø=Ø=kor katkemat*  
 girl TOP-COP that Ishikari half 3.A=3.O=have woman  
*Ø=Ø=kor mat-ikor*]<sub>INST</sub> *a=Ø=e-resu*  
 3.A=3.O=have woman-treasure IND.A=3.O=with.APPL-raise  
 ‘As for the girl, I raised her with the treasures from that woman from Upstream  
 Ishikari (i.e. by using the protective power of the treasures).’ (K8109171UP.100)

A prototypical applicative construction provides an alternative to oblique realization of a semantically and syntactically peripheral participant, i.e. there must exist a non-applicative (base) construction, and it does not involve Agent and Theme/Patient participants because they are non-peripheral. Addition of an Agent typically results in causativization; likewise, addition of a Theme/Patient results in non-causative transitivization. Although Ainu is often included as an example of a language with a prototypical applicative construction (Spencer 1991: 253), a great deal of Ainu applicatives deviate considerably from the prototype lacking non-applicative paraphrases (*e-* applicatives (Content, Theme) and *ko-* (Recipient)) or adding a participant with a Theme role (*e-* applicatives) or both (45b).

- (45) a. *ainu or un iteki ikka yan*  
 Ainu place at PROH steal IMP.POL  
 ‘Don’t steal in the land of the Ainu!’ (KI 117)
- b. *a=e=e-ikka hene a=e=ray-ke yak-ka*  
 IND.A=2.SG.O=APPL-steal and IND.A=2.SG.O=die-CAUS if-even  
 ‘Even if you were taken away and killed . . .’ (K7908032UP)

The thematic applicative (45b) is very unusual cross-linguistically and I suggest that it originated in the instrumental applicative and was based on the three-argument schema. In fact, most thematic applicatives involve derived intransitive bases containing either noun incorporation or valency-decreasing antipassive *i-*/anticausative *si-*, which are often fossilized as in *ikka* ‘steal’ (45a).

For some applicatives there have been reported corresponding non-applicative paraphrases, but they hardly ever appear in actual texts. Thus, the dative applicative *ko-* in principle allows a paraphrase with dative *eun*, which is employed for marking obliques such as Recipient, Addressee, or animate Goal, but the frequency of *eun*-marked constructions in texts is rather low as it is preferably used with those few verbs for which there are, for some reason, no corresponding Recipient/Addressee applicatives (e.g. *se* ‘carry sth on one’s back’, *ye* ‘tell sth; tell sb’). If there is a Recipient/Addressee applicative in *ko-*, the *eun*-marked construction, as in (46a), has a low chance of occurring in texts, and no chance if Recipient is encoded by a pronominal argument as in (46b). The last speakers seem to use the paraphrase more frequently under Japanese influence.

- (46) a. [*aynu Ø=nimar-a*]<sub>A</sub> [*iskar emko Ø=Ø=kor nispa*]<sub>R</sub>  
 human/Ainu 3.A=half-POSS Ishikari upstream 3.A=3.O=have rich.man

*eun* [*nea nispa*  $\emptyset=\emptyset=kor$  *pa p*]<sub>T</sub> *opitta*  $\emptyset=\emptyset=rura$  *pa*  
 DAT that rich.man 3.A=3.O=have PL NMR all 3.A=3.O=carry PL  
 ‘The (remaining) half of the Ainu carried all of that rich man’s belongings  
 to the rich man at Upstream Ishikari.’ (K8109171UP.234)

- b. *aynurakkur kamuy a=karku*  $\emptyset=an$  *kusu-keray-po*  
 Ainurakkur god IND.A=nephew.POSS 3.S=exist.SG because-thanks-DIM  
 [*a=\emptyset=kor* *tures-i*]<sub>T</sub> *a=i<sub>R</sub>=ko-rura*  
 IND.A=3.O=have younger.sister-POSS IND.A=IND.O=to.APPL-carry  
 ‘Thanks to my divine nephew Ainurakkur my younger sister has been carried  
 to me.’ (KI 267)

Most typological studies of applicatives (e.g. Peterson 2007) present them as an alternative way of introducing (peripheral) arguments, but Ainu applicatives often appear to be the *only* way.

Another typologically unusual property is the ability of so-called unaccusative intransitives to host applicative *e-* and *ko-*, first pointed out in Shibatani (1990: 66). GB framework (Baker 1988: 254) claims that there is a restriction on the formation of applicatives from a particular set of intransitives, namely unaccusatives, i.e. those whose subject originates as an object in deep structure (So: *fall*, *die*; opposed to unergatives *run*, *work*), as presented by Perlmutter (1978) in his Unaccusative Hypothesis. Indeed, Ainu data appear suggestive of the possibility of applicative formation from unaccusative bases of the following classes; applicatives from internally caused verbs of change of state are lacking.

(i) applicatives from verbs of change of state

- (47) a. *e-sik* ‘be(come) full with sth’ (T 123) (Instrument) < *sik* ‘be(come) full’  
 b. *ko-onne* ‘age/live long together with sb’ (Co-patient) < *onne* ‘be(come) old’  
 c. *o-rer* ‘sink in sth’ (Location) < *rer* ‘sink’

(ii) applicatives from verbs of inherently directed motion

- (48) a. *e-ek* ‘come (SG) for (doing sth)’ (Purpose) < *ek* ‘come (SG)’  
*e-san* ‘descend (SG) somewhere’ (KI 518) (Location) < *san* ‘descend (SG)’  
*e-sirepa* ‘arrive at some place’ (B 133) (Location) < *sirepa* ‘arrive’  
 b. *ko-ek* ‘come (SG) to sb’ (Goal) < *ek* ‘come (SG)’  
*ko-hemesu* ‘climb somewhere’ (Goal) < *hemesu* ‘climb’  
*ko-hokus* ‘fall with sb/sth’ (Co-patient) < *hokus* ‘fall’  
*ko-yan* ‘ascend (SG) at sb’s place’ (Goal) < *yan* ‘ascend/come (SG) ashore’  
 c. *o-ek* ‘come (SG) to sb’ (Goal) < *ek* ‘come (SG)’  
*o-hemesu* ‘climb somewhere’ (Goal) < *hemesu* ‘climb’  
*o-yan* ‘ascend (SG) somewhere’ (Goal) < *yan* ‘ascend/come ashore (SG)’

(iii) applicatives from verbs of existence and appearance

- (49) a. *e-an* ‘live/exist (SG) somewhere’ (T 71) < *an* ‘live/exist (SG)’  
*e-isam* ‘die at some time’ (T 167) (Location) < *isam* ‘not exist; die’  
*e-ray* ‘die (SG) at some place’ (N1 99) (Location) < *ray* ‘die (SG)’  
*e-rok* ‘live/reside (of a God) somewhere’ (T 120) (Location) < *rok* ‘sit (PL)’  
 (lexicalized)  
*ko-rewsi* ‘stay overnight with sb’ (Co-patient) < *rewsi* ‘stay’

- b. *o-rewsi* ‘stay overnight at some place’ (T 481) (Location) (vt) < *rewsi* ‘stay’  
*o-sirepa* ‘arrive at some place’ (N 117) (Location) < *sirepa* ‘arrive’

However, no one has convincingly shown that these verbs are actually unaccusative in Ainu. In Bugaeva (2010a: 787), I have shown that they are likely to pass a telicity test for unaccusativity, but since there are no tests available for checking the degree of agentivity in Ainu – an important component of unaccusativity – there is reason to suspect that they might not be completely non-agentive.

Remarks should be made on applicative use in different text types. As Refsing (1986: 32) notes, applicative constructions appear extremely common in Kindaichi’s and M. Chiri’s material collected in the 1920s/1930s, but are much less common in her own material collected in the 1980s: The speaker<sup>29</sup> she worked with preferred postpositions, which she explains as a sign of the decline of Ainu and influence of Japanese. Although a language contact-induced change might have played a role, it might not have been the only reason. Kindaichi’s (1993 [1931]) and M. Chiri’s (1974 [1936], 1973 [1942]) grammars are based on the study of Ainu oral literature, mainly the poetic genre of *yukar* ‘heroic epic song’ and *kamuy yukar* ‘divine epic song’, which exhibit highly sophisticated language characterized by a high degree of polysynthesis, while Refsing’s grammar is based on colloquial Ainu. I suggest that the difference in applicatives’ frequency is also due to the difference in genre. Very frequent use of (monosyllabic) applicative morphology in poetic Ainu may simply be the way of marking case relations that fits the meter (4/5 syllables a line) and melody of songs better than postpositionals. In colloquial Ainu and oral literature of prosaic genres, applicatives are used less frequently and seem to encode greater discourse salience. Applicatives in poetic Ainu might not necessarily exhibit the discourse function associated with prototypical applicatives and should ideally be studied separately from “genuine applicatives.”

#### 15.4.4.1.2 Causatives

“The causative construction generally represents a linguistic expression which denotes a complex macro-situation consisting of two micro-situations or component events: (i) the causing event in which the causer does or initiates something in order to bring about a different event (i.e. the caused event), and (ii) the caused event in which the causee carries out an action or undergoes a change of condition or state as a result of the causer’s action” (Song 2001: 257, which originates in Nedjalkov and Silnitskij 1969: 5). Syntactically, the causative construction usually (but not necessarily: indefinite causative suffix *-yar/-ar* does not change valency) involves valency increase by addition of a new argument in A function (the causer) and the original S argument (the causee) goes into O function in the new transitive clause. In the causative of a transitive, the causer always becomes A; the original A (the causee) and/or O arguments generally have their syntactic functions reassigned (Dixon 2000: 31).

Ainu has a number of morphological causatives but only the productive *-re/-e/-te* and the indefinite causative *-yar/-ar* have been regarded as causative markers in Ainu studies (Table 15.4). I have shown that less productive derivations in *-V/-ke/-ka*, which have traditionally been treated as transitives (slot V), may also be regarded as causatives with respect to their syntax and function since they fall within the above definition of causatives (Bugaeva 2008c, 2010b). However, it should be noted that there is a semantic difference between causative constructions marked by different suffixes: *-V/-ke/-ka* mark

direct causation, *-re/-e/-te* direct/indirect causation, and *-yar/-ar* indirect causation (both factitive and permissive).

The causation processes within each of the unproductive lexical causatives *-V* (*-a/-e/-u/-o/-i*, not fully predictable), *-ke*, *-ka* are organized according to the semantics of the base verbs (for details see Bugaeva, forthcoming a): *-V* (50a) – for base verbs denoting processes and states (111 verbs), e.g. *mos* ‘be awake’ → *mos-o* ‘wake sb up, jolt sb awake’; *-ka* – for verbs denoting spontaneous processes/states, emotions and sensation (30 verbs), e.g. *uhuy* ‘burn’ → *uhuy-ka* ‘make sth burn’; *iyunin* ‘be painful’ → *iyunin-ka* ‘hurt sth’; *-ke* (50b) – for motion verbs and some verbs denoting processes/states (10 verbs), e.g. *ray* ‘die (SG)’ (vi) → *ray-ke* ‘kill sb (SG)’ (vt); *yan* ‘land.SG’ (vi) → *yan-ke* ‘land sth.SG’ (vt). Most causatives in *-V* and *-ka* are formed from intransitives, while causatives in *-ke* are formed from intransitives only.

Ainu allows for double causatives which are derived by combinations of direct and indirect causative markers, viz. unproductive *-V/-ke/-ka* + productive *-re/-e/-te* (50c), or two productive suffixes *-re/-e/-te*.

- (50) a. *or-o-wa* [caycaye]<sub>oc</sub> *poro-n-no* *a=Ø=tuy-e*  
 there-POSS-ABL twig big-EP-ADV IND.A=3.O=cut-CAUS<sub>1</sub>  
 ‘Then we cut many twigs.’ (AB 305)
- b. *pira hontom* *a=Ø=ran-ke* *kor* [tus]<sub>s</sub> *Ø=tuy*  
 cliff middle IND.A=3.O=descend-CAUS when rope 3.S=cut  
 ‘When (the boat) was dropped to the middle of the cliff, the rope broke.’  
 (T3 24)
- c. [*hat punkar*]<sub>oc</sub> *a=e<sub>02</sub>=tuy-e-re* *wa karip*  
 grape vine IND.A=2SG.O=cut-CAUS<sub>1</sub>-CAUS<sub>2</sub> and hoop  
*a=Ø=kar*  
 IND.A=3.O=make  
 ‘“I made you cut a grape vine and made hoops (which I put around our house to kill the enemies . . . ,” said my grandmother.)’ (K7803233UP.144)

Just like trivalent applicatives, double causatives or any causative from bivalent transitives, including causatives from applicatives (51), have two unmarked objects, which are, more or less, similar in terms of their encoding and behavioral properties.

- (51) a. *arpa* ‘go (SG)’ → *ko-arpa* ‘go (SG) to sb/some (SG) place’ → *ko-arpa-re* ‘send sth/sb to sb/some place’ (lit. ‘make sth/sb (SG) go to sb/some place’) (vd)
- b. *so-uk* ‘borrow’ (lit. ‘take a debt’) → *e-so-uk* ‘borrow sth’ (vt) → *e-so-uk-te* ‘lend sth to sb’ (lit. ‘make/let sb borrow sth’; lexicalized) (vd)

#### 15.4.4.1.3 Antipassives

An antipassive is a syntactic construction in which the underlying A becomes S (Dixon and Aikhenvald 2000: 9); it is not limited to ergative languages, as is often thought, but may refer to any overtly marked construction that deletes an object (A → S)<sup>30</sup> in a language of any grammatical system. Antipassive in Ainu is marked with the derivational prefix *i-* which originated in the pronominal indefinite O marker *i-* (the latter may probably be traced back to the bound noun/nominalizer *hi/i* ‘place/time/thing’). In Ainu studies, the antipassive is traditionally referred to as the marker of a generalized object ‘(indefinite) person/thing’ (Tamura 1988: 67). It is

a valency decreasing construction in which the object is blocked completely and no oblique expression is possible, thus antipassives from bivalent transitives result in intransitives.

- (52) a. *ora-no nani usey Ø=Ø=kar nea iwatarap*  
 then-ADV immediately hot.water 3.A=3.O=make that baby  
*Ø=Ø=huraye a Ø=Ø=huraye a*  
 3.A=3.O=wash ITR 3.A=3.O=wash ITR  
 ‘Right away, my wife boiled some water, and washed that baby carefully.’  
 (K8109171UP.168)
- b. *ontaro or un Ø=i-huraye*  
 tub place ALL 3.S=APASS-wash  
 ‘She did laundry in a tub.’ (lit. ‘washed things’) (T 218)

Antipassives from trivalent transitives are much less common, but possible; they do not trigger intransitivization, just a valency decrease, as in (53b) and (54b). The base trivalent verbs are usually semantic ditransitives, as in (53a), (54a), i.e. verbs of transfer which contain in their role frame a Theme (T) and Recipient (R). According to Malchukov (Malchukov *et al.* 2010: 31), in the case of ditransitives, it is usually the Theme argument that can be antipassivized, as in (53b), while antipassivization of the Recipient argument is extremely rare cross-linguistically; nevertheless, it is possible in Ainu (54b) (for detailed discussion see Bugaeva 2011b: 248–9).

- (53) a. [*pirka us-ke*]<sub>T</sub> [*Ø=ona-ha*]<sub>R</sub> *Ø=Ø=ko-puni*  
 be.good place-POSS 3.A=father-POSS 3.A=3.O=to.APPL-raise/extend  
 ‘(She) served the good parts of (meat) to her father.’ (K8010291UP.152)
- b. [*a=Ø=kor ekasi*]<sub>R</sub> *hoski-no [nea okkaypo utar]*<sub>A</sub>  
 IND.A=3.O=have grandfather be.early-ADV that man PL  
*ko-i*<sub>T</sub>-*puni*  
 to.APPL-APASS-raise/extend  
 ‘The young men served my grandfather first.’ (K7803231UP.072)
- (54) a. [*a=sa-utar-ih*]<sub>R</sub> [*nep ka*]<sub>T</sub> *a=Ø=e-kasuy*  
 IND.A=elder.sister-PL-POSS what even IND.A=3.O=with.APPL-help  
*ka somo ki no*  
 even NEG do.AUX and  
 ‘I did not help my sisters with anything.’ (K7908032UP)
- b. *tokaci wa Ø=ek Ø=pewre kur [monrayke]*<sub>T</sub>  
 Tokachi ABL 3.S=come.SG 3.S=be.young man work  
*Ø=Ø=e-i*<sub>R</sub>-*kasuy*  
 3.A=3.O=with.APPL-APASS-help  
 ‘A young man who came from Tokachi helped *people* with work.’ (SK 160)  
 From Ishikari dialect.

Generally, antipassives in Ainu are much less productive than applicatives or causatives. They are derived only from verbs of certain classes such as verbs of Pursuit/Affected Subject/Perception/Cognition and Interaction (Bugaeva, forthcoming b). There are many lexicalized antipassives, e.g. *ku* ‘drink sth’ (vt) → *i-ku* ‘drink *alcohol*’ (vi).

## 15.4.4.1.4 Reciprocals and sociatives

The reciprocal meaning is commonly expressed by the prefix *u-*. In most cases (except object-oriented and “indirect” reciprocals), reciprocal derivation involves intransitivization (55); for details of reciprocals and sociatives in Ainu see Alpatov *et al.* (2007).

- (55) a'. *káni anak húci ku=Ø=koyki* (vt)  
 1SG TOP granny 1SG.A=3.O=bully  
 ‘I bullied granny.’ (OI)
- a". *húci (káni) en=koyki* (vt)  
 granny 1SG 1SG.O=bully  
 ‘Granny bullied me.’ (OI)
- b. (*cókay*) *u-koyki=as* (vi)  
 1PL.EXC REC-bully=1PL.EXC.S  
 ‘We bullied each other.’ (OI)

Sociative derivation does not involve any valency change and has no specialized marker. It is commonly expressed by the reciprocal *u-* + polysemous applicative *ko-*. From a purely morphological viewpoint, sociatives are reciprocals derived from applicatives. It is tempting to treat generally all sociatives as reciprocals derived from comitative applicatives.<sup>31</sup> In most cases, however, the combined prefix *uko-* (less frequently *ue-*) functions as a single sociative morpheme, i.e. these derivations are semantically related immediately to the non-applicative base verb without the prefix *ko-* rather than to the formally intermediate applicative form in *ko-* (this latter form may be lacking). Thus, semantically, (56b) relates to (56a) rather than to (56c):

- (56) a. *pakoat* ‘to be accused of a crime’ (T 507; OI) (vi)  
 b. *uko-pakoat* ‘for everyone to be accused of crime together’ (T 760; OI) (vi)  
 cf. c. *ko-pakoat* ‘to get/become involved in sth’ (T 331) (vt)

*uko-* is also productive as a marker of object-oriented reciprocals, e.g. *ninu* ‘sew sth’ (vt) (T 419) → *uko-ninu* ‘sew sth (two or more things) together’ (vt) (T 759).

Another means of expressing sociative meaning is reciprocal *u-* in combination with causative *-re/-e/-te* (less commonly *-ka*), e.g. *mina* ‘laugh’ (vi) → *u-mina-re* ‘laugh together’ (vi) (OI).

The reciprocal prefix combines with certain nouns as a marker of the dual number, e.g. *u-tek* ‘both hands’. It is also used to denote family relations on both nouns and verbs, cf. *ona* ‘father’ → *u-ona-kor* ‘to be (related as) father and child’. There are reciprocal forms from relational nouns and postpositions, cf. *sam* ‘proximity’ → *u-sam* ‘next to each other’, *tura* ‘with sb/sth’ → *u-tura* ‘with each other, together’.

## 15.4.4.1.5 Reflexives and anticausatives

Most Ainu grammars mention two reflexive prefixes, *yay-* and *si-*, with a note that *yay-* expresses an intentional action and *si-* an unintentional action, i.e. a kind of middle voice (Tamura 2000 [1988]: 204). This was first pointed out in Kindaichi (1993 [1931]: 280), but he gives only one minimal pair, *yay-pusu* ‘float up (by one’s own effort)’ (lit. ‘draw out oneself’) vs. *si-pusu* ‘float up (by itself)’ (lit. ‘draw out by oneself’). There have been made a few recent attempts to uncover the difference between *yay-* and *si-*. Kirikae (1994: 316) suggests that *yay-* is “the agentive self” and *si-* is “the patientive

self,” which is close to Kindaichi’s original interpretation. Satō (2007a: 31) tries to explain the difference in terms of the notions of “direct reflexive” and “indirect reflexive”: “In the direct reflexive (*yay-*), the coreferential subject participates in the activity directly, while in the indirect reflexive (*si-*), it participates in the activity only indirectly either in the sense that the subject ‘does’ the activity only with the help of some other person or in the sense that the subject is related to the activity merely locationally,” because there is also a reflexive-causative use of *si-* and a few other cases that are hardly unintentional, e.g. *kasuy* ‘help sb’ (vt) → *kasuy-re* ‘make/let sb help sb’ (vd) → *si-kasuy-re* ‘make/let sb help oneself’.

I suggest that *yay-* has a major reflexive function and also a number of other functions usually expressed by reflexive markers in other languages (Geniušienė 1987): reflexive-possessive, e.g. *ko-yupu* ‘fasten sth to sth/sb’ (vb) → *yay-ko-yupu* <REFL-to.APPL-fasten> ‘fasten sth to oneself’ (vt) (AB 47), anticausative, e.g. *osura* ‘throw sth/sb’ → *yay-osura* ‘throw oneself’ (NV 93), and, occasionally, anticausative functions, e.g. lexicalized *nu* ‘listen to sth’ (vt) → *yay-nu* <REFL-hear> ‘think’ (vi) (AB 47).

*si-* has a major anticausative function (original O turns into S), implied by its traditional interpretation as middle/unintentional reflexive. However, it may also be used productively as anticausative, e.g. *etaye* ‘pull sth’ (vt) → *si-etaye* <REFL-pull> ‘pull oneself in (go back inside)’ (vi) (AB 47), and reflexive-causative. It may also mark reflexivity on relational nouns, e.g. *osmak* ‘behind’ → *si-y-osmak* ‘behind oneself’ (T 671), which is probably indicative of its reflexive origin; the anticausative and other “middle” functions appeared later, presumably under the functional pressure of a newer reflexive marker *yay-*.

*si-* is not the only anticausative marker. The “intransitive suffix” *-ke* (Table 15.4, slot V)<sup>32</sup> may also be regarded as an anticausative marker (although a much more lexical one than *si-*). There is an alternation between anticausative *-ke* and causative *-V* on bound stems such as *yas-* in (57a). Contrary to Refsing (1986: 188), who is dealing with a different dialect, the present research has revealed that the derivations in *-ke* in Southern Hokkaido Ainu are not limited to only a few verbs; there are 24 verbs in my list, all with very typical anticausative meanings (Bugaeva 2010b).

- (57) a. [*hon-i*]<sub>O</sub>     $\emptyset$ =*yas-a*                    *wa*     $\emptyset$ =*cipor-i*                     $\emptyset$ =*san-ke*  
 belly-POSS    3.S=tear-CAUS    and    3.A=roe-POSS    3.S=descend-CAUS  
 ‘(He) tore off the (salmon’s) belly and took off its roe.’ (N 394)
- b. *kotan noski wa*     $\emptyset$ =*poro* [*kem tak*]<sub>S</sub>     $\emptyset$ =*hopuni wa*  
 village middle ABL    3.S=big blood clot 3.S=fly and  
  
*kotan enka ta*     $\emptyset$ =*yas-ke*  
 village over LOC    3.S=tear-ACAUS  
 ‘A big clot of blood flew from the middle of the village and tore apart above the village.’ (AB 177)

#### 15.4.4.1.6 Noun incorporation

Noun incorporation is where a core argument (subject or object) of a clause becomes “attached to” or “incorporated into” the verb. There are three major types of syntactic noun incorporation in Ainu: O, S, and non-agentive A. Noun incorporation is not so common in the world’s languages and A-incorporation is hardly attested elsewhere.

Functionally, incorporation is characterized as a backgrounding process (Hopper and Thompson 1980: 254), i.e. when the event is of greater interest than its participants, and is unlikely to apply to arguments of high discourse salience, high animacy, specificity, etc. It is often noted that O-incorporation leading to an intransitive root is most commonly used when referring to conceptually unitary and noteworthy cultural activities (Mithun 1984).

O-incorporation is by far the most common type in Ainu and cross-linguistically, leading to a decrease of valency by one; cf. the first *turep* as an incorporated O in (58) and the second as a free noun, and the change of a S personal affix for A.

- (58) *turep-ta=as kus paye=as wa poro-n-no*  
 wild.lily-dig=1PL.EXC.S since go.PL=1PL.EXC and many-EP-ADV  
*turep ci=Ø=ta wa sap=as*  
 wild.lily 1PL.EXC.A=3.O=dig and descend.PL=1PL.EXC.S  
 ‘We went to dig wild lily (bulbs), so we came with a lot of wild lily (bulbs).’  
 (ST1 220)

A-incorporation (59) also results in valency decrease: the original A is removed from an argument position by incorporation, while the original O promotes to S.

- (59) a. *nis Ø=en=reve-re*  
 cloud 3.A=1SG.O=crawl-CAUS  
 ‘The clouds carry me (slowly).’ (ST 197)  
 b. [*káni*]<sub>s</sub> *ku=nis-reve-re*  
 1SG 1SG.S=cloud-crawl-CAUS  
 ‘I am (slowly) carried by clouds.’ (lit. ‘I am cloud-carried’) (ST 197)

S-incorporation may either decrease valency by one and result in a zero-argument verb (60), or rearrange the valency by introducing a new S argument (61b) when the noun (usually a bodypart or relational noun) is incorporated in its possessive form.

- (60) *tan-to sir-pirka siri!*  
 this-day appearance-be.good VIS.EV  
 ‘It is fine weather today.’ (NN 38)
- (61) a. *asinuma anak-ne a=kema Ø=pase*  
 IND TOP-COP IND.A=leg.POSS 3.S=be.heavy  
 ‘My legs became heavy.’ (lit. ‘As to me, my legs became heavy.’) (ST 197)  
 b. *tane anak-ne kema-pase=an pe ne kusu*  
 now/already TOP-COP leg.POSS-be.heavy=IND.S NMR COP because  
 ‘I already grew old, so . . .’ (lit. ‘There was already leg-heaviness, so . . .’)  
 (N 173)

O-Incorporation is very common with applicative verbs in which it may not only apply to base objects of transitives (62), but also to applicative objects of base bivalent (63a) and trivalent (64a) transitives with particular roles of applicative objects, viz. *e-* applicatives with Instrument and Cause/Purpose roles, *ko-* applicatives with Co-patient (64) and Goal roles, and *o-* applicatives with Location and Goal (63) roles (Bugaeva 2010a: 788–92).

- (62) *sisam anak-ne kampi cipetpap Ø=Ø=e-kamuy-nomi*<sup>33</sup>  
 Japanese TOP-COP paper torn.paper 3.A=3.O=with.APPL-god-honor

*ki p ne*  
do.AUX NMR COP

‘The Japanese really honor gods with torn paper.’ (AB 357)

- (63) a. *ar-kamuyasi or-o a=Ø=o-arpa ruwe ne*  
complete-devil place-POSS IND.A=3.O=to.APPL-go.SG INF.EV COP  
‘(If one dies at the same place of a devil), one will go to the place of devils.’ (N 20)
- b. *kamuy-or-o-arpa=an ka e-aykap korka*  
god-place-to.APPL-go.SG=IND.S even of.APPL-be.unable.AUX but  
‘I couldn’t even go to the other world [lit. to the land of gods].’  
(K8109171UP.121)
- (64) a. *a=Ø=uhuy-ka pa wa cise*  
IND.A=3.O=burn-CAUS PL and house  
  
*a=Ø=ko-uhuy-ka pa*  
IND.A=3.O=with.APPL-burn-CAUS PL  
‘(After that, since they did that to us, let’s) burn them down, (let’s) burn them  
down with the (entire) house.’ (K7908032UP)
- b. *Ø=hotke kurka ta ponyaunpe*  
3.S=sleep top LOC Ponyaunpe  
  
*a=Ø=cise-ko-uhuy-ka*  
IND.A=3.O=house-with.APPL-burn-PL  
‘We burned Ponyaunpe down with the house at the place where he had slept.’  
(O4 12) Shizunai dialect.

One of the objects, either the causee (65a) or original object (65b), can be incorporated in the case of a trivalent causative verb.

- (65) a. *sut-ketusi a=Ø=san-ke wa,*  
grandmother-chest IND.A=3.O=descend.SG-CAUS and  
  
*a=Ø=tek-kus-pa-re wa inkar=an akusu,*  
IND.A=3.O=hand-pass.through-PL-CAUS and look=IND.S then  
‘I took down my dowry chest and put my hands into it and looked in, . . .’  
(AB 153)
- b. *hoski Ø=ek ay . . .*  
before 3.S=come.SG arrow  
  
*a=Ø=si-y-oka-kus-te*  
IND.A=3.O=REFL-EP-behind-pass.through-CAUS  
‘I sent the arrow which came first behind myself.’ (AB 348)

In the case of trivalent causative-applicatives such as *sik* ‘be full’ (vi) → *e-sik* ‘be full of sth’ (vt) → *e-sik-te* ‘fill sth with sb’ (vd), there are examples of incorporation of both objects with slight lexicalization.

- (66) *a=po-utar-i ne yak-ka Ø=[po]<sub>oa</sub>-[sir]<sub>oc</sub>-e-sik-te*  
IND.A=son-PL-POSS COP if-even 3.S=son-world-by.APPL-be.full-CAUS  
‘My sons also had many children [lit. filled the world with sons].’ (ST2 29)

There is only one general restriction on O-incorporation: objects which are animate, specific, and referential at the same time cannot be incorporated. As a result, we never

find incorporated causee objects in the case of indirect causatives (both (65a) and (66) are direct causatives) and any applicative objects with the Recipient/Beneficiary role.

O-incorporation is frequent in Ainu (S/A-incorporation is rare), especially in oral literature. It is quite favored in the case of base trivalent verbs and can be regarded as a convenient strategy for avoiding the heavy double-object construction. In the case of base bivalent transitives, O-incorporation is most common with verbs of Pursuit/Affected Subject/Perception/Cognition and Interaction (Bugaeva, forthcoming b), just like antipassives (15.4.4.1.3).

#### 15.4.4.1.7 Valency classes in Ainu

As we have seen, Ainu possesses a number of valency-changing alternations characterized by high combinability but, of course, with limits. Most restrictions have a semantic basis, i.e. certain alternations tend to apply only to verbs of certain semantic classes. Thus, valency-changing alternations appear helpful in uncovering underlying principles of verb-categorization. Here is a tentative classification of Ainu valency classes arrived at by a systematic (MPI) database-related study of the behavior of 70 verbs in terms of their coding properties and accessibility to valency-changing alternations (Bugaeva, forthcoming b, c). Reciprocals and reflexives were not included.

- I Transitive (vt):
  - a. Effective Action (~ ACAUS **-ke**);
  - b. Contact (APPL **e-**).
- II Middle (vt & deponent vi with APASS **i-**):
  - a. Pursuit/Affected subject/Perception/Cognition (vt: APASS **i-**, APPL **ko-**)
  - b. Interaction (deponent vi with APASS **i-**: APPL **ko-**, **e-**, double APPL **e-ko-/ko-e-**);
- III Intransitive (vi):
  - a. Emotion/Human Affect/Ability (vi & deponent vt with APPL **e-**, **ko-**) (vi: APPL **e-**; vt with APPL **e-**: APPL **ko-**);
  - b. Motion/Location/Existence (APPL **e-**, **ko-**, **o-**);
  - c. Spontaneous (actions/processes/states) (direct CAUS **-V**, **-ke**, **-ka**, and **-re/-e/-te**).
- IV Avalent (v0):
  - a. Meteorological (none).

Interaction verbs are most prone to valency-changing alternations, e.g. *ruska* ‘be angry with’ (42). Trivalent transitive verbs do not align themselves as a separate class because all of them are derived verbs. The valency classes proposed for Ainu largely comply with the general hierarchy proposed in Tsunoda (1985) and two-dimensional hierarchy proposed in Malchukov (2005).

#### 15.4.4.2 Complex VPs

Ainu verbs may be marked for aspect, modality, and evidentiality, but there are no pure tense markers. These markings occur in complex VPs such as auxiliary verb constructions “V + auxiliary verb” and pair verb constructions “V + conjunction + supplementary verb.”

## 15.4.4.2.1 Auxiliaries: aspect, modality, and evidentiality

Auxiliary verbs always follow notional verbs, and generally originate from lexical verbs. Most form a single accented phonological unit with the verb. Only the verb takes pronominal affixes, while auxiliaries are left unmarked. A verb may have more than one auxiliary if there is no semantic interference (see Fukuda (Tamura) 2001 [1960]: 57–68). Here is a list of major auxiliaries from Tamura (2000 [1988]: 111–21) with my interpretations; Tamura’s interpretations are in quotes. The choice of SG/PL form of the auxiliary is determined by the number of the subject of the predicate (i.e. notional verb).

1. perfect aspect *a* (SG)/*rok* (PL) (“PAST/PAST-PL”);
2. durative (with atelic verbs) or iterative (with telic verbs) aspect . . . *a* (SG)/*rok* (PL) . . . *a* (SG)/*rok* (PL) (“emphatic continuative”);
3. admirative mood *aan* (SG)/*rok’oká* (PL) (cf. *an’an* in Chitose (Bugaeva 2004: 76)) (“determined”);
4. desiderative mood *rusuy* (“want”), see (15d) and (19b);
5. iterative or habitual aspect *ranke* (“repetition”), see (14) and (52);
6. conjunctive probabilitive mood *nankor* (“perhaps”);
7. recent perfect aspect *nisa* (“just finished”);
8. deontic mood *esirki* (“must”), and six other auxiliaries (see Tamura 2000 [1988]: 117–21).

Kindaichi (1993 [1931]: 293) terms *a* a perfective marker, M. Chiri (1974 [1936]: 157) a perfect (*kanryō*) or past (*kako*) marker, and Tamura (2000 [1988]: 111) uses the glossing PAST in spite of her very clear description of *a* as a perfect marker in Fukuda (Tamura) (2001 [1960]: 61) as “a past situation described by the verb(s) marked by *a* (those preceding *a*) has some consequences which are relevant for a present situation described by the verb(s) following *a*”, so I prefer the perfect interpretation.

- (67) *a*=∅=*roski*                      *a*      *inaw*                      *opitta*      *kuca* . . .                      *or*      *ta*  
 IND.A=3.O=stand.PL      PERF      prayer.stick      all      hunting.hut      place      LOC  
*i*=*hekota*                                      ∅=*hácir*                      *wa*      ∅=*okay*  
 IND.O=in.the.direction.of      3.S=fall.down      and      3.S=exist.PL  
 ‘All the *inaw*-whittled prayer sticks (used as offerings to gods) I *had* stood at the hunting hut fell down in my direction.’ (AB 187)

More insights into *a* (SG)/*rok* (PL) may be found in Satō (2007b), who notes that *a* denotes only “actional perfect” and, for this reason, does not co-occur with negative forms of verbs, the latter require the use of “stative perfect” which is encoded with a pair verb construction (same as that of *hácir* ‘fall down’ in (67)), see 15.4.4.2.2.

There are also verbs expressing Aktionsart, ability, knowing, etc. which may function either as auxiliaries or as lexical verbs, e.g. *okere* ‘finish’, *tunas* ‘fast’, *eramiskari* ‘not know’. As auxiliaries they allow for insertion of an emphatic particle *ka* and some other words, indicative of their intermediate status between verbs and auxiliaries.

Another set of undergrammaticalized auxiliaries are evidentials. Tamura (2000 [1988]: 227) describes them as: “The expressive nominalizers *ruwe* eEVD, *hawe* eSAID, *siri* eSEEN, and *humi* eFELT can be placed after sentences [=clauses – A.B.] that end with verb phrases, where they nominalize the sentence [=clause], and the copula *ne* is placed afterwards to complete the phrase,” which is suggestive of a nominalized clause analysis. As Refsing (1986: 236) points out, this construction is very similar to Japanese *n(o)-desu/da* (cf. also *yotei-da*, *yō-da*, *mono-da*, *hazu-da*, etc. – A.B.) constructions, and its

frequency in both languages might be either an argument in favor of their affinity or of a syntactic loan from Japanese (or the strengthening of an otherwise not very prominent feature in Ainu). Moreover, this construction is uncommon in the world's languages, attested so far only in Ryukyuan, Korean, Mongolian, Turkish, and some Tibeto-Burman languages (Tsunoda 1996: 158), and not straightforward in terms of syntax.

In Ainu, it is clear that the nominalized construction in question originated in a modifying clause, in which the head nouns (nominalizers) may be traced back to common nouns (see 15.3.1. and literal translations in (68)–(71)). At the present stage, the nominalized construction is on the verge of turning into a complex predicate construction consisting of verb + auxiliary verb (noun + copula), *ruwe/hawe/siri/humi-ne*.<sup>34</sup> In fact, the auxiliary analysis is taken in Kindaichi (1993 [1931]: 326–36) and M. Chiri (*joshi*, 1974 [1936]: 132–3, 155–7). Although, synchronically, I prefer Tamura's nominalized clause analysis, the respective constructions are discussed here due to a high possibility of their future reanalysis as auxiliaries.

(68)–(71) illustrate the contrast between inferential *ruwe ne* 'it is a fact that', reportative *hawe ne* 'it is said that', visual *siri ne* 'it looks that', and non-visual (= semblative) *humi ne* 'it feels that'.

- (68) *tokey ku=Ø=sak ruwe ne*  
 clock 1SG.A=3.O=lack INF.EV COP  
 '(I infer that) I don't have a clock.' (lit. 'It is **the trace of** me lacking a clock.')
- (69) *yaysama, Ø=wen hawe ne*  
 damn 3.S=be.bad/poor REPEV COP  
 'Damn! It's bad! (assumed from what has been said)' (lit. 'It is **the voice of** its being bad.')
- (70) *húci Ø=ek kor Ø=an siri ne*  
 grandmother 3.S=come.SG while 3.S=exist.SG VIS.EV COP  
 'The elderly woman is here (visible)/(I see that) Grandmother is coming.' (lit. 'It is **the sight of** grandmother coming.')
- (71) [*eytasa Ø=tópen humi*] *ne pekor ku=yay-nu*  
 too.much 3.S=sweet NONVIS.EV COP in.that.way 1SG.S=REFL-hear  
 'I think it is too sweet.' (lit. 'I think it is **the sound of** being too sweet.')

In addition to proper evidential functions, there developed a number of epistemic functions such as an assertive function of *ruwe* and probability function of *hawe*, etc., which have not received special attention so far.

#### 15.4.4.2.2 Pair verbs: aspect, modality, and miscellaneous

A few aspectual, modal and other meanings are expressed by a special type of construction "V + conjunction + supplementary verb" (Tamura 2000 [1988]: 46, 151, 184), referred here as a "pair verb" construction since similar constructions are registered in Dravidian (Steever 1988), Tibeto-Burman, and even Russian (Hook, Peter, pers. comm.). The first verb (V) expresses an action/state, the second (supplementary) verb brings in an aspectual/modal meaning, but both verbs have full finite forms, take pronominal affixes (72), and do not necessarily involve coreferential subjects (72h). Only a few conjunctions are used to

link the two verbs (*wa* ‘and’, *kor* ‘when/while/if/and’, rarely *hine* ‘and’); their use is largely fixed within certain constructions.

Most frequently used are the stative perfect (= perfect resultative) . . . *wa an/oka* ‘exist.SG/PL’ and progressive . . . *kor an/oka* ‘exist.SG/PL’; recall that there is a separate auxiliary-based construction for actional perfect; all three constructions are expressed with *-te iru* in Japanese. The aspectual pair verb constructions are often used as a test for dividing Ainu verbs into actions and states/processes (Nakagawa 1981; elaborated in Satō 2006). Regardless of their syntactic transitivity value, action verbs, e.g. *ye* ‘say’ (vt), *monrayke* ‘work’ (vi) (72a), naturally enter the progressive construction while processes, e.g. *ray* ‘die’ (vi) (72g), *kor* ‘have sth/sb’ (vt) (72e), do not; the latter allow only for the perfect resultative (72d,f). On the other hand, intransitive action verbs, e.g. *monrayke* ‘work’ (72b), cannot occur in the perfect resultative (but in the actional perfect instead (72c)), but transitive action verbs that imply a change of state can, the resultant state referring to that of their object as is clear from the non-coreferential subject marking in (72h). The progressive construction also has habitual and inceptive functions.

- (72) a. *tane ku=monrayke kor k=an ma*  
 now 1SG.S=work and 1SG.S=exist.SG FIN  
 ‘I am working now.’ (ST1 195)
- b. \**tane ku=monrayke wa k=an ma*  
 ‘I have worked.’ (my constructed example)
- c. *ku=monrayke a*  
 1SG.S=work PERF  
 ‘I have worked.’ (my constructed example)
- d. *tu tamasay ku=∅=kor wa k=an ma*  
 two bead.necklace 1SG.A=3.O=have and 1SG.S=exist.SG FIN  
 ‘I have two bead necklaces.’ (ST1 195)
- e. \**tu tamasay ku=∅=kor kor k=an ma*  
 lit. ‘I am having two bead necklaces.’ (my constructed example)
- f. (*isepo*) *∅=ray wa ∅=an*  
 hare 3.S=die and 3.S=exist.SG  
 ‘(The hare) died.’ (ST1 195)
- g. \*(*isepo*) *∅=ray kor ∅=an*  
 ‘(The hare) is dying.’ (ST1 195)
- h. *inaw k=∅=asi wa ∅=an ruwe ne*  
*inaw* 1SG.A=3.O=stand.SG and 3.S=exist.SG INF.EV COP  
 ‘I have put an *inaw* (whittled prayer stick).’ (lit. ‘I put *inaw* and it was.’)  
 (ST1 195)

There are other pair verb constructions that have equivalents in Japanese and may have developed under its structural influence: perfective *wa isam* (‘disappear’; J *-te shimau*), benefactive *wa kor-e* (‘give’; J *-te kureru*), preparatory *wa anu/ari* (‘put.SG/PL’; J *-te oku*), ‘try . . .’ *wa inkar/inu* (‘see/hear’, (65a); J *-te miru*); also *wa ek/arki* (‘come.SG/PL’; J *-te kuru*), *wa arki/paye* (‘go.SG/PL’; J *-te iku*), and polite imperative *wa pirka* (cf. J *-te ii (yo)*).



There are many ways of rendering commands with affirmative sentences (with auxiliaries, pair verbs, etc.), which vary in the degree of politeness; see (Tamura 2001 (1977) 22–141) and “periphrastic imperative expressions” in Bugaeva (2004: 20–93).

#### 15.4.5.3 Questions: polar, content, and alternative

Questions do not involve any distinctive word order but do involve rising intonation at the end. Polar questions may be formally unmarked or may optionally be marked with final particle *ya* which makes the question softer. In polar questions marked for evidentiality, the final copula *ne* is omitted (cf. 68).

- (77) a. *nisatta nupurpet or un e=arpa ya?*  
 tomorrow Noboribetsu place ALL 2SG.S=go.PL Q  
 ‘Will you go to Noboribetsu tomorrow?’ (NN 28)
- b. *yosiko e=ne ruwe?*  
 Yoshiko 2SG.A=COP INF.EV  
 ‘Are you Yoshiko?’ (NN 22)

Content questions involve the use of question-words which occupy the same position as the corresponding non-question constituents, e.g. *hunna* ‘who’ (78), *hemanta/hnta* ‘what’, *hunak/hinak* ‘where’ in Saru (Tamura 2000 [1988]: 134–5).<sup>35</sup> The particle *ya* is used in content questions comparatively rarely (Tamura 2000 [1988]: 236). For content questions with the copula, *ne* is often replaced with the existential verb *an* ‘to exist.SG’ (vi), and in evidential clauses it is obligatorily replaced with *an* (Tamura 2000 [1988]: 236–7), as in (78b).

- (78) a. *eani hunna e=Ø=ko-isoytak?*  
 you.SG who 2SG.A=3.O=to.APPL-talk  
 ‘Who are you talking to?’ (O1)
- b. *núman hunna Ø=ek ruwe an?*  
 yesterday who 3.S=come.SG INF.EV exist.SG  
 ‘Who came yesterday?’ (T2 237)

An alternative question consists of two juxtaposed questions; the second may be partially truncated. The particle *he* is placed after each questioned constituent which is the focus of the question. For further details on questions see Tamura (2001 [1978] 142–54).

- (79) *matci he e=Ø=kor rusuy tampaku he e=Ø=kor*  
 match FOC 2.SG.A=3.O=have DESID cigarette FOC 2SG.A=3.O=have  
*rusuy*  
 DESID  
 ‘Do you want the matches or the cigarettes?’ (T2 234) (the glossing is mine – A.B.)

#### 15.4.6 Clause linking: coordination, subordination, complementation

Coordination and subordination are marked with conjunctions. The most commonly used coordinating conjunction is *wa* ‘and’ (20a, 50c, 57a/b, 58, 64a), which can also be used in the case of non-coreferential subjects. Other coordinating conjunctions are *híne* ‘and’, *ayne* ‘and finally’ (21), *awa* ‘and then’, *akusu/akus* ‘then’ (65a), *hike* ‘and’, *hike ka*

'but', *kor* 'and' (21), and, presumably, *korka* 'but' (19b, 63b). There are many subordinating conjunctions, viz. *kusu* 'because' (15d, 18b, 24a), *kuni(ne)* 'in order', *kor* 'when/while/if' (50b), (70), *yak* 'if' (17a), *yak-un* 'if', *ciki* 'if' (17b), *yak-ka* 'even if' (45b), (66), *hike ka* 'in spite of', *pe/p (ora)* 'although', and others (Tamura 2000 [1988]: 157–73; no discussion of coordination vs. subordination).

Complement clauses in Ainu may have the syntactic functions of S, O, and rarely A. The link between the complement and main clause is usually marked by a complementizer, which may have a neutral meaning like English 'that' or a specific meaning that adds epistemic overtones to the overall meaning of the complement clause; the latter type has drawn the attention of typologists only recently (Frajzyngier 1995). Most commonly used complementizers are neutral *hi/i* (< 'thing/place'), *pe/p* (< 'thing/person') or epistemic *kun-i(-hi)* <going.to/maybe/surely/should-thing(-POSS)>; the choice of a complementizer depends on the context and semantic class of the main clause predicate (24a) (for details see Bugaeva 2008a: 67–8).

## 15.5 LEXICON

Ainu has few words to express abstract concepts or the products of modern civilization, but it is rich in vocabulary on plants, animals, hunting, fishing, and gathering. It possesses an interesting numeral system based on ten and twenty. There are non-composite numbers from 'one' to 'six' (*sine-p*, *tu-p*, *re-p*, *ine-p*, *asikne-p*, *iwane-pe*; and 'six' naturally means 'many'), and also for *wan-pe* 'ten' and *hot-ne-p* 'twenty', but no number for 'hundred'. Numbers from 'six' to 'nine' are composed by how far they are from 'ten', e.g. *tu-p-esan-pe* <two-thing.CL-APPL-ten-thing.CL> 'eight' is lit. 'two to ten'. Numbers over 'ten' but less than 'twenty' are built up on 'ten', while all numbers over 'twenty' are built up on 'twenty' by using multiplicative, subtractive, and additive combinations (Tamura 2000 [1988]: 255), e.g. *wan-pe e-tu-hot* <ten-thing.CL APPL-two-twenty> 'thirty' (H 262), implied reading 'forty minus ten', but lit. 'together with ten forty/on ten – forty' (H 262).

Most loanwords in Southern Hokkaido Ainu are from Japanese, e.g. *sippo* 'salt' (< J *shio*), *umma* 'horse' (< J *uma*), *tiki* 'wine cup' (< J *saka-zuki*), *kampi* 'paper' (< J *kami*). There are also a few loans from Ainu into Japanese, primarily introduced with the items referred to, e.g. *sake/shake* 'salmon' (< A *sak-ipe* <summer-food>?), *rakko* 'otter', *shishamo* 'smelt' (< A *súsam*). Ainu has had no effect on the grammar of Japanese; however, Ainu grammar has been significantly influenced by Japanese, which will be briefly discussed next.

## 15.6 CONCLUDING REMARKS: AINU VS. JAPANESE

On a deep level, Ainu and Japanese are structurally very different languages, manifested by the presence of such properties in Ainu as pronominal verbal marking, no case marking on core arguments, mixed (but basically tripartite) alignment, head-marking possessive construction, no tense marking, but a great number of coded valency alternations (including applicatives) and noun incorporation, all lacking in Japanese. However, despite these profound structural dissimilarities, most Japanese learners of Ainu (non-linguists) have an impression that the two languages are quite similar and, indeed, they generally master Ainu much quicker than native speakers of European languages. A crucial issue here is the similarity of phonological systems and the same constituent order (in positive affirmative sentences), and also an apparent similarity and even intertranslatability of some

analytic grammatical constructions which were probably acquired or transformed in Ainu due to contact with Japanese.

Here is a list of constructions which were presumably developed in Ainu under the structural influence of Japanese: (a) relative clauses and other noun-modifying constructions (15.4.3), and (b) evidential constructions (15.4.4.2.1); (c) pair verb constructions (15.4.4.2.2), which often copy the semantics of various Japanese converbial constructions; and (d), in the case of impersonal passive, the possibility of expression of the demoted Actor with the oblique allative phrase (which may be regarded as the influence of the Japanese passive) (26b).

In the case of the last Ainu speakers, the structural (and phonological, see 15.2.1) influence of Japanese is even more prominent; recall an increased use of postpositions and decline in the use of applicatives (15.4.4.1.1) and also the use of benefactive pair verb construction *wa kore* (36) (cf. *J -te kureru*) instead of *ko*-applicatives (22, 53a/b) which have no postpositional benefactive paraphrase. The issue of contact-induced grammatical change in Ainu has not received special attention so far and requires further consideration.

## TAGS

Grammatical tags other than those listed at the start of this volume: person/number indexing 1SG, 2SG, 1PL, 1PL.EXC(lusive), 2PL, 3, IND(efinite); A = transitive subject or possessor, S = intransitive subject, O = object; plus: ACAUS = anticausative, ADV = adverbial, ALL = allative, APASS = antipassive, APPL = applicative (often more specifically with.APPL, etc.), CC = copular complement, CS = copular subject, CL = classifier, COM = comitative, DESID = desiderative, DIM = diminutive, EP = epenthetic consonant, FIN/FP = final particle, HOR = hortative, INF.EV = inferential evidential, ITR = iterative, NONVIS.EV = non-visual evidential, O<sub>a</sub> = applicative object, O<sub>c</sub> = causee object, P = patient, PERF = perfect, POSS = possessive, R = recipient, REC = reciprocal, REFL = reflexive, REP.EV = reportative evidential, SG = singular, VIS.EV = visual evidential.

## NOTES

\* I would like to thank Emma Geniušienė for advising me on reflexivity, Iku Nagasaki on Japanese phonology, Peter Hook on pair verbs, and the editor Nicolas Tranter for his careful work. My deepest heartfelt thanks go to the late Vladimir Nedjalkov (1928–2009), who was my mentor, and to the late Ainu language consultant Mrs. Ito Oda (1908–2000). None of the above-mentioned researchers necessarily share my views expressed here, nor are they responsible for any errors.

- 1 A collection of works dealing with the Indo-European hypothesis of the origins of Ainu may be found in Refsing (1998).
- 2 *Moshiogusa* was compiled by Kumajiro Uehara and Chōzaburō Abe in 1792 and contains 2,740 words and several texts; it was republished in Kindaichi (1972).
- 3 Proto-Austroasiatic is reconstructed on the basis of proto-Mon-Khmer, proto-Viet-Muong, proto-Wa, and proto-Monic.
- 4 The details of the spiritual life and world-view of the Ainu may be found in Munro (1962), Fitzhugh and Dubreuil (1999), and Ainu Minzoku Hakubutsukan (2009 [1993]).
- 5 A famous Ainu textbook Tamura (1979, 1983a, 1984) with authentic audio recordings is not mentioned among the community oriented textbooks because it lacks *katakana*

- transcriptions of Ainu which are necessary for the community use; this textbook is now available online (<http://dspace.wul.waseda.ac.jp/dspace/handle/2065/873>). Nowadays, converting a Latin transcription of Ainu into a *katakana* transcription and vice versa is no longer a problem because there is special software for this; the converting is possible online free of charge: <http://www.geocities.jp/aynuitak/WEBhenkan/chiyu.htm>.
- 6 There are regular courses on Ainu in the following Japanese universities: Chiba University, Hokkaido University, Hokkaido Gakuin University, Sapporo Gakuin University, Waseda University, and occasional courses at Tokyo University and Tokyo University of Foreign Studies.
  - 7 Available online at <http://www.stv.ne.jp/radio/ainugo/index.html>.
  - 8 A brief description of early written records of Ainu may be found in Tamura (2000 [1988]: 7–9). There is also a comprehensive collection of early European records of Ainu edited by Refsing (1996).
  - 9 Most articles on Ainu grammar written by Japanese researchers in the second half of the twentieth century are reprinted in the book *Ainugo-kō, Bunpō* I–II, vols 4, 5 (2001), Tokyo: Yumani Shobō.
  - 10 The northeastern dialects (e.g. Tokachi, Obihiro, etc.) would probably be more representative of “Ainu” as a whole because they are spread over a much vaster area.
  - 11 Refsing (1986: 69) also suggests that there are diphthongs in Ainu, but she is dealing with a different dialect (Shizunai; HNE) of which I have no direct knowledge.
  - 12 Vovin (1993: 175) reconstructs for proto-Ainu twelve short and six long vowels. Personally I find it very puzzling that long vowels (five of them) are retained only in Sakhalin Ainu because in all other respects these dialects are much “newer” than those of Hokkaido.
  - 13 For the Rachishka dialect of Sakhalin (SWC), Murasaki (2009: 6) notes that it also has a distinguishable pitch accent which is, however, not distinctive.
  - 14 In some lexicalized causatives, the suffix *-re/-e/-te* does affect accentuation, e.g. *kor* ‘have’ > *ko.r-é* ‘make/let have, give’ (in accordance with the accentuation rules).
  - 15 Satō’s (2008: 3) classification is similar to that of Tamura (2000 [1988]: 36), but he does not single out conjunctions and pronouns as separate word classes, including the former in the class of nouns and the latter in particles; there is also a difference in his classification into subclasses.
  - 16 “Spatial and conceptual demonstratives” appear on pp. 260–4, but on pp. 92–3 they are referred to as “demonstrative and referential adnominals.”
  - 17 The abbreviation S stands for intransitive subject, A for transitive subject, and O for object of transitives; none of them indicates any particular semantic role (Dixon and Aikhenvald 1997: 72).
  - 18 The locative ‘be/exist’ is expressed by a special locative copula *an/oka* ‘exist, appear (SG/PL)’, which is an intransitive verb.
  - 19 Especially the third person pronouns almost never occur in texts; common nouns are used instead.
  - 20 Also referred to as the impersonal passive (Bugueva 2011a), cf. also (26b). The impersonal passive analysis was adduced by Satō (1995).
  - 21 Traditionally, these functions of the indefinite have been treated as homonymous forms expressing different meanings (Kindaichi 1993 [1931]; M. Chiri 1974 [1936]; Kubodera 1977; Tamura 1970, 2001 [1972a]; Shibatani 1990; Satō 2004). The unified approach to the analysis of the indefinite person accepted in this chapter was adduced in Refsing (1986: 219) for Shizunai (HNE), which is quite different from Southern Hokkaido Ainu, and later taken over in Tamura (1988 [2000]) for Saru (HSW) and Nakagawa (1995, 1997b) for Chitose (HSW).
  - 22 The etymology of *asinuma* is rather transparent: it consists of the pronominal indefinite marker *a=*, followed by the third person singular pronoun *sinuma*

- ‘(s)he’, which, in my view, suggests something like ‘self’, cf. Tamura’s early interpretation of *asinuma* as “a speaker, a person in question” (1972a: 376).
- 23 Here and henceforth, the indefinite has a logophoric function because the examples constitute a part of the larger quotational context of a folktale.
- 24 The choice of the respective conceptual/possessive forms of relational nouns (e.g. *or* ‘place’ or *or-o* ‘the place of’) is dictated by the animacy properties of their preceding nouns: animates require the possessive forms (26b) and inanimates most commonly co-occur with the conceptual forms (23), (24) (less commonly with the possessive forms).
- 25 A general term referring to all types of clausal modifiers (i.e. not just relative clauses in a narrow sense) adduced in Matsumoto (1997).
- 26 A comparative analysis of noun-modifying constructions and nominalizations in Japanese and Ainu may be found in Okuda (1989), but constructions like (40) are not included in the discussion.
- 27 Note that the indefinite causative suffix *-yar/-ar* does not affect valency because the causee is indefinite and cannot be introduced as an argument.
- 28 A detailed discussion of the encoding and behavioral properties of trivalent transitives, both applicatives and causatives, may be found in Bugaeva (2011b).
- 29 If one looks at the poetic genres of oral literature recorded from the same speaker (Suteno Orita, Shizunai dialect; HNE) that Kirsten Refsing has worked with, there is no impression that applicatives are infrequent (see Okuda 1991–1995).
- 30 In fact, antipassivization is just the opposite of applicativization: the former deletes an object and the latter adds it.
- 31 A kind of analogy to this is the derivation of sociatives in Kabardian (Kazenin 2007).
- 32 The suffix *-ke* is employed for both decreasing and increasing valence (cf. 15.4.4.1.2); a similar phenomenon has been attested in other languages, e.g. Manchu-Tungus languages and Korean (Nedjalkov, I.V. 1991: 29).
- 33 *kampi cipetpap ekamuyynomi* ‘honoring gods with torn paper’ implies *J gohei* ‘wand with hemp and paper streamers’ used in Shinto ceremonies. The Japanese ritual is opposed to the Ainu way of honoring gods with *inaw* prayer sticks and *tónoto*-wine, regarded by Ainu as the only true way of honoring *kamuy*, including all animals. Ainu believed that only when treated well, i.e. killed and worshipped properly, *kamuy* may return to life and come back to the human, bringing their meat and skin as gifts.
- 34 Evidence for the auxiliary reanalysis may be found in other Hokkaido Ainu dialects, in which the respective “nouns” show a much more advanced stage of grammaticalization into auxiliaries as they have undergone loss in phonetic substance (erosion) and the copula has, presumably, lost its accent, e.g. an inferential evidential *ru-ne* (presumably, no accent) in Ishikari (SK: 29; Central, HNE), cf. *ruwe né* (Okuda, Osami, pers. comm.) in Chitose. The same phenomenon is observed in Southern Hokkaido Ainu in the case of some former nouns/nominalizers for which no free nouns any longer exist and we can only speculate on their original meanings, e.g. *kus-ne* ‘going to/intend to’ < *kusu* ‘intention (?)’ + *ne* <COP>, classified as an auxiliary in Satō (2008: 86).
- 35 Question-words are subject to a great dialectal variation; they are even partly different in Saru and Chitose, cf. Bugaeva (2004: 87).

## SOURCES

- AB Bugaeva, Anna (2004).  
 B Batchelor, John (1889 [1938, 1995]).  
 H Hattori Shirō (ed.) (1964).

- K Nakagawa Hiroshi and Bugaeva, Anna. (forthcoming). *A web-accessible corpus of folktales of the Saru dialect of Ainu by Mrs. Kimi Kimura (1900–1988)*. ELDP, SOAS, University of London. (<http://lah.soas.ac.uk/projects/ainu/>). The numbers code the recording details, e.g. for K7803232UP.035, year=(19)78, month=03, day=23 and the number of the tape on that day=2; UP means *uwepeker* ‘prosaic folktale’ genre of Ainu oral literature; .035 means the line number.
- KI Kubodera Itsuhiko (1977).
- KII Kubodera Itsuhiko (ed.) (1992) *Ainugo nihongo jitenkō: Kubodera Itsuhiko ainu shūroku nōto chōsa hōkokusho* [A manuscript of Ainu–Japanese dictionary: Reporting on Itsuhiko Kubodera’s Ainu records]. Sapporo: Hokkaidō Kyōiki Iinkai.
- KS Bugaeva, Anna and Endō Shiho (eds). Kurokawa Setsu (speaker) and Nathan, David (multimedia developer) (2010).
- N Nakagawa Hiroshi (1995).
- NN Nakagawa Hiroshi and Nakamoto Mutsuko (1997a).
- NV Nevskij, Nikolaj (1972).
- O4 Okuda Osami (1994).
- OI Oda Ito (1908–2000), informant of the Chitose dialect of Ainu; the author’s fieldnotes.
- SK Sunazawa Kura (1983) *Watashi no ichidai no omoide Ku sukup opuspe* [My lifestory]. Sapporo: Miyama Shobō.
- ST Satō Tomomi (1992) ‘“Hōgō”-kara Mita Hoppō-no Shogengo’ [The languages of the Northern peoples from the viewpoint of ‘incorporation’], in Miyaoka Osahito (ed.) *Kita-no gengo: ruikei to rekishi*. Tokyo: Sanseido. 193–201.
- ST1 Satō Tomomi (2008).
- ST2 Satō Tomomi (1997) ‘Hebinomi onna’ [A snake-swallowing woman], in Watanabe, Hitoshi, Ōshima, Minoru; Kirikae, Hideo and Satō, Tomomi (eds), *Ainu minzoku bunkazai chōsa hōkokusho* 17. Sapporo: Hokkaido kyōiku iinkai. 44–66.
- T Tamura Suzuko (1996).
- T1 Tamura Suzuko (1984).
- T2 Tamura Suzuko (2000 [1988]).
- T3 Tamura Suzuko (1984) *Ainugo shiryō* (1) [Ainu audio materials (1)]. Tokyo: Waseda daigaku gogakukyōiku kenkyūjo.
- T4 Tamura Suzuko (2001 [1972a]).

## REFERENCES

- Ainugo-kō, Bunpō* I–II [Ainu studies, Grammar I–II]. Tokyo: Yumani Shobō.
- Ainu Minzoku Hakubutsukan (ed.) (1993/2009) *Ainu bunka no kiso chishiki* [Ainu culture basic knowledge]. Tokyo: Sōfūkan.
- Alpatov, Vladimir M., Bugaeva, Anna and Nedjalkov, Vladimir P. (2007) ‘Reciprocals and sociatives in Ainu,’ in Nedjalkov, Vladimir P. (ed.) *Reciprocal constructions*. (Typological Studies in Language 71), Amsterdam, Philadelphia: John Benjamins, 1,751–1,822.
- Asai Tōru (1969) ‘Ainugo no bunpō. Ainugo Ishikari hōgen no gairyaku’ [An Ainu grammar. An outline of the grammar of the Ishikari dialect of Ainu], in *Ainu Minzokushi* 1, Tokyo: Daiichi Shuppan, 768–99.
- (1974) ‘Classification of dialects: cluster analysis of Ainu dialects,’ in *Bulletin of the Institute for the Study of North Eurasian Cultures* 8, Sapporo: Hokkaido University, 45–136.

- Austerlitz, Robert (1976) 'L'appellation du renne en japonais, ainou et surtout en ghiliak,' in Heissig, Walter (ed.) *Tractata Altaica: Denis Sinor; sexagenario optime de rebus Altaicis merito dedicata*, Wiesbaden: Harrassowitz, 45–9.
- Baker, Mark (1988) *Incorporation: A theory of grammatical function changing*, Chicago: University of Chicago Press.
- Batchelor, John (1889/1938/1995) *An Ainu-English-Japanese dictionary*, Tokyo: Iwanami Shoten.
- Bugaeva, Anna (2004) *Grammar and folklore texts of the Chitose dialect of Ainu (Idiolect of Ito Oda)*, ELPR (Endangered Languages of the Pacific Rim) Publications Series A-045, Suita: Osaka Gakuin University.
- (2007) 'Ainugo Mukawa hōgen tekisuto: Araida Seino san no shinyō [Texts of the Mukawa dialect of Ainu: devine epics by Mrs. Seino Araida]', in Tsumagari Toshiro (ed.) *Languages of the North Pacific Rim*, vol. 14, Sapporo: Hokkaido University, 65–85.
- (2008a) 'Reported discourse and logophoricity in Southern Hokkaido dialects of Ainu', *Gengo Kenkyū*, 133: 31–75.
- (2008b) 'Speech report constructions in Ainu', in Vajda, Edward J. (ed.) *Subordination and coordination strategies in North Asian languages*, (Current Issues in Linguistic Theory 300), Amsterdam and Philadelphia: John Benjamins, 17–29.
- (2008c) 'Ainu causatives,' *Nihon Gengogakkai Yokōshū* [Proceedings of Linguistic Society of Japan], 136: 420–6.
- (2010a) 'Ainu applicatives in typological perspective,' *Studies in Language*, 34(4): 749–801.
- (2010b) 'Anticausative/causative verb alternation in Ainu,' paper presented at *Societas Linguistica Europaea (SLE), 43rd Annual Meeting*, Vilnius University, Lithuania, September 2.
- (2011a) 'A diachronic study of the impersonal passive in Ainu,' in Malchukov, Andrej and Siewierska, Anna (eds) *Impersonal constructions cross-linguistically* (Studies in Language Companion Series 124), Amsterdam and Philadelphia: John Benjamins, 517–46.
- (2011b) 'Ditransitive constructions in Ainu,' *Sprachtypologie und Universalienforschung (STUF)*, 64(3): 237–55.
- (forthcoming a) 'Causative constructions in Ainu,' in Kulikov, Leonid and Serzants, Ilja (eds) *Transitivity and voice in Indo-European and beyond: A diachronic typological perspective*, Amsterdam, Philadelphia: John Benjamins.
- (forthcoming b) 'Valency classes in Ainu,' in Comrie, Bernard and Malchukov, Andrej (eds) *Valency classes cross-linguistically*, Berlin, New York: Mouton de Gruyter.
- (forthcoming c) 'Ainu valency classes database,' in Hartmann, Iren, Haspelmath, Martin and Taylor, Bradley (eds) *Valency classes across languages*, Munich: Max Planck Digital Library. (Available online at <http://www.valency-classes.info/Ainu>)
- Bugaeva, Anna and Endō Shiho (eds) (2010) Kurokawa, Setsu (speaker) and Nathan, David (multimedia developer). *A Talking dictionary of Ainu: A new version of Kanazawa's Ainu conversational dictionary*, ELDP, SOAS, University of London. (<http://lah.soas.ac.uk/projects/ainu/>)
- Bugaeva, Anna and Matsumoto Yoshiko (2009) 'Noun-modifying constructions in Ainu,' paper presented at the *1st meeting of the PFIH – NMC project* headed by Matsumoto Yoshiko, Stanford University, July 28.
- Chamberlain, Basil Hall (1887) *Aino fairy tales* (n. 1, 2). Tokyo: Kobunsha.
- Chiri Mashihō (1936/1974) Ainu gohō gaisetsu [An outline of Ainu grammar], in *Chiri Mashihō chosakushū*, vol. 4, Tokyo: Heibonsha, 3–197.
- (1937/1981) *Ainu Mintanshū* [Ainu folktales]. Tokyo: Iwanami Bunko.

- (1942/1973) *Ainu gohō kenkyū – karafuto hōgen o chūshin to shite* [Studies in Ainu grammar – with an emphasis on the Sakhalin dialect], in *Chiri Mashiho chosakushū*, vol. 3, Tokyo: Heibonsha, 455–586.
- (1953, 1954, and 1962/1975, 1976) *Bunrui ainugo jiten* [A classificational dictionary of the Ainu language], vols 1–3, reprinted by Tokyo: Heibonsha.
- Chiri Yukie (1923/1978) *Ainu shin'yōshū* [Ainu songs of gods], Tokyo: Iwanami Bunko.
- Dixon, Robert M. W. (2000) 'A typology of causatives: form, syntax and meaning', in Dixon, Robert M. W. and Aikhenvald, Alexandra Y. (eds) *Changing valency. Case studies in transitivity*. Cambridge: Cambridge University Press, 30–89.
- (2010) *Basic linguistic theory*, vol. 2, Oxford: Oxford University Press.
- Dixon, Robert M. W. and Aikhenvald, Alexandra Y. (1997) 'A typology of argument-determined constructions,' in Bybee, Joan, Haiman, John, and Thompson, Sandra A. (eds) *Essays on Language Function and Language Type*, Amsterdam, Philadelphia: John Benjamins, 1–113.
- (2000) 'Introduction' in Dixon, Robert M. W. and Aikhenvald, Alexandra Y. (eds) *Changing valency. Case studies in transitivity*. Cambridge: Cambridge University Press, 1–29.
- Dobrotvorskiĭ, Mikhail (1875) *Ainsko-russkij slovar'* [An Ainu-Russian dictionary]. Kazan': Kazan' University Press.
- Fitzhugh, William W. and Dubreuil, Chisato O. (eds) (1999) *Ainu. Spirit of a Northern People*, Washington: Arctic Studies Center. National Museum of Natural History Smithsonian Institution in Association with University of Washington Press.
- Frajzyngier, Zygmunt (1995) 'A functional theory of complementizers', in Bybee, Joan and Fleischman, Susanne (eds) *Modality in grammar and discourse*, Amsterdam: John Benjamins, 473–502.
- Fukuda (Tamura) Suzuko (1955) 'Ainugo no dōshi no kōzō' [The structure of verbs in Ainu], *Gengo Kenkyū*, 30: 46–64.
- (1960/2001). 'Ainugo Saru hōgen no jodōshi' [The auxiliary verbs of the Saru dialect of Ainu], *Minzokugaku Kenkyū*, 24(4): 67–78. Reprinted in: *Ainugo-kō, Bunpō I*, vol. 4, Tokyo: Yumani Shobō, 370–92.
- Geniušienė, Emma (1987) *The typology of reflexives*, Berlin, New York, Amsterdam: Mouton de Gruyter.
- Givón, Talmy (1983) *Topic continuity in discourse: A quantitative cross-language study* (Typological Studies in Language 3), Amsterdam, Philadelphia: John Benjamins.
- Gjerdman, Olof (1926) 'Word-parallels between Ainu and other languages', in *Le Monde Orientale*, 20, facs. 1–3, 29–84.
- Hattori Shirō (1959) *Nihongo no keitō* [Origins of the Japanese language], Tokyo: Iwanami Shoten.
- (ed.) (1964) *An Ainu dialect dictionary*, Tokyo: Iwanami Shoten.
- Heine, Bernd and Kuteva, Tania (2007) *The genesis of grammar. A reconstruction*, Oxford: Oxford University Press.
- Hokkaidō Utari Kyōkai (ed.) (1994) *Akor itak* [Our language], Sapporo: Kurūzu.
- Hopper, Paul and Thompson, Sandra A. (1980) Transitivity in grammar and discourse, *Language*, 56: 251–99.
- ICHL20 International Symposium (2011) 'Historical linguistics in the Asia-Pacific region and the position of Japanese,' National Museum of Ethnology, Osaka, July 30.
- Kannari Matsu and Kindaichi Kyōsuke (1959–1975/1993) *Ainu jojishi yūkarashū* [A collection of Ainu heroic epics], vols 1–9, Tokyo: Sanseido.
- Katayama Tatsumine (2003) '*Aynu shin'yōshū*' o yomitoku [Commentary on 'Ainu songs of gods'], with English translation by Julie Kaizawa, Tokyo: Sōfukan.
- Kayano Shigeru (1974/2005) *Uwepeker shūtaisei* [A compilation of folktales] +1CD, Tokyo: Nihon dentō bunka shinkōzaidan.

- (1996) *Kayano Shigeru no ainugo jiten* [An Ainu dictionary by Kayano Shigeru], Tokyo: Sanseido.
- (1998) *Kayano no ainu shinwa shūsei* [A collection of Ainu myths by Kayano Shigeru], vols 1–10, Tokyo: Heibonsha.
- Kazenin, Konstantin (2007) ‘Reciprocals, comitatives, sociatives, and reflexives in Kabardian,’ in Nedjalkov, Vladimir P. (ed.) *Reciprocal constructions* (Typological Studies in Language 71), Amsterdam, Philadelphia: John Benjamins, 739–72.
- Keenan, Edward L. and Comrie, Bernard (1977) ‘Noun phrase accessibility and universal grammar,’ *Linguistic Inquiry*, 8: 63–99.
- Kindaichi Kyōsuke (1931/1993) *Ainu yūkara gohō tekiyō* [An outline grammar of the Ainu epic poetry], *Ainu jojishi yūkara no kenkyū* 2, Tokyo: Tōyō Bunko, 3–233. Reprinted in: *Ainugogaku kōgi* 2 (Lectures on Ainu studies 2), *Kindaichi Kyōsuke zenshū. Ainugo I*, vol. 5, Tokyo: Sanseido, 145–366.
- (ed.) (1972) *Moshiogusa* [Seaweeds] by Uehara Kumazirō, in *Ainugo shiryō sōsho*, Tokyo: Kokusho kankōkai.
- Kirikae, Hideo (1994) ‘Some aspects of reflexivity in Ainu’, in Shibatani, Masayoshi (ed.) *A comprehensive study of the function and typology of language*, pp. 310–17. Kobe: Kobe University.
- (1996) ‘Ainugo Tokachi hōgen ni yoru mukashibanashi “Shima o hiite oyogu Otasut no shōnen no monogatari” no jiten to bunpō’ [Lexicon and grammatical sketch of the Ainu folktales ‘An Otasut boy who pulled an island swimming’], in *Gakuen Ronshū*, Hokkai gakuen daigaku gakujutsu kenkyūkai, 123–286.
- (2003) *Ainu shin’yōshū jiten. Tekisuto, bunpō kaisetsu-tsuki* [Lexicon to Yukie Chiri’s Ainu *shin’yōshū* ‘songs of gods’ with text and grammatical notes], Tokyo: Daigaku shorin.
- Kubodera, Itsuhiko (1977) *Ainu jojishi shin’yō seiden no kenkyū* [The study of Ainu heroic epics and songs of gods], Tokyo: Iwanami Shoten.
- Lindquist, Ivar (1960) ‘Indo-European features in the Ainu language. With reference to the thesis of Pierre Naert’, in *Lunds Universitets årsskrift*, 54(1): 3–67.
- Malchukov, Andrej (2005) ‘Case pattern splits, verb types and construction competition’, in Amberber, Mengistu and de Hoop, Helen (eds) *Competition and variation in Natural languages: The case for case*, Amsterdam and Tokyo: Elsevier, 73–117.
- Malchukov, Andrej, Haspelmath, Martin, and Comrie, Bernard (2010) ‘Ditransitive constructions: A typological overview’, in Malchukov, Andrej, Haspelmath, Martin and Comrie, Bernard (eds) *Studies in ditransitive constructions: A comparative handbook*, Berlin, New York: Mouton de Gruyter, 1–67.
- Matsumoto Yoshiko (1997) *Noun-modifying constructions in Japanese: A frame-semantic approach* (Studies in language companion series 35), Amsterdam, Philadelphia: John Benjamins.
- Meiji-mae nihonkagakushi kankōkai (ed.) (1971) *Meiji-mae nihonjinruigaku, senshi-gakushi* [Japanese anthropology before the Meiji era, prehistory], Tokyo: Nihon gaku-jutsu shinkōkai.
- Mithun, Marianne (1984) ‘The evolution of noun incorporation,’ *Language*, 60: 847–94.
- Munro, Neil G. (1962) *Ainu creed and cult*, London: Routledge and Kegan Paul.
- Murasaki, Kyōko (1976) *Karafuto ainugo* [Sakhalin Ainu], Tokyo: Kokushokan-kōkai.
- (1979) *Karafuto ainugo. Bunpō-hen* [Sakhalin Ainu. Grammar volume], Tokyo: Kokushokan-kōkai.
- (2001) *Asai Take kōjutsu Sakhalin ainu no mukashibanashi* [Sakhalin Ainu folktales narrated by Take Asai], Tokyo: Sōfūkan.
- (2009) *Karafuto Ainugo nyūmon kaiwa* [First step for the Sakhalin Ainu language], Kushiro: Ryokugeisha.
- Murayama Shichirō (1992) *Ainugo no kigen* [Origins of the Ainu language], Tokyo: San’ichi Shobō.
- (1993). *Ainugo no kenkyū* [Studies in Ainu], Tokyo: San’ichi Shobō.

- Naert, Pierre (1958) 'La situation linguistique de l'Ainou et Indoeuropeen', in: *Lunds Universitets årsskrift*, 53(4): 1–234.
- Nakagawa Hiroshi (1981) 'Asupekuto-teki kanten kara mita ainugo dōshi' [Ainu verbs from a perspective of aspect], *Gengogaku ensū '81*, Tokyo: Tokyo University, 141–155.
- (1984/2001) 'Ainugo no meishi to basho hyōgen' [Ainu nouns and expression of place], *Gengogaku ronshū '84*, Tokyo: Tokyo University. Reprinted in: *Ainugokō, Bunpō II*, vol. 5, Tokyo: Yumani Shobō, 211–29.
- (1988) 'Ainugo chitose hōgen no ninshō daimeishi to sono rekishi-teki ichi' [Personal pronouns in the Chitose dialect of Ainu and their historical position], in *Tokyo daigaku gengogaku ronshū '88*, Tokyo: Tokyo daigaku bungakubu gengogaku kenkyūshitsu, 239–53.
- (1995) *Ainugo chitose hōgen jiten* [A dictionary of the Chitose dialect of Ainu], Tokyo: Sōfukan.
- (1997) *Ainu mukashibanashi no sekai* [The world of Ainu fairytales], Tokyo: Heibonsha.
- (2000–2011) 'Ainugo kōshōbungei tekisutoshū 1–11' [Ainu folklore text 1–11], in *Yūrashia gengo bunka ronshū 3–13*, Chiba: Chiba University.
- (2007) *Kamuy yūkar de ainugo o manabu* [Learning Ainu by yūkar], Tokyo: Hakuuisha.
- (2009) 'Ainugo no settōjidō' [Degree of prefixation in Ainu], in Tsumagari Toshiro (ed.) *Sakhalin no gengo no sekai*, Sapporo: Hokkaido daigaku daigakuin bungaku kenkyūka, 63–70.
- Nakagawa Hiroshi and Bugaeva, Anna (forthcoming) *A web-accessible corpus of folktales of the Saru dialect of Ainu by Mrs. Kimi Kimura (1900–1988)*, ELDP, SOAS, University of London. (<http://lah.soas.ac.uk/projects/ainu/>)
- Nakagawa Hiroshi and Nakamoto Mutsuko (1997) *Ekusupusso ainugo*, Tokyo: Hakuuisha.
- Nedjalkov, Igor V. (1991) 'Recessive-accessive polysemy of verbal suffixes,' *Languages of the World*, 1: 4–32.
- Nedjalkov, Vladimir P. and Silnitskij, Georgij G. (1969) 'Tipologija morfoložičeskogo i leksičeskogo kauzativov' [Typology of morphological and lexical causatives], in Xolodovich, Alexander A. (ed.) *Tipologija kauzativnykh konstruksij. Morfoložičeskij kauzativ*, Leningrad: Nauka.
- Nevskij, Nikolaj (1972) *Ajniskij folklor* [Ainu folklore], Moscow: Nauka. Translated in Japanese by Uoi Kazuyoshi (1991) *Ainu fōkuroa*, Sapporo: Hokkaidō shuppan kikaku sentā.
- Okuda Osami (1989/2001) 'Nihongo to ainugo ni okeru rentai shūshoku to bun no meishika' [Noun-modifying and clausal nominalization in Japanese and Ainu], in *Waseda daigaku gogaku kyōiku kenkyūjo kiyō 39*. Reprinted in: *Ainugo-kō, Bunpō II*, vol. 5, Tokyo: Yumani Shobō, 243–56.
- (1991–1995) *Shizunai hōgen no denshō. Orita Suteno no kōshōbungei 1–5* [Folktales in the Shizunai dialect. Oral tradition by Suteno Orita 1–5], Shizunai: Shizunai Board of Education.
- Patrie, James (1982) *The genetic relationship of the Ainu language* (Oceanic Linguistics. Special Publications 17), Honolulu: University of Hawai'i.
- Payne, Thomas (2006) *Exploring language structure*, Cambridge: Cambridge University Press.
- Perlmutter, David M. (1978) 'Impersonal passives and the unaccusative hypothesis,' in *Proceedings of the fourth annual meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, University of California, 157–89.
- Peterson, David A. (1999) 'Discourse-functional, historical, and typological aspects of applicative constructions,' PhD diss, Berkeley University.
- (2007) *Applicative constructions*, Oxford: Oxford University Press.

- Pilsudski, B. (1912) *Materials for the study of Ainu language and folklore*, Cracow: Imperial Academy of Sciences. Reprinted in: Majewicz, Alfred F. (ed.) (1998) *The collected works of Bronislaw Pilsudski*, vol. 1, Mouton de Gruyter.
- Refsing, Kirsten (ed.) (1986) *The Ainu language. The morphology and syntax of the Shizunai dialect*, Aarhus: Aarhus University Press.
- (ed.) (1996) *Early European Writings on the Ainu Language*, 10 vol, Richmond, Surrey: Curzon.
- (ed.) (1998). *Origins of the Ainu language: the Ainu Indo-European controversy*, 5 vol, Richmond, UK: Curzon.
- Satō Tomomi (1995) ‘Ainugo no judōbun ni kansuru hitokōsatsu’ [A consideration of the Ainu passive], *Hokudai bungakubu kiyō* 44(1): 1–18.
- (1995–1998) ‘Tori ni ongaeshi o sareta yūbetsu no ikuresuye no hanashi’ [A story of Ikuresuye from Yūbetsu whose favors were returned by a bird] and other folktales by Ueda Toshi, in Watanabe Hitoshi, Ōshima Minoru, Kirikae Hideo, and Satō Tomomi (eds) *Ainu minzoku bunkazai chōsa hōkokusho* 15–18, Sapporo: Hokkaido kyōiku iinkai.
- (1997) ‘Possessive expressions in Ainu,’ in Hayashi Tōru and Bhaskararao, Peri (eds) *Studies in Possessive Expressions*, Tokyo: Tokyo University of Foreign Studies, 143–60.
- (2003) ‘Phonological status of the epenthetic glides in the Chitose dialect of Ainu,’ in *Hokkaidō-ritsu ainu minzoku bunka kenkyū sentā kenkyū kiyō* 9, Sapporo: Hokkaidō-ritsu ainu minzoku bunka kenkyū sentā, 11–34.
- (2004) ‘Ainu bungaku ni okeru ichininshōtai no mondai’ [Some problems with the first person narration in Ainu], *Hokudai bungaku kenkyūka kiyō*, 112: 171–85.
- (2006) ‘Ainugo chitose hōgen no asupekuto -kor an, wa an o chūshin to shite’ [Aspect in Ainu: Note on *kor an* and *wa an* in the Chitose dialect of Ainu], *Hokkaidō-ritsu ainu minzoku bunka kenkyū sentā kenkyū kiyō* 12, Sapporo: Hokkaidō-ritsu ainu minzoku kenkyū sentā, 43–67.
- (2007a) ‘Ainugo chitose hōgen no saiki settōji’ [Reflexive prefixes in Chitose Ainu dialect], *Ninchi kagaku kenkyū* 5, Muroran: Muroran ninchi kagaku kenkyūkai, 31–9.
- (2007b) ‘Futatabi ainugo chitose hōgen no asupekuto ni-tsuite: toku ni kanryō o arawasu keishiki o megutte’ [Aspect in the Chitose Ainu revisited: With a focus on perfect forms], *Hokkaidō-ritsu ainu minzoku bunka kenkyū sentā kenkyū kiyō* 13, Sapporo: Hokkaidō-ritsu ainu minzoku bunka kenkyū sentā, 1–14.
- (2008) *Ainugo bunpō no kiso* [The basics of Ainu grammar], Tokyo: Daigakushorin.
- Shibatani Masayoshi (1990) *The languages of Japan*, Cambridge: Cambridge University Press.
- Shitaku Yae no Denshō Kankōkai (ed.) (2007) *Shitaku yae no denshō: Tomizu Keiichi sairoku. Kayō sanbun* [Shitaku Yae’s folklore, recorded by Tomizu Keiichi. Songs and prose], Kushiro: Shitaku Yae no Denshō Kankōkai.
- (ed.) (2011) *Shitaku yae no denshō: Tomizu Keiichi sairoku. Sanbun 1* [Shitaku Yae’s folklore, recorded by Tomizu Keiichi. Prose 1], Kushiro: Shitaku Yae no Denshō Kankōkai.
- Song, Jae Jung (2001) *Causatives and causation: A universal-typological perspective*, London: Longman.
- Spencer, Andrew (1991) *Morphological theory*, Oxford, UK, Cambridge, MA: Blackwell.
- Steever, Sanford B. (1988) *The serial verb formation in the Dravidian languages*, Delhi: Motilal Banarsidass.
- Strong, Sarah (2011) *Ainu spirits singing: The living world of Chiri Yukie’s Ainu Shin’yōshū*, Honolulu: University of Hawai’i Press.
- Sunazawa Kura (1983) *Watashi no ichidai no omoide. Ku sukup opuspe* [My life story], Sapporo: Miyama Shobō.
- Tamura Suzuko (1964/2001) ‘Ainugo saru hōgen no meishi (sono 1)’ [Nouns in the Saru dialect of Ainu (part 1)], in *Waseda daigaku gogaku kyōiku kenkyūjo kiyō* 3. Reprinted in: *Ainugo-kō, Bunpō* 1, vol. 4, Tokyo: Yumani Shobō, 107–33.

- (1966/2001) ‘Ainugo saru hōgen no meishi (sono 2): Shozokukei no keisei’ [Nouns in the Saru dialect of Ainu (part 1). Formation of possessive forms], in *Waseda daigaku gogaku kyōiku kenkyū kiyō* 3. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 134–54.
- (1970) ‘Personal affixes in the Saru dialect of Ainu,’ in Jakobson, Roman and Kawamoto, S. (eds) *Studies in general and oriental linguistics*, presented to Hattori Shirō, Tokyo: TEC Company Ltd, 577–611.
- (1972a/2001) Ainugo Saru hōgen no ninshō no shurui [Types of grammatical person in the Saru dialect of Ainu], *Gengo Kenkyū*, 61: 17–39. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 370–92.
- (1972b/2001) ‘Ainugo saru hōgen ni mirareru chōfukuhō’ [Reduplication in the Saru dialect of Ainu], in *Ajia afurika bunpō kenkyū* 1. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 357–69.
- (1973/2001) ‘Ainugo saru hōgen no goseidōshi no kōzō’ [The structure of compound verbs in the Saru dialect of Ainu], in *Ajia afurika bunpō kenkyū* 2. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 407–28.
- (1977/2001) ‘Ainugo saru hōgen no meirei hyōgen’ [Commands in the Saru dialect of Ainu], in *Ajia afurika bunpō kenkyū* 5. Reprinted in: *Ainugo-kō, Bunpō* II, vol. 5, Tokyo: Yumani Shobō, 22–141.
- (1978/2001) ‘Ainugo saru hōgen no gimon hyōgen’ [Questions in the Saru dialect of Ainu], in *Ajia afurika bunpō kenkyū* 6. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 142–54.
- (1979) *Aynu Itak. Aynugo nyūmon* [The Ainu language. An introduction into Ainu], *Goken kyōzai sensho* 21, Tokyo: Waseda daigaku gogaku kyōiku kenkyūjo.
- (1982/2001) ‘Ainugo saru hōgen ni okeru jyōge o arawasu ichimeishi’ [Relational nouns indicating upper and lower positions in the Saru dialect of Ainu], *Gengo kenkyū* 82. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 155–82.
- (1983a) *Aynugo kiso goi* [Basic lexicon of Ainu], *Goken kyōzai sensho* 31, Tokyo: Waseda daigaku gogaku kyōiku kenkyūjo.
- (1983b) ‘Syntactic patterns of position nouns in Japanese, Ainu, Canadian Eskimo and Basque,’ in *IKER-2: Piarres Lafitteri Omenaldia*, Bilbao: Euskalzaindia.
- (1984) *Aynu Itak. Aynugo nyūmon. Kaisetsu* [The Ainu language. An introduction into Ainu. Commentary], *Goken kyōzai sensho* 32, Tokyo: Waseda daigaku gogaku kyōiku kenkyūjo.
- (1984–2000) *Ainugo shiryō* 1–12 [Ainu audio materials 1–12], Tokyo: Waseda daigaku gogakukyōiku kenkyūjo.
- (1988/2000) ‘Ainugo’ [The Ainu language], in Kōji Takashi, Kōno Rokurō and Chino Eiichi (eds) *Gengogaku daijiten* [Encyclopedia of linguistics], Tokyo: Sanseido. English translation: *The Ainu language*, (ICHEL Linguistic Studies 29), Tokyo: Sanseido.
- (1993/2001) ‘Ainugo saru hōgen ni okeru “naka” o arawasu mitsu no meishi: *or, onnay, tum*’ [Three nouns with the meaning ‘inside’ in the Saru dialect of Ainu: *or, onnay, tum*], in *Waseda daigaku gogaku kyōiku kenkyūjo* 30 *shūnen ronbunshū*. Reprinted in: *Ainugo-kō, Bunpō* I, vol. 4, Tokyo: Yumani Shobō, 269–92.
- (1996) *Ainugo Saru hōgen jiten* [A dictionary of the Saru dialect of Ainu], Tokyo: Sōfūkan.
- Teramura Hideo (1992) ‘Rentai shūshoku noshintakusu to imi (sono1)’ [The syntax of noun-modifying constructions (part 1)], in *Teramura Hideo ronbunshū I Nihongo bunpō-hen*, Tokyo: Kuroshio, 157–207.
- Torii Ryūzō (1903) *Les Ainou des Iles Kouriles*. In the series: Etudes Archaeologiques Ethnologiques, Tokyo University.
- Tsunoda Tasaku (1985) ‘Remarks on transitivity,’ *Journal of Linguistics* 21: 385–96.

- 
- (1996) Taigenjumbun [Noun-concluding constructions], in Suzuki Tai and Tsunoda Tasaku (eds) *Nihongo bunpō no shomondai: Takahashi Tarō sensei koki kinen ronbunshū*, Tokyo: Hitsuji Shobō, 139–61.
- (forthcoming) ‘Mermaid construction in Japanese,’ in Tsunoda Tasaku (ed.) *Mermaid constructions cross-linguistically* (Typological Studies in Language), Amsterdam, Philadelphia: John Benjamins.
- Vovin, Alexander (1993) *A reconstruction of proto-Ainu*, Leiden, New York, Köln: E. J. Brill.

# INDEX

Notes are referenced in the form “346n17” (page 346, note 17). Page numbers for tables are in **bold**.

- accent (= tone) 7, 26–7, 127–8, 174–5, 177, 180, 191, 215–16, 252–3, 271–2, 291–2, 316–18, 331–4, 351, 353–4, 384, 413, 415, 466–7
- addressee honorifics *see* speech levels
- adjectives of emotion, Japonic 234, 289, 426–8, 440–1
- adversative passive 9, 262, 298, 326, 400
- afterthoughts 213, 239–40, 286, 288
- Agarisuji = Kuroshima island/variety of Ryukyuan 382
- Ainu 1, 5, 6, 7, 10, 12, 461–509
- alienable vs. inalienable possession 478, 480
- Altaic 6, 28, 461
- Amami-Okinawan *see* Northern Ryukyuan
- animacy 290, 318, 323, 357, 364, 370, 388, 490
- annexation of Korea by Japan (1910) 19
- answers 150, 261, 295, 397
- anticausative 481, 483, 488–9
- antipassive 481, 483, 486–7
- apophony *see* noun stem variation
- applicatives 7, 471, 479, 481–5, 488, 490, 491–2, 498
- Aragasuku island/variety of Ryukyuan 382, 383, 384
- areal relationship 6–19, 28
- Arte da Lingoa de Iapam* 246, 248, 249–53, 256
- aspect *see* tense and aspect
- assimilation, in compounds 130–2, 214, 250, 257, 272–3; *see also* Law of Initials; *rendaku*; *renjō*; *sandhi*
- Asuka period (592–710) 189, 198, 242
- ateji* 253
- Austronesian 461
- auxiliary verbs, quasi-aspectual 10, 92, 148, 233, 291–2, 393–4, 436, 494–5, 499
- Azuma 4; *see also* Eastern Old Japanese
- Basque 461
- benefactive 10, 261, 298–9, 372, 377, 495, 499
- bikago* (= euphemism) 302
- bound nouns 34, 51, 57, 62, 64, 65, 67, 69, 85–6, 103, 134, 157, 287, 469, 470, 480, 486
- bungo* 212, 268, 310
- Buton island, Indonesia 16
- case *see* particles, noun-following; zero-marking of nouns
- causative 61, 68, 91, 94, 107–8, 138, 153, 206–7, 219, 221, 236–7, 255, 262, 282, 283, 299–300, 326, 343, 358, 364–5, 376–7, 386, 401–2, 403, 419, 422, 445–6, 467, 481–2, 485–6, 489
- ch’aja [chaja] 41, 42, 49–51, 59, 62, 67
- Chinese character script 12–19, 41, 46–51, 63, 73–5, 82–3, 118, 123, 132, 161–2, 193–4, 218, 243, 253, 264, 273–5, 276–7, 292, 310, 455
- Chinhan [Cinhan] confederacy 55
- Ch’oe Malli [Choy Manli] 12, 83–4
- Chosŏn [Cosen] dynasty (1392–1910) 73, 83–4
- Christian materials, fifteenth/sixteenth-century Japan 246, 248–9
- chữ nôm 15–16
- Ciacia 16
- Classical Japanese 212–45, 268, 293, 443
- classifiers 6, 12, 99–100, 134, 141–2, 198–9, 226, 256–7, 258, 284–6, 287, 288, 292, 357, 387–8, 429–30, 471
- clefts and pseudo-clefts 152, 261, 297, 400
- coda consonants 249–50, 465, 467; *see also* moraic nasal; moraic /Q/; moraic /r/; moraic /v/
- commands (= imperative) 9, 59, 65–6, 104, 138, 150–1, 173–9, 184, 205, 235–6, 261, 296, 325, 336–7, 342, 359, 373, 397, 422, 443, 496–7
- comparison 88–9, 136, 221, 259, 278, 386, 399
- complementation *see* subordination
- contraction, vowel 57, 193, 314, 383
- converbs 10, 58–9, 66, 157; *see also* coordination; subordination
- coordination of sentences 10, 156, 209–10, 225, 229, 230, 238–40, 263, 302, 303–4, 362, 404, 407, 497–8; of nouns *see* *dvandva*; particles, noun-following
- copula 61, 67, 69, 84–5, 91, 94, 102, 138, 140, 145–6, 196, 221, 224, 225, 229–30, 233, 235, 237, 257, 258–9, 268, 279–81, 282, 286, 289–90, 291, 292, 300, 340, 341, 355, 364, 372, 374, 387, 392, 424–5, 469, 471, 473–4, 497

- dakuten *see* diacritics on kana  
 debitive *see* modality, deontic  
 deletion of word-final vowels 328–9, 332, 335, 338–9, 340  
 desiderative *see* volition  
 devoicing 130, 270, 272, 314, 315, 328, 330, 383, 385, 413  
 diacritics on kana 219, 253, 275  
 dialects, Japanese: Chūgoku 340, 346n17; Kyoto 215–16, 249, 251, 252, 334, 346n13; Kyūshū (Amakusa islands, Fukuoka, Kagoshima, Kumamoto, Nagasaki, Ōita) 313–14, 328–45; Nagano and Niigata 252; Osaka 271, 334, 346n13; Shikoku 330, 334, 340, 346n17; Tōkoku (Akita, Aomori, Iwate, Yamagata) 313–28, 345; Tokyo 215–16, 247, 252, 368, 370, 371  
 dialects, Korean: 123–4, 131, 161, 168–73; central (Ch'ungch'ōng [Chwungcheng], Hwanghae [Hwanghay], Kangwŏn [Kangwen], Kyōnggi [Kyengki]) 175–6; Cheju [Cecwu] island 5, 81, 82, 178–9; North vs South Korean 123, 124, 129, 163, 180–4; northeastern (Hamgyōng [Hamkyeng]) 82, 174–5; northwestern (P'yōngan [Phyengan]) 173–4; southeastern (Kyōngsang [Kyengsang]) 78, 79–80, 82, 93, 164, 177–8; southwestern (Chōlla [Cenla]) 78, 79–80, 127, 176–7  
 diminutive 356, 386, 417  
 diphthongs 78, 82, 126–7, 170, 173, 178, 190, 413, 466  
 discourse markers 294, 355; *see also* particles, sentence-final  
 distributive 58, 64–5, 90, 287, 390–1  
 double-subject construction 145–6, 288–9  
 dvandva 221, 273  
  
 Early Middle Chinese 194, 242, 273  
 Early Middle Japanese 26, 212–13; *see also* Classical Japanese  
 Early Old Korean 41–2, 48, 51–6  
 Eastern Old Japanese 4, 26, 189–90, 204, 213  
 Edo period (1603–1868) 11, 212, 268, 269, 308, 310  
 emphasis 58, 64–5, 90, 107, 127, 143, 152, 205–6, 214, 221, 223, 237, 261, 275, 288, 297–8, 326, 343, 374, 392, 399–400, 413, 435, 443–5, 493, 496  
 endangerment, language 5, 351, 409, 413  
 English: in Japan 20, 277, 308; in Korea 20, 124, 163–5  
 Eskimo 461  
 euphonic avoidance of consonants 216  
 evidentiality 234, 260, 293, 303, 324, 361, 376, 394, 425, 435–41, 447, 470, 492–4, 497, 499  
 exclamation 11, 56, 65–6, 88, 95, 101, 105, 109, 184, 205, 225, 236, 261, 294–6, 398, 400; *see also* particles, sentence-final  
 existential verb, verb of existence 34, 93, 138, 146, 147, 149, 153, 230, 258–9, 260, 290, 323, 340, 364, 373, 374, 392, 394, 424, 473, 484, 492, 497  
  
*Feiqe monogatari* 246  
 female vs male: language 11, 268, 278, 289, 294, 295, 301, 302; writers 212–13, 219  
 focus 91, 100–1, 107, 137, 143, 152, 205–6, 236, 261, 279, 297–8, 355, 368–9, 372, 373, 374, 398–400, 419, 428, 443–4, 496  
 folktales, Ainu 464, 475–6  
 Foundation for Research and Promotion of Ainu Culture 463  
 fractions 222, 256–7, 286, 389  
 furigana 275, 310  
  
*gairaigo* 17, 264, 307; *see also* loanwords  
*oeraeō* [oylaye] 17; *see also* loanwords  
*Gekanshū* 253  
 geminates, word-initial 351, 352–3, 381–2  
 gemination 214, 248, 250, 270, 285, 414, 453n1  
*genbun itchi* 268  
*Gendai kanazukai* 268  
*Genji monogatari* 212, 263–4  
 glide insertion 468–9  
 glottalization 351, 413–15  
 glottal stop 79, 130, 351–3, 382, 404, 465–6, 467, 468  
 go'on lexical stratum 242–3, 273, 285  
 grammaticalization 10, 32, 61, 62, 95, 146, 149, 158, 223, 240, 243, 260, 275, 291–2, 318, 319, 335, 340, 362, 378, 396, 397, 405, 470, 473, 476, 493  
*gyarumoji* 278  
  
 Hachijōjima dialect 4, 5  
 handakuten *see* diacritics on kana  
 han'gūl [hankul] xvi, 12, 14–15, 16, 26, 56, 73, 76, 77, 78, 79, 81, 82–4, 123–4, 125, 132–4, 161, 249, 251  
 Han'gūl [Hankul] Society 123  
 Hateruma island/variety of Ryukyuan 3, 383, 387, 388, 389, 395, 408  
 Hatoma island/variety of Ryukyuan 4, 381–411  
 hearsay (= reportative) 149, 158, 234, 260, 286, 289, 292–3, 324, 394, 493–4  
 Heian period (794–1185) 16, 192, 193, 212–15, 217, 218–19, 228, 230, 232, 237  
*Heike monogatari* 8, 16, 213, 219–20, 246, 253  
 hentaigana 214, 218–19, 220, 231, 253, 273, 276–7  
*hentai kanbun* *see* pseudo-Chinese  
 hiragana 13, 14, 218, 219, 253, 273, 274–6, 277, 278, 310  
 Hirara city/variety of Ryukyuan 382, 384, 390  
*hitotsugana* 314, 330; *see also* *yotsugana* issue  
*Hōgen bunpō zenkoku chizu* 313  
*Hōjōki* 212–13, 219–20, 246, 250

- Hokkaido 461ff  
*Hokkaido Former Aborigines Protection Act* (1899) 462–3
- honorific (= subject-honorific), honorification 7–9, 11, 34, 55, 60–1, 64, 66–9, 70, 85, 86, 88, 92, 94, 96, 97–8, 109–10, 111, 114, 135, 138, 146, 150, 153–4, 176, 184, 194, 207–8, 213, 217, 219, 221, 230, 232, 236, 237, 253, 255, 257, 261, 262, 264, 278, 282, 300–2, 356, 357, 361–2, 363–4, 367, 387, 402–3, 423, 431, 446, 474
- hortative *see* proposals
- humble (= object-honorific) 7–9, 11, 61, 68–9, 97, 110–11, 144, 153, 154, 207–8, 216, 219, 230, 237, 262, 278, 282, 300, 301–2, 363–4, 372, 402
- hundokcha [huntokca] 42, 47, 49–51, 56
- hun'gaja [hunkaca] 13, 49–50
- hunmin chǒng'ŭm [hwunmin cengum] 14, 73, 83
- Hunmin chǒng'ŭm [Hwunmin cengum]* 79, 83, 132
- Hunmin chǒng'ŭm haerye [Hwunmin cengum haylyey]* 77, 83
- hyangch'al [hyangchal]* 13, 41–2, 44–5, 47, 50, 57, 58, 60, 63, 65–8, 76, 82
- hyangga [hyangka]* 16, 42, 44–5, 48, 57–61, 73–4, 76
- Hyangyak kugŭppang [Hyangyak kwukuppang]* 42, 51, 63, 74
- ideophones, mimesis, phonaesthesia, onomatopoeia, sound symbolism 160–1, 124–5, 128, 139, 155, 214, 248, 262, 264, 274, 303, 308, 408
- idu [itwu] 8, 14, 15, 16, 33, 41–4, 48, 49–51, 55–6, 57–61, 76, 82–3, 84
- Ikema island/variety of Ryukyuan 381–2, 385, 408
- imperative *see* commands
- inchoative 152, 230, 299–300, 394, 428–9
- inclusive vs. exclusive pronouns 33, 391, 434, 471, 472
- incorporation 481, 489–92, 498
- indefinite person 471–2, 474, 475–6, 486
- Indo-European 461
- interrogative *see* questions
- intervocalic voicing 315–16, 322, 330–1, 340, 345, 465, 466
- intransitive:transitive verb pairs 237, 279, 299–300, 401–2, 445–6; *see also* causative; passive
- Irabu island/variety of Ryukyuan 3, 363, 384
- Iriomote island/variety of Ryukyuan 3, 381, 382, 402, 408, 409
- Irop'a [Ilopha]* 246, 249, 251
- Ishigaki island/variety of Ryukyuan 3, 381, 383, 385, 386, 387, 390, 395, 408, 412, 413
- Japanese, Modern 3–4, 6, 7, 8, 15, 214, 215, 217, 219, 236, 239, 242, 247, 248–52, 253, 268–312, 314, 315, 316, 317, 320–2, 323, 324, 325, 326, 328–32, 333–4, 336–7, 345, 376, 408, 409, 413
- Japanese–Korean language family 5, 7, 24–5, 30–4
- Japanese-Ryukyuan (= Japonic) 3–5, 7, 24–7, 30–4, 216, 351, 353, 357, 367, 378, 461
- Jilin lèishi see* *Kyeylim yusa*
- jindai moji* 12
- jinmeiyō kanji 273–4
- Jōmon culture 462
- jōyō kanji 273–4
- Kabira island/variety of Ryukyuan 381, 383
- Kagerō nikki* 212
- Kagoshima prefecture/dialect 313, 328–45
- kaida 14, 416, 430, 455, 455–7
- kakari-musubi* 190, 206, 225, 236, 241, 254, 258, 261, 369, 373–4, 443–4
- Kamakura period (1192–1333) 212–13, 219, 224, 229, 230, 246, 250, 254, 258, 259, 260
- Kamchatka 461
- Kamo no Chōmei 212
- kana 16, 123, 213, 218–19, 251, 253, 264, 268–9, 270, 273–8, 292
- kanamajiribun see* mixed script
- Kanamojizukai* 253
- kanbun* 219, 230, 253
- kanbun kundoku* 219, 230
- kanji *see* Chinese character script
- kan'on lexical stratum 242–3, 273, 285
- Karimata island/variety of Ryukyuan 382, 383, 384, 404
- katakana 14, 44, 218–19, 220, 250, 253, 273–7, 310, 330, 463, 464
- keigo *see* honorific; humble
- Ki no Tsurayuki 212
- kinterms 9, 33, 109, 144, 153–4, 162, 287, 290, 335, 356, 370, 385, 386, 431, 451–2, 478
- kogo* 212
- kōgo* 212, 268, 283, 310
- Koguryō [Kokwulye] 5, 41, 42, 49, 51, 53–4, 55–6
- Kohama island/variety of Ryukyuan 383
- Kojiki* 15, 32, 48, 118, 189, 192, 194, 218
- kokuji see* made-in-Japan/Korea characters
- Komi island/variety of Ryukyuan 383, 395
- Konjaku monogatari-shū* 213, 219–20, 230
- kō-otsu* vowel distinction, in Old Japanese 190–1, 192
- Korean, Contemporary/Modern 6, 7, 8, 15, 45–6, 51, 54, 55, 56–7, 60, 61, 62, 63, 64–5, 67, 70, 123–67
- Koryō [Kolye] period (10thC–1392) 41, 44, 48, 51, 73–4, 77, 118
- Koryōmar [Kolyemal] 179–80
- Koryōsa [Kolyesa]* 42

- kugyōl [kwukyel] 8, 14, 16, 41, 42, 44, 46–8, 50–1, 52, 57, 63–70, 74, 76, 82, 120n1; dot-type 13, 46, 47–8, 69
- Kume island/variety of Ryukyuan 431
- Kunigami variety of Ryukyuan 5
- Kurile islands 461, 462
- Kwanggaet'o [Kwangkaytho] stela 42, 49
- Kyerim yusa [kyeylim yusa] 29, 42, 48, 63, 64–5, 73–6, 78
- kyōgen 246
- kyōiku kanji 273
- labial lenition, Japanese 214, 248, 253
- language attrition 409, 413, 462–3
- language isolates 3, 5, 24, 461
- language revitalization 463
- Late Middle Chinese 191–2, 194, 242–3, 273
- Late Middle Japanese 11, 213, 246–65, 268, 310n4
- Late Old Korean 41–2, 44–5, 48, 57, 58, 63–70
- Law for the Promotion of the Ainu Culture and for the Dissemination and Advocacy of the Traditions of the Ainu and the Ainu Culture* (1997) 463
- Law of Initials, Korean 129, 132, 134, 164, 171, 173, 180, 182
- liaison: in Korean, 131–2, 133–4; in Japanese *see* renjō
- loanwords 126, 134, 139, 141, 161, 163–4, 269, 270, 271, 273, 274, 276, 277, 302, 307, 328, 383, 387, 403; from Ainu 7, 498; from Chinese 13, 17–19, 62, 63, 161–3, 210, 214, 215, 242–3, 378–9, 408; from Dutch 265, 274, 408; from English 20, 124, 164–5, 378–9, 408; from or via Japanese 7, 17–19, 124, 163, 164–5, 181, 352, 378–9, 388, 408, 429, 453n1, 467, 498; from Khmer 265; from Korean 7; from Northern Ryukyuan 408; from Old or proto-Korean 25, 32; from Paekche 210–11; from Portuguese 246, 264–5; from Russian 181; from Spanish/Latin 264–5
- logophoric person 474, 475–6
- made-in-Japan/Korea characters 10, 13–14, 19, 63, 273
- Mahan confederacy 55
- Makura no sōshi* 212
- man'yōgana 13, 14, 189, 193–4, 218
- Man'yōshū* 4, 16, 189–90, 193–4, 218
- Meiji period (1868–1912) 19, 268, 277, 308
- Middle Chinese 12, 18, 79, 214, 250; *see also* Early Middle Chinese, Late Middle Chinese
- Middle Korean 7, 11, 18, 25, 28–29, 41, 73–122, 126, 127, 128, 169
- Mid Old Korean 41–2, 43, 45, 50, 56–63
- mimetic, mimesis *see* ideophones
- Minamoto no Takakuni 212
- Minna island/variety of Ryukyuan 381, 383
- missionaries, Christian 164, 246, 264–5, 463
- mixed script 4, 16, 123, 132, 219, 253, 273
- Miyako: group of islands/Ryukyuan varieties 3, 4, 351, 363, 376, 381, 382, 384, 385, 386, 396, 406, 408–9; island/variety of Ryukyuan 3, 5, 351, 381, 384
- Miyara island/variety of Ryukyuan 386, 390, 395
- modality, epistemic/deontic 11, 60, 67, 149, 204–5, 234–5, 260, 292–4, 324–5, 341, 376, 394–5, 441–2, 493–4 (*see also* evidentiality); explanatory 287, 294
- modality of utterance 292, 294, 251–2
- Mongolian 5, 10, 11, 15–16, 271, 494
- Mongolic 6–7, 10, 11, 28
- monophthongization 31, 57, 176, 193, 250, 251–2, 314, 322, 328
- moraic nasal 214, 248, 269, 270, 315, 382, 414
- moraic /Q/ 269, 270
- moraic /r/ 382
- moraic /v/ 382
- Murasaki Shikibu 212
- Muromachi period (1392–1573) 246, 254, 258, 259–60, 263
- Nagahama island/variety of Ryukyuan 382
- Nakachi island/variety of Ryukyuan 381, 382, 383, 385, 386, 387, 388, 389, 390, 391, 392, 394, 395, 396, 399, 404, 406, 409
- Nara period (710–794) 4, 189, 192, 210, 242
- nasal drift *see* renjō
- negation 62, 69–70, 105–6, 148–9, 204, 224–5, 233, 259–60, 281–284, 292, 320–2, 323–4, 336–7, 338, 340–1, 359–61, 366, 373–4, 386–7, 422–3, 426, 496
- neo-classical coinages 17, 19, 303, 309
- neologism 124, 134, 161, 163, 164, 165, 264, 308, 309
- Nihon gengo chizu* 313
- Nihon shoki* 15, 48, 49, 51, 118, 189, 191–2, 194, 198, 218
- NINJAL (National Institute for Japanese Language and Linguistics) 313–14
- Nishihara island/variety of Ryukyuan 381, 382, 394, 396
- Nivkh 461, 462
- nominalization, nominalizer 34, 55, 59, 66–7, 69, 70, 96, 115, 158–9, 210, 227, 240–1, 263, 287, 327, 377–8, 387, 389–90, 470, 493–4
- Norito* 189, 194, 207
- noun stem variation 120n9, 174, 221, 272, 470, 478, 480, 490
- nūshū 13
- numerals 33, 75, 95–6, 141, 198–9, 226, 256–7, 284–6, 357, 387–9, 429–30, 498
- numerals, interrogative 98, 226, 284, 357, 387
- numerals, ordinal 96, 141, 198, 256, 285–6, 388
- Ōgami island/variety of Ryukyuan 382, 383, 384

- Okinawa: island/variety of Ryukyuan 3, 5, 16, 351, 379, 380, 405, 408, 409; US occupation of 379, 409
- Okinawa Hōgen Shinbun* 16
- Okinawan *see* Shuri Ryukyuan
- (w)okoto-ten 13; *see also* kugyōi, dot-type okurigana 269, 277
- Old Japanese 4, 12, 25, 49, 51, 54, 189–211, 213, 214, 218, 222, 223, 248
- Old Korean 12, 25, 28, 41–72, 73, 77, 120n1
- Omoro sōshi* 4
- onbin 213, 214, 215, 216, 217–18, 224, 282–3
- ōnmun [enmun] 14, 120n8
- onomatopoeia *see* ideophones
- orthography 123, 129, 132–4, 189, 192, 193–4, 218–19, 253, 273–8; *see also* names of scripts and styles of writing
- Ōura island/variety of Ryukyuan 384
- Paekche [Paykcey] 5, 41, 42, 48, 49, 51, 53–4, 55, 56, 168, 193–4, 210–11
- pair verbs *see* auxiliary verbs, quasi-aspectual
- palatalization 35n7, 80, 130, 171–2, 173, 175, 176, 177, 178, 214, 251, 269, 352, 381, 413, 414, 465
- panmal 9, 146, 151, 154–5, 159, 177
- particles, noun-following 6, 34, 46, 56, 57–8, 64–5, 87–91, 135–7, 194–6, 221–3, 254, 278–80, 318–20, 334–6, 338–9, 355, 368–71, 374, 385–6, 417–21, 476–7; sentence-final 236, 260–1, 289, 294–6, 394, 396–8, 443, 469, 497
- passive 9, 60, 91, 108, 136, 138, 140, 152, 206–7, 219, 221, 223–4, 234, 236–7, 255, 259, 260, 261–2, 278, 279, 282, 283, 298–9, 300, 302, 326, 343, 345, 358, 364–5, 370, 376, 386, 387, 400–1, 419, 422, 442, 445–6, 477, 499
- pejorative 195, 262, 278, 302
- periodization: Japanese 246; Korean 40, 120n1
- ‘phags-pa 14
- phonaesthesia *see* ideophones
- phonograms, phonography 13–16, 32, 49–51, 62–3, 73–4, 82, 191–2, 193–4, 198, 218
- plural: nouns 67–8, 96–7, 134–5, 182, 195, 219, 253, 278, 356, 385–6, 417, 432; verbs 471, 474, 481, 496
- politeness *see* speech levels
- polite particle 9, 155, 279, 281, 300, 403, 443, 446
- polysynthesis 6, 461, 485
- possessive affixes 470, 478
- possessive vs. conceptual noun forms, Ainu *see* noun-stem variation
- postpositions: particles *see* particles, noun-following; quasi-grammatical nouns 56, 238, 303, 389, 469–70, 477, 489
- potential mood 62, 70, 105–6, 149–50, 234, 236, 260, 262, 279, 282–3, 325, 342, 376, 387, 395–6, 422, 423, 442, 493
- prefixation 7, 86, 148, 194, 207, 221, 223, 257, 262, 264, 302, 367, 387, 390–1, 461–2, 467, 469, 470, 473, 474, 478, 482, 486, 488
- prenasalization, sequential *see* rendaku
- prenasalized consonants 27, 191, 206, 213, 214, 217, 222, 250, 315
- pro-drop 286, 287, 472
- prohibitive *see* commands; negation
- proposals (= (co/ex)hortative) 65–6, 104–5, 138, 151, 174, 175, 176, 178, 179, 235, 260, 296, 326, 359, 397–8, 421–2
- proto-Ainu 462, 465, 467
- proto-(Mainland)Japanese 215, 345
- proto-Japanese-Korean 5, 33–4
- proto-Japanese-Ryukyuan 4, 7, 24–7, 30–4
- proto-Korean 7, 24, 27–34
- proto-Ryukyuan 26, 216, 408
- pseudo-Chinese (= *hentai kanbun*, *pyōnch’e hanmun*) 15, 42, 218, 253
- pseudo-Chinese characters *see* made-in-Japan/Korea characters
- pyōnch’e* [*pyenchey*] *hanmun see* pseudo-Chinese
- Pyōnhan [Pyenhan] confederacy 55
- punctuation 46, 268, 277, 310
- /Q/-final morphemes 328–30
- questions (= interrogative) 9, 58–9, 66, 138, 150–1, 173–9, 205, 235, 260–1, 295–6, 325, 342, 359–61, 373, 396–7, 442–3, 497
- quotation *see* evidentiality; subordination
- Rakuyōshū* 253
- reciprocal voice 232, 481, 488
- recollection 289, 294, 324
- reduplication, plural 219, 253, 278, 385, 390; *see also* ideophones
- reflexive voice 481, 488–9
- relational nouns *see* postpositions, quasi-grammatical nouns
- relative clause, relativization 6, 70, 75, 93, 112–4, 137, 143–4, 200–1, 227–8, 257, 286–7, 327, 343, 378, 389, 430, 478–9; internally headed 6, 227–8, 257, 286, 287, 389
- rendaku (= sequential prenasalization, sequential voicing) 193, 217, 222, 247, 272–3, 275, 308, 329–30, 367
- renjō (= liaison) 214, 247
- requests 9, 59, 127, 261, 296, 301, 303, 307, 397, 443
- ruby *see* furigana
- Ruiju myōgishō* 192, 215–16
- Ryukyuan 3–4, 5, 9, 16, 24, 32, 34, 189, 313, 334, 351, 370; *see also* individual varieties
- Ryukyuan, Northern = Amami-Okinawan 4–5, 351–80, 408
- Ryukyuan, Southern = Sakishima 4–5, 381–411; *see also* Yonaguni

- Ryukyu kingdom/dynasty (1429–1879) 3, 19, 351, 379
- Sakhalin 461, 464, 467, 470
- Sakishima *see* Southern Ryukyuan
- salience hierarchy, genitive/subject marking 222, 370, 385–6
- Samguk sagi* [*Samkwuk saki*] 5, 25, 42, 50, 51, 53, 73, 76
- Samguk yusa* [*Samkwuk yusa*] 42, 44–5, 50, 51, 53, 57, 74
- Samhan period/states 41, 51, 54, 55
- sandhi: in Ainu 467–8; in Korean 129–32; in Japanese *see* onbin; rendaku; renjō
- Sarahama island/variety of Ryukyuan 381
- Sarashina nikki* 8, 212
- scrambling 96, 100, 107, 286
- Secret History of the Mongols* 15–16
- sei-daku *see* prenasalized consonants
- Sei Shōnagon 212
- Sejong [Seycong], King 12, 14, 73, 79, 81, 83–4, 123, 133
- selective 236, 419, 432, 443–5
- semantograms, semantography 13–16, 49–50, 62, 82, 191, 193, 198, 218
- Senmyō* 190, 195, 207
- Setsuyōshū* 248, 253
- Shiraho island/variety of Ryukyuan 383
- Shittan kuden* 246, 248, 251
- shōmono* 246
- Shuōwén jiēzì* 218
- Shuri variety of Ryukyuan 3, 4, 6, 9, 11, 16, 216, 251–380, 412, 431, 461
- Silla [Sinja] 5, 28, 41, 42–3, 48, 49, 51, 53–6, 61, 63, 73, 76, 77, 118, 168
- sinicization of Korea, eighth-century 5, 12, 63
- Sino-Japanese stratum 7, 17–19, 124, 163, 226, 242–3, 247, 248, 249, 250–1, 252, 253, 256, 257, 262, 264, 270, 271, 272–3, 274, 277, 284–5, 302, 316, 331, 352, 357, 378–9, 387–8, 408, 429–30; *see also* loanwords, from Chinese
- Sino-Japanese War (1894–1895) 19
- Sino-Korean stratum 7, 12, 17–19, 28, 63, 64, 83, 118, 123, 129, 132, 134, 141, 148, 161–3, 165, 171, 181–2, 183; *see also* loanwords, from Chinese
- Sino-Paekche 194
- slang 275, 309
- sociative 488
- Sonai variety of Ryukyuan 382, 383, 408, 410
- sōrobun* 219, 237, 268, 310
- speech levels, addressee-based (= politeness) 9, 109–11, 150–1, 153–5, 183–4, 237, 262, 281–3, 300–1, 342, 357–60, 363, 364, 443–44; *see also* polite particle
- standardization 124, 181, 184, 276, 277, 283
- subordination 10, 56, 66, 116, 157–8, 209–10, 240, 263, 305–7, 327, 343–5, 377–8, 404–7, 447–51, 497–8
- switch reference 230, 238–40, 241, 263
- syllabary *see* kana
- tabooed vocabulary 264, 277, 302, 408
- Taketomi island/variety of Ryukyuan 382, 383–4, 395, 399, 408, 416
- Taketori monogatari* 219, 243
- Tarama island/variety of Ryukyuan 3, 381, 382, 383, 384, 385, 408
- telicity 203, 485, 493
- tense and aspect 60, 67, 94–5, 101, 113, 138, 143–4, 146–8, 203–4, 230–3, 254–6, 259, 279, 290–2, 321–3, 336–8, 340, 357–61, 392–4, 422–4, 426, 436–40, 471, 492–5
- tensed consonants 77–8, 79, 80, 83, 125, 129–30, 131–2, 162
- t*-final morphemes in Japanese 249–50
- Three Kingdoms period (–7thC) 5–41, 42, 48, 49–50, 51, 54, 55, 70, 118, 168
- t'o [tho] marks 43, 45, 46–8, 50–1, 56, 63, 66
- Tōhoku region/dialects 313ff
- tone *see* accent
- Tongnip Sinmun* [*Toklip Sinmun*] 123
- tō'on lexical stratum 264
- topic, topicalization 58, 64–5, 90, 106, 136–8, 151–2, 205–6, 221, 236, 254, 261, 279, 296–7, 326, 338, 342–3, 374, 398, 420, 443
- Tosa nikki* 16, 212, 217, 219–20
- total availability 18–19, 20
- tōyō kanji 268, 273
- Tsurezuregusa* 212, 213
- Tungusic 6–7, 10, 28, 501n32
- Turkic 6–7, 10, 494
- typology 5–7, 11, 24, 461–2, 470, 484, 498
- Uji shūi monogatari* 213, 230
- ūmdokcha [umtokca] 42, 49–50, 55–6, 63
- ūmgaja [umkaca] 13, 49–50
- Unified Silla [Sinja] period (7thC–10thC) 41, 73, 118
- unwitnessed past 230, 239
- uwepeker* *see* folktales, Ainu
- Uyghur 5, 10
- verbalizer on adjectives 365, 368
- verb splitting (light verb) 225–6, 229, 233, 236, 237, 398, 404
- verbs, analogical attrition of Japanese vowel stems 320–1, 336–7
- Vietnamese 15, 16, 18
- Vocabulario da Lingoa de Japan* 246, 247, 248, 249, 250, 251, 252, 263–4
- voicing, sequential *see* rendaku
- volition (= desiderative) 150, 234, 255, 260, 289, 290, 303, 325, 341, 395, 442, 493
- vowel copying 466, 478
- vowel harmony 11, 28, 30, 57, 58, 81, 87, 107, 125, 128, 134, 160
- vowel shortening 328, 383–4

*wabun* 219, 253

*wakan konkōbun* 253

*Wamyō ruijushō* 192

*wasei kanji* see made-in-Japan/Korea characters

Western Old Japanese see Old Japanese

word classes 84, 86, 97, 111, 119, 353, 354–5, 357, 469–70, 471

word order 6–7, 15, 42, 56, 63, 100, 143, 150, 151, 193, 206, 257, 260, 286, 288, 290, 322, 339, 435, 461, 497

writing: in Japan before Chinese characters 12, 193; Japanese writing for Okinawan 16, 351; see also individual scripts and writing styles

xingsheng characters 12, 14

Yaeyama group of islands/Ryukyuan varieties 351, 381, 382, 384, 395, 396, 406, 409, 412, 413, 431

Yanbian autonomous prefecture 168

Yonaguni island/variety of Ryukyuan 3, 4, 5, 9, 14, 236, 241, 351, 386, 412–57

*Yongbi ōch'ŏn'ga* [*Yongpi echenka*] 8, 16, 77, 83, 84

Yoshida Kenkō 212

*yotsugana* issue 249, 330

*yukar* see folktales, Ainu

zero-marking of nouns 7, 8, 10, 11, 12, 13, 14























Taylor & Francis

# eBooks

## FOR LIBRARIES

ORDER YOUR  
FREE 30 DAY  
INSTITUTIONAL  
TRIAL TODAY!

Over 23,000 eBook titles in the Humanities, Social Sciences, STM and Law from some of the world's leading imprints.

Choose from a range of subject packages or create your own!

Benefits for  
you

- ▶ Free MARC records
- ▶ COUNTER-compliant usage statistics
- ▶ Flexible purchase and pricing options

Benefits  
for your  
user

- ▶ Off-site, anytime access via Athens or referring URL
- ▶ Print or copy pages or chapters
- ▶ Full content search
- ▶ Bookmark, highlight and annotate text
- ▶ Access to thousands of pages of quality research at the click of a button

For more information, pricing enquiries or to order a free trial, contact your local online sales team.

UK and Rest of World: [online.sales@tandf.co.uk](mailto:online.sales@tandf.co.uk)

US, Canada and Latin America:  
[e-reference@taylorandfrancis.com](mailto:e-reference@taylorandfrancis.com)

[www.ebooksubscriptions.com](http://www.ebooksubscriptions.com)



Taylor & Francis eBooks  
Taylor & Francis Group



A flexible and dynamic resource for teaching, learning and research.