CAMBRIDGE STUDIES IN LINGUISTICS 92

Tone Sandhi

Patterns Across Chinese Dialects

MATTHEW Y. CHEN

CAMBRIDGE

more information - www.cambridge.org/0521652723

This page intentionally left blank

Matthew Chen's landmark study offers the most comprehensive analysis to date of the rich and complex patterns of tone used in Chinese languages. Chinese has a wide repertoire of tones which undergo often surprising changes when they are connected in speech flow. The term tone sandhi refers to this tonal alternation. Chen examines tone sandhi phenomena in detail across a variety of Chinese dialects. He explores a range of important theoretical issues such as the nature of tonal representation, the relation of tone to accent, the prosodic domain of sandhi rules, and the interface between syntax and phonology. His book is the culmination of a ten-year research project and offers a wealth of empirical data not previously accessible to linguists. Extensive references and a bibliography on tone sandhi complete this invaluable resource which will be welcomed as a standard reference on Chinese tone.

Having taught for many years at the University of California, San Diego, MATTHEW CHEN is Professor and chair of Linguistics and Dean of the Humanities and Social Sciences faculty at the City University of Hong Kong. He has written numerous articles on linguistics and is Associate Editor of the *Journal of Chinese Linguistics*, University of California, Berkeley.

In this series

52 MICHAEL S. ROCHEMONT and PETER W. CULICOVER: *English focus constructions and the theory of grammar*
53 PHILIP CARR: *Linguistic realities: an autonomist metatheory for the generative enterprise*
54 EVE SWEETSER: *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*
55 REGINA BLASS: *Relevance relations in discourse: a study with special reference to Sissala*
56 ANDREW CHESTERMAN: *On definiteness: a study with special reference to English and Finnish*
57 ALLESSANDRA GIORGI and GUISEPPE LONGOBARDI: *The syntax of Noun Phrases: configuration, parameters and empty categories*
58 MONIK CHARETTE: *Conditions on phonological government*
59 M.H. KLAIMAN: *Grammatical voice*
60 SARAH M.B. FAGAN: *The syntax and semantics of middle constructions: a study with special reference to German*
61 ANJUM P. SALEEMI: *Universal grammar and language learnability*
62 STEPHEN R. ANDERSON: *A-Morphous morphology*
63 LESLEY STIRLING: *Switch reference and discourse representation*
64 HENK J. VERKUYL: *A theory of aspectuality: the interaction between temporal and atemporal structure*
65 EVE V. CLARK: *The lexicon in acquisition*
66 ANTHONY R. WARNER: *English auxiliaries: structure and history*
67 P.H. MATTHEWS: *Grammatical theory in the United States from Bloomfield to Chomsky*
68 LJILJANA PROGOVAC: *Negative and positive polarity: a binding approach*
69 R.M.W. DIXON: *Ergativity*
70 YAN HUANG: *The syntax and pragmatics of anaphora*
71 KNUD LAMBRECHT: *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*
72 LUIGI BURZIO: *Principles of English stress*
73 JOHN A. HAWKINS: *A performance theory of order and constituency*
74 ALICE C. HARRIS and LYLE CAMPBELL: *Historical syntax in cross-linguistic perspective*
75 LILIANE HAEGEMAN: *The syntax of negation*
76 PAUL GORRELL: *Syntax and parsing*
77 GUGLIELMO CINQUE: *Italian syntax and universal grammar*
78 HENRY SMITH: *Restrictiveness in case theory*
79 D. ROBERT LADD: *Intonational phonology*
80 ANDREA MORO: *The raising of predicates: predicative noun phrases and the theory of clause structure*
81 ROGER LASS: *Historical linguistics and language change*
82 JOHN M. ANDERSON: *A notional theory of syntactic categories*
83 BERND HEINE: *Possession: cognitive sources, forces and grammaticalization*
84 NOMI ERTESCHIK-SHIR: *The dynamics of focus structure*
85 JOHN COLEMAN: *Phonological representations: their names, forms and powers*
86 CHRISTINA Y. BETHIN: *Slavic prosody: language change and phonological theory*
87 BARBARA DANCYGIER: *Conditionals and prediction: time, knowledge and causation in conditional constructions*
88 CLAIRE LEFEBVRE: *Creole genesis and the acquisition of grammar: the case of Haitian Creole*
89 HEINZ GIEGERICH: *Lexical strata in English: morphological causes, phonological effects*
90 A.M.S. MCMAHON: *Lexical phonology and the history of English*
91 KEREN RICE: *Morpheme order and semantic scope: word formation and the Athapaskan verb*
92 MATTHEW Y. CHEN: *Tone sandhi: patterns across Chinese dialects*

Supplementary volumes

LILIANE HAEGEMAN: *Theory and description in generative syntax: a case study in West Flemish*
A.E. BACKHOUSE: *The lexical field of taste: a semantic study of Japanese taste terms*
NIKOLAUS RITT: *Quantity adjustment: vowel lengthening and shortening in early Middle English*

Earlier issues not listed are also available

CAMBRIDGE STUDIES IN LINGUISTICS

Tone Sandhi

To PL, sine qua non

TONE SANDHI

Patterns across Chinese dialects

MATTHEW Y. CHEN

City University of Hong Kong

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published in printed format 2000

ISBN 0-521-65272-3 hardback

Contents

Preface

Although applied to the description of languages as diverse as Ewe of Ghana (Clements 1978), Mixtec of Mexico (Hunter and Pike 1969), and Kairi of Papua New Guinea (Newman-Petterson 1990), the term **tone sandhi** refers, first and foremost, to the kind of tonal alternations one typically finds in languages spoken in China and surrounding areas of the southeast Asian mainland. These languages share certain areal characteristics, notably highly developed tonal repertoires matched, in many cases, by even more complex tonal alternations in connected speech. The Sanskrit root of *sandhi* means junction, connection, combination, or liaison (cf. Allen 1962, Andersen 1986). Tone sandhi, therefore, *sensu stricto* describes phonetically conditioned morphotonemic alternations at the juncture of words or morphemes. Over time, however, tone sandhi has been extended to cover a number of related phenomena, including allotonic variations, intonational effects, and morphologically or syntactically conditioned tone changes. Such a broad and loose usage of the term has one advantage over any strict definition, as Andersen (1986:2) argues in a different context, in that it gives us a convenient label without prejudging the issues involved.

Even though the tonal/registral systems and their historical developments have been investigated extensively in such language groups as Tibeto-Burman, Miao-Yao, Kam-Tai, and Mon-Khmer, tone sandhi has not received as much attention[1] – with the notable exception of the Sinitic group. As a consequence, our study of tone sandhi has a distinctly Chinese focus and flavor, although I will not shy away from drawing on other languages for facts and inspiration.

[1] For instance the nearly 900-page report on the Kam-Tai languages spoken in China (J. Wang et al. 1984) devotes exactly two short paragraphs to tone sandhi in only one out of the eight languages surveyed, all of which have highly developed tone systems. It is not clear whether the Kam-Tai languages are tone-rich but sandhi-poor, or display sandhi phenomena that have gone underreported as in the early days of Chinese dialectology.

The investigation of tone sandhi has a long tradition in Chinese linguistics. The fourteenth-century pronouncing dictionary *Zhongyuan Yinyun* contains the earliest intimation of tone sandhi I am aware of. There the author, Zhou Deqing, noted that in versification "it would be best to avoid a pair of Rising tones or a pair of Departing tones." The Standard Mandarin tone sandhi rule as we know it today already found a clear formulation by the sixteenth-century Korean scholar Cui Shizhen, which I quote in part:

> If both syllables are in the Rising tone, then the circumstances make it difficult to retain the original tone. In this case, pronounce the first syllable like the voiced variety of the Level tone, and then the second syllable can retain its original tone when pronounced. (from Mei 1977:238f.)

A similar statement is found in the very first Western language grammar of Chinese written by Francisco de Varo, O.P. in 1682, whose statement is reproduced here:

> *mai*, en tercera tonada es comprar . . . mas juntandola con otra tercera, como *mai ping*, pronunciandolas juntamente el *mai* ba quasi a ser primera ["*mai*, in the third tone, means 'to buy' . . . but next to another third tone, as in *mai ping*, it nearly becomes the first tone when pronounced together"; tr. MYC] (Francisco de Varo, 1682, *Arte de la Lengua Mandarina*, p. 9)

In modern times, systematic descriptive work on Chinese dialects dates back to the 1920s. The early surveys tended to focus on the static tonal systems, and were sparse in information on the dynamic interaction of tones in context. Two early works by Chiu (1931) and Luo (1930) – both on the south Min dialect of Xiamen – are exceptional in their extensive treatment of sandhi phenomena and close attention to phonetic details. These are among the true pioneers of the study of tone sandhi. Since then important works, especially doctoral dissertations, have been devoted to this specialized topic. But it was the launching of the journal *Fangyan* in 1979 that marked the beginning of a dramatic explosion of empirical knowledge about the range and diversity of sandhi phenomena, some of which occur in obscure and hard-to-reach dialects. We get a glimpse of the overall picture from A. Hashimoto (1987 [1980]), Chen (1991a [= 1985]), and Ballard (1988). Since then our knowledge about the subject matter has broadened and deepened considerably, and the time is ripe for a new synthesis.

Two leitmotifs underscore the study of tone sandhi with particular reference to Chinese. The first concerns the internal structure of tone. Different hypotheses regarding tonal features and their geometrical arrangements make different predictions about the typology of possible tonal processes. Tone sandhi, therefore, serves as an effective diagnostic probe into the anatomy of the complex entity we call tone. The second recurrent theme of tone sandhi studies concerns the scope or domain of sandhi rules. Unlike most segmental phenomena, which tend to be localized,[2] tonal processes are notorious long-distance runners, sometimes spanning entire phrases and sentences. The precise definition of tone sandhi domains, therefore, raises intriguing questions about the interface between phonology and grammatical structure. There is a third, far less well-developed issue, namely the interplay among sandhi processes. Given a tonal string A–B–C, the sandhi form of the whole is typically the composite result of the elementary processes operating on the substrings A–B and B–C. How exactly the elementary processes interact to produce the ultimate sandhi output is a topic that has not been heretofore explored in depth.

Thematically, this book is organized as follows: after the introductory and stage-setting chapter, the book is divided into two parts. The first part, comprising chapters 2 to 6, deals with various types of tone sandhi phenomena. Chapter 2 focuses in particular on tonal geometry and the typology of sandhi rules. Chapters 3 and 4 are devoted to issues regarding how a sandhi process is implemented (e.g. directional iteration) and how one process may interact with another. These issues are of particular significance from the perspectives of Optimality Theory (Prince and Smolensky 1993, inter alia). Chapters 5 and 6 together constitute an in-depth investigation of one particular dialect, New Chongming. One of the startling discoveries is that this northern Wu dialect is well on its evolutionary path toward a classic accentual system.

The question of sandhi domains is taken up in part two, consisting of chapters 7 to 11. The scope within which tone sandhi rules operate ranges from a sublexical stress-foot to a phonological word, phonological phrase, and intonational phrase. One surprising finding is that the "Minimal Rhythmic Unit," which circumscribes the scope of tone sandhi in Beijing Mandarin (chapter 9), is not commensurate with any of the categories we know of conventional prosodic hierarchy. This and other findings are

[2] Vowel harmony and nasalization being two well-known exceptions.

summarized in the Concluding Remarks, followed by References and a Subject Index.

I take for granted not only generative phonology with its subtheories including autosegmental phonology, feature geometry, metrical phonology, and prosodic phonology, but also Optimality Theory, which by the time this book sees the light of day, will no doubt have become every practicing linguist's stock in trade. Where alternative descriptions of the facts are equivalent, I generally couch my analysis in conventional, generative terms. I will not hesitate to exploit the insights and formalisms of Optimality Theory when it sheds light on the issue at hand.

Much of the empirical data underpinning the present study is in the public domain, published in journals or being circulated through informal channels (unpublished doctoral dissertations, manuscripts, etc.). However, given the nature of the theoretical issues under investigation, published sources are often silent on critical aspects of the problem; as a consequence we had to develop new *kinds* of data, involving a variety of larger and more complex constructions. This is true even of relatively well-known and well-documented dialects, such as Standard Mandarin or Xiamen. I use myself as the principal informant for Xiamen and Mandarin, consulting other native speakers where subtle judgments may differ. In data sampling, I have resisted the butterfly collector's temptation, and opt for in-depth analysis of a few dialects selected for certain structural properties they serve to illuminate. I have collected a fair amount of original data on the following dialects: New Chongming and Old Chongming (northern Wu), Wenzhou (southern Wu), Tianjin (northern Mandarin), and Pingyao (Jin).

Acknowledgments

From the earliest conception to the final stages of redrafting this book, I have benefited from the inspiration and insights of countless friends and colleagues. But most of all, I am indebted to my past and current students who have nurtured and shared with me an abiding interest in the wondrous ways tones behave in connected speech. Chilin Shih, Tony Hung, Yuchau Hsiao, Hongming Zhang, and Huichuan Hsu and Lily Chan have all completed their doctoral theses on tone sandhi. They have shaped my thinking as much as I have theirs.

My thanks go to the informants who patiently sat through interminable sessions. In particular, I wish to name two informants who are linguists

in their own right: Pan Wuyun and Zhang Huiying. Pan Wuyun not only provided the raw data, but also offered some of the insights on which my analysis of Wenzhou (chapter 11) is based. Xu Baohua of Fudan University was my most gracious host in the fall of 1986 and the summer of 1990 during which most of the fieldwork on Wu dialects was carried out. Several of his colleagues, especially Pan Wuyun, Qian Nairong, Zhang Hongming, and Chen Zhongmin helped with various aspects of data elicitation and transcription. With patience and remarkable thoroughness, Lily Chan and Karanda Tang assisted me with the myriad minutiae of the final draft.

While the bulk of the present draft was written between 1994 and 1996 (during which the author was granted two quarters of sabbatical-in-residence), the underlying research took much longer. I have presented some of the preliminary results at various conferences and gatherings, including: a lecture series delivered at Centre de Recherches Linguistiques sur l'Asie Orientale (Paris, June 1990), talks delivered at the LSA Summer Institute (University of California, Santa Cruz, July 1991), Berkeley Linguistics Society Meeting (February 1992), Workshop on the Psychobiological Basis of Language (Taipei, December 1993), the Tilburg conference on "Derivational Residue" (Tilburg, October 1995), and on several occasions at East-Asian Linguistics Workshop (University of California Irvine, October 1993, October 1995), North-American Conference on Chinese Linguistics (Cornell University, May 1991; University of Delaware, May 1993; University of Southern California, May 1994; University of Wisconsin, June 1995; University of Illinois, May 1996), International Conference on Chinese Linguistics (Singapore, July 1992; Paris, June 1993), International Symposium on Chinese Languages and Linguistics (Taipei, July 1991, July 1992), Pan-Asiatic Linguistics Symposium (Chulalongkorn University, January 1992; Mahidol University, January 1996), and various colloquia (Berkeley, November 1990; November 1991; University of California, Los Angeles, April 1991; October 1995). I thank the organizers and the audiences of these meetings for their comments and criticisms.

An National Science Foundation grant (BNS-8608374, 1986–90) and a fellowship awarded by the Wang Institute of Graduate Studies (Tyngsboro, Mass., 1986–87) gave the initial impetus to this decade-long project by providing the time and wherewithal to make several trips to China and to bring informants to this country for extensive fieldwork. Over the years, the University of California, San Diego, Academic Senate has made smaller grants to pay for research assistance and related costs. In

grateful acknowledgment of their generous support, this book comes as a long overdue final report.

I wish to thank Richard Attiyeh, Vice-Chancellor of Research and Dean of Graduate Studies, University of California, San Diego, Joseph C-Y. Chen, Director of Wei-Kung Institute, and Yuchau Hsiao of National Chengchi University for their generous publication subsidy made to Cambridge University Press.

Notational conventions

Symbols	Meaning
T	Tone, tone root
T′	A modified or derived tone.
o	Zero tone, or unspecified for tone.
33, 24, 51 . . .	Tone letters devised by Y-R. Chao (1930), where digits indicate the pitch value on a five-point scale, 5 being highest. Thus 33, 24, 51 represent, a mid level, high rising, and a falling tone respectively.
H, M, L	High, mid, low. H, M, L and tone letters are used interchangeably throughout this book. Where appropriate, a dot separates tone digits and H, M, L, for instance to distinguish HM.L (= HM + L) from H.ML (= H + ML), or 3.52 (= 3 + 52) from 35.2 (= 35 + 2).
T1, T2 . . .	Tonal categories: tone 1, tone 2, etc. in a given system.
CVN	"Smooth" or "legato" syllable, i.e. ending in a vowel, offglide or a nasal.
CVq	"Checked" or "staccato" syllable, i.e. ending in an obstruent coda -p,t,k or a glottal stop -q.
MC	Middle Chinese
I, II, III, IV	MC tonal categories: *ping, shang, qu, ru*; tone IV cooccurs only with checked or staccato syllables CVq.
Xa,b . . .	Middle Chinese tonal category X, register *a* (typically high) or *b* (low). Middle Chinese Ia, Ib, IIa, IIb . . . are sometimes referred to as T1, T2, T3, T4 . . . , with odd / even numbers indicating register a and b respectively.
Tq	Checked tone, i.e. a tone associated with a checked syllable (CVq); thus 5q and 23q stand for a high and a low rising tone linked to CVq. Smooth or legato tones, i.e. tones linked to CVN are unmarked.

Symbols	Meaning
Hr, Lr	High register, low register, equivalent to [+upper] and [–upper]
h, l	Terminal tone segments, equivalent to [+raised] and [–raised] Lower case tone letters [h, m, l] are also used to indicate default pitch values H, M and L.
E, R, F	Even (= level), rising, falling
σ, μ	Syllable, mora
ϕ, p-word	Foot, prosodic word
p-phrase	Phonological phrase
IP	Intonational phrase
(x .), (. x)	Left-, right-prominent metrical unit
$C_1 \gg C_2$	Constraint 1 ranks above constraint 2
X } Y	Candidate X is more "harmonic" than candidate Y
☞	Picks out the winning candidate

Phonetic transcription

Examples are usually cited in the form they are found in the sources. Occasionally, the same pronunciation is transcribed differently by different authors. For instance, the Shanghai pronunciation of the word *Australia* is given as *qo.ta.li.ya* (Duanmu 1993a) and *ɔ.tʌ.li.iʌ* (Xu et al. 1988). I have made no systematic effort to standardize the phonetic transcriptions or phonemicizations across the board. The inconsistencies from one source to another are seldom crucial for our purpose; where they are relevant, they will be duly noted. Here are some recurrent symbols used here and their interpretations:

q	glottal stop
h̲	voiced / murmured h
y, w	glides of i,u
ü	front-rounded u
C′	aspirated
ng	velar nasal
ɤ	mid, back, unrounded vowel
E, I	(in small cap), lax vowels

Where phonetic transcriptions are not provided in the sources, I use the standard Pinyin system, which is also the accepted convention for transcribing Standard (Beijing) Mandarin. Pinyin differs from the IPA system chiefly in the symbols used for consonants. Here are the Pinyin symbols and their phonetic interpretation:

		labial	dental	retroflex	palatal	velar
stops	plain voiceless	b	d			g
	voiceless aspirated	p	t			k
affricates	plain voiceless		z	zh	j	
	voiceless aspirated		c	ch	q	
fricatives		f	s	sh	x	h
sonorants		m	n, l	r		ng

Occasionally when clarity demands, I separate syllables as well as tones by a dot. Thus *tian.an.men* "The Gate of Heavenly Peace" is syllabified as indicated (rather than *tia.nan.men*). More pertinently, MH.L = MH + L, while M.HL = M + HL; correspondingly: 35.1 = 35 + 1, but 3.51 = 3 + 51.

In citing examples, I often adopt the following format:

red	ten character		←	literal gloss
hong	[*shi* *zi*]	"the Red Cross"	←	translation
(MH.LM)(LM)			←	input, base form
(MH. o)	(LM)	Deletion		
(M. H)	(LM)	Spread		
(M. H)	(MLM)	M-Insertion	←	output, sandhi form

Here the semantic relation between "the Red Cross" and its component parts (lit. "the character for the word ten") is somewhat opaque.[3] I omit the gloss in those cases where the mapping between the literal (morpheme-by-morpheme) gloss and the English translation is transparent. The square brackets indicate morphosyntactic constituency, while the parentheses mark prosodic units, in this case the metrical feet. The sandhi rules that generate the intermediate or final outputs are named on the right by their labels given in the main text.

[3] The graphic representation of the root morpheme "ten" happens to be shaped like the cross: ✚.

Rule formalism

Occasionally I rotate the rewrite arrow clockwise by 90 degrees so that the target (which I underline for clarity) and the environment of a phonological rule would be right next to each other, rendering the rule more transparent. For instance:

$$\begin{array}{l} \underline{55}.\ 33 \\ \downarrow \\ 53 \end{array}$$

is equivalent to

$$55 \quad \rightarrow \quad 53\ /\ \_\_\_\ 33$$

1 *Setting the stage*

This introductory chapter is intended to provide the necessary background for our investigation of tone sandhi. After a brief description of the genetic grouping of the languages of China, from which we draw the bulk of our primary data (section 1), I give a thumbnail sketch of the tone system of Middle Chinese (circa AD 600) and its evolution into the diverse patterns we see in modern dialects (sections 2–3). Historical tonal categories furnish us with a common frame of reference as we move from one dialect to another. I then set tone sandhi in the context of various types of tonal perturbations in connected speech, including tonal coarticulation, intonational effects, and morphologically conditioned tone changes (section 4). Tone sandhi processes often strike the analyst as arbitrary and totally lacking in phonetic or functional motivation. Section 5 shows that we can make sense of, if not explain, certain puzzling synchronic facts if we look at them from a diachronic perspective. This chapter closes with some terminological clarification (section 6).

1 Languages and dialects of China

According to *Major Statistics of the 1982 Census*, published by the People's Republic of China State Statistics Bureau (Beijing, October 1982), China (including Taiwan) has a population of 1,026 million.[1] Of these, 977.2 million or 95.2% speak one form or another of Chinese. The remaining 46.2 million are distributed over a wide variety of language families/stocks, spoken mostly on the periphery of China, with a high concentration of speakers of "minority" languages across the southwestern provinces. *Language Atlas of China* (Longman, Hong Kong 1987), compiled by the Australian Academy of the Humanities and the Chinese Academy of Social

[1] Quoting official statistics, *Language Atlas of China* (A-1) puts the population at 1.1 billion by April 1989.

Sciences, affords us a glimpse of the linguistic diversity within the political boundaries of China, that includes Sinitic as well as Tibeto-Burman, Kam-Tai, Miao-Yao, Austronesian (Formosan), Mon-Khmer, Altaic, and even Indo-European languages.

Of more immediate interest is the classification of the Chinese languages, more commonly referred to as "dialects." We can make meaningful typological generalizations not only about individual dialects, but about dialect groups. For instance, the "southern" dialects typically have larger tonal inventories than the Mandarin group (see Cheng 1973b, 1991 for statistical data). More importantly, the Jin, Wu, Min, Hakka, and some Mandarin dialects display highly complex tone sandhi, while Xiang, Gan, and especially Yue show only limited tonal alternations. Furthermore, sandhi processes take different forms in different dialect (sub)groups: tone deletion and tone spread, widely attested in Wu, are all but unknown among Mandarin and Min dialects.[2] It has been often noted that while northern Wu has a left-prominent prosodic structure, Min, Mandarin, and southern Wu exhibit a right prominence. This difference in rhythmic organization entails far-reaching consequences in tone sandhi behavior. As these and other generalizations hold across groups of dialects, it is often useful to identify the group membership of a particular dialect under discussion.

While Yuan (1960) still serves as a standard reference and most informative overview of Chinese dialectology, more recent surveys can be found in Egerod (1967), Norman (1988), and You (1992). Intensive research in the genetic classification and geographical distribution over the last two decades or so has culminated in *Language Atlas of China* (1987). The *Atlas* divides Chinese dialects into ten groups as shown in table 1.1. Mandarin, spoken by roughly 65 percent of the entire population of China, covers the largest area – basically the entire region north of the Yangzi river and the southwestern provinces (Yunnan, Guangxi, Guizhou, Sichuan). The compilers of the *Atlas* have separated the dialects spoken in Shanxi and adjacent regions of Hebei and Shaanxi from the surrounding Mandarin dialects, and put them under the Jin group. The remaining eight groups – sometimes collectively known as the "southern" dialects – are all concentrated in the southeastern corner. Aside from its main "homeland" located at the borderland where Jiangxi, Fujian, and Guangdong meet, pockets of Hakka speakers are found in Guangxi, western Guangdong, Taiwan, and

[2] That is, outside of the well-known so-called "neutral tone" phenomena.

Table 1.1. *Chinese dialects*

Group	Speakers (in millions)	Location (Provinces)	Representative Dialects
Mandarin	662.2	north of Yangzi rivers, and south-west provinces	Beijing, Tianjin, Ruicheng
Jin	45.7	Shanxi, north Shaanxi, west Hebei	Pingyao, Changzhi
Wu	69.8	south Jiangsu, Zhejiang, south-east Anhui	Shanghai, Suzhou, Danyang, Chongming, Zhenhai, Tangxi, Wenzhou, Wenling
Hui	3.1	south-east Anhui, west Zhejiang	Tunxi
Gan	31.3	Jiangxi, east Hunan	Nanchang
Xiang	30.9	Hunan	Changsha
Min	55.1	Fujian, Taiwan, east Guangdong, Hainan (south-east Asia)	Fuzhou, Xiamen, Chaozhou, Taiwanese, Wenchang
Yue	40.2	Guangdong, east Guangxi (south-east Asia, Americas)	Cantonese, Taishan
Pinghua	2.0	south Guangxi	Nanning
Hakka	35.0	south Jiangxi, west Fujian, east Guangdong, parts of Taiwan	Meixian, Changting, Pingdong

scattered over a large area of Sichuan. Even more far-flung is the Min (super)group. Specifically, varieties of southern Min are spoken not only on the mainland (Fujian and eastern Guangdong), but have spread over much of the islands of Taiwan and Hainan, and the Leizhou peninsula in southwestern Guangdong.

Citing *Renmin Ribao* (Overseas edition, March 11, 1989), R. Li (1989: 164) estimates overseas Chinese population at somewhere between 26.8

and 27.5 million, most of them living in southeast Asia (25 million) and the Americas (1.8 million), with the rest being scattered over Europe (380,000), Oceania (180,000) and Africa (80,000). Most of the overseas Chinese speak one or other Chinese dialect, in descending order of numerical strength: Yue (including Cantonese, 11 million), Southern Min (including Xiamen, Chaozhou, 8.6 million), Mandarin (3.5 million) and Hakka (0.75 million).

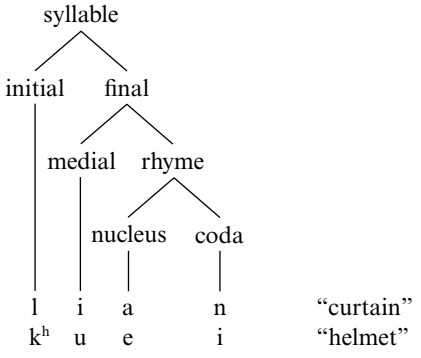
2 Historical background

For reasons that will become apparent (see section 5), an elementary knowledge of historical Chinese phonology is indispensable as a background to the understanding not only of tone sandhi, but of tonal systems in modern dialects. Not only do traditional tonal categories provide us with a handy common frame of reference that holds relatively constant across dialects, but in some cases tone sandhi rules are unstatable without reference to the same classical categories (see section 5.1).

2.1 Middle Chinese tonal categories

We start with the syllable. In most Chinese dialects, the maximal syllable consists of CGVX, that is, a consonantal onset, a prenuclear onglide, the nucleus, and a coda (which can be either an offglide, a nasal, or a voiceless stop). The hierarchical structure of the syllable is captured by the following diagram, labeled with the traditional terms commonly used in Chinese philological literature:

(1)



The "initial" is, of course, the onset. The "final" includes the medial (onglide) as well as the rhyme. The rhyme consists of a nucleus and a

coda, which can be either an offglide or a consonantal ending.[3] The nucleus is the only obligatory syllabic element: thus bare vowels [i] "to heal," [u] "house," [ü] "rain," [a] (a prefix), [ə] "hungry" etc. all constitute legitimate syllables in Standard Mandarin. There remains some ambiguity regarding the status of the medial: whether the prenuclear glide belongs with the onset, or forms part of the final. For the on-going debate regarding subsyllabic constituency, I refer the reader to Lin (1989), Duanmu (1990a), and Bao (1990b, 1996a).

For our purposes, it suffices to note that syllables fall into two classes: (i) "**checked**" **syllables**, namely syllables ending in an occlusive coda (-p,t,k, often reduced to a glottal stop -q); (ii) "**smooth**" or "**slack**" **syllables**, namely either an open syllable CV (possibly with an offglide), or a syllable closed by a nasal stop. This dichotomy, whose tonological significance will become transparent immediately below, is quite robust not only in Chinese, but across other tone languages of southeast Asia, notably Kam-Tai and Miao-Yao, where the two syllable types are known by more colorful and expressive terms such as ***staccato*** vs. ***legato***, or ***dead*** vs. ***live*** syllables (cf. Gandour 1974, M. Hashimoto 1984, Thongkum 1987, Thurgood 1992). For short, I will sometimes use CVq and CVN to symbolize these two types of syllables.

Four tonal categories, referred to by their traditional nomenclature as *ping*, *shang*, *qu*, and *ru*, have been firmly established since Middle Chinese (hereafter MC; approximately from AD 200 to 900), as reflected in the pronouncing dictionary *Qieyun* (AD 601), a landmark in the history of Chinese phonology.

(2) Middle Chinese tone categories

	traditional name	gloss
I	*ping*	"level"
II	*shang*	"rising"
III	*qu*	"departing"
IV	*ru*	"entering"

Tone IV occurs exclusively with checked syllables, while tones I, II, and III are associated with smooth syllables. This cross-classification of tones

[3] Sometimes both. Thus, the northern Min dialect of Fuzhou has in its syllable inventory words like [souŋ] "sour," [keiq] "orange," etc. (-q = glottal stop).

and syllable types is motivated by the observation that checked syllables tend to have an impoverished tonal inventory, and exhibit markedly different sandhi behavior compared to smooth syllables, as will become amply evident in the ensuing chapters. For this reason, Chinese linguists often talk about "**smooth tones**" (*shu sheng* = tone I, II, III) and "**checked tones**" (*ru sheng* = tone IV). Throughout this book, I will suffix the symbol -q to the tone letters representing a "checked tone," while leaving the "smooth tones" unmarked. Thus, 55q, 13q etc. stand for a high level and low rising tone linked to a checked syllable CVq.

Based chiefly on the Japanese monk Annen's description in *Xi-tan zang* (or *Shittan zô*; written in AD 880), Mei (1970:109–110) reconstructs the following tonal values for MC (around 8th century):

(3) Middle Chinese tone values

	categories	reconstructed phonetic values
I	Level (*ping*)	long, level, and low (with two allotones)
II	Rising (*shang*)	short, level, and high
III	Departing (*qu*)	longish, probably high and rising
IV	Entering (*ru*)	short, with uncertain pitch and contour

Some of the descriptive terms for pitch height and contour are taken directly from Annen, who characterized tone I and II as "*zhi di*" (straight and low) and "*zhi ang*" (straight and high).[4] The hypothetical durational distinction is based primarily on the ancient buddhist practice of using tone II and III syllables to transcribe Sanskrit short and long vowels, respectively. As noted above, entering tone syllables end in an oral stop -p,t,k. Not surprisingly, Annen describes the entering tone as "*jing zhi*" (abruptly stop), a "checked" quality that is still readily observable in those modern dialects that have preserved the old p,t,k codas (often reduced to a glottal stop, hereafter symbolized as -q). The reconstruction of tone III is somewhat more speculative: it is inferred from the fact that tone II syllables with a voiced obstruent onset had merged with tone III.[5] Since merger presupposes a certain phonetic affinity, and since tone II is

[4] Pulleyblank (1978:178) interprets *zhi ang* as "straight rising" instead, citing as evidence another contemporaneous document *Yuanhe Yunpu* (806–827), in which tone II is described as "*li er ju*" (stern and rising), where *ju* (lit. to lift up) clearly denotes a rising pitch movement.

[5] This is clearly indicated in Annen's statement that tone II only occurs with *qing* "light" syllables, i.e. syllables with voiceless (and sonorant) initials.

known to be high, it stands to reason to assume that tone III also had a high pitch at the time the merger took place.[6]

Hirayama (1974, 1975) and Ting (1984) have made attempts at reconstructing the tonal values of Proto-Min and Proto-Wu, respectively.

Our thumbnail sketch of ancient Chinese tonology would not be complete without an aside on the hypothesis first put forward by Haudricourt (1954a, b, 1961), now generally referred to by the broader term "tonogenesis."[7] Haudricourt advanced the theory that the archaic Chinese tonal system arose through the loss of certain final consonants, in an evolution that parallels Vietnamese. Specifically, Haudricourt maintains that Archaic Chinese tone II, III and IV originated from CVq, CVs and CVk, respectively (-q represents a glottal stop, -s is a sibilant, and -k stands for any of the full oral stops -p,t,k). Crucially, Pulleyblank (1978) extends this hypothesis down to the more recent historical period of Middle Chinese. Specifically, he claims that the so-called "tones" actually corresponded to different syllable types prevailing in Late Middle Chinese (8th century), which still retained the old consonantal desinences. The Haudricourt–Pulleyblank hypothesis has found both supporters (Mei 1970, Sagart 1986) and skeptics (Ting 1981, 1996, Ballard 1985, 1988).

2.2 *Tone split*

The four Middle Chinese tones have undergone various splits and mergers. Tone split is sensitive to various phonological conditions, most notably the voicing contrast in the syllable onset,[8] as illustrated by the northern Wu dialect of **Songjiang** which, like all other Wu dialects, still maintains the voiced/voiceless contrast in the onset. Each of the MC tonal categories is split neatly into a high and a low register – known in traditional terminology as *yin* and *yang* – yielding a perfectly symmetrical eight-tone system (data from *Jiangsusheng he Shanghaishi Fangyan Gaikuang* 1960). In each

[6] For comparison, here is Ting's (1996:152) reconstruction of the MC tonal values, based on *Xi-tan zang* and other evidence (including comparative):

I	Level (ping):	level, probably low
II	Rising (shang):	high-rising
III	Departing (qu):	falling, probably mid-falling
IV	Entering (ru):	abrupt and short

Ting rejects any length contrast among tones I, II, and III.

[7] Coined by Matisoff (1970, 1973); see Hombert (1975, 1978), Hombert, Ewan, and Ohala (1979), Mazaudon (1977), and references cited therein. Utsat, a Chamic Austronesian language spoken on the Hainan island, instantiates a particularly transparent case of transition from an atonal to a fullblown tonal system. See Thurgood (1992) for details.

[8] Yip argues at length that what is crucial in conditioning tone split is not [voicing] but [murmur]. For details, I refer the reader to Yip (1980:211–242, and 1993b:249–254).

case, the **yang** register with a voiced onset has a lower pitch value than the corresponding **yin** register. Songjiang is fairly typical of Wu dialects.

(4) Songjiang register split

register		tone			
		I	II	III	IV
a.	high (*yin*)	53	44	35	5q
b.	low (*yang*)	31	22	13	3q

5q, 3q indicate checked tones

The pitch values are indicated by the familiar tone digits first introduced by Y-R. Chao (1930). The tonal space is idealized as a five-point vertical scale, where 5 and 1 represent the highest and the lowest pitch respectively.[9] Thus 53 and 13 stand for a high-falling and a low-rising tone respectively. Examples are given below:

(5)	I	a.	*ti*	53	"low"
		b.	*di*	31	"lift'
	II	a.	*ti*	44	"bottom"
		b.	*di*	22	"younger brother"
	III	a.	*ti*	35	"emperor"
		b.	*di*	13	"field"
	IV	a.	*paq*	5q	"hundred"
		b.	*baq*	3q	"white"

Needless to say, the voice-sensitive split into two registers is not always as neat or symmetrical. Take **Beijing Mandarin**. The correspondence between the MC tonal categories and their modern phonetic values is summarized in the following table (based on Chen 1976:152).

(6) Beijing Mandarin

MC onset	MC tones			
	I	II	III	IV
voiceless	55	213	51	55, 35, 213, 51
sonorant	35			51
voiced obstruent		51		35

The leftmost column indicates the three types of Middle Chinese onsets: voiced and voiceless obstruents, and sonorants (including liquids, nasals,

[9] Exactly the opposite of the convention that prevails in African and Amerindian tonological literature.

and ∅-initials). Notice that tone I splits along the familiar [±voiced] division. Tone III remains a single cohesive category. Tone II also bifurcates along the voicing line, but in this case sonorants side with the voiceless rather than the voiced obstruents. Furthermore, the voiced obstruent onset syllables that split off from tone II have merged with tone III syllables. Finally, tone IV words (originally associated with checked syllables) are redistributed among the other tonal categories, conditioned by the three-way contrast between voiceless, sonorant, and voiced obstruent initials.[10] Note that two sweeping historical changes have occurred in Beijing Mandarin: all voiced obstruents have become voiceless, and all checked syllables (CVq) have lost their stop endings entirely. This means that both voicing and smooth vs. checked syllable contrasts are now recoded in purely tonal terms. In short, the evolution from MC to the tonal system of Beijing as we know it today entails the following historical processes:

(7) a. Register split of tone I
b. Tone IIb merges with tone III
c. Redistribution of tone IV among other tonal categories
d. Devoicing and, in some cases, aspiration of voiced obstruents
e. Loss of obstruent codas

Needless to say, (a, b, c) must precede (d) since the former are voice-sensitive, a distinction that is neutralized by the latter. Furthermore, (c) must pre-date (e) since the defining characteristic of tone IV is CVq, with an oral stop coda (symbolized by -q), which has dropped out via (e). Some examples follow:

(8)

	MC	Standard Mandarin		
I	*tang*	*tang*	55	"ought to"
	lang	*lang*	35	"wolf"
	dang	*t'ang*	35	"sugar"
II	*tang*	*tang*	214	"party"
	lang	*lang*	214	"bright"
	dang	*tang*	51	"to swing, sway"
III	*tang*	*tang*	51	"to pawn"
	lang	*lang*	51	"wave"
	dang	*tang*	51	"to procrastinate"
IV	*t'ak*	*t'uo*	55	"to entrust"
	lak	*luo*	51	"to fall"
	dak	*tuo*	35	"to stroll, pace"

[10] Tone IV words with a voiceless initial are scattered unpredictably among all four tonal categories in modern Beijing.

Linguists have long noted the pitch-depressing effect of voiced obstruents, and sought to explain the cross-linguistic patterns in physiological terms.[11] For a general discussion and critical review see Hombert (1978), Hombert, Ohala, and Ewan (1979), and references cited there. For our purposes, it should be noted that subsequent historical changes – in particular, devoicing – may intersect and obscure the phonetically motivated partition of tonal categories into a high and a low register. It is not uncommon for the *yang* or b-register to show a high tone in a modern dialect instead of the expected low register, in a process sometimes referred to as "register flip-flop." A. Hashimoto (1986) sampled 997 dialects, and found 340 cases of register reversal. For this reason, we will often simply refer to register a (*yin*) and b (*yang*), to dissociate the relatively constant tonal categories from their often unpredictable phonetic values.

Somewhat less well known, but nonetheless quite common among Sino-Tibetan languages, is tone split along the dividing line between plain and aspirated onsets. Ye (1983) reports a three-way split of MC tones resulting in a perfectly symmetrical twelve-tone pattern in the Songling variety of **Wujiang**, also a northern Wu dialect.

(9) Wujiang three-way tone split

		I	II	III	IV
voiceless	plain	55	51	412	5q
	aspirated	33	42	312	3q
voiced[12]		13	31	212	2q

Since Wujiang has retained voicing and aspiration, the multiple splits merely produce allotonic variations rather than giving rise to new tonal categories. Examples illustrating the allotonic distribution within the four MC tonal categories follow:

(10)	I	55	*tiɩ*	"fall, topple"
		33	*tʻiɩ*	"day, sky"
		13	*diəu*	"head"
	II	51	*tø*	"short"
		42	*tʻi*	"body"
		31	*dɛ*	"light, insipid"

[11] Related is the blocking effect of voiced consonants on H-spread. Conversely, voiceless consonants tend to block L-spread (see Hyman and Schuh 1974).

[12] Including sonorants and ∅ or vocalic onset.

III	412	*tɛ*	"toward"
	312	*tʼɛ*	"to withdraw"
	212	*dəu*	"big"
IV	5q	*tiəq*	"drop"
	3q	*tʼiəq*	"iron"
	2q	*doq*	"to read"

Aspiration-triggered tone split has been widely attested both within the Sinitic family and beyond (cf. Ho 1989). Shi (1994) cites no less than 111 languages within the Sino-Tibetan phylum that instantiate such a tonal development. In addition to 23 and 22 instances in the Gan and Wu groups of Chinese, there are 66 other cases: 44 in Dong (belonging to the Kam-Tai branch), and 22 in the various Miao-Yao languages, both considered by some to be part of the Sino-Tibetan family. A plausible phonetic explanation is offered in Hombert et al. (1979) and Shi (1994).

Tone splits along other phonological parameters have been reported in the literature: breathy voice, prenasalization, fortis vs. lenis consonants, vowel height, length and tensity, etc. (Hombert 1978, Endô 1994). Surprisingly, despite the well-known intrinsic pitch variations associated with vowel height,[13] tone split along the high/low vowel distinction is so rare that Hombert et al. (1979:52) state flatly: "It would seem that the interaction between tones and vowel height works in only one direction: tone can affect vowel height, but not vice-versa."[14]

2.3 Tone mergers

MC tones have undergone two sweeping mergers. First, tone IIb has splintered off from IIa, and falls together with tone III across all dialect groups, suggesting an early onset of this sound change. The exact membership of IIb is defined somewhat differently from dialect to dialect, owing to the "amphibious" nature of the sonorants. Specifically, the sonorants behave sometimes as voiced obstruents, sometimes as voiceless (more

[13] Attested in tonal (Itsekiri, Yoruba, Ewe) as well as non-tonal languages (English, Danish, German, Japanese, Korean, French, Serbo-Croatian, Hungarian); see Beckman (1986:129) for sources and references. Similar data on Chinese can be found in Wu and Cao (1979).

[14] Fuzhou instantiates a classic case of tone-on-vowel influence: a low-to-high tone change in sandhi contexts induces a concomitant vowel change:

ei	→	i
ou		u
øy		y
ɔi		øy
ai		ei
au		ou

precisely, as unmarked for voicing). This somewhat complicated picture is summed up in the following diagram, where p, m, and b stand for the three classes of initial consonants: voiceless obstruents, sonorants, and voiced obstruents:[15]

(11)

p	m	b	
IIa	IIb	IIb	Suzhou, Changsha, Guangzhou, Fuzhou
IIa	IIa	IIb	Nanchang, Shuangfeng, Xiamen, all northern dialects

However defined, tone IIb has merged with tone III in most dialects – with the notable exception of Wenzhou (Wu), Guangzhou (Yue) and Chaozhou (Min), which have kept IIb apart as a distinct tonal category.

The second major merger stemmed from the weakening and loss of MC -p,t,k endings, marked by the entering tone IV. We can distinguish four stages of this development that must have taken centuries to run its course.

a. The original state of affairs is still visible in many Yue and southern Min dialects, which have preserved a full-fledged series of obstruent codas.
b. The occlusive endings -p,t,k have weakened into an undifferentiated glottal stop -q in most Wu dialects.
c. In yet others, this -q has dropped out altogether. However, the original CVq syllables have stuck together as one cohesive tonal category. This is the case with Changsha (of the Xiang group), which no longer has CVq syllables, but has maintained a separate class identifiable by means of a distinctive tone contour [24].
d. Finally, the stop endings have disappeared without a trace, segmentally or otherwise.

For most purposes, the only relevant distinction is between checked and smooth tones, regardless of the degree to which the original occlusive endings have been preserved or reduced. For this reason, in the rest of this book I will distinguish only between T and Tq (where T stands for

[15] This table is derived from the computerized dialectological corpus DOC [Dictionary on Computer], based on the first edition of *Hanyu Fangyin Zihui* (Beijing University, 1962), which did not include Yangjiang and Jian'ou. For a recent description of DOC, see Cheng (1994a).

For a plausible historical account of the "amphibious" behavior of sonorants with respect to tone splits and mergers, see Yip (1980:240f.).

any tone). Thus 5q means that the high level tone corresponds to an MC tone IV, originally associated with the CVq syllable type.

What is striking about the evolution of MC tones is the stability of categorical membership in the face of phonetic diversity. In other words, tone A in dialect X corresponds with remarkable regularity with tone B in dialect Y, regardless of the phonetic shapes of tones A and B. Take the modern reflexes of tone Ib. It has seven distinct phonetic shapes {35 : 42 : 24 : 213 : 31 : 55 : 34} in the seven Mandarin dialects represented in *Hanyu Fangyin Zihui* (second edition, Beijing 1989). The inter-dialectal correspondence is as categorically systematic as it is phonetically heterogeneous. How a single subcategory tone Ib evolved into such a wildly disparate set of modern reflexes in closely related languages is still poorly understood. For an overview of the complex diachronic developments of MC tones across the full spectrum of Chinese dialects, see the large scale study conducted by Chang (1975), based on all the sources available at the time.

3 Tone patterns in present day dialects

The historical processes of splits and mergers have given rise to a wide variety of tonal systems in the modern dialects of Chinese, ranging from three to ten tonal categories, according to Cheng (1973b:96). Synchronic tone patterns are traditionally described by specifying two parameters, namely pitch *height* (high, low, mid, etc.) and tone *shape* (even, rising, falling, falling-rising, or rising-falling).[16] Take the **Songjiang** case alluded to in section 2.2 for illustration (data from *Jiangsusheng he Shanghaishi Fangyan Gaikuang* [Nanjing, 1960], p. 11).

(12) Songjiang tone system:

σ type	onset	even	rising	falling
CVN	voiceless	44	35	53
	voiced	22	13	31
CVq	voiceless	5q		
	voiced	3q		

"Voiced" includes both voiced obstruents and sonorants

[16] Falling-rising and rising-falling are also known as "concave" and "convex" tones respectively after Wang (1967).

It has eight phonetic tone shapes. However, since Songjinang, like most other Wu dialects, has preserved the voiced/voiceless onset as well as the CVq/CVN contrast, the eight tones can be reduced to three contrastive categories, differentiated only in terms of tone shape (level, rising, falling), with predictable pitch height. Thus, the level tone has four allotones {44, 22, 5q, 3q} in phonetically definable complementary distribution. However, the prevailing practice is to treat the checked tones as if they constituted distinctive categories apart from their smooth counterparts, not without some justification, in view of the peculiar sandhi behavior and distributional restrictions of checked tones. Note in passing that Songjiang is also typical in restricting contour tones {35, 13, 53, 31} to smooth (CVN) syllables.

By a similar process we can reduce the 12 tones of Wujiang (see section 2.3) neatly into a three-tone system, by exploiting the redundancies implicit in the two-way contrast between CVN and CVq on the one hand, and the three-way opposition between voiced, plain voiceless, and voiceless aspirated onsets, on the other.

However, the picture is somewhat less symmetrical and clear-cut in some cases when we look more closely. **Shanghai**, a related but far better known Wu dialect, serves to illustrate the point. This dialect has five citation tones {53, 34, 23, 5q, 12q}, corresponding to Middle Chinese categories Ia, IIIa, b and IVa, b. Their distributions among syllables with the three classes of initial consonants are summarized in the following table, where "+" and "–" indicate the cooccurrence or its absence, of the relevant classes of tones and initials.

(13) Shanghai tone system

onsets	Ia	IIIa	IIIb	IVa	IVb
	53	34	23	5q	12q
voiceless[17]	+	+	–	+	–
sonorant	+	(+)	+	(+)	+
voiced obstruent	–	–	+	–	+

There is no question that the two high register tones [53] (MC tone Ia) and [34] (tone IIIa) stand in contrast, as exemplified by [tɔ 53] "knife" vs. [tɔ 34] "island" and countless other minimal pairs. On the other

[17] Including ∅-initial.

hand, it is possible to collapse the checked tones [12q] and [5q] with [23] and [34] respectively.[18] The question concerns whether [34] contrasts with [23] and [5q] with [12q]. A check of the syllabary in Xu et al. (1988:16–23, 78–100) yields an ambiguous answer: the high register tones [34, 5q] in principle occur only with voiceless initials, while the low register tones [23, 12q] only go with voiced onsets.[19] However, there are sporadic exceptions forming a few minimal doublets and triplets as exemplified below:

(14)	a.	*nu*	34	IIIa	"diligent"
			23	IIIb	"sweet rice"
	b.	*wAq*	5q	IVa	"to dig"
			12q	IVb	"slippery"
	c.	*mE*	53	Ia	"every"
			34	IIIa	"beautiful"
			23	IIIb	"slow"

The parenthesized (+) in (13) (shaded cells) is intended to signal the sporadic and exceptional nature of the cooccurrence of IIIa and IVa with voiced initials (sonorants). We may discount a handful of minimal pairs like those cited above, or rephonemicize the minimal pairs like "diligent" and "sweet rice" as /*nu* LH/ and /*nhu* LH/ respectively, where h stands for murmur or breathy voice. Likewise, "to dig" and "slippery" are given the lexical representations /*waq* LH/ and /*whaq* LH/. The phonetic values of LH are then determined by the segmental composition of the tone-bearing syllables. This is the practice adopted, for instance, in Sherard (1972), and followed by Yip (1980, 1993b), Duanmu (1990a), and Jin (1995). Either way, we could in principle compress Shanghai into a two-tone system, where the rising tone has four allotones, predictable (by and large) on the basis of voicing and syllable type:

[18] In principle, [5q] could go with the high falling [53] as well. However, most analysts group [5q] with the high rising [34] instead because they display a similar behavior in sandhi contexts.

[19] Which is actually a convenient oversimplification. It has been well known since Chao (1928, 1935) that voiced stops are voiced only in word-medial positions; in initial positions they are actually voiceless (by the standard test of voice onset time; cf. Lisker and Abramson 1967), contrasting with their voiceless counterparts only in terms of the closure time, longer for voiceless and shorter for voiced stops. For a critical survey of previous studies and original experimental data concerning voicing in Wu dialects, see Shen and Wang 1995.

(15)

		rising /LH/	falling /HL/
CVN	voiceless	34	53
	voiced	23	
CVq	voiceless	5q	
	voiced	12q	

Primary sources and secondary literature are not consistently explicit on the phonemic status of the various phonetic tones. Fortunately, it is not often the case that analysis crucially hinges on the phonemic status of the tones in question. In most cases, it suffices to take the citation tones (i.e. tones that appear with monosyllabic root morphemes pronounced in isolation) on their face value, and note their alternations in various sandhi contexts.

Cantonese (spoken in Hong Kong, Guangzhou, and by many overseas Chinese) has one of the more complex tonal systems, characteristic of the Yue group to which it belongs. The nine tones of Cantonese are customarily cross-classified as follows (based on Hashimoto 1972:92 and *Hanyu Fangyin Zihui*, 2nd ed. 1989):[20]

(16) Cantonese tone system

	level		rising	falling
	CVN	CVq		
high (*yin*)	33	5q; 3q	35	53 (~ 55)
low (*yang*)	22	2q	23	21

The high falling tone [53] has [55] as a free (stylistic) variant. Ancient voiced obstruents have undergone devoicing in Cantonese. As a consequence, *yin* and *yang* registers, originally derived from the voiced/voiceless dichotomy, now contrast exclusively in terms of pitch height. On the other hand, the stop endings (-p,t,k) are preserved intact. This means that [5q, 3q] and [2q] are in complementary distribution with [33] and [22]. Nevertheless, as noted before, the checked tones [5q, 3q, 2q] are treated as separate tonal categories in accordance with the descriptive conventions adopted by most dialectologists in China. Finally, the bifurcation of the *yin* register

[20] I adopt the tone values of *Hanyu Fangyin Zihui* in the interest of cross-dialect comparability.

checked tone into [5q] and [3q], attested in many other Yue dialects, is conditioned by the opposition between what A. Hashimoto (1972:176f.) refers to as lax (short) and tense (long) vowels, for instance [ɪ] vs. [ɛ:] and [a] vs. [a:].[21] The nine tones of Cantonese are illustrated by the following three sets of monosyllabic root morphemes:

(17) Cantonese

MC categories	tone values						
Ia	53 (~ 55)	si	"poetry"	fu	"husband"	wan	"warm"
Ib	21	si	"time"	fu	"to hold"	wan	"cloud"
IIa	35	si	"to send"	fu	"bitter"	wan	"to look for"
IIb	23	si	"market"	fu	"wife"	wan	"to allow"
IIIa	33	si	"to try"	fu	"rich"	wan	"to shut up"
IIIb	22	si	"affairs"	fu	"father"	wan	"to transport"
IVa-i	5q	sɪk	"to know"	fat	"sudden"	wat	"twisted"
IVa-ii	3q	sɛ:k	"lead"	fa:t	"law"	wa:t	"to dig"
IVb	2q	sɪk	"to eat"	fat	"to punish"	wat	"pit" (of fruit)

Bobai, also a Yue dialect (spoken in Guangxi) and one of the earliest tone languages reported in the classic work of Wang Li (1932), is often cited as the Chinese dialect endowed with the richest tonal repertoire, boasting ten tonal categories. Although Bobai has phonetically quite different tone shapes from Cantonese, in terms of tonal categories it differs from Cantonese only in that tone IVb is further split into two subcategories.[22]

Interpretation of phonetic data

From a broader perspective, Cheng (1973b) surveyed 737 contemporary Chinese dialects, and found 69 different phonetic shapes. The phonetic transcriptions call for judicious interpretation. With rare exceptions, phonetic transcriptions are based on aural judgment, and vary according to different practices and implicit assumptions on the part of the fieldworkers. Furthermore, the five-point pitch scale specifies a far greater number of tone shapes than one would ever need to describe any one language, as a consequence, forcing arbitrary choices upon the fieldworker in many cases. Take the tone shape [54] (attested in 57 dialects). One cannot tell a

[21] Hashimoto's (1972) inverted [a] : [ʌ:] correspond to Yuan's (1960) [a] : [a:]. For typographical simplicity, I adopt Yuan's transcription.

[22] The bipartition of tone IVb is also attested in Yangjiang, another Yue dialect.

priori whether it is basically a high level tone [55] with a slight declination effect, or a variant of [53], or for that matter, [454], and so forth. By the same token, if a dialect has only one rising tone, whether one transcribes it as [24], [34], [35] etc. depends as much on personal preferences and (implicit) theoretical assumptions as on the objective phonetic reality. For instance, if one partitions the pitch range into high (4–5) and low (1–2) registers – interpreting 3 as either high or low as the case may be – and if one disallows, for theoretical reasons, cross-register contour tones, one would reject [24] in favor of [34] or [35]. The analyst has to make judgment calls of this sort on a case-by-case basis.

One concrete example will suffice to illustrate the need for a judicious interpretation of the raw phonetic data. According to Norman (1973), **Jianyang** has the following citation tones:

(18) Jianyang citation tones

even	rising	falling
33	35	53
		43
		32
		31
		21

On the face of it, we have an oddly skewed system, with one even, one rising, but five falling tones. It not only displays a lopsided preponderance of contour tones, but is highly marked, if not impossible given a classificatory scheme like Yip's (1980, 1989) or Bao's (1990a), which allows only a maximum of four level tones, two rising and two falling tones. I agree with Yip (1980:206) who suggests that in all likelihood, three of the five falling tones, namely [43, 32, 21], ought to be considered basically level tones [44, 22, 11] with a slight downdrift or final fall in pitch, a predictable phonetic effect noted in the tonological and intonational literature.[23] On this view, the Jianyang system should be revised as follows:

[23] On the phenomena of "declination" and boundary low cf. Pierrehumbert (1980), Liberman and Pierrehumbert (1984), and, with particular reference to Chinese, Tseng (1981) and Shih (1988, 1991). Tseng (1981:143) notes, for instance, that the terminal fall is a special case of breath-group intonation.

(19) Jianyang tone system

even	rising	falling
44	35	53
33		
22		31
11		

In interpreting the raw phonetic data, we need to keep the overall sound pattern in mind and, where appropriate, follow the eminently sensible heuristic principles proposed by Maddieson (1978b:45f.) in discounting non-distinctive phonetic details. One common non-distinctive feature is the final lowering already noted above in relation to Jianyang. To cite another example, one may judiciously disregard certain finer details resulting from the undershooting of phonological targets. Take the two pitch curves in **Standard Thai** which we annotate as [214] and [451] in Chao's tone digits. Maddieson interprets them as basically rising [14] and falling [51] tones in Standard Thai, with two and not three tonal targets each; in his view, the initial portions of the pitch curves transcribed as [2] and [4] represent nothing more than imprecise approximations toward the intended targets [1] and [5].

4 Tones in context

Rich and highly developed as tonal systems have become in Chinese, they are surpassed in many instances by even more complex and intricate sandhi processes, which often drastically alter the phonetic shape of adjacent tones, when they come into contact with each other in connected speech. This tonal alternation in connected speech is what has been referred to as **tone sandhi**, and constitutes the subject matter of our inquiry.

Before we delve into the specifics of tone sandhi, it is useful to place tone sandhi in the context of the various types of tonal modifications. Tone shapes may be subject to the influence not only of the neighboring tones, but also of the overall intonation. Furthermore, morphotonemic alternations may be triggered not only by strictly phonological environments, but also by morphological contexts.

4.1 Tone sandhi

Beijing Mandarin exemplifies tone sandhi in its classical, canonical form. Recall that Beijing has a four-tone system illustrated by the oft-cited standard quadruplet:

(20)	T1	*ma*	[55]	"mother"
	T2	*ma*	[35]	"hemp"
	T3	*ma*	[214]	"horse"
	T4	*ma*	[51]	"to scold"

The rule in question turns a "dipping" i.e. falling-rising tone [214] into high rising [35] when followed by another [214]. Since [214] is traditionally labeled Tone 3, this processes is known as third tone sandhi, or **T3 Sandhi** for short. T3 Sandhi is illustrated by the following examples. (For clarity, a dot separates two tones represented by two to three tone digits.)

(21)	a.	*xiao*	"small"
		214	
		xiao gou	"small dog, puppy"
		214. 214	base form
		35. 214	sandhi form
	b.	*mai*	"buy"
		214	
		mai ma	"to buy a horse"
		214. 214	base form
		35. 214	sandhi form
	c.	*mai*	"to bury"
		35	
		mai ma	"to bury a horse"
		35. 214	base form = sandhi form

Attention has been drawn to this phenomenon since the earliest descriptions of Chinese, presumably because this drastic change in tone shape is both perceptually salient and functionally significant – for instance, as a consequence of this systematic morphotonemic alternation, the expression "to buy a horse" becomes indistinguishable from "to bury a horse" (cf. Wang-Li 1967). But T3 Sandhi is but one of the many ways the idealized tone shape may be modified under the influence of neighboring tones and the overarching intonation pattern. Based largely on aural discrimination, Y-R. Chao (1968:26–29) mentions three others. First, [214] is obligatorily reduced to a simple low fall [21] everywhere except in utterance-final positions, where it may, but does not necessarily, show up

in its full form, with an upglide at the end. This means that, in conjunction with T3 Sandhi, tone 3 has three allotones, distributed as follows:

(22) Allotones of T3

		allotones 214	21	35
Sandhi contexts	before [214]	–	–	+
	utterance-final	+	+	–
	elsewhere	–	+	–

Here plus "+" and minus "–" indicate the occurrence or systematic absence of an allotone in a given context. Note that both the full [214] and the reduced [21] forms of T3 may occur in the utterance-final position; hence the notation "21(4)" in the examples that follow. Take the three occurrences of *xiao* "small," highlighted in bold face below.

(23) a. *dan* ***xiao*** "coward" (lit. gall + small)
214. 214 base form
35. 21(4) sandhi form

b. ***xiao*** *gou* "small dog, puppy"
214. 214 base form
35. 21(4) sandhi form

c. ***xiao*** *mao* "small cat, kitten"
214. 55 base form
21. 55 sandhi form

In utterance-final position, it shows up with the underlying [214], optionally reduced to [21] (case a); before another [214] it obligatorily turns into [35] (case b); elsewhere it may assume only the truncated form [21], also referred to as "half-third tone" (case c). Correspondingly, the rule responsible for this alternation is known as the **Half T3 Sandhi**.

Second, a rising tone [35] becomes a high level [55] when preceded by [55, 35] and followed by any tone. We shall have more to say about this rule (see chapter 7, section 2.3), generally referred to as **T2 Sandhi**, since [35] is the second tone in Standard Mandarin. A small handful of examples will suffice for our immediate purpose.

(24) a. *tian wen tai* "observatory" (lit. astronomy + platform)
55. 35. 35 base form
55. 55. 35 sandhi form

b. *ren min bi* "renminbi" (Chinese currency)
35. 35. 51 base form
35. 55. 51 sandhi form

There is a third rule, whereby neutral-toned syllables assume a pitch level that is determined by the preceding full tone. Specifically: neutral tone is predictably high [4] after [214], mid [3] after [35] and [55], and low [1] after [51]. This neutral tone rule – or **NT Sandhi** – is illustrated by the allotones of the inherently toneless enclitic -*de*, which functions, among other things, as the nominalizer:

(25) a. *zi de* "the purple one"
214. 4

b. *hong de* "the red one"
35. 3

c. *xin de* "the new one"
55. 3

d. *da de* "the big one"
51. 1

Finally, apart from the above phonologically determined sandhi processes, certain sandhi alternations are morphologically conditioned or lexically marked. One such alternation is referred to in Chang (1992:166f.) as the *Yi-bu-qi-ba* rule, because it applies specifically to a closed list of the four lexical items *yi* "one," *bu* "not," *qi* "seven" and *ba* "eight." This rule obligatorily changes the tone of these syllables into [35] when followed by another falling [53] tone syllable. As a consequence, we have a minimal pair such as this:[24]

(26) a. *bu dui* "not correct"
53.53 base tone
35.53 *Yi-bu-qi-ba* rule
cf. *bu hao* "not good"
53.214

b. *bu dui* "troops"
53.53 no change

Bu "not" has the lexical tone /53/ as demonstrated by *bu.hao* [53.214] "not good." It takes on a rising tone [35] in (a) via the *Yi-bu-qi-ba* rule, which does not extend to case (b) where the homophonous *bu* represents a different morpheme altogether, with the meaning "division" (of an army).

[24] I owe this minimal pair to Kratochvil (1987:256).

Table 1.2. *Tonal coarticulation in Beijing*

fu-		**-ji**							
		T1	259–258	T2	209–262	T3	211–153	T4	291–162
T1	266–266	258–270 #	269–276	274–291 #	209–235	273–290 #	219–144	268–286 #	296–176
T2	211–253	219–**245** #	266–273	216–257 #	217–249	223–281 #	**247**–144	216–**243** #	281–177
T3	214–154	213–176 #	263–272	223–168 #	**188**–258	(= T2 + T3)		225–178 #	284–175
T4	287–159	299–**218** #	262–266	300–**227** #	207–252	310–**238** #	201–141	300–**227** #	278–174

Hyphenated numbers indicate the pitch value (in Hz) at the two endpoints of a tone shape
separates two adjacent tones
Bold letters mark coarticulation effects highlighted in Shih (1988)

4.2 *Tonal coarticulation*

In the brief sketch given above, Beijing exemplifies what has been thought of as tone sandhi in its canonical form: contextually determined tonal alternation in its simplest and purest form. However, when one examines the actual pitch curves of connected speech with the aid of laboratory instruments, one soon realizes that the tone sandhi rules formulated above produce at best an intermediate level of linguistic abstraction, which stands at a considerable distance from the pitch envelope in actual phonetic records. The full specification of pitch curve in natural, fluent speech entails at least two other types of processes. First, **tonal coarticulation**. I will limit myself to only those aspects of tonal coarticulation highlighted in Shih (1987, 1988, 1991).[25] To factor out segmental influence on tones, she constructed disyllabic expressions [*fu* + *ji*], that encode meaning contrasts exclusively in tonal terms. Thus [*fu.ji* T1.T1] means "to hatch chicks," while [*fu.ji* T1.T2] signifies "husband's residence," and so forth. Given the four-tone system of Beijing Mandarin, there are 16 such disyllabic sequences. The acoustic measurements of the pitch values of these disyllabic strings are summarized in table 1.2.

The four rows of the leftmost column correspond to the four tones of the first syllable [*fu*], pronounced in isolation; the two hyphenated numbers represent the average pitch value (in Hz) at the two end points of each tone shape. Likewise, the isolation tone values of the second syllable [*ji*] are given across the top. The phonetic shapes of the two tones in connected speech are specified in the cells where the relevant rows and columns intersect. "#" separates the two adjacent tones. Recall that

[25] Other recent experimental works on tonal coarticulation in Beijing Mandarin include Wu (1982, 1985), X. Shen (1990a, b, 1992).

Table 1.3. *Tonal coarticulation in Beijing*

fu-		**-ji**							
		T1	HH	T2	MH	T3	ML	T4	HL
T1	HH	HH # HH		HH # MH		HH # ML		HH # HL	
T2	MH	M**H**$^{-}$ # HH		MH # MH		MH # **M**$^{+}$L		M**H**$^{-}$ # HL	
T3	ML	ML # HH		ML # **M**$^{-}$H		(= T2 + T3)		ML # HL	
T4	HL	H**M** **#** HH		H**M** # MH		H**M** # ML		H**M** # HL	

T3 + T3 → T2 + T3 via the T3 Sandhi rule, hence the shaded cell. Shih (1988) used bold numbers to highlight some of the most salient and consistent tonal coarticulation effects. To bring out more clearly the essential points, I have translated table 1.2 into table 1.3, where gradient pitch values are reduced to three idealized discrete targets H(igh), M(id), and L(ow), with superscript + and – representing up- or down-shift. Thus HH, ML, and MH^{-} stand, respectively, for a high level, a low-falling, and a rising tone that undershoots its H target.

Shih singled out the following points for special comment. (i) HL falls to the baseline of the pitch range only in the final position; elsewhere it falls only to the mid-range. (ii) The low-falling ML assimilates to the preceding rising MH. (iii) The low ML depresses and delays the rise of a following tone. (iv) The rising MH does not quite reach its H target when it is following by a Hx, which abbreviates {HH and HL}, namely a tone shape with H as its first element. These tonal coarticulation rules can be stated informally as follows:

(27) i. T4 + any T — $\underline{\text{HL}}.\text{T} \rightarrow \text{HM}$

ii. T2 + T3 — $\text{MH}.\underline{\text{ML}} \rightarrow \text{M}^{+}\text{L}$

iii. T3 + T2 — $\text{ML}.\underline{\text{MH}} \rightarrow \text{M}^{-}\text{H}$

iv. T2 + {T1, T4} — $\underline{\text{MH}}.\text{Hx} \rightarrow \text{MH}^{-}$

(For clarity and simplicity, I rotate the rewrite arrow "→" by 90 degrees; the target of tonal modification is underlined. Accordingly, rule (i) is equivalent to: HL → HM / ___ T.) As we can see immediately, (i–iii) are all assimilatory in nature, and all have the effect of smoothing the transition between tonal targets. The sharp fall HL is reduced to HM by (i), since whatever tone follows, its initial target is either H or M. Rules (ii) and (iii) raise or lower the M portion of the second tone, depending on whether the first tone ends in H or L. Rule (iv) may appear to be dissimilatory in nature, but lends itself to an alternative interpretation, as suggested by Shih (1988:6f.). What is going on in (iv) is that the H target of T2 is deleted when followed by another H target. The phonetic H^{-} actually represents the transition tone between M and H. To put it differently in the more familiar format:

(28)	*fu.ji*	"medication taken orally"[26]
	MH.HL	base form
	M_. HL	Absorption
	MH^{-}. HL	Tone Interpolation

Tone Absorption, a process widely attested in African tone languages,[27] fuses two adjacent like tones. Tone Interpolation inserts a transitional pitch H^{-}, intermediate between M and H.

X. Shen (1992) proposes various diagnostics to distinguish tonal coarticulation from tone sandhi including: (a) only assimilation is considered coarticulation, but tone sandhi may be both assimilatory and dissimilatory; (b) tonal coarticulation obeys only language-independent biomechanical constraints, while tone sandhi may be subject to language-specific morphological and phonological conditions; (c) tone sandhi may effect tonemic change, while tonal coarticulation involves only allotonic variations.

It is not clear whether it is desirable or even possible to segregate tonal coarticulation from tone sandhi proper. For one thing, it is not clear that late phonetic coarticulation is exclusively assimilatory. Shih (1987:10) observes that in Standard Mandarin, the high rising [35] and falling [53] tones are phonetically *higher* when preceding a low [214] tone, apparently a dissimilatory coarticulation effect.

[26] As opposed to externally applied ointment or the like.
[27] For instance, Bamileke, Mende, Kikuyu, Hausa, Ngizim, among others, cited in Hyman and Schuh (1974).

As for Shen's second diagnostic, suffice it to say that certain late phonetic tonal perturbations may be grammatically controlled. To cite one example, Wu (1985:79f.) produces acoustic data to show that the pitch curve associated with a rising tone MH in medial position actually turns into a falling contour HM when it is followed by a low or mid tone, as illustrated by the following example

(29)	[*zhan lan*] *guan*	"exhibition hall"
	L L L	
	MH.MH.L	T3 Sandhi
	MH.HM.L	Tonal coarticulation

This fairly subtle contextual tonal perturbation has eluded most students of Chinese tonology, and is detectable only instrumentally, thanks to Wu (1982). Remarkably, such a "low-level" phonetic rule is sensitive to morphosyntactic structure: it seems to be excluded from right-branching constructions like

(30)	*he* [*leng shui*]	"drink cold water"
	H L L	base tone
	H MH.L	T3 Sandhi
	–	Tonal coarticulation

Furthermore, as we will see immediately below, what have been usually regarded as late phonetic intonational effects are also grammatically controlled.

Finally, tone sandhi literature typically covers allotonic as well as morphotonemic alternations. Take the prototypical **T3 Sandhi** in Beijing discussed above:

(31) Beijing T3 Sandhi

T3 → T2 / ___ T3

T2 = MH

T3 = L

The conventional view is that T3 merges with T2 in the specified sandhi context. Thus perceptual tests conducted by Wang-Li (1967) established that *qima* "at least" from underlying T3 + T3 becomes indistinguishable from *qi ma* "to ride a horse" from T2 + T3; likewise *fenchang* T2 + T3 "graveyard" and *fenchang* T3 + T3 "flour factory" are homophonous in actual speech. This established view has been challenged by Zee (1980) and Kratochvil (1987). Zee (p. 101) shows that underlyingly T3 preceding

another T3 has an overall lower Fo (fundamental frequency) than the corresponding T2 in the same environment. This would seem to suggest that T3 Sandhi does not entail a paradigmatic substitution of T3 by T2, but rather only turns T3 into some allotone, say T3′, in the sandhi context, and hence maintains the categorial distinction between T2 and T3′. Whatever difference may persist between T2 and T3′, it is not clear whether such difference is reliably perceived by a native speaker (Wang-Li 1967). In any event, T2 and T3′ are functionally equivalent with respect to other sandhi processes. Recall that Beijing has another sandhi rule, which I have referred to as T2 Sandhi. It turns a non-final [35] into [55], when preceded by [55] or [35]. What is important here is that this T2 Sandhi rule treats underlying [35] (T2) and derived [35] (T3′, via T3 Sandhi) alike. This is demonstrated below:

(32)	"observatory"	"watershed"	
	a. *tian wen tai*	b. fen shui ling	
	H.MH.MH	H. L. L	base tones
	– – –	H. MH.L	T3 Sandhi
	H. H. MH	H. H. L	T2 Sandhi

Note that both of the medial syllables *wen* (underlyingly /MH/ or T2) and *shui* (originally T3 or /L/) undergo T2 Sandhi.

I will therefore take the position that there is no essential difference between tone sandhi and tonal coarticulation, except that tone sandhi processes are perceptible to the (trained but) unaided ears,[28] and therefore more likely to be reported by fieldworkers and integrated to a greater extent into the phonological component of the grammar. In practical terms, the sandhi phenomena we are about to investigate are circumscribed only by their accessibility: with the notable exception of Standard Mandarin and a small handful of better known dialects, published reports provide only more or less standardized phonetic transcriptions without

[28] Clements (1979:549) conjectures that the minimal perceptible interval is in the order of one semitone on the diatonic scale; his hypothesis is confirmed by experimental evidence: Gaoba Dong, a Kam-Tai language spoken in Guizhou, China, distinguishes five level tones (in addition to three rising and one falling tone), ranging from 129 to 259 Hz, covering roughly an octave (C_2 to C_3). Interestingly, the interval between the two lowest tones is 9 Hz, or about one semitone (see Shi et al. 1987).

In terms of discrimination between tone shapes, Dreher and Lee (1968) report that a minimum pitch change (upwards or downwards) of 2 semitones is necessary for Mandarin speakers to distinguish a rising or falling tone from a level tone. Comparable figures are 2.5 semitones for French speakers (Rossi 1971) and between 1.5 and 3 semitones in the case of Dutch subjects ('t Hart 1974). The just noticeable difference can be considerably smaller under controlled conditions, cf. Klatt (1974), Gandour (1978).

supporting acoustic data.[29] As a consequence, for all practical purposes, the "low-level" phonetic coarticulation effects of the sort illustrated above are beyond the scope of our investigation.

4.3 Intonational effects: declination and catathesis

Apart from localized coarticulation effects, tone also interacts more globally with the overall intonation pattern of the utterance. Despite its intuitive and poetic appeal, Y-R. Chao's (1968:39) metaphor of the intonational overlay upon lexical tones as "small ripples riding on large waves" is too impressionistic to make testable predictions. Since then, however, acoustic studies on tone and intonation in Chinese, basically limited to Standard Mandarin, have multiplied (Tseng 1981, Wu 1982, 1985, Gårding 1984, 1987, J. Shen 1985, Shih 1987, 1988, 1991, X. Shen 1990a, b, 1992, Liao 1994). I will only mention two types of intonational effect on tone – declination and catathesis – again liberally drawing on data from Shih.

Declination refers to the gradual downtrend of pitch over the course of an intonational phrase that is "blind to the phonological sequence of accents or tones" (Pierrehumbert and Beckman 1988:11), claimed by some to occur universally in natural fluent speech (cf. Ohala 1978, Ladd 1984, Levelt 1989). **Catathesis** or (automatic) downstep, on the other hand, is Fo lowering due to specific tonal combinations, typically the lowering effect of a L on a subsequent H in a sequence of interspersed Hs and Ls (Liberman and Pierrehumbert 1984).[30] Catathesis is attested in English (Pierrehumbert 1980), Japanese (Pierrehumbert and Beckman 1988), and Chinese. The effects of both declination and catathesis are demonstrated in an experiment that is as simple as it is elegant. Each of the four pentasyllabic sentences below has a high tone H in the odd positions, interspersed with varying tones in the even-numbered positions (Shih 1987, 1988).

[29] Notable exceptions include Shanghai (Zee and Maddieson 1980), Fuzhou (Chan 1985), Xiamen/Taiwanese (Du 1988), Wuxi (Chan and Ren 1989), and Zhenhai (Rose 1990). I have made original recordings of connected speech in the following dialects: Tianjin (north Mandarin), Pingyao (Jin), New Chongming (northern Wu), and Wenzhou (southern Wu).

[30] Catathesis or automatic downstep gives rise to phonemic downstep when the conditioning L disappears. Thus HLH → H!H, where downstepped !H contrasts with the plain H. Since Welmers (1959) drew attention to this widespread tonal phenomenon in African languages, it has been studied extensively. See Clements (1979), Hyman (1979, 1986, 1993), Clements and Ford (1981), and references cited therein.

Table 1.4.

context	average Fo of		
	H_1	H_3	H_5
Sentence a	284	265	252
Sentence b	286	255	239
Sentence c	299	260	236
Sentence d	280	259	241

Sentences a, b, c, d refer to (33)
$H_{1, 3, 5}$ refer to the high tones in the first, third, and fifth position of the sentence

(33) a. H_1.**H**.H_3.**H**.H_5
ji shi xiu tuo che "the mechanic fixes the cart"

b. H_1.**MH**.H_3.**MH**.H_5
gong ren shou fang zu "the worker collects the rent"

c. H_1.**ML**.H_3.**ML**.H_5
jing li he guo zhi "the manager drinks juice"

d. H_1.**HL**.H_3.**HL**.H_5
shang dian chu jiu shu "the store old books"

By measuring the average pitch values of the three odd-numbered H tones in each sentence,[31] we can detect the impact of both declination and catathesis, if any occurs. The results are shown in table 1.4. Sentence (a) consists exclusively of a string of Hs. Nevertheless the third and the fifth Hs (= H_3 and H_5) are pronounced on a gradually declining pitch, suggesting an appreciable declination effect.[32] An intervening low (or non-high) target tends to induce a steeper decline on the following Hs. This downstep effect is particularly robust in the case of sentence (c), where the Hs are interrupted by a low-falling ML. Thus the pitch declines by an average of 39 Hz from H_1 to H_3, and by 24 Hz from H_3 to H_5, in comparison with context (a), where the downdrifts at the same intervals are on the order of 19 Hz and 13 Hz respectively.

[31] Each sentence was recorded in random order, four times in natural speech and four times in "reiterant speech." Table 1.4 represents the averages of natural speech tokens.

[32] Curiously, neither Tseng (1981) nor Liao (1994) found a consistent declination effect in their samples consisting of mostly naturally occurring speech. Shih (1987, 1988, 1991), on the other hand, drew her conclusions from sentences specifically designed to test various hypotheses. I suspect that the sample control may have something to do with the discrepancies in their findings.

In a subsequent paper, Shih (1991) further demonstrated that declination is not a purely physiological phenomenon, but is controlled at least in part by the grammar. Specifically, she showed that declination is neither gradual nor constant, but steeper and more pronounced at word boundaries. In other words, Fo declines more gently within words, and more precipitously between words.

We are limited in our access to the relevant phonetic data on the tone–intonation interplay outside of Beijing Mandarin. A serious study of tonal behavior that integrates the results of intonational studies across dialects will have to await another occasion.

4.4 Tone sandhi vs. tone change

One final class of processes affecting tones in connected speech calls for some clarification. In all the cases we have examined so far, the context triggering tonal modification is phonological in nature, whether local (a neighboring tone, segmental influence) or global (intonation). Sometimes what conditions tone change is morphological in nature, as exemplified below:

(34) a. Taishan

ngwoi	33	"I"
ngwoi	22	"we"

b. Zhongshan

hy	22	"go"
hy	35	"gone perfective"

c. Wenling

huɔ̃	31	"yellow"
huɔ̃	15	"yolk"
baq	1q	"white"
baq	51	"albumen, egg white"

Clearly, morphotonemic alternation functions as a "process morpheme," inflectional in the case of Taishan and Zhongshan (both of the Yue group), and derivational in the case of Wenling (Wu, from R. Li 1979). Chan (1989), to whom I owe the Yue examples, observes that marking verbal aspect by tonal means is a productive process in Zhongshan. Similar morphotonemic alternations have been documented for Bobai (also Yue; Kam 1980) and other dialects groups as well, including Xiang (Tang 1960) and Min (Ting 1983), among others.[33] Sinologists generally use the

[33] For similar processes, now no longer productive, of derivation by tone change in Thai (Siamese), see Kam (1980) and work by Prapin Manomaivibool (1976) cited there.

term **tone change** (*bianyin*) to refer to morphologically conditioned tonal modification – to distinguish it from phonologically conditioned **tone sandhi** (*biandiao*). One can think of tone change thus defined as an analog of ablaut and umlaut in English functioning as both inflectional (*foot* ~ *feet*, *sing* ~ *sang* ~ *sung*) and derivational devices (*food* ~ *feed*).

The best known and most fully documented cases of tone change are Beijing Mandarin and Cantonese. A brief sketch of these two case studies will further elucidate the form and function of this phenomenon. I will suggest that one should recognize various types of "tone change," some of which are reducible to generic tone sandhi, while others are not.

4.4.1 Tone change in Beijing Mandarin

The first type of tone change is exemplified by Beijing Mandarin, which has inherited from ancient Chinese a fair number of pairs of cognate words that are related in meaning and differ only tonally. These cognates[34] reflect a once productive morphological process. Downer (1959) gathered over 200 such cognate sets from *Jingdian Shiwen*, a seventh century AD collection of readings. In each case, the etymological base carried MC tone I, II, or IV, and the derived form tone III, corresponding to modern Beijing T4 [51]. The derived forms fall into various grammatical/semantic categories, such as noun, verb, causative etc. In other words, tone III marks a deverbal noun, a denominal verb, or a causative verb derived from an adjective, and so forth. Some canonical examples in their present-day pronunciation are listed below for illustrative purposes:

(35)	a.	*zhong*	55	"center"
			51	"to hit the center of a target"
	b.	*zhong*	214	"seed"
			51	"to plant"
	c.	*yin*	55	"shade"
			51	"to shelter"
	d.	*hao*	214	"good"
			51	"to like"
	e.	*heng*	35	"horizontal"
			51	"cross-grained, hard to deal with"

Each of these doublets also constitutes a homographic pair, i.e. two lexical roots sharing one Chinese character. Others have diverged orthographically, as in

[34] Sometimes referred to as "word families."

(36)	a.	*jing*	55	"to pass through"
			51	"path"
	b.	*ming*	35	"name"
			51	"to name"
	c.	*mai*	214	"to buy"
			51	"to sell"

Yet others have parted ways phonologically, as in

(37)	*[tšʰuan]*	35	"to transmit"
	[tšuan]	51	"record, biography"

where aspiration was once tonally conditioned: voiced stops and affricates became voiceless aspirates in syllables taking MC tone I (corresponding to [55, 35] in Beijing); elsewhere they turned into plain voiceless stops or affricates.

Clearly, it would be inappropriate to apply the term "tone sandhi" to describe the phenomenon sketched above. It would be misleading to treat the alternation between [*hao* 214] "good" and [*hao* 51] "to like" as anything but vestigial relics of a morphological process that has long faded into the dim past.[35] Sinologists typically refer to such morphotonemic alternations by the term "derivation by tone change," apart from tone sandhi. It behooves us to maintain this conceptual and terminological distinction.

4.4.2. Tonal morpheme in Cantonese

More problematic is a second type of tone change instantiated by Cantonese that has received considerable attention in tonological literature (Chao 1947, Whitaker 1955–56, A. Hashimoto 1972:93–100, 180–187, Kam 1977, 1980, Yip 1980:60–65, Wong 1982, Bao 1990a:182–193). Recall that Cantonese has a nine-tone system (see (17)), which I reproduce below, translating the numerical notation [53, 21] etc. into the more familiar symbols [HM, ML] and so forth, for reasons that will become immediately obvious.

[35] Haudricourt (1954b) speculated that Archaic Chinese had an -s suffix, with various functions, including "transitive, causative" and the like. On this hypothesis, the [hao 214 ~ 51] alternation would have evolved in the following way:

Archaic Chinese *xâu "good" → MC xâu II = Beijing [hao 214] "good"
Archaic Chinese *xâu-s "good + transitive/causative" → MC xâu III = Beijing [hao 51] "to like"
Here II and III stand, as usual, for rising and departing tones respectively.

(38) Cantonese tone system

	level		rising	falling
	CVN	CVq		
high (*yin*)	M	Hq; Mq	MH	HM (~ H)
low (*yang*)	L	Lq	LM	ML (~ M)

To oversimplify matters considerably, the most productive type of tone change in Cantonese is one that turns a tone into a high-rising [MH] or level [H] tone in certain contexts:

(39) a. {HM, Hq} → H
b. {M, Mq, L, Lq, ML, MH, LM} → MH

I will single out three of the common contexts, illustrated with examples drawn chiefly from Hashimoto (1972). First, after the "vocative" prefixes [*a* M] and [*lou* LM]:[36]

(40) a.	*a tsæng*	M.HM	→	M.H	"old Zhang"
b.	*a ts'an*	M.ML		M.MH	"old Chen"
c.	*a tsiu*	M.L		M.MH	"old Zhao"
d.	*lou lei*	LM.LM		LM.MH	"old Li"
e.	*lou kʷɔ:k*	LM.Mq		LM.MHq	"old Guo"
f.	*lou mak*	LM.Lq		LM.MHq	"old Mo"

Second, in the second syllable of a reduplicated adjective, with or without the marker -*tei* MH. Adjective reduplication has an intensification effect: AA = very A, while AA-*tei* translates as "a little A." Relevant examples follow:

(41) a.	*hung hung*	ML.ML	→	ML.MH	"very red"
b.	*yit yit*	Lq.Lq		Lq.MHq	"very hot"
c.	*p'a p'a tei*	M.M.MH		M.MH.MH	"a little scared"
d.	*lang lang tei*	LM.LM.MH		LM.MH.MH	"a little cold"

Finally, certain lexically marked items undergo tone change with concomitant shift in meaning:

[36] Both with the connotation of familiarity and informality, roughly like "old" as in "Old Joe." Tone values are regularized to conform with *Hanyu Fangyin Zihui* (2nd edition, 1989).

(42)	a.	*tˈɔng*	ML	"sugar"	→	MH	"candy"
	b.	*nøy*	LM	"female"		MH	"daughter, girl"
	c.	*min*	M	"face"		MH	"face" (metaphorical)[37]

These morphotonemic alternations cannot be attributed to any apparent phonologically definable contextual influence. The question arises as to whether or not the term tone sandhi should be extended to encompass this second type of tone change. Pulleyblank (1991:448f.) advocates a narrow construal of the term sandhi, and excludes alternations exemplified in (40–42) from the scope of tone sandhi, which is more appropriately – on etymological grounds – restricted to the liminal changes arising out of the juxtaposition of tones.

Unfortunately, the dichotomy between morphologically and phonologically conditioned alternations is not as clear cut as one might wish. From a diachronic perspective, both types of processes often spring from the same source. Take English *foot* ~ *feet* and *food* ~ *feed*. Synchronically speaking, the vowel mutation is obviously not phonologically conditioned, but merely encodes the associated morphological information. From a historical perspective, however, umlaut (unlike ablaut) was presumably triggered by the vowel quality of the inflectional/derivational endings that have since disappeared. Thus, *feet* had its origin in pre-historic Old English **foːt-iz* (cf. Pyles and Algeo 1982:114f.). From this angle, the morphological process of vowel mutation is not substantially different from the canonical phonological process of vowel harmony or anticipatory assimilation.

To return to the Cantonese case, one could argue, after Yip (1980:60f.), that what appears to be a morphologically conditioned tone change can be brought in line with the canonical form of phonological rules, if we make allowance for the theoretical device of floating tones. For concreteness, take the "familiar vocative" case. This construction has the following configuration:

(43)
```
a  X
|  |
M  T <H>
```

where [a] is the M-toned vocative prefix; X, together with its lexically assigned tone T, is a place holder for the proper name, and the floating <H> represents a purely tonal morpheme – in Y-R. Chao's preautosegmental

[37] As in the surface or side of something, or as in "to lose face."

and remarkably prescient words, a "non-syllabic and non-segmental suffix" (Chao 1956:1). [*a tsœng*] "old Zhang" and [*a ts'an*] "old Chen" would then have the following as input to the phonological component:

(44) a. *a tsœng*
M HM <H>

b. *a ts^{h}an*
M ML <H>

Docking of the floating <H> on *tsœng* and *ts^{h}an*, each of which carries its lexically assigned [HM] and [ML] respectively, results in complex tones, thereby triggering a tone-simplification process. Cantonese prohibits complex tones (i.e. falling-rising or rising-falling). When complex tones arise from the linking of floating tone, they are simplified by "smoothing," namely by eliminating a medial tone segment, while preserving the two endpoints. In other words:

(45) a. *a tsœng* → *a tsœng*
M. HM<H> M. HH

b. *a ts^{h}an* → *a ts^{h}an*
M ML<H> M MH

which are exactly the pronunciations we want ([HH] on *tsoeng* is non-distinct from [H]). As for *a tsiu* "old Zhao" and *a lei* "old Li," linking the floating <H> to the proper names would create [LH] and [LMH], the latter simplifying to [LH].

(46) a. *a tsiu*
M. L<H>

b. *a lei* → *a lei*
M. LM<H> M LH

At this point several logical moves are open to us. The most theory-neutral option is to say that Cantonese distinguishes only two rising

tones: high [MH] or low [LM]; as a consequence, the *tertium quid* [LH] is reduced to one or the other permissible contours. As it turns out, Cantonese opts for [MH] as the alternative of choice.[38]

The treatment of the "familiar vocative" can be extended, *mutatis mutandis*, to cover adjective reduplication and lexically marked tone changes. Notice, however, that this account of the "familiar vocative" makes crucial use of an underlyingly floating <H>. There is, for instance, no plausible dummy syllable occupying the slot after the proper name, that could undergo deletion leaving behind a floating <H>.[39] The question then turns on whether it is justifiable, synchronically speaking, to posit a {familiar vocative} morpheme that is purely tonal, i.e. without syllabic mooring, in the *underlying* representation.

The answer to this question is somewhat ambiguous. While purely tonal morphemes are certainly atypical in Chinese,[40] the legitimacy of underlyingly floating tones is unassailable on theoretical and cross-linguistic grounds. Close analogs of the {familiarity} morpheme in Cantonese are not hard to find outside Chinese. Perhaps the best known example is the associative particle in the Chadic language of **Ga'anda** as described in Kenstowicz (1994:363f.) based on Ma Newman (1971). The essential facts are given below:[41]

(47) a. *al* "bone"
M

b. *cunewa* "elephant"
L L L

c. *al cunewa* "bone of elephant"
M H L L

[38] Another option is to ban "cross-register" contours, including LH. This move is theory-dependent to the extent that it presupposes a tonal geometry along the lines proposed by Yip (1980). See chapter 2.

[39] Cantonese does have a suffix-like [-a] which is, however, an unlikely candidate as the dummy syllable. Consider for instance:

a tsoeng a → *a tsæng a* "old Zhang"
M HM M M HH M

There are two decisive reasons why this final particle -a cannot be the trigger of tone change: (i) tone change is independent of -a deletion: HM changes to H regardless of whether -a is present or absent; and most importantly, (ii) this final particle [-a] carries a M not a H tone when it does appear.

[40] Other hypothetical cases of floating tone are Standard Mandarin "vivid" adjective reduplication and Xiamen triple reduplication. See Yip (1980:66f.) for details.

[41] [ɓ] stands for implosive [b].

(48) a. *bar* "bark"
H

b. *puno* "maize"
MH

c. *bar puno* "husk of maize"
M H!H

One could say that the first syllable of the second member of an associative construction shifts to a high tone – by a syntactically or morphologically conditioned tone change. This would account straightforwardly for "bone of elephant," but would not, as Kenstowicz (1994:367) points out, explain the downstep (!H) on *puno* "maize" in the same syntactic environment. A far better account is to posit an associative morpheme, which is phonetically a floating <H> without a syllabic mooring (symbolized by the empty [ ] on the segmental tier in the examples). This unanchored <H> simply docks on the closest syllable to the right. The downstep in *puno* results as a by-product of the displacement of the mid-tone M by the associative H. Finally, <M>H is phonetically interpreted as downstepped H, in the customary fashion.

(49) c′. al [ ] cunewa "bone of elephant"
M <H> L

(50) c′. bar [ ] puno "husk of maize"
M <H> M H

→ bar [ ] puno
M H <M> H = M.H.!H

Presumably the associative H originated historically in a full-fledged high-toned syllable. Synchronically, however, this associative morpheme manifests itself exclusively in tonal terms.

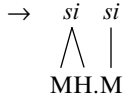
Other examples of tonal morphemes include the subordinate clause marker in Igbo (Goldsmith 1976b), tense/aspect markers in Tiv (McCawley 1970, Goldsmith 1976a:36f.), and Burmese nominal modification (Yip 1980:12). Note parenthetically that floating autosegments are not limited to tones. For instance, floating [palatal] and [labial] have been attested in Japanese (Mester and Itô 1989) and Chaha (McCarthy 1986, Rose 1994).

In short, whether or not the second type of tone change as instantiated by Cantonese familiar vocative should come under the rubric of tone sandhi depends on the justifiability of resorting to a theoretical construct such as floating tones. For our purposes, I will construe the term "tone sandhi" broadly, in keeping with the established usage of the term "sandhi" in the West. As Andersen (1986:2) puts it:

> Such a broad (and loose) understanding of the term has one advantage over any strict definition, the advantage that has helped the term to survive for so long: it makes the term useful as an informal preliminary label which can be used – unlike any strictly defined term – without prejudging the issues that a given set of data might give rise to.

The term "tone change" (*bianyin*) is extended, improperly in my opinion, to cover secondary tonal modification arising out of segmental deletion (e.g. Wong 1982, Bai 1989). This type of so-called tone change is illustrated by the Cantonese example given below:

(51) *si yat si* "give it a try" (lit. try one try)



What we have here is obviously a classic case of tonal stability: the ellipsis of *yat* leaves behind a floating H, which docks on the preceding syllable *si* to form a rising tone. This phenomenon is more properly treated elsewhere as a regular sandhi process (chapter 2, section 2).

5 Synchronic relevance of diachrony

The brief and unavoidably cursory preamble on the highly complex diachronic tonology of Chinese (section 2) was prompted by the fundamental observation that it is often difficult if not impossible to make sense of tonal systems and sandhi alternations in present-day dialects without reference to the historical tonal categories. Specifically, one needs to keep the diachronic perspective in mind in order to understand if not explain[42] what may appear at first as utterly unnatural rules and to make meaningful cross-dialectal comparisons. I will illustrate this point with one

[42] For the distinction between explanatory and "exegetic" adequacy, see Anderson (1981).

single set of data on four closely related southern Min dialects described in Dong (1960).

5.1 Unnatural classes

First of all, it is assumed that phonological rules operate on and are conditioned by natural classes of inputs or environments. It often turns out that natural classes can be defined only in terms of historical categories. Consider the **Jieyang** dialect, spoken in eastern Guangdong. As in other southern Min dialects, each morpheme or syllable[43] in Jieyang is associated with two morphotonemic alternants, generally referred to as citation tone and sandhi tone. **Citation tone** is one that appears when the tone-bearing syllable is pronounced in isolation, for instance in response to the elicitation question: "How do you say ___?" In most Mandarin and Min dialects the citation tone is identical to the final or prepausal variant. (Since the citation monosyllable is bounded by silence on both sides, it is by definition final or prepausal.) **Sandhi tones**, on the other hand, occur in connected speech, in juxtaposition to other tone-carrying syllables. In the case of Mandarin and Min, this usually means pre-final positions.[44] Correspondingly we will often talk about citation vs. sandhi contexts. The positionally conditioned morphotonemic alternations in Jieyang are summarized as follows:

(52) Jieyang morphotonemic alternations

citation tone	sandhi tone
55	11
13	
22	
44	33
42	25
313	31

It is obvious that three of the lexical tones {55, 13, 22} collapse into a single tonal category in the sandhi context. Since neutralization is irreversible, the sandhi rule is unidirectional, namely from citation to sandhi form:

(53) Jieyang neutralization
{55, 13, 22} → 11 in sandhi position.

[43] With rare exceptions, morphemes and syllables are coextensive in size; there are few polysyllabic or subsyllabic morphemes in Chinese. For expository simplicity we ignore checked tones and neutral tones, which do not bear on the issue at hand.

[44] More precisely, pre-final relative to the relevant "domain." We defer the question of what exactly delimits the appropriate "domain" of tone sandhi until later chapters.

Notice in the examples that follow *t'au* "head" carries a high level tone [55] when spoken in isolation, but a low level tone [11] when pronounced together with another toned syllable as in *t'au teng* "top." Likewise, *lau* "old" and *ts'e* "to look for" have two variant readings each depending on the context.

(54)	a.	*t'au*	55	"head"
		t'au teng	11.42	"top" (lit. head + top)
	b.	*lau*	13	"old"
		lau nang	11.55	"old man"
	c.	*ts'e*	22	"to look for"
		ts'e nang	11.55	"to look for someone"

Now, what intersection of tonal features uniquely picks out the class of citation tones that undergo the neutralization rule? The input set includes rising [13], high [55], and low [22] – curiously to the exclusion of the mid tone [44]. If we cross-classify the tonal system of Jieyang by the two parameters of pitch height (high, mid, low) and melodic shape (even or level, rising, falling, and falling-rising), the input set forms a strange and odd-looking class, scattered over the shaded boxes below:

(55) Jieyang tone system (a)

	E	R	F	FR
high	55			
mid	44		42	
low	22	13		313

shaded boxes = input tones to the neutralization rule
E, R, F, RF = even (level), rising, falling, fall-rising

The mystery evaporates as soon as we rearrange the Jieyang tones by placing them in the slots on a grid defined over MC categories:

(56) Jieyang tone system (b)

	MC tones		
	I	II	III
register a (*yin*)	44	42	313
register b (*yang*)	55	13	22

Shaded boxes = input tones to the neutralization rule

I, II and III stand for the three MC smooth tones (namely *ping*, *shang*, and *qu*). Recall that register a and b derive from voiceless and voiced initials in Middle Chinese. Unfortunately we cannot key the neutralization rule to this voicing contrast in the onset, simply because this contrast has long since been lost, so that the original phonation-type distinction has been replaced by exclusively tonal contrasts, for instance:

(57)	Ia	*tseng*	44	"true"
	Ib	*tsai*	55	"talent, skill"
	IIa	*sai*	42	"feces"
	IIb	*siõ*	13	"to think"
	IIIa	*tiam*	313	"shop"
	IIIb	*tiã*	22	"fixed, certain"

Jieyang is by no means unique in this respect. It parallels the **Pinghua** dialect spoken in the Heng county of Guangxi province, reported in Bi (1994) and discussed in Ho (1996). Pinghua has a virtually identical rule of the form: {13, 22, 2q, 42, 42q} → 21 in sandhi position. Again, the input tones (in shaded cells below) form a natural class only when defined in terms of their corresponding MC categories. In particular, the input set includes high (42, 42q), low (13, 22, 2q), and rising (13) as well as falling (42, 42q).

(58) Pinghua tone system

	I	II	III	IV
register a (*yin*)	44	33	55	4q, 33q
register b (*yang*)	13	22	42	2q, 42q

As in Jieyang, Pinghua has undergone devoicing, which precludes the possibility of redefining the input tones by reference to MC voiced obstruent onsets.

What this suggests is that the neutralization processes we have observed in Jieyang and Pinghua must have arisen at a time when either one or both of these two conditions still held true: (i) either the initial consonants still contrasted in voicing, or (ii) the b-register tones still shared some common denominator, for instance a phonetically low register. The original naturalness of the sandhi rule is overlaid and obscured by the subsequent historical changes of devoicing and/or raising of the low register tones.

5.2 *Unnatural rules*

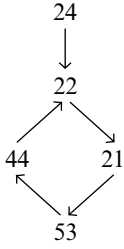
Rules can appear to be unnatural not only because they operate on seemingly unnatural classes or are conditioned by apparently disjointed sets of environments; tone sandhi rules often strike the analyst as bizarre because they seem to relate or map one tone to another in an essentially arbitrary and whimsical manner. This is so often because whatever phonetic or functional motivation there might have been has been wiped out by succeeding waves of changes. Consider the famous tone circle of **Xiamen**, another southern Min dialect described in Dong's (1960) study.[45]

(59) Xiamen morphotonemics

citation form	sandhi form
44 }	22
24 }	
22	21
21	53
53	44

Notice that citation tones [44, 24] merge into [22] in sandhi positions. For this reason, the citation form is generally taken as basic. The sandhi pattern of Xiamen is referred to as a tone circle because of the musical-chair pattern produced by the replacement of tone A by tone B, which in turn is replaced by tone C, and so forth. More graphically, the pattern of alternations can be diagrammed as follows:

(60)



Here I take a few examples from Chen (1987a) to illustrate how this chain substitution works:

[45] I have changed the pitch values to conform to the standard reference *Putonghua Minnan Fangyan Cidian* compiled by Xiamen University (1982).

(61)	a.	*p'ang*	44	"fragrant"
		p'ang tsui	22.53	"perfume" (lit. fragrant + water)
	b.	*we*	24	"shoes"
		we tua	22.21	"shoe laces"
	c.	*pĩ*	22	"sick"
		pĩ lang	21.24	"patient" (lit. sick + person)
	d.	*ts'u*	21	"house"
		ts'u ting	53.53	"roof" (lit. house + top)
	e.	*hai*	53	"ocean"
		hai kĩ	44.24	"ocean front"

The phonetics of Xiamen tone sandhi has stubbornly resisted all attempts to date at giving it a unified and explanatory account. Wang (1967) breaks down the sandhi alternations into two elementary processes:

(62)	a.	[high] switch:	high to non-high:	44, 24 → 22
			non-high to high:	21 → 53
	b.	[falling] switch:	falling to non-falling:	53 → 44
			non-falling to falling:	22 → 21

By a remarkable tour de force, Wang then combines these two into a single alpha-switching rule, the specifics and mechanical details of which need not concern us here.[46] Despite the vast formal differences, Yip's (1980:319–337) analysis boils down to roughly the same thing: her Register Switch corresponds to Wang's [high] switch, and her "Recessive Deletion" and Dissimilation basically perform the same function as Wang's [falling] switch. It doesn't matter whether we take Wang or Yip's rule formulations as the best approximation to the facts, both strike students of human language as essentially arbitrary and fortuitous (cf. Anderson 1978:158f.; Schuh 1978: 248f.; M. Hashimoto 1982).[47]

[46] Here is Wang's celebrated formulation for the curious:

$$[\alpha\text{high}, \beta\text{fall}] \rightarrow [\beta\text{high}, -\alpha\text{fall}]$$

The reader can satisfy himself that this rule generates from citation forms the desired sandhi outputs by assuming the following feature specifications for the tones involved:

	44	24	22	21	53
high	+	+	−	−	+
falling	−	−	−	+	+

The feature [rising] which distinguishes 24 from 44 plays no part in the tone sandhi rule.

[47] Anderson (1978) went on to say that since tone sandhi rules of the Xiamen type are essentially arbitrary suppletions, they are irrelevant to the question of tonal features and geometry.

Ballard (1988:26f.) makes a similar comment on Wenzhou tone sandhi. In his view, whatever "Chomskyesque" sandhi rules one might infer from the synchronic alternations are neither learnable, nor productive, in fact "not part of the speakers' grammars, but historical artifacts." Instead, each sandhi pattern is learned individually along with each lexical item.

I suspect that we are not likely to make any headway as long as we limit our field of vision to the phonetics of Xiamen tone sandhi. To see why this is so, let us compare Xiamen with **Longxi** (also known as Zhangzhou), another southern Min dialect spoken barely thirty miles inland from the coastal city of Xiamen.

(63) Longxi morphotonemics

citation form	sandhi form
14, 313	33
33	11
21	41
52	25

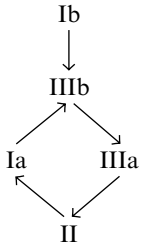
On the face of it, there is little resemblance between the two systems. But, let us first factor out some low level allotonic variations: [11, 41] and [25] in sandhi (prefinal) positions are non-distinct from [21, 52] and [14] in citation forms respectively. We then factor in the relevant historical information, and juxtapose the two systems:

(64) Morphotonemics in two southern Min dialects:

Xiamen			**Longxi**		
citation form	→	sandhi form	citation form	→	sandhi form
44 (Ia), 24 (Ib)		22 (IIIb)	14 (Ia), 313 (Ib)		33 (IIIb)
22 (IIIb)		21 (IIIa)	33 (IIIb)		11 = 21 (IIIa)
21 (IIIa)		53 (II)	21 (IIIa)		41 = 52 (II)
53 (II)		44 (Ia)	52 (II)		25 = 14 (Ia)

As usual, the digits symbolize the phonetic shapes; their corresponding MC tonal categories (Ia, Ib, IIIb, etc.) are indicated in parentheses. $T_i = T_j$ means that T_i is a positionally determined variant of T_j. The striking point that emerges from this juxtaposition is the exact and systematic point-by-point correspondence between the two sets of morphotonemic alternations – that is, when we restate the Xiamen tone circle in terms of MC categories:

(65)



Thus restated, one identical tone sandhi rule operates in both Xiamen and Longxi: the operative tonal categories hold constant across dialects, but are filled by phonetic values that vary from one dialect to the other. The probability that the two dialects have converged by chance in having developed independently of each other such a bizarre sandhi rule is vanishingly small. I take this to constitute incontrovertible evidence showing that the tone sandhi rule had arisen before the two dialects diverged phonetically. Whatever phonetic or functional explanation there might have been must be sought in the tonal system that was in place before the two cognate languages bifurcated. Unfortunately, apart from Hirayama (1974, 1975) and T. Cheng (1983) we know very little about the phonetic values of the proto-system from which the two dialects have evolved. By looking at any one particular dialect synchronically, whatever explanation we might uncover is likely to be ultimately illusory and destined to fail as soon as we try to extend it to another dialect, since by the very nature of things, the explanation rests on precarious grounds, held hostage as it is to the whims of subsequent phonetic changes that can wipe out any clue to the original motivation.

5.3 Irreversible mergers and radical restructuring

There is another reason why we want to keep an eye on the diachronic picture while performing a synchronic analysis. In some cases only a historico-comparative study gives us a clue to what appears to be highly complex patterns of morphotonemic alternations. We draw on the same source (Dong 1960) for illustration. **Quanzhou**,[48] another southern Min dialect, exhibits the following morphotonemic alternations that closely parallel those of Jieyang presented earlier:

[48] Referred to as Jinjiang in Dong (1960). For distinctness I have chosen the name of the county seat instead.

(66) Quanzhou morphotonemic alternations

MC categories	citation tone	sandhi tone
Ib	24	
IIb	22	11
IIIb	**31**	
Ia	44	44
IIa	55	35
IIIa	**31**	55

Notice that the set of {24, 22, 31} that merge into [11] in sandhi position is definable not as a natural phonetic class, but by reference to the MC category: namely register b tones. But we need not belabor this point. The new wrinkle in this case concerns the behavior of tone IIIb and tone IIIa (in bold face). These two categories have merged in citation form, both appearing as [31] in prepausal position; however, their historical differences have survived synchronically in sandhi contexts, as illustrated by these examples:

(67)	a.	*puã*	31	"half"	(from MC tone IIIa)
		puã lɔ	55.31	"half way"	
	b.	*png*[49]	31	"(cooked) rice"	(from MC tone IIIb)
		png si	11.24	"rice spoon"	

Notice that *puã* "half" and *png* "rice" share the tone shape [31] in citation form, but diverge in prefinal sandhi contexts: [31] from MC IIIa alternates with [55], but [31] from MC IIIb alternates with [11].

Let us look at the picture first from the diachronic perspective. There is no doubt that Quanzhou and Jieyang have evolved along the same path. In the interest of cross-dialectal comparison, I reproduce below the table showing Jieyang morphotonemic alternations, revised to include the relevant MC categories:

(68) Jieyang morphotonemics

MC categories	citation tone	sandhi tone
Ib	51	
IIb	13	11
IIIb	22	
Ia	44	33
IIa	42	25
IIIa	313	31

[49] *ng* represents a syllabic nasal.

When we put (66) and (68) side by side, their basic identity becomes transparent despite phonetic differences. Specifically, Quanzhou and Jieyang must have shared these two synchronic rules at some point in their history: [50]

(69) a. Register-b tones → [11] in sandhi position
b. IIa, IIIa → [35, 55] (in Quanzhou) or [25, 31] (in Jieyang) respectively in sandhi position

I specify the inputs to these rules as "register-b tones" or "IIa, IIIa" because we do not know the exact phonetic shapes of these tonal categories at the time these rules came into existence. More importantly, notice that, given the irreversibility of a neutralization rule like (a), we must take the citation tones as basic, and derive the sandhi tones by rule. In turn, consistency of analysis dictates the same directionality for rule (b), even though in principle we could take either the citation or the sandhi form as underlying in the latter case. This uniform directionality is also consistent with the canonical rule schema that holds across southern Min dialects, including Longxi and Xiamen (see above), to wit:

(70) Citation T → sandhi T′

(Here T' stands for some derived tone.) Thereafter Jieyang and Quanzhou parted their ways: most of the tones have taken on new phonetic guises depending on the dialect. Most importantly, a new rule emerged in Quanzhou:

(69) c. IIIa and IIIb → [31] in citation/prepausal form.

At this juncture, Quanzhou simply tacked rule (c) onto the grammar, crucially ordering it after rule (b). The derivation of "half way" and "rice spoon" in Quanzhou then looked as follows:

(71)

"half"	"half way"	"rice"	"rice spoon"	
puã	*puã lɔ*	*png*	*png si*	
IIIa	IIIa.31	IIIb	IIIb.24	underlying tones
–	–	–	11.24	rule (a)
–	55.31	–	–	rule (b)
31	–	31	–	rule (c)

This diachronic analysis makes perfect sense and lays bare the common pattern underlying the two affiliated dialects. From the synchronic point of view, however, the supervenient rule (c) wreaked havoc with the

[50] For our purposes we can ignore the alternation between [44] and [33] for tone Ia in Jieyang, since there the two variants do not contrast tonemically.

then-existing grammar of Quanzhou. Since new generations of Quanzhou speakers have no tribal memory of ancestral categories like IIIa and IIIb, both having merged into [31] in the citation form; there is no way for the rules of (69a, b) to survive in their original form. (69a, b) say in effect:

(72) a. [24], [22], and some [31] (derived from MC IIIb) → [11] in sandhi position
b. other [31] (derived from MC IIIa) → [55] in sandhi position

This means that we have to reverse course, and instead take [55] and [11] – the sandhi forms of IIIa and IIIb respectively – as the underlying representations for *puã* "half" and *png* "rice" respectively, and derive the citation forms [31] by a neutralization like:

(73) Prepausal neutralization
{55, 11} → 31 in prepausal position

However, this entails other analytical moves. Consider the following:

(74)	a.	*puã*	31 (< /55/)	"half"	(from MC IIIa)
		puã lɔ	55.31	"half way"	
	b.	*kau*	55 (< ?)	"nine"	(from MC IIa)
		kau tsap	35.35q	"ninety" (nine + ten)	

If *puã* "half" carries the underlying or base tone /55/, then *kau* "nine" obviously must take a different base tone, otherwise both should show up with the same surface tone in sandhi positions. This means that we have to take the alternant [35] for *kau* "nine" in sandhi position as the underlying. Note however that in the case of [24 ~ 11] and [22 ~ 11] alternations (corresponding to MC tones Ib and IIb), the citation forms must be taken as underlying, since neutralization is irreversible. The net result is that the morphotonemic alternations in Quanzhou are asymmetric: the arrows that link underlying tones to their corresponding morphophonemic alternants must point erratically in both directions, specifically:

(75) Quanzhou morphotonemics

	citation forms		sandhi forms
Ib IIb	/24/ /22/	→	[11]
IIIb IIIa	[31]	←	/11/ /55/
IIa	[55]	←	/35/
Ia	[44]	=	[44]

To summarize, Quanzhou calls for the following rules:

(76) a. citation tones [24, 22] → [11] in sandhi positions
b. sandhi tones [11, 55] → [31] in citation/prepausal positions
c. sandhi tone [35] → [55] in citation/prepausal positions

The rules stated above are illustrated by the sample derivations given below:

(77) Sample derivations in Quanzhou:

"half"	"half way"	"rice"	"rice spoon"	"nine"	"ninety"	"door"	"door opening"	
puã	*puã lɔ*	*png*	*png si*	*kau*	*kau tsap*	*mng*	*mng k'au*	
55	55.11	11	11.24	35	35. 35q	24	24.55	
–	–	–	–	–	–	–	11.55	rule (a)
31	55.31	31	–	–	–	–	–	rule (b)
–	–	–	–	55	–[51]	–	–	rule (c)

What emerges is a drastically different synchronic picture from Jieyang. A single late rule of IIIa/b merger has forced a radical restructuring on Quanzhou, and completely obscured the erstwhile uniform directionality of derivation from citation form to sandhi form, which is still visible in the case of Jieyang. If we approach Quanzhou and Jieyang exclusively from the synchronic perspective, first we see no commonality whatsoever between the two otherwise closely affiliated dialects, and second we find the erratic and asymmetric directionality of derivation in Quanzhou completely arbitrary and baffling.

6 Citation tone, base tone, sandhi tone

In view of the intersecting synchronic and diachronic dimensions, a few words in terminological clarification are in order. In the foregoing sections we alluded to citation vs. sandhi tones. Typically, it is the **citation tones** (in Chinese *danzidiao*, lit. "single word tone") associated with syllables spoken in isolation that preserve the most contrasts, while **sandhi tones**, i.e. tones associated with syllables in connected speech tend to merge. This is illustrated by a triplet from **Fuzhou** (Chan 1989).

(78) a. *sing* 55 "new"
b. *sing* 52 "grown, complete"
c. *sing* 12 "holy"

[51] The prepausal *tsap* 35q does not turn into 55, because rule (c) operates only on "legato tones," i.e. tones associated with smooth/legato syllables CVN.

(79)	*sing. ing*	55.52	a.	"bride" (newly wed person)
			b.	"adult" (grown person)
			c.	"sage" (holy person)

Spoken in isolation, the words for *new*, *grown*, and *holy* are phonetically distinct, carrying as they do different tones. However, in combination with the nominal head [*ing* 52], which means "person, human," the original tonal contrasts disappear, so that the disyllabic expression [*sing.ing* 55.52] is three-way ambiguous between *bride*, *adult*, and *sage*. It is obvious that we have to take the citation tones /55, 52, 12/ as basic, and derive their sandhi form [55] via a rule of tonal neutralization, roughly {55, 52, 12} → 55 in prefinal position.

The Fuzhou case is fairly typical among Chinese dialects. For this reason, analysts generally take citation tones to be the **underlying** or **base tones** (*jidiao* lit. "base tone" or *diceng diaoxing*, lit. "underlying tone shape"); by the same token, the sandhi tones are often considered the **surface** or **derived tones** (*biandiao*).

However, there are cases where underlying contrasts are neutralized in citation forms, but resurface in sandhi contexts. While atypical, neutralization of underlying distinctions affecting citation tones is by no means uncommon. We have already seen above the case of Quanzhou, where MC IIIa and IIIb merge into [31] in citation forms, but surface with different forms in sandhi positions. **Wenling** (R. Li 1979) is another, better known example. In this southern Wu dialect, both "skin" and "blanket" are pronounced as *bi* [31] in isolation. Despite their identical citation forms, they behave quite differently in sandhi contexts, as illustrated below:

	citation form		sandhi form	
(80)	*bi* 31	a.	*bi.li* 13.31	"inside the skin"
		b.	*bi.li* 31.31	"inside the blanket"
(81)	*bi* 31	a.	*boq.bi* 1q.15	"thin skin"
		b.	*boq.bi* 1q.51	"thin blanket"

Such puzzling facts can be understood, if not explained, in diachronic terms: "skin" and "blanket" once belonged to two distinct tonal categories, namely Ib and IIb respectively. Obviously, historical tone Ib and IIb remain separate tonal categories even in modern Wenling, and must be

encoded in the underlying forms, notwithstanding their neutralization in citation forms.

Cases like this can be multiplied ad infinitum in virtually every major dialect group. One of the chief reasons for separating Jin from other northern dialect groups is the peculiar fact that Jin dialects typically have one single category for MC tone I, while this so-called *ping* tone has split into two – Ia and Ib – in virtually every modern dialect. For instance, when pronounced in isolation, both *tou* "steal" (MC Ia) and *tao* "peach" (MC Ib) carry the tone [13] in the **Pingyao** dialect (Hou 1989:95). But this cohesiveness of tone I is only an illusion. The contrast between tone Ia and Ib resurfaces in sandhi context, witness:

(82)	a.	*tou ting*	"to eavesdrop" (lit. stealthily + listen)
		13 13	citation form
		31 35	sandhi form
	b.	*tao hong*	"peach red"
		13 13	citation form
		13 13	sandhi form

The change [13.13 → 31.35] takes place only if the first syllable carries a [13] derived from MC Ia. It is all but certain that the split of MC tone I into a- and b-registers was truly pan-dialectal and that the merger of Ia and Ib in citation form is a late development that took place in individual dialects (e.g. the Min dialect of Jian'ou) or dialect groups (e.g. Jin).[52]

It is therefore important to realize that the two axes of citation vs. sandhi tones and base/underlying vs. derived tones run diagonally to each other. Whether one must take the citation tone or the sandhi tone to be underlying depends on the observed patterns of tonal alternations.[53]

[52] This "hidden" Ia/Ib contrast that comes to light only in sandhi positions has misled some linguists into dubious historical reconstructions. For instance, C. Cheng (1991:107) constructs a most parsimonious Stammbaum model, where Jin (represented in his corpus by Taiyuan) branches off from the rest of the northern Sinitic stock at the very earliest point in time, thereby escaping the otherwise pan-dialectal tone I split. This conclusion is improbable, to say the least, from the geopolitical, historical, and cultural as well as linguistic perspectives. The tone sandhi facts strongly suggest that Jin, like all other dialects groups, has undergone tone I split; subsequently, Ia and Ib merged, but only in citation forms. In all fairness, I hasten to add that *Hanyu Fangyin Zihui*, on which Cheng (1991) based his quantitative study gives no clue to the sandhi behavior of tone Ia and Ib.

[53] The analytical move involved is entirely analogous to segmental morphophonemic alternations. Thus, there is no a priori reason to take the German bare stems [bunt] "colorful" and [bunt] "bond, union" – equivalent to "citation" forms – as the underlying representations, given that their true identity surfaces only in "sandhi" positions: [bunt-es] "colorful, neuter singular" vs. [bund-es] "union, genitive singular."

In this connection, it is best to discard altogether the confusing term **original tone** (*bendiao*), which is used variously to refer to (i) the citation tone, or (ii) the underlying tone in synchronic analysis, or (iii) the historically earlier tone value. For further clarification on this point, see Chan (1989), M. Hashimoto (1982), and Ting (1984).

2 *Tonal representation and tonal processes*

Tone sandhi come in all sizes and shapes. In our survey of the multifarious types of sandhi processes, we must be selective by necessity. Recent tonal studies have shifted focus from the formulation of sandhi rules to the representation of tones in the hope that what form sandhi rules eventually take will follow naturally from an appropriately enriched representation of tone.[1] Accordingly, I will focus on those sandhi phenomena that shed light on the nature and representation of tone.

1 Tonal representation

In virtually all Chinese linguistic literature, tonal categories are consistently classified by two sets of descriptive terms: one denoting pitch *height* (high/low or "*yin*"/"*yang*"), the other pitch *movement* (rising, falling, dipping etc.). This practice implies that tone consists of two independent, orthogonal dimensions: register and pitch contour – or to use a musical analog: key and melody. The same tune or musical phrase can be transposed an octave up or down, or from one key to another (say from C to E^b) without losing its melodic identity. This bi-partition of tonal properties is already implicit in Wang (1967), where [high, central, mid] constitute one subset of features in contraposition to [rising, falling, concave, convex]. But internally unstructured feature matrices, consistent with the prevailing descriptive model at the time, provided no formal means to segregate these two subsets of features. With Yip (1980) this implicit dichotomy between pitch height (register) and pitch movement (contour) entered into the basic vocabulary of theoretical discourse.[2]

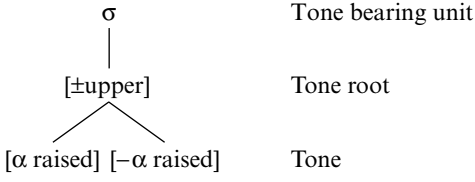
[1] For an extensive survey of tone sandhi phenomena with ample data and specific rule formulations, with particular reference to the southern dialects of Wu and Min, see Ballard (1988).

[2] However, the term "register" is used in quite different and potentially confusing senses in Asianist and Africanist traditions: in the former register is a fundamental component of

With particular reference to contour tones – a salient feature of East-Asian-type tone languages (cf. Pike 1948) – such entities can be seen as unitary melodic gestalts, where the dynamic pitch movement is a linguistic primitive; alternatively, a contour tone can be seen as a string of two pitch levels, where the pitch glide is nothing more than an interpolation, a slur or portamento, as it were, between two discrete musical notes. Both conceptions are implicit in the traditional discourse on Chinese tones. On the one hand, Chinese phonologists consistently talk about tones as *gao-jiang* "high falling," *di-sheng* "low rising," *jiang-sheng* "dipping," and so forth. On the other hand, Sinological literature has long enshrined Chao's (1930) widely adopted tone letters/digits, which precisely dissect contour tones into their constituent pitch levels. These two implicit views of tones eventually found their way into explicit descriptive frameworks, the former in Wang (1967), the latter in Woo (1969), Yip (1980, 1989), and virtually all recent literature.

Various distinctive features and geometrical arrangements thereof have been proposed. These competing proposals have been reviewed critically and in considerable detail in Anderson (1978), Hyman (1986, 1993), Snider (1988), Bao (1990a), Duanmu (1990a), L. Chang (1992), Tsay (1994), and Chen (1996), among others. There is no need to cover the same ground here (for a summary of the various classificatory schemes, see Appendix: Tone Features). Instead, I will take as a point of departure the model proposed by Yip (1989):

(1)



Basically, Yip first bisects the tonal space into two discrete **registers**: the upper ([+upper]) and the lower ([–upper]) portions of the vocal range;

the tonal category itself, while in the latter, register stands apart from the basic classificatory scheme, and acts as a "value-setting" device that shifts the pitch range upwards (upstep) or downwards (downstep) across a certain tonal span. The register in Chinese is strictly local and discrete, but is global and cumulative (incrementally higher or lower) in the Bantu languages. To further complicate the matter, "register" is used to refer to phonation types in Kam-Tai and Mon-Khmer languages. For clarification on this point, see Inkelas and Leben (1990), Yip (1993) and Hyman (1993).

within each register, the pitch is fine-tuned by the feature [±raised],[3] yielding four possible even (i.e. level) tones:

(2)

+upper	+raised
	−raised
−upper	+raised
	−raised

In a two-tone language, presumably [±upper] alone suffices for taxonomic purposes. The mid-level tone in a three-tone system can be specified either as [+upper, −raised], or [−upper, +raised], depending on how the mid tone behaves: whether it patterns like a high or a low tone in either undergoing or conditioning a sandhi rule. **Contour** tones are represented geometrically as a branching root node with two oppositely valued terminal nodes. Thus, a high-rising and a low-falling tone have the internal structure of (3a) and (b) respectively:

(3) a. high-rising: [+upper] = MH

[−raised] [+raised]

b. low-falling: [−upper] = ML

[+raised] [−raised]

In principle, one could represent complex tones in like manner, say:[4]

(4) c. high falling-rising:

[+upper] = HMH

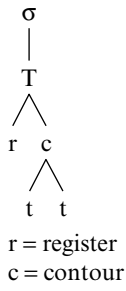
[+raised] [−raised] [+raised]

As we shall see in section 3, certain sandhi phenomena strongly suggest a refinement of this model in one respect. In Yip's conception, the register directly dominates the terminal nodes. Bao (1990a) proposes an alternative geometrical configuration, where the register and the contour hang from the tone root as sister nodes:

[3] Originally [±high] (Yip 1980), revised after Pulleyblank (1986:125) to avoid terminological confusion with vowel height and to avoid any suggestion that [+high (tone)] implies a H autosegment.

[4] Bao (1990a) imposes an upper limit of two terminal nodes per tone on a priori grounds. I leave this question open.

(5)



In Bao's system, high and low registers are featurally specified as [+stiff] and [–stiff] respectively, while the terminal nodes are labelled [+slack] (= [–raised]) or [–slack] (= [+raised]). The articulatory correlates of these features adopted from Halle and Stevens (1971) are somewhat controversial (cf. Hombert 1978:81–82).[5] I will continue to use the more abstract, non-committal terms [upper] and [raised].

A geometrical representation like (5) makes explicit a number of empirical claims, the most important of which are: (i) tone is an autosegment, linked to its segmental host by an association line; (ii) the tone-bearing unit is the syllable; (iii) contour tones are decomposable into discrete pitch levels (rising = MH, LM, falling = HM, ML); (iv) contour tones have a structural unity, represented by a branching node "c"; (v) register (pitch height) and contour (pitch movement, melodic shape) constitute two orthogonal, mutually independent dimensions of tone.

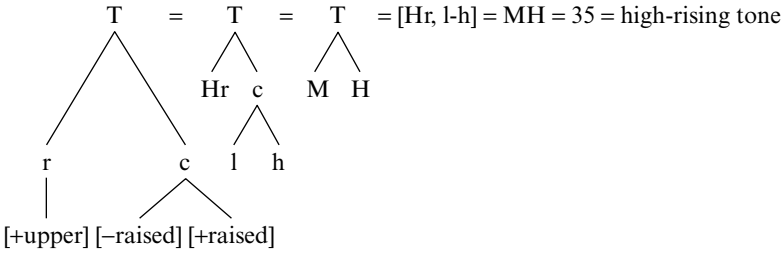
Bao's model serves as a reasonable frame of reference. The full implications of such a system and potential problems related to it will become clearer as we proceed.

As a notational shorthand, I will used the following conventions:

T	tone, tone root
t	terminal node
r	register
c	contour
Hr / Lr	high/low register, or [+upper] / [–upper]
h / l	high/low pitch, or [+raised] / [–raised]
H	[+upper, +raised]
M	[+upper, –raised] or [–upper, +raised]
L	[–upper, –raised]
MH	high-rising
ML	low-falling, etc.

[5] Interestingly, Bao (1990a) and Duanmu (1990a) have exactly the opposite interpretations of the underlying physiological mechanisms for controlling pitch associated with [stiff, slack].

H, M, and L are used interchangeably with Y-R. Chao's tone digits, allowance being made for non-distinctive overspecificity in the latter. Thus [55, 35, 31] = H(H), MH, ML, and so forth. Unless theoretical or expository considerations dictate otherwise, I follow the notation used in the sources without modification. For all practical purposes, these notations are equivalent:



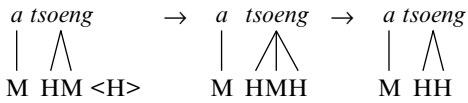
The choice of the notation used on each occasion depends on the degree of explicitness that is deemed appropriate for the discussion at hand. Occasionally, I will use the lowercase h, m, l to designate high, mid, and low tones derived through spread or default rules, and associated with prosodically weak syllables which are either inherently toneless or have undergone tonal deletion.

2 The autosegmental status of tone

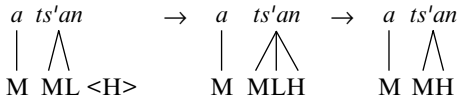
As noted above, the tonal representation adopted here characterizes tone as an autonomous entity, separate from its segmental host. Since Goldsmith (1976a), tone has been firmly established as the prime example of autosegments. The autosegmental status of tones is hardly controversial (cf. arguments presented by Yip 1980:84f.; and Bao 1990a:175f.), and various types of sandhi processes bear out this claim. Recall that in **Cantonese** the vocative form "*a X*," roughly translatable as "old X," where X stands for the addressee's surname, entails a tone change: it turns the tone associated with X into a rising tone. This is best described (after Yip 1980) by positing a purely tonal (i.e. segmentless) morpheme, consisting of a floating <H>, which docks on the syllable to the left. Where appropriate, the resulting tonal strings are simplified in accordance with the tonotactics of Cantonese (for further details, see chapter 1, section 4.4.2; also see discussion below).

(6) floating <H> (Cantonese)

a. "old Cheung"



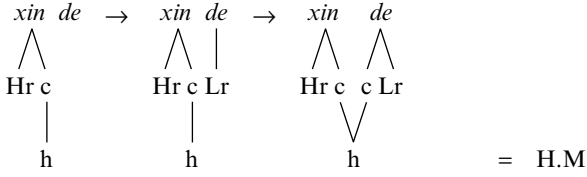
b. "old Chen"



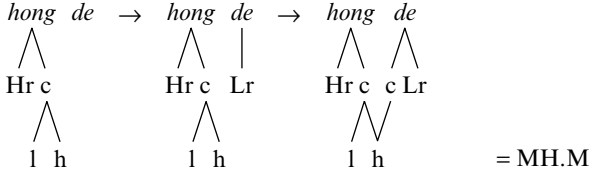
The mutually independent existence of tone and tone-bearing units (TBU) is also evidenced by the inherently toneless syllables like *de* in **Mandarin**, which functions as a relative clause marker (cf. chapter 1, section 4.1). This toneless *de* assimilates to the tone value of the lexical host on the left. Specifically, the pitch height is predictable by positing a rule that spreads rightwards the terminal node h, or l (= [+raised], [–raised]). The register of the atonic enclitic is determined by what appears to be a polarity principle: it is Lr, if preceded by Hr, Hr, otherwise, along the lines suggested by Chang (1992:108f.). Examples follow:

(7) Toneless syllables (Mandarin)

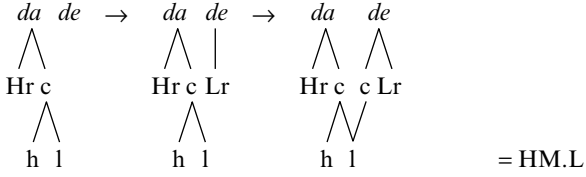
a. "the new one"



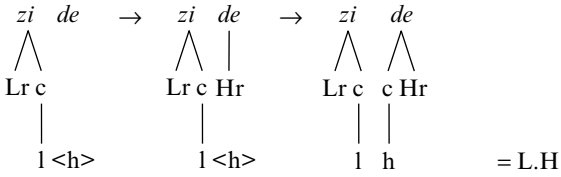
b. "the red one"



c. "the big one"



d. "the purple one"



In the last instance, the conventional analysis postulates a floating <h> for T3 words like *zi* "purple." This <h> surfaces only in phrase-final positions – or when followed by a potential TBU not otherwise already taken up by a lexically designated tone. The predicted pitch values correspond closely to Chao's (1968:35f.) and Z-J. Wu's (1985) descriptions.

Tonal autonomy also manifests itself in tonal deletion that leaves the TBU intact. The **Shanghai** case is typical. Without going into details (for which see chapter 7, section 3), the tone shapes of polysyllabic compounds in this Wu dialect are determined uniquely and exclusively by the initial tones. Specifically, a tone-deletion rule eliminates all non-initial tones without affecting the tone-bearing syllables. The initial tone then spreads rightwards to the now-toneless syllables. This is illustrated by the following examples:

(8) a. HL + T

"gutter"	"deficit"	
in kɤ	*k'uɛ k'ong*	
HL.HL	HL.MH	base tone
HL. o	HL. o	Deletion
H. L	H. L	Spread

b. MH + T

"palm"	"pretty"	
sɤ çin	*p'iɔ liã*	
MH.HL	MH.LH	base tone
MH. o	MH. o	Deletion
M. H	M. H	Spread

The opposite, namely TBU deletion while preserving the associated tones, also occurs with regularity. I will illustrate this process with some well-known facts from ellipsis in Cantonese and Nanzhuang and syllable contraction in Xiamen.

Cantonese, like many other dialects, permits the ellipsis of certain morphemes in certain constructions. Thus *yat* "one," <*ts*> (aspect marker), and <*hai*> "in, at" enclosed in angled brackets are optionally skipped over in examples such as these (from Wong 1982:17f.; cf. Whitaker

1955–56; Hashimoto 1972:93–100; 180–187, Yip 1980:87–91; Bai 1989; Bao 1990a:199–203):

(9) a. *yat t'iu* <*yat*> *t'iu* "strip by strip" (lit. one strip one strip)
Hq.ML. Hq. ML

b. *si* <*yat*> *si* "give it a try" (lit. try one try)
M. Hq.M

c. *mai* <*tsɔ*> *lak* "has sold" (lit. buy + aspect + sentence particle)
L MH Mq

d. *fan* <*hai*> *sü* "sleep here"
M MH M

<σ> = deleted syllable

However, syllable elision entails certain concomitant tonal modifications, as seen below:

(10) a. *yat tiu tiu*
Hq.ML.ML → Hq.MH.ML "strip by strip"

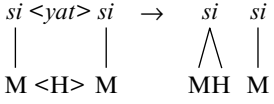
b. *si si*
M.M → MH.M "give it a try"

c. *mai lak*
L Mq → MH.Mq "has sold"

d. *fan sü*
M M → MH.M "sleep here"

One could think of this type of tone change as syntactically conditioned, and simply stipulate that the tone of certain syllables changes into MH just in these constructions. But this morphosyntactic account would fail to shed light on the connection between ellipsis and tone change. On the other hand, this intimate link is built into an account that takes cognizance of a high or rising tone originally associated with [*yat* Hq] "one," [<*ts*> MH] "aspect marker," and [*hai* MH] "in, at," which are then set afloat by syllable elision, consistent with the principle of tonal stability under segmental loss. This floating tone is then reassociated with the preceding host, as depicted below.

(11)

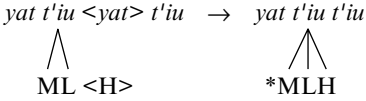


The generalized form of the process involved is statable simply as a rule that links a floating tone <T> with a syllabic anchor to the left:

(12)

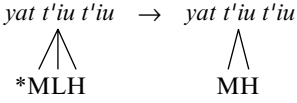
subject to independently established tonotactic conditions of Cantonese. For instance, Cantonese disallows complex tones, that is falling-rising or rising-falling contours. As a consequence, the docking of <H> set afloat by the deletion of <*yat*> in

(13)



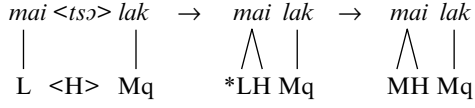
creates an illegitimate complex tone *MLH, which is simplified to MH, via a tonal simplification rule which retains only the tone segments that mark the target pitch at the two endpoints:

(14)



As for *mai lak* "sold," the docking of the free-floating H from the aspect marker <*ts*> would create an otherwise non-occurring tone in Cantonese, namely *LH; as a consequence, LH is further modified before surfacing as MH.[6]

(15)



It is clear that these cases of so-called tone change are best treated as tonal modification occurring in strictly phonologically defined contexts, indistinguishable from other sandhi processes.

Floating tones arising out of syllable loss and triggering tone change are by no means unique to Cantonese. Tone-preserving syllable contraction also occurs in the southern Min dialect of **Xiamen**. Here are some examples based on my own speech:[7]

[6] One might attribute the non-occurrence of LH to a cross-register constraint, which prohibits a high registered tone (H) from cooccurring with a low-registered tone (L). Cf. Yip (1980), Bao (1990a).

[7] Yip (1991:89) cites other examples from Robert L. Cheng (p.c.) without gloss. For clarity I employ H, M, L instead of tone letters.

(16) a. *tsa hng*[8] → *tsang* "yesterday evening"
M H **MH**

b. *tsa k'i* → *tsai* "morning"
H. HM **HM**

c. *ka lang me* → *kang me* "to scold someone"
L M M **LM**. M

d. *lai k'i hia* → *lai hia* "come over there"
L H MH **LH**.MH

Syllable contraction basically works by dropping the initial consonant of the second syllable, and fusing the two rhymes into one, subject to syllable wellformedness conditions, the details of which need not detain us here. The pitch values to the left of the arrows represent not the underlying tones, but the surface representations that have already undergone the well-known tone sandhi rule of Xiamen described elsewhere in this volume (chapter 10). Thus, *tsa* "yesterday" and *tsa* "early" in (a) and (b) carry the underlying tones [MH] and [HM] respectively. What is of interest here is the observation that the composite syllable combines the tonal properties of both constituent syllables as well. Thus [M + H] = [MH], [H + HM] = [HM], and so forth.

A perhaps even more dramatic case has come to light in Ma's (1990) report of the northwestern Mandarin dialect of **Nanzhuang** spoken in Shaanxi province. This dialect has a standard four-tone system, common in many Mandarin dialects, consisting of {11, 24, 53, 44}. Certain syllables are atonic, marked with "o." Nanzhuang has regular tone sandhi rules, including one that turns [11.11] into [24.11]. But what appears at first totally baffling is the occurrence of complex tone shapes not attested elsewhere except in examples like the following:

(17) a. *kɷ̃-puo kua ts'ɛ liᴇ* "the arm is bleeding"
11 o **4431** 53 o (gloss: arm hang color part)

b. *tio kɤ ngɤ* "I was the only one left"
1121 o 53 (gloss: leave-behind CL me)

c. *xiə̃ tso t'iɤ̃ šɤ sɷ̃* "I heard that"
1121 44 24 11 o (gloss: heart just hear say that)

d. *xiɛ̃ fã ti* "it belongs to the first wife"
11 **2431** o (gloss: first chamber possessive)

[8] /-ng/ represents a syllabic velar nasal.

From the context, it is clear that both verbs [*kua* 44] "to hang" and [*tio* 11] "to cast away, leave behind" have the perfective meaning: the arm has been wounded, and I have been left behind etc. One might say that Nanzhuang has a residual inflectional morphology, whereby the tense/aspect is encoded by a process morpheme that triggers the tone change of [44] and [11] to [4431] and [1121] respectively. This fails to account for the tones associated with [xiə̃ 1121] "heart" and [fã 2431] "chamber." More significantly it leaves unexplained the fact that these complex toned syllables carry extra length.[9] Finally, on this account, it would be a mere coincidence that precisely after these contour-toned syllables certain enclitics are expected to appear (and routinely do). In other words, instead of [kua, tio, xiə] and [fã], we have:

(18) a. *kua liE* "hang Asp"
44 o

b. *tio liE* "leave-behind Asp"
11 o

c. *xiə̃ li* "heart in"
11 o

d. *fã əʒ* "house affix"
24 o

The solution to the problem is by now obvious. Clearly an optional ellipsis rule is at work: prosodically weak enclitics like aspectual markers, postpositions, suffixes etc. are elided, leaving behind intact both a low, possibly falling, tone and a timing unit, say a mora. Nanzhuang affords us a fairly dramatic instantiation of tonal/durational stability under ellipsis. In Ma's (1990:99) words: "Ce qui est remarquable dans ces examples, c'est que la fusion des syllabes est complète au niveau segmental et ne laisse aucune trace de la syllabe absorbée alors qu'au niveau prosodique, le contour tonal et la durée de cette dernière est resté intacte."

3 Tonal geometry and the typology of spread/shift rules

I have alluded to the dual nature of contour tones: on the one hand, they are decomposable into discrete pitch levels; on the other hand, they form a structural unity and not simply a cluster of disparate tone segments. This dual nature of contour tones comes through quite clearly in sandhi

[9] Ma (1990:100) observes that these complex toned syllables are longer than the norm.

processes. On the one hand, the endpoints of a contour tone can act with certain freedom, on the other hand, the melodic contour moves as an indivisible whole. I shall refer to these two aspects of contour tones as "**edge**" and "**integrity**" effects, and demonstrate them with data from Zhenhai, corroborated with supporting evidence from other dialects.

Zhenhai, a Wu dialect located 20 km northeast of Ningbo in Zhejiang, has been described by Rose (1990) with extraordinarily rich phonetic details. It has a six-tone system:

(19)

	CVN		CVq
	rising	falling	
high register	323 MH	441 HL	5 Hq
low register	213 LM	231 ML	23 Lq

For our purposes, the tone digits given in Rose are reducible to a simpler notation using only H, M, and L. Disyllabic expressions (mostly compounds) fall into two patterns: SW (strong-weak) and WS (weak-strong). The accentual patterns are determined partly by the tonal combinations (e.g. low-register tones tend to be weak), and partly by syntactic boundaries (Rose 1990:7). Once the accentual patterns are ascertained, the sandhi forms of such disyllabic expressions are predictable on the basis of the initial tone.

3.1 Edge effects

Consider first the SW (strong-weak) patterns. The relevant facts are summarized in the following table:

(20) SW disyllabic patterns

			second σ		pattern
			Long σ: CVN	Short σ: CVq	
first σ	323	MH	334–51	334–5	a
	213	LH	114–51	114–5	b
	441	HL	44–31	44–32	c
	231	ML			d
	5q	Hq	5–51	5–5	e
	23q	Lq			f

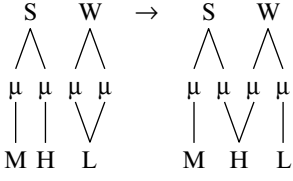
shaded areas = systematic gaps

The first syllable may carry any one of the six citation tones indicated on the leftmost column in Chao's tone letters. Recall that -q stands for a glottal stop, and therefore a "checked" or short tone. The disyllabic tone pattern is predictable from the source tone of the first syllable and the syllabic type of the second. Thus, if the first syllable carries [323], then the disyllabic pattern is [334–51] if the second syllable is long, [334–5] if the second syllable is short (= pattern a). Shaded boxes represent systematic gaps: no SW pattern corresponds to the underlying tonal sequences /231 + T/ and /23q + T/; such sequences give rise only to WS patterns to be discussed in the next subsection. Appropriate examples follow (21a, b, c, e correspond to the patterns of the table above):

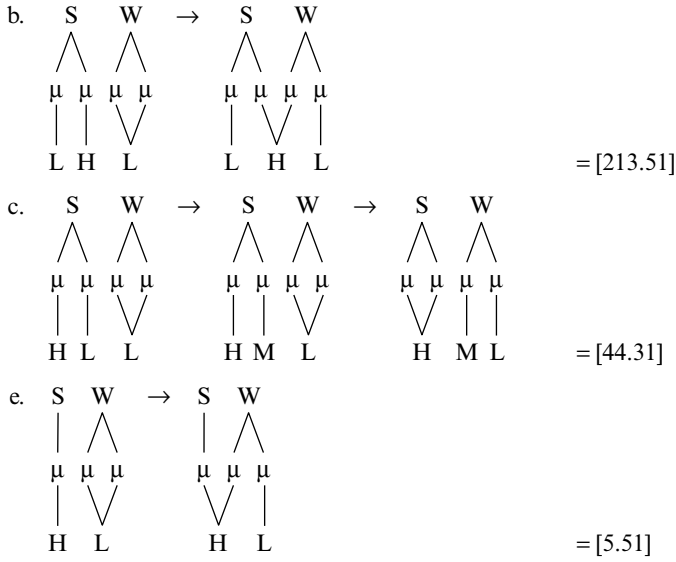
(21)	a.	"scissors"	a′.	"simply"	
		tçi ta		*tçĩ dziq*	
		323.441		323.23q	base tone
		334.51		334.5q	sandhi form
	b.	"way"	b′.	"America"	
		ta lu		*mei kaq*	
		213.213		213.5q	base tone
		114.51		114.5q	sandhi form
	c.	"lake Tai"	c′.	"strength"	
		tha wu		*tçhi lıq*	
		441.231		441.23q	base tone
		44.31		44.32q	sandhi form
	e.	"to invent"	e′.	"neighbor"	
		fa ming		*ka pıq*	
		5q.231		5q.23q	base tone
		5q.51		5q.5q	sandhi form

Details aside, what is going on in Zhenhai is transparently a rightward migration of the tone contour from the prominent to the weak syllable which has lost its lexically associated tone. To better bring out the underlying mechanism of the sandhi process, let us reduce the tone letters to a notation employing only H, M, and L next to the tone letters. The contrast between long (legato) and short (checked) syllables is translated into bimoraic vs. monomoraic units in the conventional fashion. We can now schematize the four attested SW patterns in the following fashion:

(22) a.



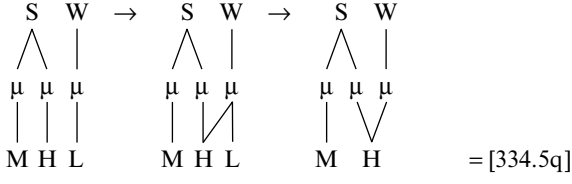
= [334.51]



Take pattern (a). The prosodically weak, unstressed syllable in the second position loses its underlying tone and acquires a default L instead. Next, the H of the prominent syllable slides into the first mora of the second syllable, creating a high-falling contour in the process. Patterns (b) and (e) operate in exactly the same fashion. Pattern (c) calls for an intermediate step involving what we may call "melodic interpolation," whereby a L between a H and a L is raised halfway to a mid-level pitch. This M, in turn, migrates rightwards, and shifts altogether into the next syllable.

We can generate the SW patterns with a short syllable in the second position in the same manner with one additional rule, namely contour simplification on short syllables. Notice that the intermediate output in pattern (a′) involves a one-mora to many-tone mismatch, which is remedied by dislodging the final L, as diagrammed below:

(23) a′.



Pattern (c) [44.31] has a minor variant (c′) [44.32q], where the short fall is not long enough to reach its target L.

What is noteworthy in our telling of the remarkable Zhenhai story is that tones do not migrate as indivisible wholes. Rather, tones slide to the right, one musical note and one mora at a time.[10] Pattern (c) demonstrates this point most dramatically. In its rightward shift, the high-falling tone disintegrates into discrete pitch levels, so that the M tone detaches itself from the preceding H, and recombines with the following L to form a new low-falling tone. The bare essence of this tone slide is reducible to (24a). Restating the process in terms of contour tones as indivisible melodic "gestalts" (24b) would obscure the true nature of the underlying mechanism.

(24) a. HM-L → H-ML
b. high-falling + low-level → high-level + low-falling

In this respect, there is no essential difference between Zhenhai, a contour tone language, and other languages usually characterized as "register-tone" systems, such as Soyaltepec Mazatec, an Otomanguean language of Mexico discussed in Goldsmith (1990:40f.):[11]

(25)

"liquid eyes"		"tears"
nta šku	→	*nta šku*
3 2 4		3 2 4

(Note that in Pike's (1948) notation used in the Mazatec example, the numbers have inverse values to those of Chao's tone letters, so that 1 = high, 4 = low, etc.)

3.2 *Integrity effects*

The composite nature of contour tones has now been widely accepted since Woo (1969), Leben (1973), and Yip (1980). Remarkably, Zhenhai also demonstrates the flipside of the picture, namely the structural unity of contour tones. If contour tones are simply a sequence of pitch levels, there is no reason to expect them to move or spread as a single unit. But this is exactly what happens in the SW patterns of disyllabic compounds. The essential facts are presented in tabular form below:

[10] The best known case of a similar tone shift is Kikuyu, reported in Clements and Ford (1979).

[11] Pike (1948:12) classifies Mazatec as "a register-tone system with a slight overlap of contour-tone characteristics."

(26) WS disyllabic patterns

			second σ			pattern
			long σ		short σ	
			Hr	Lr		
first σ	323	MH				a
	213	LM	11.334	11.24	11.4q	b
	441	HL	33.441		33.4q	c
	231	ML	11.441		11.4q	d
	5q	Hq				e
	23q	Lq	1q.T		1q.T	f

shaded areas = systematic gaps
T = base tone, unchanged

(27) b. "bedroom"
fã.kɛ
213.441
11.334

"last year"
tçju nĩq
213.231
11.24

b′. "abroad"
nga kaq
213.5q
11.4q

c. "spring"
tshɤng thĩ
441.441
33.441

c′. "western calendar"
çi lıq
441.23q
33.4q

d. "coal mine"
mei khwã
231.231
11.441

d′. "fellow student"
tang aq
231.23q
11.4q

f. "mooncake"
yı pıng
23q.323
1q.35

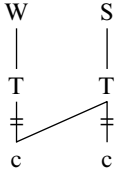
f′. "day"
njɛ tçjaq
23q.5q
1q.4q

Take example (27f), corresponding to pattern (f). The low checked tone 23q (= Lq) deletes in WS patterns. The initial toneless syllable assumes

a default L, transcribed as [1q] in Rose's notation. The second syllable in the strong position basically retains its underlying tone (symbolized by T). Notice that /323/ represents a high registered rising tone or /MH/ in our system, in contrast to the low registered rising /213/ or /LM/. It is not surprising that this MH surfaces as [35] on *pɪng* under stress in the WS pattern.[12] Pattern (f′) is unremarkable in this respect.

We now turn to the remaining patterns, where the initial syllable is of the legato or smooth type (i.e. CVN or CV). Basically, what happens is that the tonal contour (rising, falling) of the initial weak syllable is retained, but realized on the prominent, second syllable, as an instantiation of the Tone-to-Stress Attraction (TSA) principle (see chapter 7, section 1.2), which I state as follows:

(28) Tone-to-Stress Attraction



T = tonal root node
c = contour
‡ = delink

TSA has the effect of detaching the contour (symbolized here as "c") of the weak syllable and relinking it to the strong syllable, in the process supplanting the original contour of the latter. Notice that nothing is said about the register, which is determined by independent principles, to wit: (i) in strong position, the register is high if it is associated with a falling tone or checked syllable, otherwise it retains its underlying specification; (ii) in weak position, the register is uniformly low before a rising tone (i.e. MH or LH), otherwise the register retains its underlying specification. Finally, the unmarked value of the terminal node is [l] or [–raised]. I will refer to these basically "feature filling" principles as Default. The derivation of patterns (c), (d) and (e) proceeds as follows:

[12] In fact, all tones occupying the prominent position in WS are phonetically raised. Thus /23q.213/ → [1q.114] as in *njɛ naq* "noisy." I will ignore this detail for expository simplicity.

(29) c. "spring" TSA Default

tshɤng *thĩ* → *tshɤng* *thĩ* → *tshɤng* *thĩ*

W S W S W S

T T T T T T

Hr c c Hr Hr c Hr Hr c c Hr

h l h l h l l h l = [33.441]

d. "coal mine" TSA Default

mei *khwã* → *mei* *khwã* → *mei* *khwã*

W S W S W S

T T T T T T

Lr c c Lr Lr c Lr Lr c c Hr

h l h l h l l h l = [11.441]

In pattern (c), the [h-l] contour shifts rightwards to the strong syllable by TSA; the final output is fleshed out by the Default values: since the first syllable retains its high register Hr, and acquires a [l] i.e. [–raised] by Default, the result is M or [33]; the second syllable attracts the falling contour, which remains high registered under accentual prominence, and therefore emerges as [Hr, h-l], that is HM or [441], allowance being made for the phrase-final fall noted by many students of tone and intonation (for references, see chapter 1, section 3). Pattern (d) works exactly the same way. Note in particular that the falling tone on *mei* "coal" started out as low registered; but when it shifts to the stressed syllable, it takes a high register, in accordance with the principles stated above as Default, clause (i).

As for WS with a rising tone on the first syllable, the derivation is equally straightforward.

(30) b. "bedroom" TSA Default

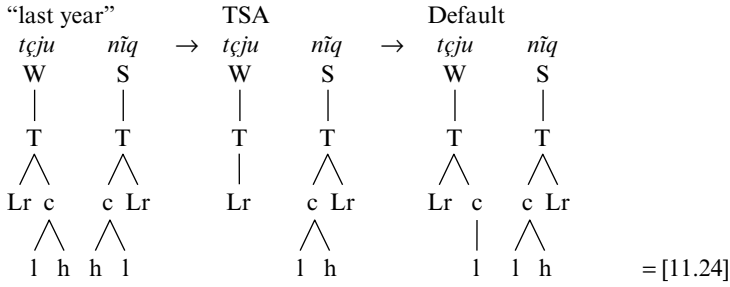
fã *kɛ* → *fã* *kɛ* → *fã* *kɛ*

W S W S W S

T T T T T T

Lr c c Hr Lr c Hr Lr c c Hr

l h h l l h l l h = [11.334]



In the case of *fã kɛ* "bedroom," TSA turns the second syllable from a falling to a rising tone; since a syllable associated with a rising tone retains its lexically specified register, *kɛ* surfaces with a high-rising tone. *tçju nĩq* "last year" parallels *fã kɛ* in every way, except that since *nĩq* "year" is underlying low registered, it emerges with a low-rising tone.

Details aside, the crucial step in our analysis of the WS patterns involves the TSA. This process is interesting in two respects. First, it moves the contour node "c" as a single unit; if the terminal nodes [h-l] or [l-h] are independent elements, there is no reason to expect both h and l to move together. This rightward contour shift is diagnostic of the structural unity of the elements that together make up the melodic contour. This argument is not different from movement as a test of constituency in syntax: only constituents move together. The second significant aspect of TSA is the independence of contour movement from register. In the cases of (29c, d), TSA must move the contour while leaving the register intact, in order to account for the contrast between a mid and a low level tone in the W position: [33.441] vs. [11.441]. In the cases of (30), only the rising contour of the first syllable migrates rightwards, leaving behind whatever register it was originally associated with, since it will eventually take on a new value for the [±upper] feature of its eventual host, the second syllable. This accounts for the contrast between [11.334] and [11.24].

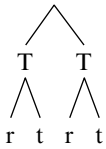
The (a′, c′, d′, f′) variants, with a checked syllable in the second position, are derivable in the same manner, with the proviso that rising or falling contour is simplified to a (high) level tone. Recall that the register associated with a checked syllable is high under stress (see clause (i) of Default stated above).

3.3 *Three models of tonal representation*

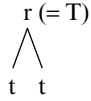
The SW and WS patterns of Zhenhai, in conjunction with data taken from other dialects, provide us with critical evidence that discriminates

among three current alternative geometrical configurations proposed for contour tones:

(31) A. (Duanmu 1990a, 1994a; cf. Clements 1985, Snider 1990)[13]



B. (Yip 1989; cf. Hyman 1993, Chen 1996)



C. (Bao 1990a)

T = tone root
t = terminal tone segment
r = register
c = contour
TBU = tone bearing unit

On the autosegmentalist view of assimilation as spreading (cf. Steriade 1982, Hayes 1986), assimilatory processes are diagnostic of constituency, since only constituents can spread. By the same token, movement rules or syntagmatic shifts have the same diagnostic value, since they differ from spread only in that the constituents that spread are further delinked from their original TBU. Under this assumption, the three models of tonal geometry make the predictions summarized in the synoptic table 2.1.

[13] For simplicity, I ignore the fact that Duanmu imposes a one-to-one relation between tone and TBU (the mora), while Clements and Snider permit a one-to-many mapping between the two. For ease of comparison, I also disregard the actual featural labels the various authors attach to the more abstract T and t, e.g. [upper, raised] in Yip, [stiff, slack] in Bao, primary (modal) and secondary (registral) H and L in Clements, Snider and Hyman, etc.

Table 2.1. *Typology of possible spread/shift predicted by alternative models of tonal geometry*

	model A	model B	model C	instantiation
1. contour spread/shift	no	no	yes	Zhenhai (Zhenjiang)
2. register spread/shift	yes	no	yes	Chaozhou (Pingyao)
3. whole contour tone spread/shift	no	yes	yes	Wenzhou (Changzhi, Danyang)
4. terminal node spread/shift	yes	yes	yes	Mandarin, Gao'an

3.4 Contour spread/shift

Model A posits two root nodes (T) for tonal contours, in effect equating a contour tone with a tonal cluster in Yip's (1989) sense. It goes without saying that model A rules out a process like the TSA we have seen in the WS patterns of Zhenhai. Model B identifies the root node with the register itself (r = T);[14] as a consequence, the only way for both of the terminal nodes (the lowercase t's) to move together is to move the root tone T as a whole. In other words, model B rules out the possibility of contour movement independent of the register. The only tonal geometry that can accommodate the facts of Zhenhai is model C. Only this model permits the terminal nodes hanging from the constituent "c" (contour) to move as a unit without taking the "r" (register) together with it.

Bao (1990a:96f.) cites Zhenjiang and Wenzhou as two cases of contour spread. On closer examination, neither case is convincing. The former case instantiates nothing more than contour simplification (all contours are leveled), while the latter probably involves as yet poorly understood paradigmatic tonal alternations (see Chen 1996). Zhenhai is the only case of independent contour spread/shift to my knowledge.

[14] Analogous to [consonant, sonorant] as the featural content of the segmental root node, cf. McCarthy (1988).

3.5 Register spread/shift

How about the converse of contour spread/shift: can register spread independently of contour? Bao (1990a:88f.) answers in the affirmative, citing Pingyao (a Jin dialect; based on Hou 1989). I have argued at length elsewhere (Chen 1996) that the Pingyao data are amenable to an alternative construal as register neutralization rather than register spread. However, a more convincing case has come to light. The dialect in question, **Chaozhou** (or Teochow), belongs to the southern Min groups, and is spoken in Guangdong and by a large number of overseas Chinese in SE Asia. The crucial data are given below (from Bao 1996a):

(32) a. goods ship

hue lung	"cargo ship, freighter"
LM. H	base tone
HM. H	sandhi form

b. goods storage

hue ts'ng	"warehouse"
LM. L	base tone
ML. L	sandhi form

c. fire handle

hue ba	"torch"
HM. HM	base tone
MH. HM	sandhi form

d. fire arrow

hue tsi	"rocket"
HM. LM	base tone
LM. LM	sandhi form

We infer from this representative sample of data that Chaozhou has a contour metathesis rule which we state simply as follows:

(33) Contour metathesis
LM → ML
HM → MH

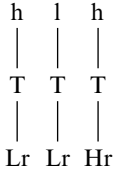
What happens next is what interests us. All tones that surface with a contour in sandhi contexts are predictable with respect to register, which is determined by the syllable to the right. Thus *hue* "goods" surfaces as HM or ML depending on whether the next syllable is high or low registered, respectively (32a, b). The same generalization holds for *hue* "fire" (32c, d). Under the standard assumption of assimilation as spreading this register assimilation is statable as follows:

(34) Register spread

tone root

Notice that the register ("r") alone spreads, leaving the contour "c" unaffected. Register spread independent of contour is unstatable in model B, since the register *is* the tone root itself, ergo the register cannot spread without taking with it the entire tonal complex, including the terminal nodes. Register spread is possible in principle in model A. For concreteness, take (32a) *hue lung* /LM.H/ "cargo ship." As it comes out of Contour Metathesis, it has the tonal string [ML.H], or equivalently, in terms of model A:

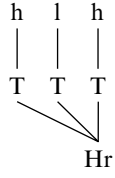
(35)



= ML.H

Which then undergoes register spread to yield the output:

(36)



= HM.H

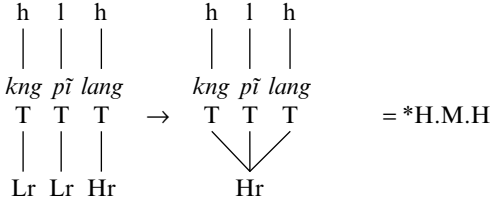
This means that register must spread leftwards over a (possibly) unbounded span. This makes wrong predictions. Extrapolating from the data presented in Bao (1996b), the following trisyllabic expression carries the underlying tonal sequence /M.L.H/:

(37) carry patient

kng pĩ lang	"to carry a patient"
M L H	base tone

Needless to say, neither M nor L is subject to Contour Metathesis. Taking /M.L.H/ as input, the unbounded leftward register spread would predict that M and L turn into H and M respectively:

(38)



We know this output is ungrammatical because Bao (1996b:33) clearly states that both M and L tones remain unchanged in sandhi (i.e. pre-final) positions.[15] One could of course stipulate that register spreads to all tautosyllabic tones, which strikes me as an oblique way of referring to contour tones as units.

3.6 Whole-tone spread/shift

We have seen above cases where either the register or the contour spreads independently of each other. How about the entire complex of contour tones spreading or moving as a whole? Since model A does not recognize contour tones as a unit, it predicts that no contour tones may spread or move as a unit. Models B and C, on the other hand, both accord contour tones a structural cohesiveness either under the root node, or under the subtonal constituent "c"; both therefore predict that contour tones may spread or move as a single entity.

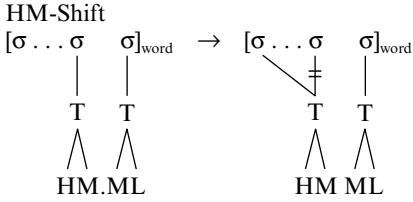
The structural unity of the contour tone and its internal geometrical organization continue to be hotly debated issues. I will not repeat the arguments for and against contour tones presented most compellingly in Yip (1989, 1992, 1995) and Bao (1990a, 1992) on the one side, and Duanmu (1990a, 1992b, 1994a) on the other. Danyang and Changzhi, once regarded as prime examples of whole-tone spread (see Chen 1991, 1996, Yip 1989, Chan 1991, 1995) have been disputed by Duanmu (1994a). Uncontroversial cases of contour spread, with or without involving the register are indeed hard to find.[16] But, if we construe syntagmatic movement as spread-cum-delink, then we have already seen a case of independent contour spread in Zhenhai. I now proceed to show what looks like a case of whole-tone dislocation or shift.

[15] Available data suggest that, although phonetically quite different, Chaozhou tone sandhi operates within the identical prosodic domain as Xiamen (see chapter 10). I have taken care to insure that the hypothetical example falls within a single tone group.

[16] Maddieson (1978a:351) observes that "In no case has a rule been found, in which a contour tone is copied through an assimilatory process" (p. 351).

Wenzhou, another Wu dialect spoken farther south in the same Zhejiang province, has a special rule, which I call HM-Shift, stated below:

(39) HM-Shift



Unlike the better known northern Wu dialects, Wenzhou displays a right-prominent prosodic pattern: the rightmost syllables tend to retain their lexically specified tones, while the syllables farther to the left undergo varying degrees of tonal neutralization (see chapter 11 for further details). HM-Shift moves a high-falling tone to the word-initial position across an arbitrary span of toneless syllables. Its effect is illustrated by the examples given below:

(40) a.	[*sun-nü*]-*xu*	"grand-daughter's husband"
	M.MH.HM	base tone
	HM.ML	Disyllabic tone sandhi
	HM. o ML	HM-Shift
b.	[*wai*-[*sun-nü*]]-*xu*	"maternal grand-daughter's husband"
	L M MH HM	base tone
	HM ML	Disyllabic tone sandhi
	HM o o ML	HM-Shift
c.	wireless telephone tube	
	[[*wu-xian*]-[*dian-hua*]]-*tong*	"radio receiver"
	ML HM L L ML	base tone
	HM ML	Disyllabic tone sandhi
	HM o o o ML	HM-Shift

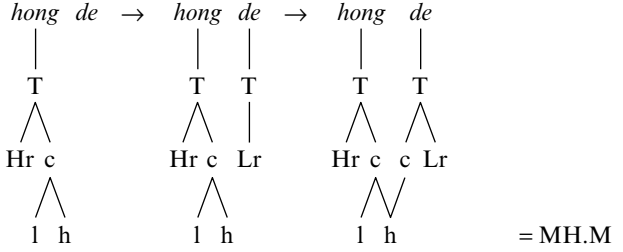
o = atonic syllables

Note that the HM.ML sequences that trigger the leftward shift can be either underlying or, as in the above examples, derived from Disyllabic tone sandhi, the specifics of which are dealt with in chapter 11 and need not concern us here. When such a sequence arises in the course of derivation, the high-falling HM automatically moves to the beginning of the word. What is of interest is of course the fact that what moves is not an individual tone segment H or M, but the tonal complex HM. Under the assumption that only constituents move, HM must constitute a structural unit. (For further details of Wenzhou, see chapter 11.)

3.7 Terminal node spread/shift

It goes without saying that the terminal node "t" (= h or l) is a constituent; not surprisingly, all models of tonal geometry[17] allow the possibility for it to spread as a unit. We have already seen an instantiation of terminal node spread in **Mandarin** in section 2. The critical example is reproduced below:

(41) "the red one"



What is noteworthy is that the toneless suffix (or particle) *de* assimilates not to the second tone segment of the preceding high-rising tone as a whole (i.e. the underlined H in the sequence MH), but only to the featural content [+raised] (symbolized here as "h"), which is then "recombined" with the default register Lr, to emerge as M (= [Lr, h] or [–upper, +raised]).

Terminal node spread is not limited to a toneless target. **Gao'an** (Yan 1981; cf. Bao 1990a:111f.), a Gan dialect spoken in Jiangxi, has the following canonical anticipatory assimilation rule that recurs in one variant or another in countless other dialects:

(42) 55 → 53 / ___ 33, 11, 3q, 1q

The effect of this rule is illustrated below:

(43) "bi-seasonal"

a. *song tçi*
55. 33 → 53.33

"egg"

b. *tçi han*
55. 11 → 53.11

"pig iron"

c. *sang tʹiɛt*
55. 3q → 53.3q

[17] Except for Wang's (1967) model or its equivalent.

"Chinese medicine"

d. *tçiung iok*
55 1q → 53.1q

Assuming the mid-level tone [33, 3q] to be [+upper, –raised] or equivalently [Hr, l], the process is straightforwardly statable in Bao's model adopted here as a terminal node spread:

(44)

Notice again, what we witness here is a *partial* assimilation, involving only the terminal node [l], which is phonetically realized as either M (= 3) or L (= 1), depending on the register it is associated with.

To recapitulate, of the three alternative geometrical representations of tonal anatomy, only model C allows for each type of the four types of spread or movement rules listed in table 2.1. To the extent that these rules are instantiated by the case studies sketched above, model C appears to accord best with the empirical facts.[18]

4 Dissimilation and substitution

Under the assumption of assimilation as spread and shift as spread-cum-delink, we have looked very closely at these two types of processes because spreading and movement rules serve as effective diagnostic probes into tonal anatomy. There are other recurring types of sandhi processes as well.

4.1 Contour vs. register dissimilation

Although assimilation and dissimilation are conceptually the converse of each other, they are implemented by quite different formal means, given the tonal representation we have adopted here. Assimilation, registral or melodic, is executed by spreading a tonal node (not necessarily the root node). Since melodic contours are decomposed into syntagmatically contrastive sequences of pitches, contour dissimilation calls for metathesis, or the permutation of the terminal nodes: x-y → y-x, where x and y are

[18] I once argued in favor of a more parsimonious model B on negative evidence (Chen 1996 [=1991]), in that all the putative cases of independent register or contour spread known at the time turned out to be dubious at best.

cover terms ranging over h and l. It is worth noting that while contour *assimilation* is quite rare, its converse, contour *dissimilation* is extremely common among Chinese dialects. It is so productive that L. Chang (1992:256) quite rightly characterizes contour dissimilation as the primary sandhi process, and cites Xian, Chengdu, Kunming, Pingyao, and Changzhi as instances.[19] We will encounter many examples of contour dissimilation as we examine individual cases. For now I will only illustrate this process with two examples from Pingyao (see next section for more details).

(45) a. "become ill"
hai bing
35 35 → 53.35

b. "gullible"[20]
er ruan
53 53 → 35.53

This can be stated as something like:

(46) [x-y] → [y-x] / ___ [x-y]

where x and y stand for the terminal nodes [αraised] and [$-\alpha$raised]. Actually, the Pingyao contour dissimilation rule is more restricted in its scope, phonologically and morphosyntactically. I defer the details until section 5.3.

Register dissimilation is far less common.[21] One possible example is the polarity principle discussed in section 2: the toneless suffix in Beijing Mandarin takes on a high register if the lexical host is low registered, and vice versa. Although Bao (1990a:281) cites Luoyang, and Chang (1992:243f.) mentions Tianjin, Chengdu, and Beijing as cases instantiating register dissimilation, it is not clear that they represent genuine and general processes. For instance, Beijing T3 Sandhi is sometimes construed as the raising of a low register tone to a high register tone before another low register tone (e.g. Cheng 1973a; cf. Hyman 1975). Recall that Beijing has

[19] However, she went on to say that "[t]he primarily dissimilatory nature of the TS rules suggests that T as a phonological property is more like stress" (Chang 1992:259). This conclusion is unwarranted in my opinion. On this view, if we were to liken – quite reasonably – rising LH and falling HL contours to iambs and trochees respectively, we would expect an iamb to turn into a trochee before another iamb, and vice versa. I am not aware of such an accentual analog to contour dissimilaton.

[20] Lit. ear + soft, i.e. easily talked into believing something.

[21] Hyman (1978:262) observes: "Very few cases of tonal dissimilation can be demonstrated to have occurred as processes (i.e. as historical changes rather than as synchronic reflexes of other changes)."

(47)

T1	55	=	H
T2	35		MH
T3	214		L?, LM?
T4	51		HL

T3 phonetically realized as [214] only in final positions, [35] before another T3, and [21] elsewhere. Now, we may represent T3 in a number of ways, for instance: (a) as L [Lr, l], or as (b) LM [Lr, l-h]. Accordingly, T3 Sandhi rule would look like either (48a) or (48b):

(48) a. L → MH / ___ L
b. LM → MH / ___ LM

Rule (a) doesn't look like a canonical register dissimilation rule at all. If anything, L should turn into H before another L. Rule (b) does look like register dissimilation, but raises the following question: if register dissimilation is the underlying mechanism at work, why doesn't the opposite occur, namely:

(49) MH → LM / ___ MH
HL → ML / ___ HL

Why the asymmetry between register and contour dissimilation? My hypothesis is this: contour dissimilation is functionally motivated by the fact that it has the effect of decreasing tonal complexity by reducing the number of tonal targets within a short span of utterance. Thus MH + MH → HM + MH = HMH, conversely HM + HM → MH + HM = MHM. Notice incidentally that tautosyllabically while HMH and MHM or their low-register counterparts (concave and convex tones) both occur, neither the double falling nor the double rising tones are attested (see table 2.5 of chapter 1). Register dissimilation generally serves no such purpose. If anything, it tends to create tonal targets that are farther apart, and therefore requires more exacting fine tuning of the underlying pitch-control mechanism. For example HM.HM → ML.HM or LM.LM → MH.LM, etc. This view of contour dissimilation as simplification is consistent with the universal tendency to flatten out tonal contours (cf. Hyman and Schuh 1974).

4.2 Paradigmatic substitution

Finally, there remains a vast assortment of tonal alternations that defy classification and description let alone explanation. As one examines one Chinese dialect after another, one is left with the baffling impression

of random and arbitrary substitution of one tone for another without any apparent articulatory, perceptual, or functional motivation. Perhaps the most famous, though by no means either the most bizarre or the most complex, case of tonal substitution is that of **Xiamen** "tone circle" alluded to earlier in chapter 1, section 5.2. Leaving aside the so-called checked tones, the "musical chair" chain substitution is summed up as follows:

(50) {44, 24} → 22 → 21 → 53 → 44

That is to say, an underlying 44 or 24 is replaced by 22, and 22 is replaced by 21, and so forth in sandhi (i.e. pre-final) position (see chapter 10 for a full treatment).[22] Suffice it for our present purpose to note that the tonal substitutions are not phonetically motivated in the sense that the complex pattern of tone changes do not contribute to a more "harmonic" output string in any plausible sense. To see why this is so, take just one link of the chain: 44 → 22 / ___ T (where T = any tone), as illustrated by the word *p'ang*, which surfaces with [44] in the final position, and [22] in the prefinal position.

(51) a. *tsin* *p'ang* "very fragrant"
T 44

b. *p'ang* *tsui* "perfume" (lit. fragrant + water)
22 T

T = any tone

It is not as if [44.T] was an illformed tonal juxtaposition. In fact [44.T] is a perfectly good tonal sequence – provided that [44.T] is derived from an underlying /53.T/, as demonstrated in (52):

(52) a. *tua* *hai* "ocean" (lit. big + sea)
T 53

b. *hai* *kĩ* "seaside, ocean front"
44 T

One might speculate that Xiamen tone sandhi serves a demarcative function in the Trubetzkoyan sense: just as fixed accents signal word boundaries (cf. Trubetzkoy 1969, Hyman 1977), the phrase ending in Xiamen is punctuated by the appearance of the base tone. But

[22] By sandhi or pre-final position I mean any position other than the final (not necessarily penult) in a domain to be defined precisely in chapter 10.

this functional explanation is ultimately illusory. Given a tonal string [. . . T.44.T . . .] we cannot tell independently of other contextual cues whether it should be parsed as (i) [. . . T.44 # T . . .], or (ii) [. . . T # 44.T . . .], where # symbolizes a phrase boundary – the interpretation of [44] is ambiguous: it is underlying in (i), derived from /53/ in (ii). In other words, the appearance of [44] does not per se disambiguate between the two parses.

In short, paradigmatic tonal substitutions such as exemplified by the Xiamen tone circle are neither functionally motivated, nor phonetically driven.

Actually, the Xiamen case is straightforward relatively speaking in that the mapping between the base tone and the sandhi output is context-free to the extent that tone A is always replaced by B, regardless of what tone precedes or follows. In some other dialects, the mapping can be far more complex. A schematic example would look like the following:

(53)

	A	B	C
A		X-T	
B			
C	Y-T		Y-T

shaded areas: no change

Let A, B, C stand for underlying tonemic categories, X, Y for sandhi tones, and T for an unaltered based tone. The input/output correspondences can be restated in rule form as follows:

(54) R1 {A, B} → X / ___ {B, C}
R2 C → Y / ___ {A, C}

The rules say that the second tone always remains unchanged (T), but A and B turn into X before B or C, and so forth. Since A-A, B-A, and C-B sequences are unaffected by either rule, these tonal combinations appear on the surface unchanged, and are represented by the shaded boxes. The important point I wish to make with this schematic example is that tonal substitutions often do not fall into neat patterns, but rather form a set of arbitrary correspondences. Specifically, B joins A to form a class as input to R1, but sides with C as environment in the same rule. By the same token, A and C form a class as the context for the C → Y correspondence in R2.

When confronted with this kind of apparently fortuitous input–output mappings, I will follow the general practice of simply summarizing the correspondences in the tabular form exemplified by (53), and forgo rule statements like those of (54) which shed little light on the underlying mechanism.

Needless to say, the arbitrariness of tonal substitutions is sometimes only apparent: what appears at first to be a crazy quilt of criss-cross mapping relations yields, if only grudgingly, to a rational analysis under closer scrutiny. One of the main challenges of the analyst is precisely to discover the principles hidden under the bewildering surface manifestations. We will deal with a particularly daunting case in chapters 5–6.

5 Neutralization and differentiation

5.1 Neutralization

No survey of tone sandhi processes would be complete without at least mentioning a recurrent theme, namely the neutralization effect of sandhi rules. We may think of Cantonese and Shanghai as two prototypes standing near the opposite poles of a continuum. Cantonese typifies a tone-rich but sandhi-poor system: it has one of the largest tonal inventories (nine contrastive tonemic categories; see chapter 1, section 3), but aside from contour tones arising out of syllabic elision (section 2) and morphologically conditioned tone change (chapter 1, section 4.4.2), the underlying tones remain virtually unchanged in connected speech. This means that the number of possible different tonal strings equals 9^n, where n = the number of syllables.[23] Shanghai, on the other hand, is tone-poor, with five tones (reducible to just two, if onset voicing and syllable structure are taken into account; see chapter 1, section 3), but remarkably sandhi-rich: within a prosodic domain (to be made more precise in chapter 7, section 3), all but the initial tonal contrasts are eliminated, so that the tone shape of the polysyllabic string is determined exclusively by the initial syllable. This means that the number of possible distinct tonal strings does not increase with the number of syllables. In other words, a polysyllabic string has one of exactly five possible tonal envelopes, regardless of syllable count. It goes without saying that this tone deletion entails sweeping neutralizations, as illustrated by a rather mundane doublet like this one:

[23] For examples, a quadrisyllabic expression in Cantonese can have any one of 6,521 possible tonal sequences (= 9^4).

(55) a. *ma mɔ* "buy a cat"
LH.HM base tone
LH. o Deletion
L. H Spread

b. *ma mɔ* "buy a hat"
LH. LM base tone
LH. o Deletion
L. H Spread

Such sandhi-created ambiguities can be multiplied ad infinitum. It is not obvious that the Shanghai speaker suffers from any greater communicative handicap than his Cantonese cousin, presumably linguistic and pragmatic contexts more than compensate for whatever contrasts are lost due to systematic sandhi rules.

Between these two poles, there is a whole spectrum of neutralization patterns. Take **Zhangping**, a southern Min dialect of Fujian (Zhang 1982). It has a seven-tone system. However, in a disyllabic tone group, the forty-nine logically possible two-tone strings (7 × 7) are reduced to twenty-one (= 3 × 7):

(56) Zhangping

		second σ	
		24, 11, 5q, 53, 53q	31, 21
first σ	24 11 21 5q	33-T	55-T
	31 53 53q	21-T	

T = base tone, unchanged

In (56), T stands for the citation/underlying tone of the second syllable that remains unchanged. The tonal categories of the first syllable are reduced to just three. For instance, {24, 11, 21, 5q} all merge into [33], when followed by {24, 11, 51, 53, 53q}, etc. It is not transparent on what basis the input tones or the contextual tones are partitioned into sets, the intersection of which determines the observed sandhi forms. For instance, /31/ goes with {53, 53q} in the input subset, but with /21/ in the context subset.

The **Suzhou** dialect as reported in Ye (1979) represents a more sweeping pattern of neutralization. Unlike Shanghai, Suzhou also has a partial neutralization rule rather than a radical tone deletion. Furthermore, the initial tone does not spread to the atonic syllables. Instead, the medial syllables retain their register, but reduced to a flat tone ([44] if originally high registered, [33] if low), while the final syllable carries a low-falling tone [31] if the syllable is "legato" (CVN), otherwise a simple low [22] (CVq). Examples follow:

(57)	a.	*pe ts'o si ka*	"les misérables"
		44.53.412.412	base tone
		44.44. 44. 31	sandhi form
	b.	*hiã mae hy tçin*	"wool scarf"
		13. 13. 13. 44	base tone
		13. 33. 33. 31	sandhi form
	c.	*le pae zaq*	"sapphire"
		13.52.3q	base tone
		13.44.22q	sandhi form

Curiously, what survives partial neutralization could be the melodic contour, though recoded as a contrast in pitch level. This is the case of **Old Chongming**, a northern Wu dialect, spoken on the outskirts of Metropolitan Shanghai (see Chen and Zhang 1997). What happens in this conservative variety of Chongming can be captured by the following diagram:

(58)

1st σ	2nd σ		3rd σ	
	E	O	E	O
↓	↓	↓	↓	↓
T, T′	H	M	H	o

In trisyllabic compounds, the first syllable retains its underlying tone (T) or its sandhi variant (T'), while the second syllable surfaces with H or M, depending on whether its underlying tone belongs to the so-called E or O class, where E represents the Middle Chinese "even," presumably level, tones (tones I and IV), while O stands for historical "oblique" or contour tones (tones II and III). The third syllable emerges with H or a neutral tone "o," again conditioned by the E/O dichotomy.[24] Here are some examples taken from H-Y. Zhang (1980):

[24] The historical E/O dichotomy does not necessarily translate into level and contour tones in contemporary dialects.

(59) a.	*tsõ*	*tçin*	*li*	"chief manager"
	O	E	O	MC categories
	HMH.	H.	LML	citation tone
	HM	H	o	sandhi form
b.	*fu*	*tsï*	*dziəq*	"vice-chairman"
	O	O	E	MC categories
	M.	HMH.	Lq	citation tone
	HM.	M	Hq	sandhi form

In this respect, the **New Chongming** variety (to be discussed at length in chapters 5–6) represents the most radical degree of neutralization I have observed in any Chinese language. In polysyllabic compounds, *all* tonal distinctions, whether registral or melodic, are completely obliterated, so that the paradigmatic contrasts are replaced entirely by the syntagmatic contrast of initial, medial, or final accent, phonetically realized simply as [H]. In other words, polysyllabic forms can take only one of the following tonal patterns:

(60) H. o. o. o . . .
o. H. o. o . . .
o. o. H. o . . .
o. o. o. H . . . etc.

Exactly where the accent (= H) falls is an intriguing question we will address in chapter 6.

5.2 *Differentiation*

However, one would be mistaken to conclude that all or most sandhi processes lead to neutralization. Tone sandhi can create phonetic contrasts that are absent in the underlying representations (typically the citation forms). Sandhi-triggered phonetic differentiation can arise in two ways. First, historically different tonal categories have merged in citation forms, but are maintained in sandhi contexts. We have seen some examples in chapter 1, section 6. Here are additional examples. Middle Chinese (MC) tones Ib and IIb share an identical citation form [13] in modern **Wenling** (data from R. Li 1979). However, in sandhi contexts (even in identical tonal sequences), their distinct historical origins resurface. This is illustrated below:

(61) a.	*zï ko*	"current price"
	13.55	citation form (*zï* "current," from MC tone Ib)
	13.55	sandhi form

	b.	*zï ko*	"market price"
		13.55	citation form (*zï* "market," from MC tone IIb)
		31.55	sandhi form
(62)	a.	*hɤq dɔ*	"spade (of cards)" (lit. black + peach)
		55q.13	citation form (*dɔ* "peach," from MC tone Ib)
		55q.51	sandhi form
	b.	*hɤq dɔ*	"organized crime" (lit. black + way of life)
		55q.13	citation form (*dɔ* "way of life," from MC tone IIb)
		33q.31	sandhi form

The (62a) and (b) examples above are homophonous in citation form, but pairwise distinct in connected speech. In (61) the same *zï* [13] before the same [55] shows up as [13] in (a), but as [31] in (b), according to their historical tone category. In (62), the same *dɔ* [13] triggers different sandhi alternants in the same morpheme *hɤq* "black" as well as surfacing with two distinct phonetic shapes [51] and [31]. It is clear that the MC categories have survived to this day in modern Wenling. As to which of the morphotonemic variants {13, 31, 51}, or some other "*quartum quid*" should be taken as underlying, that is an analytical decision that must be made in the context of the system of sandhi alternations as a whole.[25]

A second way in which divergent sandhi forms may arise from identical citation/base categories is structure-sensitive tone sandhi. The simplest case is one where tone sandhi applies to a tonal string of one construction type but blocks in an identical tonal configuration of a different morpho-syntactic structure. The best known example occurs in Shanghai, where tonal deletion and spread is limited to lexical compounds, giving rise to such classic minimal pairs as this one:[26]

(63)		fry rice	
	a.	*tsʼɔ vɛ*	"to fry rice" (phrase)
		MH. LH	base tone (tone sandhi not applicable)
	b.	*tsʼɔ vɛ*	"fried rice" (compound)
		MH. LH	base tone
		MH. o	Tone Deletion
		M. H	Tone Spread

We find a far more dramatic and intricate situation in **Old Chongming**, which has not only two subsystems of sandhi rules for lexical and phrasal

[25] For a specific proposal of Wenling underlying tones, see Ting (1984).
[26] We have a great deal more to say about Shanghai in chapter 7, section 3.

constructions, but also different sandhi rules keyed to different postlexical construction types, e.g. one applying to number + measure expressions, another to verb + directional complements, and so forth. Here are some illustrative examples taken from Chen and Zhang (1997):

(64) Lexical vs. phrasal sandhi alternation

a.	*fang-xing*	"to let pass through customs" (lexical)
	M LM	base tone
	HMH. H	sandhi form
b.	*fang ping*	"to lay flat" (phrasal)
	M LM	base tone
	M H	sandhi form

(65) Construction-specific sandhi alternation

a.	*si dun*	"four meals" (number + measure word)
	M M	base tone
	M H	sandhi form
b.	*ci ci*	"every time" (reduplicated nouns, lit. time + time)
	M M	base tone
	HMH.M	sandhi form
c.	*jin qu*	"go in" (verb + directional complement)
	M M	base tone
	HMH.H	sandhi form

(64a) and (b) are a near-minimal pair, with identical base tones /M.LM/. Lexical and phrasal sandhi rules give rise to two distinct tonal sequences. (65a, b, c) all start out with an underlying /M + M/ sequence, but diverge in their sandhi forms, due to the fact that the postlexical sandhi rules are construction-specific, a peculiar characteristic not only of Wu dialects, but also of the Jin group, including Pingyao (see below).

Although unusual in its complexity, Old Chongming is by no means unique in exploiting construction-specific tone sandhi. Structure-sensitive tone sandhi has created massive new contrasts that separate lexical compounds with different meanings from each other, and one construction type from another across a wide spectrum of Chinese languages, especially of the Wu group. **Zhoushan** (represented by the Dinghai varietey), spoken off the north-eastern coast of Zhejiang, is quite typical in these respects. According to Fang's (1987) description, Zhoushan distinguishes three types of tone sandhi: A is generic, applying to all lexical compounds, B applies to verb–object, and C to modifier + head constructions. As a consequence,

identical underlying tonal strings give rise to different sandhi forms depending on which of the three sandhi rules is involved:

(66)

	input string		sandhi form	rule type	example	
a.	53.13	→	33.35	A	*tçi dẽ*	"egg" (lit. chicken + egg)
			33.13	C	*sã dẽ*	"to lay eggs"
b.	34.53	→	53. o	A	*tçi tẽ*	"simple"
			33.53	C	*çia sang*	"to paint/sketch from nature" (lit. write + nature)
c.	34.22	→	53.11	A	*sɤ ngai*	"water buffalo" (water + cow)
			44.44	B	*çiɔ hiã*	"sheep" (small + goat)

It would not do to simply say that Zhoushan segregates lexical compounds from syntactic phrases, as (a) might lead one to believe, because given the idiosyncratic meaning of an expression like *çia sang* "to paint/sketch from nature" (as in an art class), it must in all likelihood be treated like a frozen wordform, a complex predicate with an incorporated NP (cf. Ackerman and Webelhuth (forthcoming) and discussion in chapter 9, section 3 and Appendix). Neither does it suffice to distinguish two construction types: modifier–head vs. verb–object, since both members of the pair in (c) instantiate a modifier–head construction. In any case, the application of different types of sandhi rules has given rise to minimal pairs with divergent meanings:

(67)

a.	black man *haq ning* 5q 22	→	i.	5q. 53	"undocumented vagrant"	(type A)
			ii.	5q. 44	"black person"	(type C)
b.	mind channel *zong tçing* 22 53	→	i.	22. 53	"nerve"	(type A)
			ii.	13. o	"mental disease"	(type C)
c.	hit hand *ta çiu* 34. 34	→	i.	35. o	"thug, hit-man"	(type A)
			ii.	33. 35	"to slap one's hand"	(type B)
d.	cover blanket *ke bi* 44.13	→	i.	44. o	"blanket"	(type A)
			ii.	33.35	"to cover oneself with a blanket"	(type B)

Sometimes different sandhi rules are mediated by structure-sensitive metrical forms. **Xiamen** is a good example. Most lexical and phrasal constructions are right-prominent. However, verb + resultative or directional complements and certain lexical compounds exhibit a left-prominent pattern, with automatic consequences in sandhi alternation:

(68) a. (. x)

kiã si	"coward" (lit. fear + death)
44.53	base tone
22.53	sandhi form

b. (x .)

kiã si	"scared to death"
44.53	base tone
44. o	sandhi form

Such accentual contrast in Xiamen has well-known analogs in many Mandarin dialects, including the following in **Mandarin**:[27]

(69) a. (. . x)

xiang qi lai	"want to get up"
214.214.35	base tone
35.214.35	sandhi form

b. (x . .)

xiang qi lai	"remember" (lit. think + rise + come)
214.214.35	base form
214. o o	sandhi form

T3 Sandhi applies as usual in (a), turning T3 into T2. However, verb + resultative constructions such as "remember" (meaning something like "to think up," "to recall to consciousness" etc.) have a left-prominent accentual pattern that triggers tonal reduction, and hence destroys the requisite context for T2 Sandhi, which applies only to contiguous T3s (see chapter 1, section 4.1, and chapter 7).

5.3 Pingyao: construction-sensitive tone sandhi

I close this section on differentiation with a brief sketch of **Pingyao**, a Jin dialect (all data from Hou 1989). The Jin group is unique among the northern dialects in (i) exploiting structure-sensitive tone sandhi rules, which are common among the Wu dialects, but unattested, to my knowledge, in

[27] 2 and 3 in the examples of (12) represent tonal *categories*, not phonetic values. They are phonetically [35] and [214] respectively.

the northern dialects, including the vast Mandarin group; (ii) having collapsed (in citation form) two MC tonal categories Ia and Ib, which have split in virtually all other Chinese dialects, both northern and southern.[28] In connected speech, however, Ia and Ib behave quite differently. Pingyao, therefore, instantiates both types of differentiation. Pingyao has the following tonal system based on citation forms:

(70)

citation form	MC categories
13	Ia, Ib
53	II
35	III
23q	IVa
54q	IVb

Though merged in citation forms, the resurfacing of the underlying contrasts between MC Ia and Ib in sandhi contexts is illustrated by such minimal pairs as:

(71)

a.	*tçʼiang piʌq*						
	13 13q	→	i.	31.35q	"pencil"	(from Ia + 13q)
			ii.	13.13q	"wall"	(from Ib + 13q)
b.	*tung tçʼyə*					
	13 35	→	i.	13.13	"rainbow"	(from Ia + 35)
			ii.	31.35	"coppersmith"	(from Ib + 35)

It is obvious that Ia and Ib have retained their separate underlying identities, even though both are pronounced with a low-rising tone [13] in citation forms. There is no reason to equate the underlying input to sandhi rules with the citation form.

The second type of tonal differentiation arises out of structure-discriminating tone sandhi. Pingyao distinguishes three types of tone sandhi: A is limited to subject–predicate or verb–object constructions; C targets only reduplicated verbs; and B applies to all other structures,

[28] Another peculiarity of Jin is the survival of the MC tone IV associated with CVq syllables. This was the single most important reason that prompted R. Li (1985) to separate Jin from the surrounding Mandarin dialects. Having preserved MC tone IV hardly justifies setting up Jin as a distinct dialect group: after all, it is well known that the south-eastern branch of Mandarin (i.e. the so-called Jiang-huai branch) is characterized precisely by its retention of CVq. However, from the dynamic perspective of sandhi processes, there is ample justification for regarding Jin as a distinct group of northern dialects. As far as I know, no Mandarin dialect displays the kind of construction-sensitive tone sandhi exemplified by Pingyao, and probably other Jin dialects as well.

including modifier–head and duplicated nouns. A few examples will illustrate this point:[29]

(72) a. open/start car
k'ae tš'ïE
13 13 → i. 13.13 "to drive a car" (type A)
(verb–object)
ii. 31.35 "open cart" (type B)
(modifier–head)

b. open open
k'ae k'ae
13 13 → i. 31.35 "idea, plan" (type B)
(reduplicated noun)
ii. 31.35 "can be opened" (type B)
(verb–complement)[30]
iii. 35.31 "to open a little" (type C)
(reduplicated verb)

Both *k'ae* "open/start" and *tš'ïE* "car, vehicle" derive from MC tone Ia, so that their underlying tonal identity is not in question. From one identical tonal string /13.13/ as input, we derive three different sandhi outputs [13.13] (type A, tone sandhi vacuous), [31.35] (type B), and [35.31] (type C).

Minimal pairs are not limited to underlying /13.13/ sequences. Here are a few doublets having other tonal combinations.

(73) a. *kung suei*
53. 53 → i. 35.513 (A) "to boil water"
ii. 53.53 (B) "boiling water"

b. *tung šəu*
35. 53 → i. 35.513 (A) "to start a scuffle" (lit. move + hand)
ii. 35.53 (B) "cold hands"

c. *ung ung*
35. 35 → i. 35.53 (B) "pot" (reduplicated noun)
ii. 53.31 (C) "to ask" (reduplicated verb)

d. *çîq çîq*
54q.54q → i. 53.13 (B) "straw mat" (reduplicated noun)
ii. 513.53 (C) "to learn" (reduplicated verb)

[29] *Tš'ïE* 13 and *k'ae* 13 in each instance is written with the same Chinese character. The morpheme *tš'ïE* 13 is glossed as "car, vehicle"; whether *k'ae* 13 with the core meaning "to open" or "to start" (as in starting an engine) constitutes one or several morphemes is an open question.

[30] Corresponding to Mandarin *kai de kai* meaning "succeed in opening" or "can be opened."

Doublet (a) constitutes a true minimal pair: the two expressions stringing together two identical words "boil" + "water" are pronounced differently depending on the grammatical relation holding between the constituent parts: [35.513], via sandhi type A, if construed as a verb–object construction, [53.53] otherwise. The examples of (b) show that sandhi A and B are not reducible to phrasal vs. lexical rules. If anything, type A resembles a postlexical rule, since it applies to more or less free and productive subject–predicate and verb–object constructions. By equating sandhi A with postlexical rules, we are hard put to explain why an expression like *tung šəu* "to start a scuffle" undergoes A-sandhi rather than B-sandhi: like the Zhoushan example cited earlier *çia sang* "to paint/sketch from nature," *tung šəu* "to start a scuffle" is highly lexicalized with an idiosyncratic meaning not transparently derivable from its constituent parts "move" + "hand."

While it is possible to summarize the three types of sandhi alternations either in tabular or rule form, the principles behind the inductive summaries are more elusive. I will attempt only a condensed account of type A sandhi. To facilitate the presentation of facts, I will make a number of simplifying assumptions. First, I will consider the checked tones 23q and 54q as allotones of their "legato" counterparts, namely 13 and 53; in fact the two sets of allotones behave exactly alike with respect to type A sandhi. Second, 53 has a contextually determined variant: in the final position, [53] is phonetically realized with an upglide: [423]. In the following account I will ignore this last phonetic detail. With these caveats in mind, we can lay out the essential facts as reported in Hou (1989) in a tabular form as follows:

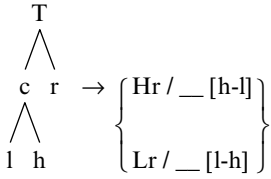
(74)

		second σ		
		13	35	53
first σ	13		31-T	35-T
	35	13-T	31-T	35-T
	53		53-T	35-T

T = base tone, unchanged
shaded areas: no change

The second syllable resists any change. The phonetic shape of the initial syllable is determined by two principles, one governing the register, the other the melodic shape, stated below:

(75) Register Neutralization



(76) Contour Metathesis
[x-y] → [y-x] / ___ [x-y, Hr]
x, y = h or l

Register Neutralization says in effect that a rising tone is high registered before a falling tone, and low registered before a rising tone. I owe the Contour Metathesis rule to Bao (1990a), but have compressed it into a simpler notation in order to bring out the underlying mechanism. It says basically that a rising tone (where [x y] = [l-h]) turns into a falling tone if followed by another *high*-rising tone; conversely, a falling tone (where [x-y] = [h-l]) assumes the opposite melodic shape when followed by another *high* falling tone. The high register condition on the second tone is symbolized by Hr as an integral part of the tonal complex associated with the second syllable. Both Metathesis and Neutralization act like WFCs to be simultaneously met by the sandhi forms. How each of the nine disyllabic tone patterns of (71) is derived via these two rules is demonstrated below:

(77)

base tone	Register Neutralization	Contour Metathesis
13.13	vacuous	n/a
13.35	vacuous	31.35
13.53	35.53	vacuous
35.13	13.13	n/a
35.35	13.35	31.35
35.53	vacuous	vacuous
53.13	n/a	n/a
53.35	n/a	vacuous
53.53	n/a	35.53

Register Neutralization applies vacuously to the input string /13.13/; since the second rising tone is low registered, Contour Metathesis is inapplicable. Consequently, /13.13/ surfaces unaltered. /13.35/ and /13.53/, on the other hand, undergo Metathesis and Neutralization respectively to emerge

as [31.35] and [35.53]. Finally, /35.35/ is subject to both Neutralization and Metathesis, consequently the initial tone /35/ is both lowered in register (→ 13) and inverted in melodic shape (→ 31).

Appendix Tone features

		55	44	33	22	11
Gruber (1964)	high	+	+		–	–
	$high_2$	+	–		+	–
Wang (1967)	high	+	+	–	–	–
	central	–	+	+	+	–
	mid	–	–	+	–	–
Sampson (1969)	high	+	+	–	–	–
	low	–	–	–	+	+
	central	–	+	+	+	–
Woo (1969)	high	+	+	–	–	–
	low	–	–	–	+	+
	modify	–	+	–	+	–
Halle and Stevens (1971)	stiff	+		–		–
	slack	–		–		+
Maddieson (1972)	high	+	+	–	–	–
	low	–	–	–	+	+
	extreme	+	–	–	–	+
Yip (1980)	upper	+	+		–	–
	raised	+	–		+	–
Clements (1985)	root-tier	H	H		L	L
	tone tier	H	L		H	L
Hollenbach (1988)	high	(+)	+	+	–	–
	central	(–)	–	+	+	–
	extreme	+	–	(–)	(–)	(–)
Hyman (1993 [1989])	tone tier	H		H,L		L
Bao (1990a)	stiff	+	+		–	–
	slack	–	+		–	+
Duanmu (1990a)	stiff	+	+	–	–	–
	slack	–	–	–	+	+
	above	+	–	–	–	–
	below	–	–	–	–	+
Chang (1992)	stiff	+	+		–	–
	sp. Glottis	–	+		–	+
	const. Glottis	u	u		u	u

Not commensurate with any of the systems listed above is Tsay (1994), who argues, very cogently and insightfully in many respects, against the conventional wisdom of cutting up the tonal space into subranges that is implicit in all current models. Instead, she employs basically Chao's tone letters. Thus a high-rising and a low-falling tone are represented simply as [35P] and [31P], where P = linguistic pitch.

3 *Directionality and interacting sandhi processes I*

1 The nature of the problem

In this chapter and the next we take up the question of how sandhi processes are implemented and how they interact with each other.

Certain types of tone sandhi operate across the board within a certain domain. Thus, the Wu dialect of Tangxi, a left-prominent and right-spread prototype, retains the initial tone, obliterates all subsequent tones, and extends the initial tone rightwards over the entire phonological word, regardless of either the number of syllables or how the syllables are structured prosodically or morphosyntactically (see chapter 7, section 2.2). Similarly, within a tone group of indefinite length, Xiamen simply replaces every non-final base tone by its corresponding sandhi tone (see chapter 10). In either case, tone sandhi generates the output in one sweep, so to speak, and there is no question about how sandhi processes interact with each other.

There are of course sandhi processes of a more "local" nature which do not interact with each other; in such cases the output simply represents the sum of the changes brought about by the individual rules. The simplest case is illustrated below. The north-eastern Min dialect of **Fuqing** reported in Feng (1993) has these two rules: (i) one turns a non-final high-falling tone HM into H, as shown in (1a) and (b); (ii) the other lowers a high tone H when followed by a low tone L, as instantiated in (1c). The trisyllabic sandhi forms of (1d) simply display the combined effects of the two elementary operations (i) and (ii). Neither operation affects the outcome of the other: the two rules do not interact.

(1)	a.	*jian tou*	"shoulders"
		HM.H	base form
		H H	sandhi form
	b.	*bao cai*	"cabbage"
		HM.L	
		H L	

c. *huang di* "the emperor"
H L
L L

d. *[bo ling] cai* "spinach"
HM H L
H L L

More often than not, however, the elementary operations on different input substrings potentially intersect each other, for instance by creating or undoing the sandhi contexts specified for the rules, thus giving rise to the familiar feeding, bleeding, and other ordering relations. Take the case of **Yantai**, a northern Mandarin dialect reported in Qian et al. (1981). This dialect has an inventory of four citation tones: /31a, 31b, 214, 55/.[1] Each citation tone has two to three different phonetic realizations, depending on the sandhi context, as stated in (2).

(2) Disyllabic tone sandhi

a.	31a	→	35 / ___ 31	
b.	31b	→	55 / ___ 31	
c.	55	→	31 / ___ 55	
d.	214	→	35 / ___ 31	(i)
			55 / ___ 214	(ii)

Now, given an input string like /214.214.31/, there are two possible sandhi forms, depending on how we scan the input string, left to right, or in the opposite direction, as shown in (3a, b). For clarity, here I adopt the notational convention of underlining the two-tone "window" that a local sandhi rule scans for possible application, and highlighting a localized change with a vertical shaft linking an input tone (top) to its corresponding output (bottom).

(3) a. 214.214.31
| by rule (2d-ii)
55.214.31
| by rule (2d-i)
55. 35. 31

[1] /31a/ and /31b/ derive from different historical sources, namely *yin ping* (high-register *ping*) and *yang ping* (low-register *ping*), respectively. Although they have merged phonetically in citation form, they maintain their underlying contrasts in sandhi positions, as clearly indicated in rule (2a) vs. rule (2b).

b. 214.<u>214.31</u>

| by rule (2d-i)

<u>214.35</u>.31 (n/a)

<u>TT</u> = two-tone local "window"
n/a = no applicable rule to local window

Surprisingly, neither (3a) nor (3b) is the sandhi form for an expression like *wu dian zhong* "five o'clock" (= (5e) below). Remarkably, what happens is that Yantai partitions its sandhi rules into two subsets: one, more context-sensitive, governs pairs of adjacent tones, while another, more sweeping in scope, applies to longer tonal strings. Accordingly, longer concatenations of tones are affected by a totally different set of rules, which have the effect of massive neutralization, in effect reducing the sixty-four combinatorial possibilities ($4 \times 4 \times 4$) to exactly three. Thus, all the nine three-tone sequences of (5a–i) collapse into one single sandhi form [31.35.31], by virtue of rule (4a). As stated in (4), the sandhi forms of trisyllabic expressions are uniquely determined by the last syllable, regardless of the underlying tonemic distinctions associated with the first two syllables.[2] For our present purpose, the important observation to make is that (2) and (4) apply to two disjoint, non-intersecting domains. Again, problems of rule interaction and implementation do not arise.

(4) Trisyllabic tone sandhi

a. T.T.31 → 31.35.31
b. T.T.214 → 55.55.214
c. T.T.55 → 33.21.55
T = any tone

(5)

			base form	sandhi form
a.	*[tuo la] ji*	"tractor"	31.31.31	31.35.31
b.	*gao [zhi biao]*	"high target"	31.214.31	same
c.	*xin [da yi]*	"new overcoat"	31.55.31	same
d.	*[mei ren] jiao*	"canna, Indian shot"	214.31.31	same
e.	*wu [dian zhong]*	"five o' clock"	214.214.31	same
f.	*da [li hun]*	"get divorced"	214.35.31	same
g.	*[jiao ke] shu*	"textbook"	55.35.31	same
h.	*nao [shui zai]*	"have a flood"	55.214.31	same
i.	*[dian hua] ji*	"telephone receiver"	55.55.31	same

Dramatic as the Yantai case is, it is hardly the norm. As one surveys the sandhi patterns across a fair sample of Chinese dialects, one inevitably

[2] Quadrisyllabic and longer expressions typically break up into disyllabic and trisyllabic chunks (Qian et al. 1981:77).

arrives at the following fundamental generalization: in the majority of cases, polysyllabic sandhi forms reflect the aggregate of incremental changes effected by the same "local" processes that operate on two-tone sequences. Hence, there arises the question: How do elementary processes combine and interact to produce the attested polysyllabic outputs?

Let me illustrate the nature of the problem with a schematic example. Let A, B, C, etc. stand for various tonal categories. The hypothetical rules (6a) and (6b) would generate different outputs depending on how they interact. Scanning the string ABC left to right, the joint operation of (6a) and (6b), in that sequence, produces XZC as output (= (7a)). Working in the opposite direction, (6b) and (6a) give rise to YZC as the end result (= (7b)).

(6) a. A → X / __ B
Y / __ Z
b. B → Z / __ C

(7) a. left to right: $\underline{\text{A B}}$ C
| by rule (6a)
X $\underline{\text{B C}}$
| by rule (6b)
X Z C

b. right to left: A $\underline{\text{B C}}$
| by rule (6b)
$\underline{\text{A Z}}$ C
| by rule (6a)
Y Z C

$\underline{\text{TT}}$ = current "window" spanning a two-tone substring (same below)

Another type of interaction is exemplified below. Rules (8a) and (8b) stand in what is referred to as mutual bleeding relation: the application of one rule preempts that of the other. We get alternative outputs (9a) or (9b) depending on which process takes precedence over the other.

(8) a. B → X / A ___
b. B → Y / __ C

(9) a. left to right: $\underline{\text{A B}}$ C
| by rule (6a)
A $\underline{\text{X C}}$
rule (8b) not applicable
A X C

b. right to left: A B C
| by rule (8b)
A Y C
rule (8a) not applicable
A Y C

It is important to note at the outset that the question of rule implementation is not reducible to the classic problem of rule ordering. (Standard) Mandarin provides a prime example to illustrate this point. The T3 Sandhi rule turns a "dipping," i.e. a falling-rising tone [214] into a rising tone [35], when followed by another dipping tone.

(10) T3 Sandhi
214 → 35 / __ 214

Since we are dealing with one single rule, the question of rule ordering cannot in principle arise. Nevertheless, depending on how we apply T3 Sandhi to input strings we get different outputs. Applied left to right, the rule yields [35.35.214] (= 11a); applied in the opposite direction, the change from [214] to [35] on the second syllable *lao* destroys the requisite context for the same operation on *zhi*, as illustrated by (11b). The broader question that arises is this: what general principles determine the mode according to which the elementary processes interact to produce the attested composite results. Put differently, what mechanism regulates the traffic of phonological rules? In the Mandarin case, the answer is obvious: tone sandhi applies cyclically on morphosyntactic constituents.[3] But more complex modes of rule application are far from unusual, as we shall see.

(11) a. tiger gall "brave"
[*lao-hu*] *dan*
214.214.214
|
35.214.214
|
35. 35. 214

b. paper tiger "coward"
zhi [lao-hu]
214.214.214
|
214. 35. 214 (n/a)

n/a = no applicable rule to current window (underlined)

[3] This is not quite true. See chapter 9 for details.

On the descriptive level, the problem confronting the linguist is brought home most vividly by what may strike most readers as a curious format widely used in reporting Chinese dialectological data. For illustration, take the Hakka dialect of Changting (to be discussed in chapter 4). Like Mandarin, this dialect has a fairly simple tone system, with five base tones. But unlike Mandarin, where tone sandhi is restricted, by and large, to a sequence of two adjacent [214]s, Changting exhibits far more intricate patterns of tonal coarticulation. Thus, corresponding to the 125 possible three-tone sequences ($5 \times 5 \times 5$), there are 69 different sandhi forms.[4] To exhaustively and relentlessly list the sandhi forms of trisyllabic expressions – let alone longer sequences – would be as futile as it is unrevealing. But, that is exactly what most dialectological reports do! Thus both Luo (1982) and Rao (1987), the two primary sources on Changting, devoted page after page to long lists of two- and three-syllable sandhi forms. (Not surprisingly, the authors gave up after trisyllabic expressions.) This format is repeated over and over for a large number of other dialects, notably of the Mandarin and Wu subfamilies. It should be noted that these fieldworkers are fully cognizant of the obvious, if not straightforwardly statable, relationship between the disyllabic base and the longer sandhi patterns. For instance, to paraphrase Luo (1982:192), the author stated quite explicitly that the trisyllabic sandhi patterns of ABC are the composites of two overlapping subpatterns, namely AB and BC. But she failed to spell out how exactly the elementary processes operating on AB and BC combine to produce the ultimate result of ABC.[5] This is, in its generalized form, the crux of the problem we will address in this chapter and the next.

The problem would be relatively trivial if the potential conflicts were resolvable by appealing to cyclicity, simultaneous application, or consistently rightward or leftward iteration, or some such straightforward principles. The tack I am taking points to a number of general constraints that are in all likelihood rooted in language processing. Specifically, I will appeal to such notions as (i) Structural Affinity, one manifestation of which is the familiar phonological cycle;[6] (ii) Temporal Sequence, which

[4] Discounting 31 combinations which do not undergo any sandhi. My figures are based on Hsu (1994).

[5] It is precisely the lack of well-understood general principles of rule interaction that forced authors of similar dialectological reports to resort to a tabular form of summarizing the facts, in keeping with the common practice in such leading journals as *Fangyan* and *Zhongguo Yuwen*.

[6] For a summary and review of the history and current issues related to cyclic rules, see Cole (1995) and copious references cited therein.

favors a left to right direction of scanning, in tandem with the temporal order of planning, organization, and execution of speech events; and (iii) Derivational Economy. One particular subcase of Derivational Economy is (iv) avoidance of what I call Backtracking. Backtracking describes a sequence of events depicted below.

(12) a. A → X / ___ B
b. B → Y / ___ C
c. X → Z / ___ Y

(13)

```
A B C
|          by rule (a)
X B C
  |        by rule (b)
X Y C
|          by rule (c)
Z Y C
```

Suppose a language has the three rules stated above as (12a, b, c). Scanning a string ABC from left to right, the substring AB first turns into XB by virtue of (a). Next, moving the two-segment window of scansion to BC, we apply rule (b), thereby producing XYC as output. At this point, we need to "backtrack" and change XY to ZY in accordance with rule (c). Presumably such a mode of processing is especially "costly" and is generally avoided.[7]

The discovery of the general principles governing rule implementation and interaction is in itself a worthy pursuit in the tradition of a long-standing debate regarding intrinsic vs. extrinsic rule ordering, and multiple application of iterative rules (for an overview of the issues, see Kenstowicz and Kisseberth 1977, chapters 4 and 5, inter alia; for a historical survey, see Iverson 1995; cf. Bromberger and Halle 1989). The enterprise takes on added significance in the context of non-derivational models of phonology, notably Harmonic Phonology (Goldsmith 1993b), Cognitive Phonology (Lakoff 1993), Two-Level rules (Koskenniemi 1983, Karttunen 1993, Kaplan and Kay 1994), and Optimality Theory (OT; Prince and Smolensky 1993, McCarthy and Prince 1993a, 1995a). Let us suppose that the outputs of (7a, b) as well as (9a, b) are all equally well-formed or optimal relative to the input. In such cases, strictly monostratal output constraints and relational faithfulness conditions will necessarily underdetermine the

[7] Cf. "Strict Cyclicity" and "Reaching Back" constraints (see Cole 1995).

choice among the competing sandhi forms. As I will argue at length in this chapter and the next, in order to account for a rich and highly complex array of tone sandhi facts, it is necessary to exploit and extend in unconventional ways some of the seminal insights of OT. Specifically, we need to entertain a "harmonic serialist" conception of OT, which makes allowance for constraints on derivations as well as constraints on outputs (see section 7).

In the sections that follow I will first consider in some detail the relatively simple case of Tianjin to motivate an analysis that appeals to constraints that are conceived as constraints not on the output per se, but on the derivation itself. We will then extend the proposed approach to the analysis of other, in some cases considerably more complex, sandhi systems in the next chapter.

The rest of this chapter is organized as follows: after sketching the relevant facts of Tianjin couched in conventional processual terms in section 2, I will present a straightforward analysis that crucially exploits the notion of constraints on derivation in sections 3–4, then consider the question of whether we can achieve the same effects through non-derivational means, first in OT (section 5), then in two-level rule phonology (section 6). Finally, we explore a serialist version of OT in section 7, and close with concluding remarks in section 8.

2 Tianjin: directionality effect

We begin by examining a fairly simple and self-contained set of facts drawn from **Tianjin** to motivate what appears to be a set of constraints on derivation and to illustrate the approach I am adopting here. Tianjin is a northern Mandarin dialect, spoken in a major metropolitan area of over 6 million residents,[8] located about 100 km south-east of Beijing. It has a four-tone system described in slightly different phonetic terms by different authors.

(14)	21	45	213	53	(according to Li and Liu 1985)
	11	55	24	53	(according to Shi 1990)
	L	H	R	F	(symbols used here)

Ignoring some of the redundant details encoded in Li and Liu's (1985) transcriptions, and relying on Shi's (1990) experimental evidence, I takes these four categories to represent low (level), high (level), rising and falling

[8] 1974 figures.

tones, and will henceforth refer to them by the mnemonic symbols L, H, R, and F respectively. With the exception of Hs,[9] two identical tones may not abut each other (a fairly common manifestation of the Obligatory Contour Principle or OCP); consequently, they must undergo a dissimilatory process informally stated as (15), illustrated with appropriate examples on the right.

(15) Dissimilation[10]

a.	LL	→	RL	*fei ji*	"airplane"
b.	RR	→	HR	*xi lian*	"wash one's face"
c.	FF	→	LF	*jing zhong*	"net weight"

In addition, Tianjin not only bars adjacent identical tones, it also disallows certain *partially* identical sequences of tone segments. In particular, it prohibits the sequence FL, which is more perspicuously spelled out as H<u>L.L</u> – with two successive lows. When such a sequence arises, it undergoes a common process called Tonal Absorption (16), which has its analogs in many West African languages.[11]

(16) Tonal Absorption

FL	→	HL	*ren zhen*	"earnest"

This much is unremarkable. The analytical task becomes more challenging and interesting when we look at tone sandhi in longer constructions. Limiting ourselves to trisyllabic sequences, we have 64 combinatorial possibilities (= 4^3). Of these, 37 combinations are sandhi-free, i.e. not subject to either Dissimilation or Absorption. Thus underlying tonal strings like LRF and RFH freely occur in their unaltered form. In fact, all sequences of the form THT never undergo tone sandhi, since H freely combines with any tone (T) in any linear order. Twenty other three-tone sequences call for only one single adjustment. For instance, /FLR/ → [HLR] (by Absorption),

[9] It is not unreasonable to assume, along with Yip (1989), that H is underlyingly unspecified, and therefore does not participate in the dissimilatory process.

[10] In principle OCP can be assured by a "progressive" dissimilation, whereby LL → LR. As a matter of fact, dissimilation is uniformly "regressive" in Tianjin (and Mandarin dialects in general). This is due to the right-prominent pattern characteristic of Mandarin (and Min) dialects (cf. Hoa 1983, Duanmu 1993a, 1995). Right prominence is diagnosed, inter alia, by the tonal stability or resistance to change of the syllable/constituent on the right.

[11] Hyman and Schuh (1974) cite Bamileke, Mende, Kikuyu, Hausa, Ngizim, among others. In Hyman and Schuh's conception, tonal absorption is implemented by rightward shift, with automatic OCP effect.

```
σ  σ        σ  σ        σ  σ
/\ |   →    |  /\   =   |  |
HL L        H  LL       H  L
LH H        L  HH       L  H
```

Table 3.1

	Input	Output	[x x] x	x [x x]	[x x x]
P1	FFL	LHL	*[si.ji] qing* "evergreen"	*zuo [dianche]* "take a tram"	
P2	RRR	HHR	*[li.fa] suo* "barber shop"	*mu [lao.hu]* "tigress"	*ma.zu.ka* "mazurka"
P3	FFF	HLF	*[su.liao] bu* "plastic cloth"	*ya [re.dai]* "subtropical"	*yi.da.li* "Italy"
P4	LLL	LRL	*[tuo.la] ji* "tractor"	*kai [fei.ji]* "pilot a plane"	
P5	RLL	HRL	*[bao.wen] bei* "thermos cup"	*da [guan.qiang]* "speak in a bureaucratic tone"	
P6	LFF	RLF	*[wen.du] ji* "thermometer"	*tong [dian.hua]* "make a phone call"	
P7	FLL	FRL	*[lu.yin] ji* "cassette recorder"	*shang [fei.ji]* "board an airplane"	

and /RFF/ → [RLF] (by Dissimilation). No further sandhi process is applicable to outputs [HLR, RLF]. The remaining seven tonal combinations potentially involve more than one sandhi process, and will therefore constitute the focus of our interest.

The seven tonal combinations in question are summarized and exemplified in table 3.1, and will be referred to as P1–7 (patterns 1–7) in our ensuing discussion. The input and output forms are given in the second and third columns. The last three columns indicate the constituency structure of the trisyllabic expressions. Left- and right-branching structures are indicated as [x x] x and x [x x]. Thus *si ji*, literally "four season," modifies *qing* "green," to yield the compound for "evergreen," i.e. "green all four seasons," etc. There are two instances of transliterations of foreign words for "mazurka" and "Italy." They are internally unstructured, and therefore represented as [x x x].

Two of the seven tone patterns are "unidirectional" in the sense that multiple application of tone sandhi can apply only in one, predictable direction, in both cases right to left. For instance, in P5, the substring RL in /RLL/ is not susceptible to tonal modification until the second half LL is changed into RL, thereby creating a sequence of two Rs, triggering the change RR → HR, in a chain-reaction as it were. The same goes for P6.

(17) a. P5 R L L
| by Dissimilation (a)
R R L
| by Dissimilation (b)
H R L

b. P6 L F F
| by Dissimilation (c)
L L F
| by Dissimilation (a)
R L F

The interplay among the sandhi processes becomes more intricate when different sandhi processes can operate on different substrings all at once. Depending on which operation takes place first, the same input string may bifurcate in its derivational paths, and wind up with quite distinct sandhi forms. This is the case with the remaining five tonal patterns P1–4 and P7. Contrast P2 and P4 below:

(18)

	⇨		⇦		examples	
P2	a.	R R R	b.	R R R	*[li fa] suo*	"barber shop"
		| by Dism-b		| by Dism-b	*mu [lao hu]*	"tigress"
		H R R		R H R	*ma zu ka*	"mazurka"
		| by Dism-b		(n/a)		
	☞	H H R				
P4	a.	F F F	b.	F F F	*[su liao] bu*	"plastic cloth"
		| by Dism-b		| by Dism-a	*ya [re dai]*	"subtropical"
		L F F		F L F	*yi da li*	"Italy"
		| by Dism-c		| by Abs		
		L L F	☞	H L F		

P1, P2 etc. stand for Pattern 1, 2 and so forth
⇨, ⇦ symbolize directionality of rule application
Dism, Abs = Dissimilation, Absorption stated earlier
☞ marks attested sandhi forms
T T = current local window
n/a = no applicable rule to current window

The crux of the problem confronting the linguist is to find some general principle or principles which govern the traffic of sandhi operations and guarantee a derivational path leading to the desired output. We can discard immediately a number of obvious but unworkable hypotheses. First of all, a consistently left-to-right or right-to-left iterative rule application fails to derive the attested outputs, as clearly demonstrated by the examples given above. Cyclic application can be ruled out for the simple reason that the directionality of sandhi operations is blind to the constituent structure (if any) of the trisyllabic expressions. For instance, P2 /RRR/ undergoes TS from left to right, regardless of whether the string has a left-branching (as in [*li fa*] *suo* "barber shop," lit. fix hair + shop), or right-branching structure (as in *mu* [*lao hu*] "tigress," lit. female + tiger), or no structure at all (as in *ma zu ka* "mazurka"). Exactly the opposite is true of /FFF/. Notions like transparency and derivational economy also fail to make the right predictions in any straightforward manner. Thus, a right-to-left application of Dissimilation to P2b /RRR/ is both transparent and derivationally simpler, while the same process working in the opposite direction in the case of P2a is derivationally more complex, and functionally opaque in the sense that the initial R turns into a H, even though, judging by the surface representation, it does not occur in front of another R, as the structural description of Dissimilation (clause b) demands. In other words, P2a instantiates a case of overapplication or counterbleeding rule relation. One would, therefore, expect transparency as well as derivational economy to dictate the choice of P2b as the correct alternative, contrary to the attested facts.

Finally, one may invoke extrinsic order as a last resort. Hung (1987b), for instance, posits the following ordering relations (X > Y means rule X precedes rule Y):[12]

(19) Dissimilation-c > Dissimilation-a
Dissimilation-a > Dissimilation-b
Dissimilation-a > Absorption
Dissimilation-c > Absorption

Aside from being merely stipulative, extrinsic order still fails to account for the observed directionality effect, which is diagonal to the precedence

[12] Tan (1987) reverses the order between Dissimilation-a and Dissimilation-c, reflecting certain idiolectal differences. Hung's order is consistent with Li and Liu (1985) and corroborated by other Tianjin informants. For details, see Chen (1987b), Hung (1987b).

relation holding between two different (sub)rules. Recall P2 /RRR/ in (18). There only one single subrule (Dissimilation-b) is relevant. Therefore, the question as to which rule precedes which does not arise. And yet, the correct output can be derived only by applying the relevant rule in one direction or the other.

All attempts to date have failed to render an account of the facts of Tianjin tone sandhi. This is, in short, what I once characterized as the ordering paradox of Tianjin tone sandhi in a paper (Chen 1986), which touched off a lively debate in the literature.[13] In a sense, ordering paradox is a misnomer. The problem goes beyond rule ordering; it relates in general to the way phonological processes are implemented.

3 A derivational account

3.1 Temporal Sequence

The most satisfactory analysis of the Tianjin facts turns out to be surprisingly simple. The analysis rests on two key notions: Well-formedness and Temporal Sequence. We noted earlier that Tianjin bars two types of tonal juxtapositions: (a) a sequence of two contiguous identical tones (except HH); and (b) a falling tone followed by a low tone, or FL. The former is a reflection of the usual OCP, where identity is defined over the entire tonal complex. Thus the sequence FR is structurally [H<u>L</u>][<u>L</u>H], consisting of two *distinct* contour tones – even though there are two abutting identical terminal tone segments (underlined). The latter, (b), may be regarded as a prohibition of *partial* identity in narrowly defined cases (i.e. FL = H<u>L.L</u>, with two adjacent lows). We may refer to (20b) as partial OCP, which I will label as OCP′. Dissimilation and Absorption jointly guarantee the satisfaction of WFC.

(20) Well-formedness conditions (WFC)
  a. OCP: no adjacent identical tones (except HH)
  b. OCP′: no FL sequence

The second key element is the principle of Temporal Sequence, which makes the default assumption that we apply rules left to right, in tandem with the planning and execution of speech. By coupling Temporal Sequence with WFC, we can express the core generalization as follows:

[13] For references, see Bibliographical Appendix, under the subheading 3.2.

(21) Generalization
By default rules apply from left to right – unless such a mode of application produces an ill-formed output, in which case the direction of operation is reversed.[14]

Exploiting the device of constraint ranking and extending the notion of constraints to include constraints on derivations as well as outputs (see discussion below), we can formulate Temporal Sequence in its stark simplicity. By ranking WFC above Temporal Sequence as in (23) we get exactly the right results.

(22) Temporal Sequence: apply rules left to right.

(23) WFC ≫ Temporal Sequence.[15]

This is illustrated by the following tableau.

(24)

					WFC	Prmpt	Temp
P1	☞	a	⇨	F F L | L F L | L H L			
		b	⇦	F F L | F H L (n/a)		*	*
P2	☞	a	⇨	R R R | H R R | H H R			
		b	⇦	R R R | R H R (n/a)			*

[14] We find a similar directionality in syllabification as reported by Noske (1993).
[15] "Con_i ≫ Con_j" reads as: constraint (i) dominates – that is, overrides – constraint (j).

(24) cont'd

					WFC	Prmpt	Temp
P3		a	⇨	F F F | L F F | L L F	*		
	☞	b	⇦	F F F | F L F | H L F			*
P4		a	⇨	L L L | R L L | R R L	*		
	☞	b	⇦	L L L | L R L (n/a)			*
P5		a	⇨	R L L (n/a) R L L | R R L	*		
	☞	b	⇦	R L L | R R L | H R L			*
P6		a	⇨	L F F (n/a) L F F | L L F	*		

(24) cont'd

					WFC	Prmpt	Temp
	☞	b	⇦	L F F | L L F | R L F			*
P7		a	⇨	F L L | H L L | H R L		*	
	☞	b	⇦	F L L | F R L (n/a)			*

Temp = Temporal Sequence
Prmpt = Preemptive clause, see below
n/a = no rule applicable to the current local window (underlining) spans a two-tone substring.

The left-to-right (symbolized as ⇨) application of Dissimilation and Absorption to P1 and P2 yields perfectly well-formed tonal strings LHL and HHR, respectively. These are then, unremarkably, the attested outputs. However, the same mode of rule application produces an ill-formed string for each of the patterns P3–6. For instance, a serial left-to-right application to P3 /FFF/ ends in the string LLF (P3a), which is inconsistent with OCP. Under such circumstance, the rules apply in the opposite direction (as indicated by the arrow "⇦"), with HLF as the end result (P3b). In other words, violation of Temporal Sequence is forced by the overriding imperative of WFC.

3.2 Preemptive clause

There is one pattern /FLL/ (= P7 in the tableau in (24)) we have not yet dealt with. A straightforward left-to-right application yields a perfectly well-formed output HRL (P7a). But surprisingly it is FRL (P7b) that is picked as the optimal candidate. What tips the balance in favor of

FRL? It cannot be derivational economy because, as P1–2 demonstrate, Derivational Economy plays a subordinate role to Temporal Sequence: in both of these cases it is the longer derivation that yields the desired outputs (P1a and P2a). What determines the choice of the two alternative paths of derivation for P7 seems to be this: when a string simultaneously violates OCP (total identity) and OCP′ (partial identity), one must *first* undo OCP violation by means of Dissimilation. Given that the two substrings of /FLL/ violate both OCP (F<u>LL</u>) and OCP′ (<u>FL</u>L), Dissimilation must apply first to undo OCP violation, yielding FRL. If such a serial operation results in a shorter derivation, that is merely serendipitous. Needless to say, this amounts to saying that Dissimilation is ordered before Tonal Absorption. For want of a better term, I will use the label "preemptive clause" stated as (25) to refer to this ordering relation between Dissimilation and Absorption as instantiated by P7 in the tableau in (24).

(25) Preemptive clause
When a string simultaneously violates OCP and OCP′, first undo OCP violation (by Dissimilation).

Needless to say, the Preemptive clause and Temporal Sequence sometimes converge in a left-to-right operation. This is illustrated by P1: P1b simultaneously violates Preemptive clause and Temporal Sequence, and hence loses out to competitor P1a.

3.3 No-Backtracking

Let us make explicit one implicit assumption. Straight rightward or leftward applications do not exhaust the logical possibilities. There is a third mode of rule application, namely "persistent" rule application: apply rules whenever applicable (cf. Chafe 1968, Myers 1991). This is illustrated by option (c) in P3 and P4 in the tableau in (26).

(26)

				WFC	NoBT	Temp
P3	a	⇨	<u>F F</u> F | L <u>F F</u> | *L L* F	*		

(26) cont'd

					WFC	NoBT	Temp
	☞	b	⇦	F <u>F F</u> | <u>F L</u> F | H L F			*
		c	BT	<u>F F</u> F | L <u>F F</u> | <u>L L</u> F | R L F		*	
P4		a	⇨	<u>L L</u> L | R <u>L L</u> | *R R* L	*		
	☞	b	⇦	L <u>L L</u> | <u>L R</u> L (n/a)			*
		c	BT	<u>L L</u> L | R <u>L L</u> | <u>R R</u> L | H R L		*	

Temp = Temporal Sequence
BT = Backtracking
NoBT = No-Backtracking
italics signal ill-formed substrings

As one scans P3 from left to right, applying Dissimilation as one goes, one gets **LL*F at the end of the line – as in P3a. This output is ill-formed, containing, as it does, an offending substring, **LL* (italicized for emphasis). But there is in principle no reason to rule out the possibility of *doubling back*, so to speak, and applying Dissimilation again, to ultimately produce an acceptable, though not attested, tonal string RLF in P3c. The same possibility is open to P4: the straight left to right rule application yields the ill-formed **RR*L. One could "repair" this malformed representation by reapplying the Dissimilation rule, and yield HRL of P4c as the final result. I will refer to the last step of the derivations of P3c and P4c as "*backtracking*." In order to forestall such a move, we posit a constraint No-Backtracking, and appropriately revise the ranking order. This correctly rules out (c) as the optimal candidate for P3 and P4 in the tableau in (26).

(27) No-Backtracking (NoBT):
Do not backtrack.

(28) {WFC, No-Backtracking} ≫ Temporal Sequence

A couple of comments on backtracking are in order. Although Backtracking calls for an extra derivational step, No-Backtracking is not reducible to the more generic constraint of economy measurable in terms of length of derivation. To see why this is so, recall P2 /RRR/ of the tableau in (24). The winning output HHR is generated by a left-to-right application of Dissimilation, even though a right-to-left mode of rule application would result in a shorter derivation, with an equally well-formed output RHR. This clearly suggests Temporal Sequence ≫ Derivational Economy. In fact, Economy exerts no overt influence as far as trisyllabic forms are concerned. On the other hand, as seen in the tableau in (26), P3 and P4 motivate the ranking No-Backtracking ≫ Temporal Sequence. To collapse together No-Backtracking and Derivational Economy would inevitably lead to a ranking paradox. The intuitive content of our analysis is this: "backtracking" involves a *qualitatively* different and more exacting processing task than merely adding one more step in a derivation (for more on this point, see section 4 below).

One additional aside: No-Backtracking is related to, but not identical with Strict Cyclicity Condition, which prohibits "reaching back" into a string wholly contained within an earlier cycle (see Cole 1995 for a recent review of issues related to strict cyclicity and non-derived environment). Two crucial differences set these two constraints apart from each other.

First of all, in our account it is No-Backtracking that eliminates candidate P4c in the tableau in (26), instantiated by the example below:

(29) pilot fly machine "to pilot an airplane"
kai [*fei ji*]
$[[L\ L]_{\alpha}\ L]_{\beta}$
|
$[R\ L]_{\alpha}$ (i)
|
$[R\ R\ L]_{\beta}$ (ii)
|
$[H\ R\ L]_{\beta}$ (iii)

But there is no meaningful way of talking about α and β in the example above as cyclic domains: α is patently a non-constituent, being made up of the verb *kai* "to pilot" plus *fei* "to fly," which is part of a compound noun *fei ji* "airplane" (lit. flying machine). Therefore, to appeal to Strict Cyclicity Condition to block reapplication of Dissimilation at step (iii) would stretch the notion of the cycle beyond any reasonable bounds. Secondly, even where it does make sense to talk about cyclic domains, as in the case of (30),

(30) *tuo la* "traction"
$[L\ L]_{\alpha}$
|
R L (i)
---
tuo la ji "tractor" (traction + machine)
$[R\ L\ L]_{\beta}$
|
R R L (ii)
|
H R L (iii)

Strict Cyclicity Condition, as commonly understood, does not block step (iii). The reason is that the underlined substring in <u>RR</u>L as it emerged from step (ii) constitutes a *new* environment, since the rising tone R in the middle does not arise until the current cycle β, even though the substring <u>RR</u> is wholly contained within a previous cycle α. Dissimilation should be allowed to apply at (iii) for the same reason that the spirantization rule (t → s / ___ i) should be allowed to apply in the classic Finnish example /vete/ "water."

(31) Finnish[16]

	"mother"	"want-past"	"water"	
root	[äi**ti**]	[halu**t**]-**i**	[vete]	
cycle	–	–	–	
word	–	–	ve**ti**	e → i / ___ #
cycle	–	halu**s-i**	ve**si**	t → s / ___ i

Notice that the requisite context of a high vowel [i] is not available until the word cycle. Therefore, spirantization applies, even though both the target [t] and the immediate environment (the vowel [i]) are entirely enclosed within the root. /vete/ contrasts with /äiti/ "mother" in that in the latter case, the substring *ti* constitutes a non-derived environment.

3.4 Summary

To summarize, the default direction of rule application is from left to right – unless this temporal sequence is overridden by the undominated WFC, No-Backtracking, or Preemptive clause. The derivational account is encapsulated in the ranking hierarchy:

(32) {WFC, No-Backtracking, Preemptive clause} ≫ Temporal Sequence

4 Constraints on derivation?

The derivational account sketched above in section 3 makes crucial use of three constraints that are distinctly derivational or processual in character: (i) The Preemptive clause or extrinsic order; (ii) left-to-right directionality (stated as Temporal sequence); (iii) No-Backtracking. In an increasingly declarative, non-derivational climate, it behooves us to consider their a priori plausibility and empirical support.

Extrinsic order is something of a last resort in any account. In the Tianjin case in question, perhaps the ordering relation postulated by the Preemptive clause is not entirely arbitrary or purely stipulative. One might speculate that the Preemptive clause reflects a gradient robustness of OCP effects: when a tonal string simultaneously violates OCP (total identity) and OCP′ (partial identity), the tendency is to undo the more egregious OCP offense (by means of Dissimilation) before worrying about the minor OCP′ violation (that can be fixed by Absorption). The weaker,

[16] Based on Kiparsky (1973). The underlying forms are motivated by such forms as *vete-nä*, *äiti-nä* (essive sg.) and *halut-a* (infinitive). Cf. Kenstowicz (1994:203ff.) and more general review and discussion in Cole (1995).

patently parochial nature of OCP′ is demonstrated by the permissible tonal sequences (33a, b, c), which show that partial identities are freely tolerated in Tianjin:

(33) a. FR = H<u>L.L</u>H ok
b. LR = <u>L.L</u>H ok
c. RH = L<u>H.H</u> ok
d. FL = H<u>L.L</u> *, → [H.L] by Absorption

The bias for left-to-right directionality accords with common sense. (I will continue to use the spatial/orthographic metaphor of left-to-right directionality to refer to temporal sequence, with apologies to Chinese and Semitic writing systems.) It stands to reason that, other things being equal, phonological processing ideally coincides with the temporal sequencing of the planning and execution of articulatory events. A right-to-left processing, on the other hand, would require buffering of long stretches of speech in order to make current decisions dependent on materials many syllables away (cf. Levelt 1989). For psycholinguistic evidence showing a left-to-right bias in speech organization (phonological encoding), see Meyer (1990, 1991). This temporal sequence manifests itself for instance in the predominantly left-to-right parsing of syllables into feet (cf. Hayes 1995). This is certainly so in Shanghai, where footing is diagnosed by means of tonal distribution:

(34) Shanghai tone/stress domain[17]

a.	*gao er ba qiao fu*	"Gorbachev"
	(HL.LH)(HL.LH) HL	left-to-right footing (left-headed); no degenerate foot
	(HL. o) (HL. o) o	tone deletion (affecting stressless syllables)
	(H L) (H L) o	tone association (left to right), ok
b.	*gao er ba qiao fu*	
	HL(LH.HL)(LH.HL)	right-to-left footing
	o (LH. o)(LH. o)	tone deletion
	o (L H)(L H)	tone association, *

The left-to-right parsing of syllabic trochees results in a rhythmic pattern with the first and the third syllables occupying the prominent positions, hence shielded from tonal deletion. One-to-one association of the surviving tone segments yields the right output in (34a). Footing in the opposite direction would give rise to the starred ungrammatical reading in (34b).

[17] Cf. Duanmu 1993a, 1995; see chapter 7, section 3 for further discussion.

As for No-Backtracking, virtually all sentence-processing models (from both production and comprehension ends) assume a left-to-right, incremental parsing of materials as soon as they are heard (rather than waiting until the end of the utterance). This occasionally gives rise to the classic garden-path phenomenon, whereby the hearer is misled into committing her/himself to a default analysis until the surprise ending, at which point s/he has to *backtrack* and reparse (for recent surveys, cf. Pritchett 1992, Clifton, Frazier, and Rayner 1994, Tanenhaus and Trueswell 1995). Backtracking represents a particularly complex processing task. A phonological analog (outside of the Tianjin case) would be the English Rhythm Rule, as illustrated by example (35), which requires backtracking at step (c): since stress clash does not arise until step (b), when the "current window" has already moved to the last two words. Unfortunately, empirical evidence for the speaker's ability to perform stress retraction on-line proves to be elusive (cf. Cooper and Eady 1986, Kelly and Bock 1988, Levelt 1989, Beckman et al. 1990, Grabe and Warren 1995, Vogel, Bunnell, and Hoskins 1995, and Shattuck and Huffnagel 1995; for possible explanation, see Hayes 1995).

(35) thirteen abstract paintings
  a. thir'TEEN ab'STRACT
  b. thir'TEEN 'ABstract 'PAINtings
  c. 'THIRteen 'ABstract 'PAINtings (xx = current window)

However, MalakMalak affords us some evidence showing an aversion to backtracking. Goldsmith's (1990) analysis of MalakMalak is paraphrased as follows:

(36) MalakMalak (Goldsmith 1990:173–177)
  a. Group syllables into left-headed feet, from right to left (weak prohibition on degenerate feet)
  b. Word-level prominence falls on the leftmost stressed syllable
  c. Stress clash resolution:
    i. *either* apply "trochaic reversal" ("restricted to *one step* in the repair," p. 177)
    ii. *or* delete the degenerate foot

One intriguing aspect of Goldsmith's analysis is the stipulation, that of the two alternative clash resolution strategies, one resorts to "trochaic reversal" only if it can repair stress clash in one single step, as illustrated by 2c in (37) below; otherwise, the degenerate foot is deleted, as exemplified by 3d. Cast in OT terms (37) 2 demonstrates the ranking Prs-σ (i.e. parse syllables into feet) ≫ Left-headedness; but 3 calls for the reverse ranking. Note in

particular, that multiple violation of Left-headedness is irrelevant under the strict domination hypothesis (cf. Prince and Smolensky 1993). The ranking paradox highlights the derivational character of the constraint against backtracking which does not readily translate into constraints on the output.

(37) MalakMalak

				Clash	Prs-σ	Lft-Hd
1 /mu.nan.ka.ra/	☞		("MU.nan)('KA.ra)			
2 /mɛl.pa.pu/		a	("MɛL)('PA.pu)	*		
		b	mɛl.("PA.pu)		*	
	☞	c	("MɛL)(pa.'PU)			*
3 /ar.ki.ni.yaŋ.ka/		a	("AR)('KI.ni)('YAŋ.ka)	*		
		b	("AR)(ki.'NI)('YAŋ.ka)	*		*
	$	c	("AR)(ki.'NI)(yaŋ.'KA)			**
	☞	d	ar.("KI.ni)('YAŋ.ka)		*	

small caps indicate stress
"σ = primary stress
'σ = secondary stress
Prs-σ = parse σ's into feet
Lft-Hd = left-headed foot
$ = expected winner
☞ = attested winner

Restated in derivational terms, the mysterious one-step repair stipulation makes sense: while both 2c and 3c of the tableau in (37) involve "trochaic reversal" (whereby trochees are turned into iambs in order to alleviate stress clash), only 3c involves backtracking. We can make this point clearer in the form of stepwise derivations (a) and (b) which generate the outputs 2c and 3c of the tableau in (37) respectively:

(38)	a.	/mɛl.pa.pu/	(= (37), ex.2c)
	→	mɛl.('PA.pu)	R-to-L footing
	→	("MɛL).('PA.pu)	R-to-L footing, Clash
	→	("MɛL).(pa.'PU)	Trochaic reversal (1), ok
	b.	/ar.ki.ni.kaŋ.ak/	(= (37), ex.3c)
	→	ar.ki.ni.('KAŋ.ka)	R-to-L footing
	→	ar.('KI.ni)('KAŋ.ka)	R-to-L footing
	→	("AR)('KI.ni)('KAŋ.ka)	R-to-L footing, Clash
	→	("AR)(ki.'NI)('KAŋ.ka)	Trochaic reversal (1), Clash
	→	("AR)(ki.'NI)(kang.'KA)	Trochaic reversal (2), *

In the former case, stress clash is resolved instantly by trochaic reversal as soon as the clash arises. In the latter case, the current window (symbolized by underlining) first moves leftwards in a right-to-left footing, then reverses course in successively changing trochees to iambs, creating a domino effect, as it were. It is this backtracking that militates against derivation (b). Under precisely this condition, defooting is the strategy of choice to resolve stress clash.

In short, there appear to be a priori plausibility and empirical support for the constraints that embody the intuitive content of the derivational account sketched here.

5 A non-derivational alternative

In the preceding sections we have sketched a fairly straightforward derivational account of the directionality effects of Tianjin. We have couched our analysis in terms of operations, and constraints on operations – i.e. constraints on how (order, directionality) sandhi processes are to be implemented. In recent years, phonologists have become increasingly skeptical about the psychological plausibility and computational feasibility of serial derivation as envisaged by classical generative phonology.[18] Thus, all functions of derivation are replaced by intra- and inter-level relations in Harmonic Phonology (Goldsmith 1993b:28), Cognitive Phonology (Lakoff 1993), and two-level rule models (Karttunen 1993). Meanwhile, OT has shifted the burden of analysis from the theory of operations to the theory of well-formedness, expressible exclusively as constraints on the output (McCarthy and Prince 1993b:8).[19] As Prince and Smolensky (1993:5) put it, from a parallel-processing perspective of GEN, "the Input → Output map has no internal structure: all possible variants are produced by GEN in one step and evaluated in parallel."[20] It would be a basic category error to talk about constraints on derivation – since there is no serial derivation to speak of. Given the conceptual appeal, tantalizing promise as well as spectacular successes of the constraint-based approach – which is most fully articulated and empirically tested in OT – it would be worth asking

[18] See, for instance, Goldsmith's *Introduction* to *The Last Phonological Rule* (Goldsmith 1993a). For an opposing view, see Bromberger and Halle (1989).

[19] Kindred in spirit is the "phonotactics and repair strategies" school of thought represented by Singh (1987) and Paradis (1988), *inter alia*.

[20] GEN is the function that "generates" from a given input a potentially infinite set of output candidates.

if the derivational account given in the preceding section could be brought more in line with the prevailing declarative frameworks.

5.1 Output constraints

One of the fundamental insights of constraint-based theories is that phonological rules do not arbitrarily deform lexically specified inputs, but are functionally motivated by the need to conform to certain output conditions. Even a derivational account requires explicit reference to output constraints stated as WFC. We also recognize that Dissimilation and Tonal Absorption are set in motion by WFC. But, notice that Dissimilation and Absorption do not merely declare a tonal string to be ill-formed, but *prescribe* specific ways of undoing or "repairing" WFC violations. For instance, Dissimilation turns an offending sequence of /FL/ into HL, and not, for instance, an equally well-formed RL. However, it is possible to reconceptualize the prescriptive rules ("do X") as a combination of strictly output constraints and faithfulness conditions. First we identify OCP and OCP′ as the overriding output constraints, repeated below. In addition, we have a generalized ban on adjacent partially identical tones, which we shall refer to as OCP″.

(39) a. OCP — no adjacent identical tones (except HH)
b. OCP′ — no *FL (= H<u>L.L</u>) sequences
c. OCP″ — no adjacent partially identical tones (*<u>L.L</u>H, *<u>H.H</u>L, *H<u>L.L</u>H etc.)

In order to create the optimal tonal targets set by the constraints of (39), one is limited to inserting or deleting a tone segment (thereby violating Dep and Max, respectively). Syllable–tone mapping (abbreviated as σT) further constrains possible insertions and deletions: one cannot attach an extra tone segment to a syllable already linked to a contour tone R or F (= LH, HL), in the process creating a complex three-tone segment melody like *LHL or *HLH; by the same token, one cannot delete a level tone H or L – under the penalty of producing a toneless syllable. Other conceivable "repair strategies" to guarantee OCP, for instance featural change (from H to L or vice versa) and contour reversal R ⇔ F by metathesis (HL ⇔ LH), are barred by (featural) Identity and Linearity of mapping. These constraints are stated as follows:[21]

[21] Constraints (Max, Dep) and (Ident, Linear) are paraphrases of McCarthy and Prince (1995b); Dep and Max correspond roughly to Fill and Parse in Prince and Smolensky (1993).

(40) Max
Maximize input-to-output correspondence
(i.e. do not delete or leave any input tonal material unparsed).

Dep
Every output tone must correspond to / be a dependent of some input tone
(i.e. do not insert or epenthesize any tonal material).

Syllable–tone mapping (σT)
A syllable must be associated with at least one, at most two, tone segments.[22]

Identity
Preserve featural identity of tone segments.

Linearity
Respect linear order of tone segments (i.e. no metathesis).

The tableau in (41) illustrates how these constraints jointly guarantee the correct disyllabic sandhi forms.

(41)

			OCP, OCP′	σT	Linear	Ident	Max	Dep	OCP″
1		/FL/							
	a	FL	*						
☞	b	HL = H<L>.L					*		
	c	LL = <H>L.L	*				*		
	d	F<u>H</u>.L = HL<u>H</u>.L		*				*	
	e	RL			*				
2		/LL/							
	a	LL	*						
☞	b	RL = L<u>H</u>.L						*	
	c	HL				*			
	d	oL		*			*		
3		/FR/							
☞	a	FR = HL.LH							*
	b	HR = H<L>.LH					*		
4		/HH/							
☞	a	HH							*
	b	F.H = H<u>L</u>.H						*	

[22] By setting σT at {0,1} tone per syllable, one describes a language which allows toneless syllables, but prohibits contour tones; likewise, by specifying {0,1} syllables per tone we are describing a language that permits floating tones but disallows doubly linked tones. See Goldsmith (1990).

(41) cont'd

			OCP, OCP′	σT	Linear	Ident	Max	Dep	OCP″
5		/FR/							
	☞ a	FR = HL.LH							*
	b	HR = H<L>.LH					*		
6		/HH/							
	☞ a	HH							*
	b	F = H<u>L</u>.H						*	
7		/FF/							
	a	HF = H<L>.HL					*		*
	☞ b	LF = <H>L.HL					*		
8		/RR/							
	☞ a	HR = <L>H.LH					*		
	b	LR = L<H>.LH					*		*
9		/LR/							
	☞ a	LR							*
	b	FR = <u>H</u>L.R						*	*
	c	RR = L<u>H</u>.R	*					*	
	d	oR		*			*		
	e	HR				*			

/T T/ input/lexical representation
<u>T</u> inserted tonal material
<T> deleted tonal material

Crucially, the ranking {OCP, OCP′} ≫ {Dep, Max} encodes the observation that the undominated[23] OCP/OCP′ are satisfied at the expense of Max (no deletion) or Dep (no insertion). Thus 1b wins out over 1a despite having deviated from the input by L-deletion. Likewise, 2b is preferred over 2a notwithstanding a H-insertion. Note further that OCP/OCP′ not only enforces insertion or deletion, but also determines which tone segment is to be inserted or deleted and where. Thus, by deleting a H from /FL = HL.L/ we get LL (= candidate 1c) – hardly an improvement. Other more "outlandish," though conceivable, ways of satisfying OCP and OCP′ – by metathesis (1e), featural change (2c), deletion resulting in a toneless syllable (2d) or insertion yielding a complex tone (1d) – are ruled out by the inviolate constraints of Linearity, Identity, and σT respectively.

[23] At least within the subcorpus under scrutiny. OCP/OCP′ violations do occur in longer connected speech, see section 8 below.

Unlike OCP/OCP′, the generic OCP″ which bans partial identity of all sorts (including [. . . H.H . . .]) must rank *below* Max/Dep. This is demonstrated by (5a) and (6a): in principle, by deleting the L portion of the falling tone in FR one can achieve a more "harmonic" tonal sequence HR (= H<L>.LH) in (5b), but this move is blocked by the ranking Max ≫ OCP″; as a consequence, OCP″ violation is tolerated in (5a); (6a) makes the same point about the ranking Dep ≫ OCP″. Summing up, we have the ranking {OCP, OCP′} ≫ {Dep, Max} ≫ OCP″.

It is important to note that the low ranking OCP″ is nonetheless decisive, when higher ranking constraints are neutral between competing candidates. This is shown in (7) and (8), where we see OCP″ at work picking the right outputs. For instance, judging by OCP″, (7b) represents a better tonal target than (7a), even though both satisfy OCP at the expense of one single Max violation.

Finally, the ranking automatically accounts for the curious fact that while OCP″ blocks a "derived" LR in (8b), an underlying LR freely occurs despite OCP″ violation in (9a). This is so because inserting a H either in front of or after the L either doesn't do anything to remove the OCP″ violation (9b) or, worse, creates a more serious OCP offense (9c). Neither is deleting the L a viable option, since it would result in a toneless syllable (9d). Lastly, changing LR to HR is barred by a higher ranking Identity (9e). In other words, no conceivable way of fixing up an underlying /LR/ improves the pattern of constraint satisfaction. As a consequence, (9a) stands despite OCP″.

In short, Dissimilation and Absorption formulated as conventional rewrite rules merely stipulate, by fiat, that LL dissimilates to RL, but RR dissimilates to HR (and not, for instance, LL to HL, or RR to FR). A constraint-based account, on the other hand, makes an explicit claim about the relative "cost" placed on alternative strategies: metathesis and featural change (barred by higher ranking Linearity and Identity) are more heavily penalized than inserting or deleting tonal materials (prohibited by the lower ranking Dep and Max), subject to permissible syllable–tone mappings (σT). In this respect, OT brings out more clearly the dynamics of interacting forces that converge on the attested disyllabic sandhi forms. On this score, there is much to commend a constraint-based analysis over its rule-based competitor.

5.2 Alignment

While we can reconceptualize and restate with reasonable ease the phonological rules of Dissimilation and Absorption in non-processual, declarative

terms, it is not as straightforward to obtain the directionality and rule-ordering effects by non-derivational means. Directionality effect is ubiquitous: languages may choose to group segments into syllables and syllables into feet either from left to right or in the opposite direction. Generalized Alignment theory (McCarthy and Prince 1993b) offers a readily available mechanism for mimicking directional syllabification and footing. This particular approach is illustrated below:

(42) directional footing (Crowhurst and Hewitt 1995)

	Pintupi ⇨	All-feet-Left		All-feet-Right		
a.	(σ σ) (σ σ) σ	0+2	☜	3+1		(measured in σ)
b.	σ (σ σ) (σ σ)	1+3		2+0	☜	
	⇦ **Yakan**					

(43) directional syllabification (Mester and Padgett 1993)

	Egyptian Arabic ⇨	All-s-Right		All-s-Left		
a.	(CVC)(Ce)(CV)	2+1+0	☜	0+2+3		(measured in μ)
b.	(CV)(CeC)(CV)	3+1+0		0+1+3	☜	
	⇦ **Iraqi Arabic**					

⇨, ⇦ encode directionality of footing or syllabification
σ, μ = syllable, mora
e = epenthetic vowel

The conventional way of stating the facts encapsulated in (42) is to say that footing "proceeds" from left to right in Pintupi (cf. Hayes 1995:62–64), and right to left in Yakan; degenerate (monosyllabic) feet being disallowed. Alternatively, one would say that instead of prescribing a procedure "do X" in a particular order – as from left to right – we simply compare the end results. In a left-to-right footing, the resulting prosodic units are "stacked" or "packed" to the left edge of a domain (say stem or p-word) in the following sense: in (42a), the first foot is smack against the left edge, and is therefore zero syllables away from the left edge; the middle foot is separated from the edge by two syllables; the last syllable is unfooted. The total score is then: 0 + 2 = 2. By the same evaluation procedure, the alternative footing (42b) ends up with the score 3 + 1 = 4. Therefore the prosodic organization represented by (42a) wins out over (42b) – if the optimal target (parameterized with respect to Pintupi) is that all feet be aligned to the *left*, informally stated as All-feet-Left. By

switching to a right-aligned parameter (All-feet-Right), we guarantee the mirror image configuration (42b) for Yakan.[24]

Directional syllabification with automatic vowel epenthesis (in the manner proposed by Itô 1986, 1989) can be recast non-derivationally in basically the same way, as suggested in (43), the difference being that, in this case, the distance from the edge is measured in terms of moras instead of syllables.[25]

Both directional footing and syllabification involve parsing a string of elements into prosodic structures with a constituent edge to be aligned with one or the other end of another morphological or prosodic category. Tone sandhi, on the other hand, does not create structures that can be oriented in one direction or the other; rather, it involves certain paradigmatic substitutions of one tone for another, as dictated by WFCs. For instance, there is no meaningful way of talking about a left- or rightward alignment of

[24] The mechanical details are a bit tricky. In particular, if degenerate feet are allowed, the arithmetics of syllable count would require a reversal in alignment. This is illustrated by the following:

	Icelandic			
	⇨	All-feet-Right	All-feet-Left	
a.	(σ σ) (σ σ) (σ)	3+1+0 ☜	0+2+3	(measured in σ)
b.	(σ) (σ σ) (σ σ)	4+2+0	0+1+3 ☜	
	⇦			
	Weri			

Thus, Icelandic syllabification proceeds from left to right, but the syllables must be aligned to the *right*. The reverse is true of Weri. This means that while both Pintupi and Icelandic display a left-to-right directionality effect, they must align their syllables in opposite directions:

Pintupi: Foot Binarity ≫ Parse-s ≫ All-feet-*Left*

Icelandic: Parse-σ ≫ Foot Binarity ≫ All-feet-*Right*

Crowhurst and Hewitt (1995) unpack the degeneracy parameter into Foot Binarity and Parse-σ (parse syllables into feet). Consequently, languages that allow or enforce degenerate feet rank Parse-σ over Foot Binarity; metrical systems that disallow degenerate feet exhibit the opposite ranking.

Parenthetically, according to Hayes (1995:188–198), the final syllable in Icelandic does not form a degenerate foot; instead, it owes its perceived secondary prominence to phonetic final lengthening. As illustrated above, whether or not degenerate feet are allowed (in weak position) materially affects the computation of alignment. I simply assumed the representation as given in Crowhurst and Hewitt (1995) for illustrative purposes.

[25] Likewise, there may be technical problems associated with Alignment as a declarative means for achieving the desired effects of directional syllabification. Davis (1995) presents evidence that while syllabification in Cairene Arabic proceeds from left to right, syncope works in the opposite direction. This presents no problem for a rule-based analysis, since in principle directionality may be rule-specific. On the other hand, if we use Alignment to specify the left or right orientation of the syllable structure resulting from both processes, we inevitably run into an Alignment paradox.

Laura Benua (p.c.) informs me that the reverse direction of syncope in Cairene Arabic may reflect an independently motivated prosodic constraint. I leave this question open.

For recent literature on directionality in OT, see Duncan 1994, Kager 1994, Klein 1993, Zoll 1994 and 1995.

two equally well-formed output strings HRL and FRL corresponding to P7 /FLL/. We must look elsewhere for an answer.

5.3 *Well-formedness, Faithfulness, and *Complex*

Let us consider some of the more obvious alternatives that readily come to mind. First of all, well-formedness pure and simple (as defined by OCP, OCP′) underdetermines the choice of alternative candidates. Even restricting ourselves to those outputs that can be generated by applying the rules left to right (with backtracking, where necessary) and right to left, for each of the patterns P1–4 and 7 there are two equally acceptable sandhi forms.[26] Thus both candidates (a) LHL and (b) FHL of P1 (in (24), section 3.1) are unimpeachable as tonal sequences. Likewise both candidates (b) HLF and (c) RLF of P3 (in (26), section 3.3) are perfectly well-formed. Finally, take P7 /FLL/ (24). Candidate (b) FRL is picked as winner, but there is nothing in the output itself that dooms candidate (a) HRL, which actually occurs as a sandhi form for three distinct inputs, as shown in (44). In other words, there is nothing wrong about HRL per se; it simply is not the optimal output given the input /FLL/.

(44)	/HRL/	=	HRL	[*niu jiao*] *jian*	"tip of a horn, hair-splitting"
	/HLL/	→	HRL	*pi* [*shu bao*]	"leather briefcase"
	/RRL/	→	HRL	*lao* [*mu ji*]	"old hen"

When we relativize output to the input, this usually means Faithfulness. Thus, in some absolute sense [ta] or perhaps [ba] is arguably the optimal, canonical syllable. But not all syllables are reducible to [ta] or [ba] without incurring massive Faithfulness infractions.[27] For the sake of simplicity, let us measure faithfulness by simply counting each instance of tonal modification (italicized) regardless of whether it involves insertion, deletion, metathesis, or featural change.[28] By this yardstick, indeed, while both P7a and P7b are well-formed, P7b wins out because it is the closest, i.e. most faithful, correspondent of /FLL/, with one rather than two points of departure. This is illustrated in the tableau in (45). Likewise, of the two well-formed candidates of P4 (i.e. consistent with OCP/OCP′), namely b and c, P4b involves the least deformation, and not surprisingly emerges as the winner.

[26] Right-to-left application and left-to-right iteration with backtracking yield identical outputs for P5 and P6.

[27] Cf. comment on this point in Chomsky (1994).

[28] Ignoring for the moment that methathesis and featural change constitute more serious violations of Linearity and Identity.

(45)

				OCP, OCP′	Faith (= Linearity, Ident; Max, Dep)
P7			/FLL/		
	a	⇨	*H*R*L*		**
☞	b	⇦	F*R*L		*
P4			/LLL/		
	a	⇨	*RR*L	*	**
☞	b	⇦	L*R*L		*
	c	BT	*HR*L		**
P2			/RRR/		
☞	a	⇨	*HH*R		**
	b	⇦	R*H*R		*
P3			/FFF/		
	a	⇨	*LL*F	*	**
☞	b	⇦	*HL*F		**
	c	BT	*RL*F		**

Italicization indicates any deviation from input

For ease of cross-reference, in this and subsequent tableaux, a, b, and c always refer to output candidates generated by a left-to-right, right-to-left, and left-to-right rule application cum backtracking, respectively, consistent with the tableaux in (24) and (26)

However, when we examine other cases in this light, Faithfulness makes counterfactual predictions. This is illustrated by P2 in (45). P2a wins out despite the fact that it deviates more drastically from the input, involving, as it does, two instances of tonal substitution (i.e. R → H), while P2b, an equally acceptable string entails only one. In fact, faithfulness measured in terms of the number of tonal substitutions plays no overt role in deciding the fate of the competing sandhi forms of Tianjin.

In this connection it is worth noting that Derivational Economy in general – Backtracking in particular – is related but not reducible to Faithfulness. Take P3 of (45). Each of the three candidates incurs two Faithfulness violations. Therefore, they are equidistant from the input, or equally (un)faithful. Needless to say, candidate (a) is eliminated because it runs afoul of WFC. But Faithfulness is neutral between the two remaining candidates (b) and (c). Candidate (c) is disfavored because, in the derivational account, it entails a particularly complex mode of rule application, namely Backtracking. The relevant differences are brought out more perspicuously in (46), which repeats P3 of (26) in a slightly different format.

(46)

		⇨		⇦		BT
P3	a.	F F F	b.	F F F	c.	F F F
		|		|		|
		L F F		F L F		L F F
		|		|		|
		L L F	☞	H L F		L L F
						|
						R L F

There is one further plausible output constraint, namely *Complex, a species of *Structure (No Structure), which disfavors contour tones (R, F), as opposed to level tones (H, L). As shown in (47), while *Complex picks the right winners in P1–3, it fails to make the correct prediction in P7.

(47)

				*Complex	Faithfulness
P1			/FFL/		
☞	a	⇨	LHL		**
	b	⇦	FHL	*	*
P2			/RRR/		
☞	a	⇨	HHR	*	**
	b	⇦	RHR	**	*
P3			/FFF/		
☞	b	⇦	HLF	*	**
	c	BT	RLF	**	**
P7			/FLL/		
	a	⇨	HRL	*	**
☞	b	⇦	FRL	**	*

5.4 *Markedness*[29]

Let us consider one last hypothesis. Suppose we introduce a notion like markedness, whereby certain tones are favored over others. Specifically, let us suppose we have the following ranking: *R ≫ *H ≫ *F ≫ *L, which amounts to saying that R is the most marked (least favored) tone, while

[29] I owe much of the discussion that follows in this section to Moira Yip, who made many valuable suggestions at the UC Irvine worship (October 21, 1995), at which I presented an earlier version of my analysis. Needless to say, she should not be held accountable for the details of analysis incorporating the notion of markedness as instantiated here.

L is the least marked (most favored). We get the right results by interposing this block of markedness constraints between the undominated constraints {OCP, OCP′, Linearity, Identity} and the violable constraints {Max, Dep}. (48) illustrates how one could exploit the notion of markedness to derive the effects of directionality and rule ordering – which is a non-trivial result.

(48)

					OCP, OCP′	Lin, Ident	*R	*H	*F	*L	Max, Dep	OCP″
P1				/FFL/								
	☞	a	⇨	LHL				*		**	**	
		b	⇦	FHL				*	*	*	*	
P2				/RRR/								
	☞	a	⇨	HHR			*	**			**	*
		b	⇦	RHR			**	*			*	*
P3				/FFF/								
		a	⇨	LLF	*				*	**	**	
	☞	b	⇦	HLF				*	*	*	**	
		c	BT	RLF		*	*		*	*	*	
P4				/LLL/								
		a	⇨	RRL	*		**			*	**	
	☞	b	⇦	LRL			*			**	*	*
		c	BT	HRL		*	*	*		*	*	
P5				/RLL/								
		a	⇨	RRL	*		**			*	*	
	☞	b	⇦	HRL			*	*		*	**	
P6				/LFF/								
		a	⇨	LLF	*				*	**	*	
	☞	b	⇦	RLF			*		*	*	**	
P7				/FLL/								
		a	⇨	HRL			*	*		*	**	
	☞	b	⇦	FRL			*		*	*	*	*

The internal rankings within the markedness block are motivated by the following observations: (i) *R ≫ *H, since P2b loses out to P2a on account of its having one extra R; (ii) *H ≫ *F, because P7b differs from P7a in having an F instead of a H in the initial position – P7b wins; (iii) likewise, *F ≫ *L: P1a starting on a L is preferred over P1b which begins on a F,

the two outputs being otherwise identical. Domination being a transitive relation, we obtain the following hierarchy {*R ≫ *H ≫ *F ≫ *L}, which we will refer to collectively as Markedness.

Needless to say, Markedness must rank over {Max, Dep}, otherwise its effect would be neutralized by the latter. This is demonstrated by the tableau in (49). A higher ranked block of {Dep, Max} would eliminate candidates P1a and P2a in favor of P1b and P2b, overriding the Markedness hierarchy. Therefore, according to this ranking, the expected output candidates would be P1b and P2b, counterfactually.

(49)

				OCP, OCP′	Lin, Ident	Max, Dep	*R	*H	*F	*L	OCP″
P1			/FFL/								
	☞	a	⇨ LHL			**!		*		**	
	$	b	⇦ FHL			*		*	*	*	
P2			/RRR/								
	☞	a	⇨ HHR			**!	*	**			*
	$	b	⇦ RHR			*	**	*			*

☞ = attested
$ = expected

While the properly ranked constraints {OCP, OCP′, Linear, Ident} ≫ Markedness ≫ {Dep, Max} ≫ OCP″ enable us to dispense with constraints of a derivational nature and thus arrive at a more elegant and parsimonious analysis consistent with the basic assumptions of OT, they raise problems of both technical and conceptual nature.

In terms of the mechanics of implementation, we have a ranking paradox to contend with. As shown in (49), Markedness crucially dominates {Max, Dep}. Now, a high-ranking Markedness would wrongly select candidate (b) over (a) in (50) – since HR is more "marked" than LR. Observing that what militates against candidate (b) may be the fact that it violates OCP″, we may resolve this ranking paradox by promoting OCP″ over Markedness, as in the tableau in (51). But this solution merely creates another ranking paradox. (51) (with the crucial ranking OCP″ ≫ Markedness ≫ {Max, Dep}) is incompatible with the hierarchy of dominance established earlier in section 5.1: {OCP, OCP′} ≫ {Max, Dep} ≫ OCP″.

(50)

				OCP, OCP′	Lin, Ident	*R	*H	*F	*L	Max, Dep	OCP″
			/RR/								
☞	a	⇨	HR = <L>H.LH			*	*!			*	
$	b	⇦	LR = L<H>.LH			*			*	*	*

(51)

				OCP, OCP′	Lin, Ident	OCP″	*R	*H	*F	*L	Max, Dep
			/RR/								
☞	a	⇨	HR = <L>H.LH				*	*			*
	b	⇦	LR = L<H>.LH			*!	*			*	*

Even if we could somehow overcome the ranking paradox, there remain other more serious problems, not the least of which is the apparently arbitrary rank order *R ≫ *H ≫ *F ≫ *L. Unlike *Complex which divides tones into two natural subsets (complex/contour and simplex/level), Markedness lacks any obvious intuitive motivation, especially in view of the fact that H is presumably the least marked in most tone languages. In a system where syllables/moras contrast only in being tone-bearing or toneless (but not in carrying one tone vs. another tone), the tone is almost invariably H. With particular reference to Tianjin, the unmarked nature of H is suggested by the fact that (i) whereas LL is banned by OCP, HH is not, and (ii) while FL (= HL.L) is barred by OCP′, its mirror image RH (LH.H) is not. In general, H is "visible" only to the weak OCP″. Finally, Markedness is keyed to the tonal repertoire of individual dialects, hence is less likely to generalize in a straightforward manner across dialects which also show robust directionality effects, as we shall see in the next chapter.

6 Cross-level constraints

The foregoing discussion in section 5 suggests that there is no straightforward mechanism within conventional, monostratal OT (Prince and Smolensky 1993) to capture the fundamental left–right asymmetry of serial rule application; nor is there an obvious way to mimic the effect

of Preemptive clause. However, the Correspondence Theory of Faithfulness developed by McCarthy and Prince (1995b) opens the door to an alternative, two-level approach. According to this view, grammar consists of a set of constraints or conditions which "declare" which pairings or mapping relations between an input and an output are licit or illicit. In this sense, phonology is "declarative" and non-derivational. One particular version of two-level rules I will explore is that developed by Kaplan and Kay (1981, 1994), Koskenniemi (1983), Karttunen and Koskenniemi and Kaplan (1985), Karttunen (1993), and Orgun (1995a).

6.1 *Licit and illicit correspondences*

A two-level analysis of the Tianjin facts, which I owe to Orhan Orgun (p.c.), is summed up in the following rules:[30]

(52) Two-level rules

```
R1  L        R2  F        R3a  R        R4a  F
    |            |             |             |
    R L          H L           H R           L F
                              R3b  R R      R4b  F F
                                   |             |
                                   H             L H
```

Two-level rules read very much like conventional phonological rules, with vertical shafts replacing horizontal arrows. There is one important difference: environments licensing or enforcing a particular correspondence can be stated at the input (top) and/or at the output (bottom) level. For instance, R3a says that R may correspond to H, if it is followed by another R at the *output*. By the same token, R3b states that R-to-H mapping is licit if R is followed by another R at the *input*. Taken together, R3a and R3b have the effect of licensing an R–H correspondence as long as another R follows at either level. R4b is the most complex of all: it stipulates an F-to-L mapping if and only if this pair is followed by another tone which is underlyingly F, but surfaces as H.

How the two-level rules account for the Tianjin facts is summarized in (53), where licit (more precisely, mandatory) and illicit correspondences

[30] When presented with the Tianjin tone sandhi facts, Orhan Orgun took it upon himself to work out a viable two-level rule account. I am grateful to him not only for his formulation of the two-level rules, but for valuable discussions which have shaped this section in important respects. I hasten to note that Orgun is not necessarily committed to the analysis embodied in (52), which he proposed by way of exploring the descriptive power of a formal system.

are symbolized as "|" and "‡" respectively. Identity relations are left unmarked. For ease of reference and comparison with the derivational account, I number the seven trisyllabic patterns exactly as before (P1–7), with the symbols "⇨, ⇦, BT" indicating the mode of rule application in a derivational account.

(53)

	⇨		⇦		BT	
P1	a.	FFL | | LHL ☜	b.	FFL ‡ | FHL		
P2	a.	RRR | | HHR ☜	b.	RRR ‡ | RHR		
P3	a.	FFF ‡ | LLF	b.	FFF | | HLF ☜	c.	FFF ‡ | RLF
P4	a.	LLL ‡ | RRL	b.	LLL | | LRL ☜	c.	LLL ‡ | HRL
P5	a.	RLL ‡ | RRL	b.	RLL | | HRL ☜		
P6	a.	LFF ‡ | LLF	b.	LFF | | RL ☜		
P7	a.	FLL ‡ | HRL	b.	FLL | FRL ☜		

"|," "‡" stand for licit and illicit correspondences, respectively
☞ indicates attested outputs

Take P1a. The first F maps onto a L by virtue of R4b, since the context satisfies conditions jointly defined over the input (F) and the output (H); the second F maps onto H by virtue of R2 since it abuts a L on the surface. P1b, on the other hand, contains an illicit correspondence, and is therefore not a viable phonetic realization of the lexically specified tone pattern /FFL/.

6.2 Conceptual problems

The table in (53) illustrates the observation made independently by Johnson (1972) and Kaplan and Kay (1981, 1994) that it is always possible to telescope serial steps of operations into a single more complex correspondence. The technique of *serial composition* – i.e. replacing two or more transducers (each of which models an elementary rule) operating in sequence with a single automaton composed of a cascade of transducers – can in principle mimic the effects of directional rule application as well as rule ordering, in effect eliminating intermediate stages in a phonological derivation (see Karttunen 1993:177ff.; Kaplan and Kay 1994). Declarative two-level rules and the conventional rewrite rules are therefore in principle descriptively equivalent for "regular" languages.

However, we can discern a number of conceptual problems associated with the two-level rule approach. First, given the nature of the sandhi processes in question (with contexts on the right), one can mimic directionality by imposing the condition of the context-sensitive rules either on the input or on the output. Imposing the condition on the *input* forces a left-to-right rule application: as long as the requisite condition is met at the input level, rules apply regardless of what happens ("subsequently") to the context. Conversely, imposing the condition on the *output* forces a right-to-left mode of implementation, since one needs to "anticipate" what eventually happens to the context on the right before one can decide whether it licenses a particular correspondence. Since the hallmark of two-level rules is their ability to stipulate at will conditions on the input and/or the output, the implication is that directionality is a rule-specific idiosyncracy. One can dramatize this point by noting that a two-level rules model predicts that we are just as likely to find a language like Tianjin except that R1 has its condition switched to the input level, while R3 is restricted to a sequence of RR on the output; in other words, R1 and R3 are replaced by the hypothetical rules formulated as R1′ and R3″ in (54). Such a scenario is highly improbable – if everything else we know about Tianjin remains constant (including WFC and No-Backtracking). Specifically, such a system would perversely dictate an ill-formed [RRL] output (55a) for /LLL/, while ruling out a perfectly well-formed [HHR] (55b) for /RRR/.

(54)

R1′	R3′
L **L**	R
|	|
R	H **R**

```
(55) a.  L L L        b.  R R R
         | |              ╪ |
         R R L            H H R
```

In contrast, on the derivational account sketched in section 3, directionality is in no sense a property that can be stipulated for a rule, but follows naturally from general principles (in particular, WFC ≫ Temporal Sequence). Consequently, the hypothetical scenario of (55) simply cannot arise.

There is a second, perhaps more serious, conceptual problem. In a derivational account, Dissimilation precedes Absorption (by virtue of the Preemptive clause), as illustrated in (56a). In other words, Absorption counterbleeds Dissimilation.

```
(56)     P1  a.  F F L                      b.  F F L
                 |        Dissimilation          |        Absorption
                 L F L                          F H L
                 |        Absorption                       Dissimilation (n/a)
             ☞   L H L                          F H L
```

In general, two-level rules handle opaque relations by stipulating conditions on the input, which is another way of signaling the fact that a correspondence/rule is not "surface-true." But one cannot simply restate R4b as R4c given in (57), simply because it would wrongly predict that P3 /FFF/ would emerge as *LLF in (58). Instead, R4b must define the condition jointly on the input *and* the output, exactly as formulated in (52). Quite apart form its aesthetic appeal (or, rather, the lack thereof), R4b is transparently a notational variant of R4d, as indicated in (59), since the only context in which a tone "starts out" as F at input but "ends up" as H at output is where this F is followed by a L (by virtue of R2 in (52)).

```
(57)     R4c   F F
               |
               L

(58)     P3    F F F
               | |
               L L F        = *

(59)     R4b   F F          =   R4d        F F L
               |                           |
               L H                         L
```

Once we "unpack" rule R4b as R4d, it brings into sharper focus three types of conceptual problems. First, one must forfeit the locality condition in order to allow the final L to act at a distance on the initial F (cf. Goldsmith 1976a, Clements 1985, Steriade 1986, McCarthy 1989, Archangeli and Pulleyblank 1994). Second, R4d redundantly repeats the environment of R2: the extra condition of a final L in R4d encodes nothing other than the context in which R2 applies. Finally, R4d telescopes two separate, elementary processes R2 and R4a into one complex correspondence. The price we pay for the telescoping of derivational stages via serial composition is increased complexity and opacity of the correspondences. As it stands, R4d renders its motivation totally opaque. From an analytical point of view, given R2 and R4a, there is no need for R4d (equivalently R4b) at all – if R2 and R4a are allowed to interact in some principled fashion (in this case, consistent with Temporal Sequence). In other words, R4d is the joint by-product of R2 and R4a. The argument here against serial composition is the flipside of McCarthy's (1993) objection against breaking up one functionally related chain shift into the two formally distinct rules of (60).[31] In this sense a two-level rule like R4b (equivalently R4d) is curiously anti-analytical and non-explanatory.

(60) Hijazi Bedouin Arabic

R1	$i \rightarrow \varnothing$
R2	$a \rightarrow i$

6.3 *Three-level rules*

We can eliminate the ungainly R4b/d from the rule system of (52) altogether if we adopt a multilevel model like Harmonic Phonology (Goldsmith 1993b) or Cognitive Phonology (Lakoff 1993). All we need is to segregate R2 (= Tonal Absorption) from all other rules (= Dissimilation), and characterize it as a W-P rule, while leaving the rest in the M-W stratum.[32] (61) illustrates how the separation of levels replicates the effect of rule ordering.

[31] Hijazi Bedouin Arabic chain shift has since been reanalysed by Orgun (1995a, b) and Kirchner (1995).

[32] M, W, and P in Harmonic Phonology and Cognitive Phonology stand for morphophonemic, word and phonetic levels of representation respectively.

(61)

	P7		P1		P3	
M:	F L L		F F L		F F F	
	|	R1	|	R4a	|	R4a
W:	F R L		L F L		F L F	
			|	R2	|	R2
P:	F R L		L H L		H L F	

Unfortunately, there is no independent justification for segregating the sandhi rules into M-W and W-P subsets. One possible argument for such a separation would be if R2 were a mere allophonic (or allotonic) distribution rule – in contraposition to the other morphophonemic (or morphotonemic) rules which relate one tonemic category to another. There is no reason to think that the F ~ H alternation stated by R2 is any more or less contrastive or any more or less redundant than any other alternations.[33] Another potential argument would be if R2 were a postlexical rule, while the others were lexical rules. Again, no independent[34] diagnostic tests (cyclicity, lexical exceptions etc.) support such a distinction. In particular both Absorption (R2) and Dissimilation (R4) apply indifferently to both lexical compounds and phrasal constructions, as illustrated in (62). As Padgett (1995:153) points out: "the positing of levels as an alternative to rule ordering is interesting exactly to the degree that we motivate them by compelling clusterings of properties."

(62)

	R4 FF → LF	R2 FL → HL
lexical	*lang fei* "wasteful"	*jiao shi* "teacher"
phrasal	*song xin* "to deliver a letter"	*diao gui* "to catch a turtle"

7 Harmonic serialism

In sections 2–4 we motivated an account of Tianjin tone sandhi facts that is predicated on the ranking of constraints that are processual or

[33] Goldsmith (1993b:32) characterizes the M-level as "the level at which expressions are structured into well-formed syllables and well-formed words, but with a minimum of redundant phonological information."

[34] That is, independent of the ordering effect, which the lexical–postlexical distinction in effect replicates.

derivational in nature, specifically: WFC, No-Backtracking, Preemptive clause ≫ Temporal Sequence. In sections 5–6 we explored the various non-derivational alternatives, each of which presents certain conceptual and/or technical problems. This seems to suggest a non-trivial role for derivation (or GEN), and the need to extend a constraint-based theory like OT to accommodate constraints on derivations rather than just constraints on outputs.

The idea of extending OT to include constraints of a derivational sort is certainly unconventional, but perhaps neither radical nor in principle incompatible with the spirit of OT. Reduced to its conceptual core, OT makes a fundamental theoretical commitment to the proposition that "grammars are defined by constraint hierarchies" (McCarthy 1995b:6). In particular, McCarthy sought to dispel the notion that "OT is inherently non-derivational." In fact, Prince and Smolensky (1993:5.2.3.3) made allowance for two alternative conceptions of GEN. In the first, parallel-processing construal – which has become the standard practice – all possible ultimate outputs are compared and evaluated at once (cf. McCarthy 1993). Accordingly, "the theory of operations," in their words, "is indeed rendered trivial" (Prince and Smolensky 1993:79). But there is an alternative, "Harmonic Serialism" version of OT, according to which one can construct the optimal candidate piece by piece, making optimizing decisions at each step of the serial derivation. To quote Prince and Smolensky (1993:79) in full,[35]

> [S]ome general procedure (Do-α) is allowed to make a certain single modification to the input, producing the candidate set of all possible outcomes of such modification. This is then evaluated; and the process continues with the outcome so determined. In this serial version of grammar, the theory of rules is narrowly circumscribed, but it is inaccurate to think of it as trivial. There are constraints inherent in the limitation to a single operation; and in the requirement that each individul operation in the sequence improve Harmony.

The serialist model envisioned by Prince and Smolensky above is depicted in (63). Proceeding from a given input, GEN can produce any number of intermediate outputs {a, b, c . . . }, so long as it performs only one single minimal modification on the input. These intermediate outputs are then

[35] Similar ideas have been delineated by Bruce Tesar (posting on OT-net, November 21, 1995).

evaluated by the constraint hierarchy. The winner candidate (in this case, c) becomes the input for the next elementary operation of GEN, which in turn produces a second batch of output candidates {e, f, g . . . } and so forth (see Prince and Smolensky 1993:5.2.3.3, with core syllabification in Berber as a test case).

(63) Input

↓	Gen
{a, b, c . . . }	Output_1; c = winner
↓	Gen
{e, f, g . . . }	Output_2: g = winner
↓	Gen
etc.	

Let us explore this alternative serialist OT approach to the Tianjin problem at hand. Suppose that GEN is restricted to one single operation stated as Tonal Absorption or any one of the subrules of Dissimilation. The intermediate outputs in a derivation (Output_1, Output_2, etc.) are then evaluated by the output conditions OCP and OCP′ (banning partial identity). This serialist approach straightforwardly accounts for P7 and P4 in (64). Thus, given /FLL/ (P7) as input, GEN is limited to performing a single tonal substitution (highlighted by underlining), thereby yielding either <u>H</u>LL (a) or F<u>R</u>L (b) as output. The choice of (b) as the winning candidate is a foregone conclusion; no further tonal replacement is possible, or necessary. P3 can be handled in the same way under the reasonable assumption that OCP dominates OCP′.[36] Notice that by replacing one or the other input F by a L, the immediate Output_1 still violates either OCP (P3a) or OCP′ (P3b). The ranking OCP ≫ OCP′ selects P3b FLF as the immediate output. Once this decision is made, the only further permissible tonal substitution on P3b is Absorption, which produces HLF as the ultimate output. The local decision made at the level of Output_1 eliminates all other subsequent logical moves that could in principle take place. In particular, it automatically precludes the following chain of events (represented by the shaded cells of (64)): FFF → <u>L</u>FF (P3a as

[36] OCP ≫ OCP′ and Preemptive clause share the same insight: OCP violation is more egregious than OCP′ infraction; therefore OCP violation dooms an output candidate (in the serialist OT account) or must be undone first (under the derivational account). See discussion in the opening paragraph of section 4.

Output$_1$, by Dissimilation) → L<u>L</u>F (Output$_2$, by Dissimilation) → <u>R</u>LF (Output$_3$, by Absorption).

(64)

	Input		Output$_1$		Output$_2$		Output$_3$	OCP	OCP′
P7	/FLL/							*	*
	↳	a	<u>H</u>LL	(Abs)				*	
	↳	b	F<u>R</u>L	(Dism) ☜					
P4	/LLL/							**	
	↳	a	<u>R</u>LL	(Dism)				*	
	↳	b	L<u>R</u>L	(Dism) ☜					
P3	/FFF/						**		
	↳	a	<u>L</u>FF	(Dism)				*	
			↳		L<u>L</u>F	(Dism)		*	
						↳	<u>R</u>LF		
	↳	b	F<u>L</u>F	(Dism) ☜					*
			↳		<u>H</u>LF	(Abs) ☜			

OCP bars adjacent identical tones
OCP′ bars FL (= HL.L)
Dism, Abs = Dissimilation, Absorption
Output$_1$, Output$_2$ . . . represent a series of intermediate outputs
↳ operation of GEN
☜ winner candidate

So far so good. But we begin to encounter problems with P5 and P6 in the tableau in (65). Given the input /RLL/ (P5), GEN is allowed only to change the L tone in the middle to R, consistent with Dissimilation (a). But the immediate output R<u>R</u>L in no way improves the "harmony" of the input; it merely shifts the locus of OCP violation. The same point can be made with P6. P1 and P2 in the same tableau present a more serious problem. Take P1: it is possible for GEN to change FFL to a well-formed output F<u>H</u>L (P1b) by means of a single elementary tonal substitution. Surprisingly, such a flawless output loses out to its competitor <u>L</u>FL (P1a, Output$_1$), which is less "harmonic" in that it entails an OCP′ violation, containing as it does a FL substring, which requires a further operation by GEN in order to emerge ultimately as F<u>H</u>L (P1a, Output$_2$), the actual winner.

(65)

	Input		Output$_1$		Output$_2$	OCP	OCP′
P5	/RLL/					*	
		a	RLL (no change)			*	
	↳	b	R<u>R</u>L (by Dism)			*	
			↳		<u>H</u>RL (by Dism) ☜		
P6	/LFF/					*	
		a	LFF (no change)			*	
	↳	b	L<u>L</u>F (by Dism)			*	
			↳		<u>R</u>LF (by Dism) ☜		
P1	/FFL/					*	*
	↳	a	<u>L</u>FL (by Dism)				*
			↳		L<u>H</u>L (by Abs) ☜		
	↳	b	F<u>H</u>L (by Abs)	$			
P2	/RRR/					**	
	↳	a	<u>H</u>RR (by Dism)			*	
			↳		H<u>H</u>R ☜		
	↳	b	R<u>H</u>R (by Dism)	$			

$ expected winner candidate
☜ attested winner

These observations suggest several conclusions. First, the decision between two alternative derivational steps allowable by GEN at any one point is not strictly "local," limited to whether or not a minimal operation improves the harmony of the immediate output; rather, the choice is "global" in the sense that EVAL picks a derivational path that *eventually* leads to a well-formed output. In other words, EVAL compares alternative derivations in their entirety. Path (a) of P5 and P6 terminates abruptly in a dead end, with an end result that involves an OCP violation, and is therefore rejected; path (b), on the other hand, does ultimately lead to a well-formed final output, and is therefore preferred – even though the intermediate output (Output$_1$) does not improve the harmonic state per se.[37]

Second, the constraints against which alternative derivational paths are weighed cannot be exclusively representational in nature – i.e. entirely determined by the output. If it were so, then any derivational path leading

[37] This is an analog to the classic traveling salesman's problem: by always picking as his next destination the closest town on the map, he may or may not come up with the best itinerary, if minimizing the total mileage is the objective.

to a well-formed output would be just as good as any other. In addition to representational economy (fewest symbols, least marked configuration etc.), we need constraints of a derivational nature (notably, fewest steps, directionality). In particular, No-Backtracking, a special case of derivational complexity, is decisive and undominated (see section 3.3). On the other hand, the "generic" constraint of derivational economy is subordinate to other constraints of a processual nature, in particular Temporal Sequence. This has been amply demonstrated by P1 and P2. Thus, as seen in (65), the derivational path P1a FFL → F<u>H</u>L satisfies the WFC in one step; P1b, on the other hand, achieves the same objective in two, FFL → <u>L</u>FL → L<u>H</u>L. The latter is preferred, because Temporal Sequence is maintained at the expense of derivational economy (see section 3.1).

It is worth noting that derivational economy is related but not reducible to faithfulness. Typically, each step of the derivation entails one degree of deviation from the input; hence, in most cases, length of derivation translates directly into the number of faithfulness violations. However, this one-to-one equivalence between faithfulness and derivational economy breaks down in some cases. Recall pattern P3 /FFF/ ((45), repeated here in a different format as (66)). Each of the three conceivable ways of applying tone sandhi rules involves two tonal substitutions, if we simply count the mismatches (in bold italics) between the input and the final output: instead of /FFF/ we have ***LL***F, ***HL***F, and ***RL***F for candidates (a), (b), and (c) of P3 respectively, each of which deviates from the input by two tonal substitutions. Faithfulness is therefore neutral among the three competing candidates. WFC eliminates candidate (a). What then arbitrates between candidates (b) and (c)? The answer has to be Derivational Economy: candidate (c) entails one extra derivational step (backtracking).[38] It is clear that Derivational Economy is irreducible to Faithfulness. This point is made most clearly in the following familiar tableau form:

[38] One might argue that RLF (= P3c) deviates from /FFF/ more drastically than its competitor HLF (= P3b) in that while the latter merely involves underparsing the terminal L, i.e. (i), the former entails metathesis, as depicted in (ii). If so, what is crucial is not derivational economy but the degree to which the output deviates from the input that matters.

(i) FLF = HL.L.F → H<L>.L.F (<L> = unparsed L)
(ii) FLF = HL.L.F → LH.L.F

There is no a priori reason to think that linearity violation (metathesis) is any more or less serious than overparsing or underparsing. For instance, New Chongming (chapter 5) allows metathesis (R ↔ F) but prohibits insertion or deletion of a terminal tone segment (H → HL, or LH → H).

(66)

P3				WFC	Faithfulness	Derivational Economy
	a	⇨	F F F | L F F | ***L L*** F	*!	**	**
☞	b	⇦	F F F | F L F | ***H L*** F		**	**
	c	BT	F F F | L F F | L L F | ***R L*** F		**	***!

! = fatal violation

In short, the derivational account sketched out in section 3 is compatible with a conception of OT which has the following properties: (i) GEN operates *seriatim*, effecting one single modification at a time; (ii) EVAL weighs alternative derivational paths against a constraint hierarchy; (iii) the relevant constraints may be representational or derivational in nature, including economy, directionality, and the like. Such a conception of OT comes close to a proposal entertained by Paul Smolensky (p.c.), in as yet unpublished works, which aims to implement the Minimalist Program (Chomsky 1992) in OT with GEN generating alternative derivations which are then evaluated by the appropriate constraints (e.g. move α the shortest distance, or make the fewest moves possible, etc.).[39]

[39] Cf. David Pesetsky's OT-net posting (September 21, 1995).

8 Concluding remarks

The facts of Tianjin tone sandhi call for a derivational/processual account (sections 2–4) that can be accommodated within a particular "harmonic serialist" conception of OT sketched in section 7.

In closing, it would be instructive to consider a major argument against a processual view. The argument in question turns on the dubious status of intermediate representations. Take an input like /RLL/ in Tianjin (= P5; (24), section 3). From the conventional derivational perspective represented here as (67) below, /RLL/ first becomes RRL, which in turn creates a "domino effect" that triggers another round of tonal adjustment to eventually emerge as HRL. There is no question about the underlying form /RLL/, which represents lexical knowledge; likewise, there is no question about the phonetic form HRL. But what about the intermediate representation RRL? Is it a mere artifact of the analysis, or does it have some empirical status?

(67) R L L (= P5, (24), section 3)
|
R R L
|
H R L

In this connection consider a small sample of longer connected speech:[40]

(68) "older-brother smokes 'Zhandou' (brand) cigarettes"

	dage chou	*zhandou yan*	
	(R L L)$_w$	(F F L)$_w$	
	(F R L)	(L F L)	
a.	(F R L)	(L H L)	attested
	(R **R L**	**L** H L)$_{IP}$	
b.	(F **R R**)	L H L)	not attested
c.	(F H R	L H L)	attested

w = phonological word
IP = intonational phrase

[40] Examples (68) and (69) correspond to (17) and (21) in Tan (1987) respectively.

Restricted to p-words as sandhi domains, Dissimilation and Absorption (stated in section 2) together yield the attested reading (a) indicated above. Applied at the intonational phrase level, the same rules predict (c) as the output, also attested. What is not attested is the intermediate form (b). This may at first cast doubt on intermediate representations in general. Now, compare (69) with (68).

(69) "telephone emits strange noise"

	dianhua	*fachu*	*guai jiao*	
	$(\underline{F\ F})_w$	$(\underline{L\ L})_w$	$(\underline{F\ F})_w$	
a.	(L F)	(R L)	(L F)	attested
	(L F	**R** **L**	**L** F)$_{IP}$	
b.	(L F	**R R**	L F)	attested
c.	(L F	H R	L F)	not attested

In more deliberate speech, p-word-bound tone sandhi produces the pronunciation of (69a) in the usual manner. In casual speech, tone sandhi operates on the intonational phrase as a well. Ideally, we expect (69c). Instead, what surfaces turns out to be precisely the "theoretically" postulated intermediate output (69b). Notice that at the start of the phrase level phonology (just below the broken line), both (68) and (69) contain a substring [. . . RLL . . .], which is identical to P5 in (67). A right-to-left direction of rule application should yield [. . . HRL . . .] (= 68c and 69c). Instead, (69b) surfaces with an offending [. . . RR . . .] substring. This suggests a "failure" in the on-line computation of sandhi processes. To appreciate the highly complicated task involved in the phonological encoding of continuous speech assembled in real time,[41] notice that the crucial substring [. . . RLL . . .] in (69) does not arise until the speaker has reached th very end of the phrase, since the change of the penult /F/ to L is crucially dependent on the context provided by the phrase-final F. This means that in order to derive (69c), the speaker must look ahead to the end of the phrase, buffer an intermediate representation [. . . RLL . . .], and work her/his way back two syllables to adjust [. . . R$\underline{\text{LL}}$ . . .] first

[41] As opposed to more or less "prefab" lexical items, idiom chunks, which may be stored in their "ready-made" sandhi forms. On issues related to "precompiled" phrasal phonology, cf. Hayes (1990).

to [. . . RRL . . .] (69b), and eventually to [. . . HRL . . .] (69c). What has happened instead, in this particular case, is that the speaker has successfully executed all the "instructions" embodied in the sandhi rules – except the last step. An occasional lapse recorded as (69b) spotlights what appears to be a partial breakdown in phonological processing, captured by "candid camera," as it were, that affords us a glimpse into the workings of tone sandhi as a process unfolding in real time.

Evidence like this strongly suggests a non-trivial dynamic dimension to phonological encoding.[42] It would be tempting to relegate constraints on real time processing to the domain of performance. But in doing so one runs the risk of closing one's mind to a potential source of explanation for systematic patterns in language. After all, phonological rules start out in their embryonic form as what Beaudouin de Courtenay (1972:161ff.) characterized as "neophonetic alternations" motivated by the ease of pronunciation. Just as phonological rules represent the conventionalization (often in their fossilized form) of articulatory forces, their mode of implementation reflects partly the chronological layering of the processes and partly general constraints on phonological processing.

In the next chapter I will continue to couch my analysis of other sandhi phenomena in conventional derivational terms. Where appropriate I will compare the derivational analysis with non-derivational alternatives.

[42] See Levelt (1989) for a synthesis on phonological processing from the production perspective.

4 *Directionality and interacting sandhi processes II*

The directionality effect we have examined at length in the preceding chapter is by no means unique to Tianjin. In this chapter we will explore similar phenomena observable in other dialects that promise to shed further light on related aspects of phonological processing.

1 Changting: preamble

The Hakka dialect of Changting, a border county between Fujian and Jiangxi province in south-east China, has been reported extensively by Luo (1982) and Rao (1987) and analyzed in depth by Hsu (1994). The table in (1) represents the underlying tone system of Changting.[1] In addition, there are two tone shapes that occur only in sandhi contexts, namely a "dipping" (i.e. falling-rising [213]), and a raised mid [44] tone. We will annotate these tones as D and M′ respectively.

(1) Changting tone system

	level		rising		falling	
high	H	55				
mid	M	33	R	24	F	42
low	L	11				

Tones undergo modifications in sandhi contexts. Unlike the Tianjin case discussed in the preceding chapter, the phonetic or structural motivations behind the sandhi changes are somewhat obscure, and some of the sandhi processes resist a straightforward formulation. Table 4.1 summarizes the two-tone patterns, with appropriate examples given in table 4.2.

[1] H and L are transcribed as 54, 21 in Luo (1982), Rao (1987). In all likelihood, the slight dip of 54, 21 in their transcription is due to the natural phenomenon of "declination," a term Liberman and Pierrehumbert (1984:161) use to refer to the overall downtrend and narrowing of pitch range towards the end of an intonational phrase. I will ignore this phonetic detail throughout the ensuing presentation.

Table 4.1. *Disyllabic tone sandhi*

	1st σ				
2nd σ	H	M	L	R	F
H					
M			L L	L R	
L	M H	M M			M F
R		M′ M	R F		
F		D M	D F	L R	M F

D, M′ = "dipping" and raised mid tones respectively

Table 4.2.

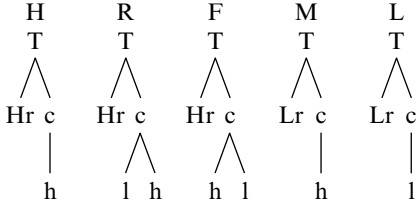
		base form	sandhi form	gloss
1	*ɔ niao*	M L	L L	"to urinate"
2	*xian hong*	M R	L R	"bright red"
3	*chi su*	L H	M H	"to be a vegetarian"
4	*hou sheng*	L M	M M	"young"
5	*gui dao*	L F	M F	"to kneel down"
6	*he cang*	R M	M′ M	"grain silo"
7	*sui bian*	R L	R F	"casual"
8	*mi kang*	F M	D M	"rice bran"
9	*shou zhuo*	F L	D F	"bracelet"
10	*qiang jie*	F R	L R	"to loot, plunder"
11	*xuan ju*	F F	M F	"to elect"

The left column and the top row of table 4.1 indicate the citation forms of the five tonal categories of the first and the second syllable respectively. The sandhi patterns, if different from the underlying forms, are indicated in the intersection cells. Where no change takes place, the cell in question is left blank. Thus, reading across the top row, the tonal combination HT remains unchanged throughout (T = any tone). The second row indicates that an underlying M is lowered to L, when followed by either L or R; conversely, the third row means an underlying L is raised to M when followed by H, M, or F. In other words, the contrast between the two low-register tones M and L is neutralized in connected speech. We can summarize the observed merger as (2), which translates straightforwardly as tone spread in standard autosegmental notation.

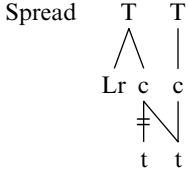
The five base tones presumably have the representations given in (3). Hr and Lr, i.e. high and low register, informally abbreviate [+upper] and [–upper] respectively.[2] The lower case letters, h and l, as usual, stand for the terminal tonal segments [+raised] and [–raised]. Assuming the tonal geometry of (3), we can restate (2) as (4). An example would be (5).

(2) M, L → M / ___ H, M, F
L / ___ L, R

(3)



(4)



(5)

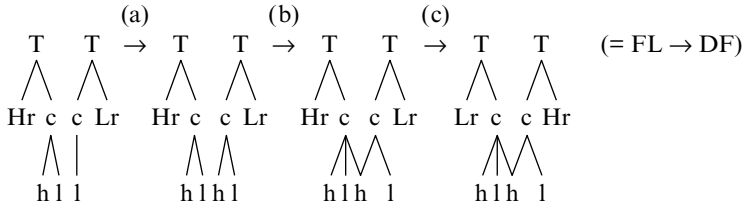
= LF → MF

example: *gui dao* "to kneel down" (= table 4.2 ex. 5)

Unfortunately, not all sandhi processes are as transparent in their function or phonetic motivation; by the same token, nor do they lend themselves to simple formal notation. To illustrate, reading across the bottom row of table 4.1, the combination of base tone F on the first syllable and L on the second shows up as D F (= ex.(9)), with both tones simultaneously undergoing modification. Assuming the same tonal geometry of (3), the process in question suggests a chain of events depicted in (6). (6) involves insertion of a high tone (step (a), perhaps a buffer high to break up two adjacent lows), h-spread (step (b)), and concomitant register flip-flop (step (c)).

[2] The rising and falling tones do not contrast in register; for expository simplicity, I simply annotate them as Hr.

(6)



example: *shou zhuo* "bracelet" (ex. 9 of table 4.2)

For reasons like this, I will simply restate table 4.1 in the form of correspondence rules, divided into three types, "forward," "backward," and "bidirectional" sandhi, for easy reference in the ensuing discussion.

(7) Backward sandhi

a. M, L → M / ___ H, M, F
L / ___ L, R

b. R → M′ / ___ M

c. F → D / ___ M
L / ___ R
M / ___ F

(8) Forward sandhi
L → F / R ___

(9) Bidirectional sandhi
F L → D F

2 Temporal Sequence and No-Backtracking[3]

Table 4.1 or, equivalently, rules (7–9) formulated in the preceding section represent the sum of tonotactics or well-formedness conditions that govern the juxtaposition of tones. Henceforth I will characterize as ill-formed any tonal substring that fits the structural description of rules (7–9) but does not undergo the appropriate sandhi change. The question is how the WFCs – individually stated as rules and synoptically represented in table 4.1 – are to be satisfied by connected speech, or equivalently, how the phonological rules are to be applied to successive two-tone substrings that make up the speech chain.

Given five tonal categories, there are 125 logically possible trisyllabic

[3] In this and following sections on Changting, I rely heavily on a very detailed and thorough analysis presented by Hsu (1994).

patterns. Obligingly, Rao (1987) exhaustively lists them all.[4] After eliminating those patterns that are unaffected by the sandhi rules, or involve only one single possible tonal substitution, we still have the sizable number of thirty-three combinations that exhibit directionality or rule-ordering effect in that they potentially give rise to different outputs depending on the sequencing of the elementary processes.

Although Changting tone sandhi is considerably more complex and opaque than that of Tianjin, we can discern the broad outlines of a reasonable analysis along the lines proposed for Tianjin. Just as in Tianjin, the default mode of rule application in Changting is also from left to right, in accord with the principle of Temporal Sequence of speech organization. This is shown by the examples given in the tableau in (10a) below:

(10a)

					NoBT	WFC	Temp	Econ
P1	☞	a	⇨	F F F | M F F | M M F				
		b	⇦	F F F | F M F | D M F			*	**
P2	☞	a	⇨	M R M | L R M | L M′ M				**
		b	⇦	M R M | M M′ M (n/a)			*	*

NoBT = No-Backtracking
Temp = Temporal Sequence
Econ = Derivational Economy
TT = current window of rule application spanning a two-tone substring

[4] All Changting examples are drawn chiefly from Rao (1987), supplemented by Luo (1982).

(11) Examples corresponding to (10a)

	Input	Output	[x x] x	x [x x]
P1	FFF	MMF	*[lao hu] jing*	*lao [gu dong]*
			"tiger well"	"antique, old fogey"
P2	MRM	LM′M	*[san lun] che*	*zheng [ying shu]*
			"tricycle"	"compete"

Given /FFF/ as input string, we derive two equally well-formed outputs MMF (P1a) and DMF (P1b), by applying the rules left to right or in the opposite direction. Candidate P1a is the attested sandhi form. This means that other things being equal, Temporal Sequence is decisive in predicting the correct output. P2 further demonstrates that Temporal Sequence overrides derivational economy in determining the mode of rule application: applying left to right, WFCs can be satisfied only at the cost of two tonal substitutions (MRM by *LM′*M, with tonal modifications in italics for clarity); proceeding in the opposite direction, WFCs are readily met by one single change (MRM → M*M′*M). Of the two conflicting constraints, it is Temporal Sequence that prevails. Notice further that Temporal Sequence dictates the mode of rule application or constraint satisfaction regardless of morphosyntactic structure: thus, even in a right-branching structure like P1 *lao* [*gu dong*] "antique,"[5] only a left-to-right iterative application (symbolized by the arrow "⇨") of tone sandhi rules yields the correct output MMF (P1a). Notice, finally, that the winning candidate P2a instantiates a case of "overapplication" in the sense that the initial M turns into L before R, even though this R in the medial position actually surfaces as M′. In other words, what we have is an opaque, counterbleeding rule relation. In short, Temporal Sequence asserts dominance over such other constraints as Cyclicity, Derivational Economy, and Surface-true transparency.[6]

However, Temporal Sequence is itself subordinate to the higher imperative of achieving well-formedness. Consider P3 and P4 of (10b).

(10b)

				NoBT	WFC	Temp	Econ
P3	a	⇨	M L F | L L F | *L M* F		*		**

[5] Bracketing indicating constituency, as usual.

[6] Transparency, or the lack thereof, has obvious learnability implications. For early discussion on transparent vs. opaque rule order see Kiparsky (1971), Kisseberth (1973), and Kenstowicz and Kisseberth (1977, esp. ch. 4) and, more recently, Iverson (1995) inter alia.

(10b) cont'd

						NoBT	WFC	Temp	Econ
	☞	b	⇦	M L F | M M F	(n/a)			*	*
P4		a	⇨	L L M	(n/a)		*		*
				L L M | *L M* M					
	☞	b	⇦	L L M | L M M | M M M				*	**

(n/a) = no applicable rule to current window
Ill-formed substrings are italicized

(12) Examples corresponding to (10b)

	Input	Output	[x x] x	x [x x]
P3	MLF	MMF	*[dui dou] bing* "bean cake"	*xin [nuo mi]* "new glutinous rice"
P4	LLM	MMM	*[mo li] hua* "jasmine flower"	*da [di fang]* "big place"

In both of these cases, a left-to-right application of disyllabic tone sandhi rules would produce an output that is opaque (counterfeeding) in that the output contains a substring that fits the structural description of a rule, and yet fails to undergo it. Take, for instance, the output of P3a contains a substring LMF (double-underlined) that is ill-formed, and is expected to change into MMF as demanded by the WFCs collectively represented by table 4.1. Likewise, scanning P4 rightwards, we find a perfectly well-formed substring LLM, and no modification is called for (indicated by (n/a)). Proceeding further to the right, the doubly underlined substring LLM is ill-formed, and turns into LMM, which in turn creates an ill-formed substring LM; output candidate P4a is therefore rejected in favor of P4b, generated by a right-to-left rule application. The generalization we can infer from P3 and P4 is that, while rules normally

apply from left to right, if this default mode of operation results in an ill-formed output, the directionality of rule application is reversed. This is captured, in OT terminology, by the now familiar ranking WFC ≫ Temporal Sequence, which enforces a right-to-left direction of rule application, as borne out by P3b and P4b.

But unlike Tianjin, Changting does not pursue WFC at all cost. The quest for well-formedness is thwarted if it calls for an exorbitant price like backtracking. This is illustrated by P5 in (10c).

(10c)

					NoBT	WFC	Temp	Econ
P5	☞	a	⇨	F F R | M F R | *M L* R		*		**
		a′	⇨BT	F F R | M F R | M L R | L L R	*			***
		b	⇦	F F R | F L R | | D *F R*		*	*	**
		b′	⇦BT	F F R | F L R | | D F R | D L R	*		*	***

Ill-formed substrings are italicized

(13) Examples corresponding to (10c)

	Input	Output	[x x] x	x [x x]
P5	FFR	MLR	*[gan cao] xie*	*da [tu hao]*
			"straw sandals"	"attack the local tyrants"

Given an input string /FFR/, neither a simple rightward (P5a) nor a simple leftward (P5b) application of sandhi rules will produce, at the end of the scan, an output that is consistent with all the conditions collectively imposed by table 4.1: both P5a and P5b contain an offending substring – MLR and DFR, respectively. To undo such violations, one must scan in the opposite direction, and reapply whatever rules are applicable. Thus, in the case of P5a, as one reaches the end of the string in a *rightward* scan, one ends up with MLR, which contains an ill-formed substring consisting of ML. Accordingly, one must reverse course and scan *leftwards* in order to turn the "deviant" MLR into an acceptable LLR (= P5a′). The same backtracking procedure – except in the opposite direction: first scanning leftwards, then rightwards – is required in P5b′. The heavy cost exacted by backtracking eliminates both candidates P5a′ and P5b′, leaving Temporal sequence to decide between two equally ill-formed candidates P5a and P5b. Not surprisingly P5a emerges the winner.

The facts presented so far can be accounted for succinctly by the following ranking:

(14) No-Backtracking ≫ WFC ≫ Temporal Sequence ≫ Derivational Economy[7]

3 Temporal sequencing vs. structural affinity

Key to our analysis of both Tianjin and Changting is the notion of Temporal Sequence, a robust though violable constraint. In particular, Temporal Sequence is subordinate to WFC, No-Backtracking, and One Step Principle. In this section we consider the interplay between temporal sequencing and other forces.

3.1 Structural affinity and cyclic effect

Recall from section 1 of chapter 3 that the same input tonal string /214.214.214/ in Mandarin surfaces as either [35.35.214] or [214.35.214] depending on the morphosyntactic structure of the tone-bearing elements, as shown in the examples repeated below (= (11) chapter 3):

[7] There remain some recalcitrant cases, which Hsu (1994) attempted to account for under the rubric of "One Step Principle." Basically, this principle bars a sandhi rule from applying to an input that has been previously altered by another sandhi rule. For details, see Hsu (1994).

(15) a. tiger gall "brave"
[*lao-hu*] *dan*
214.214.214
T3 Sandhi
35. 214.214
T3 Sandhi
35 35 214

b. paper tiger "coward"
zhi [*lao-hu*]
214.214.214
T3 Sandhi
214. 35. 214
T3 Sandhi

T3 Sandhi: 214 → 35 / __ 214

It is obvious that the two readings (a, b) are derived by means of the familiar cyclic mode of rule application.

It is worth noting that cyclicity is a reflection of a more general principle, which I will refer to as Structural Affinity: morphosyntactically most closely related items are processed together (e.g. for tonal coarticulation) before joining more distant elements into larger constructions. Structural Affinity (hereafter simply Affinity) manifests itself in other ways as well. In this connection consider the case of **Huojia**, another northern Mandarin dialect fairly extensively reported in He (1989). This otherwise little-known northern Mandarin dialect of Henan province has /33, 31, 53, 13 and 3q/ as underlying tones.[8] Of these, /13, 3q/ do not participate in tone sandhi; they invariably show up as [13, 3q]. The remaining three tonal categories should yield nine logically possible two-tone sequences (3 × 3); in fact only two sandhi patterns are attested: [31.13] and [53.13]. The correspondence between base tones and sandhi tones is summed up in (16). What is going on is, of course, a sweeping neutralization of tone shape contrasts: all two-tone sequences exhibit the pattern F-R, namely a falling tone followed by a rising tone. Furthermore, the contrast in register is only partially preserved: thus, /53/ retains its high register in the initial position for the most part,[9]

[8] Recall that -q stands for a tone associated with a "checked" syllable ending in /p,t,k/ or /q/.

[9] I have oversimplified the picture somewhat. For our purposes I will ignore the fact that /53.53/ → [31.13] unpredictably in some cases, [53.13] in others. Also, in a number of sporadic cases, /53.53/ → [31.53]/. Furthermore, whether a disyllabic compound undergoes tone sandhi at all is lexically determined. For instance, /33.33/ → [31.13] in the case of *dong-xi* "width" (lit. east + west), but remains unchanged [33.33] in the case of *yin-tian* "overcast" (lit. shade + sky). For details, see He (1989:33ff.).

but becomes uniformly low registered [13] in the final position – that is, a rising tone is predictably low registered.

(16)

	2nd tone		
1st tone	33	31	53
33	31.13	31.13	31.13
31	31.13	31.13	31.13
53	53.13	53.13	53.13

Abstracting away from the high vs. low register contrast, the relevant generalization, as far as (16) is concerned, is expressible as a sort of two-tone template:

(17) Template
T + T → F + R
Key: F, R = falling, rising

It takes only a moment's reflection to realize that when extended to trisyllabic expressions, Template would, by the very nature of things, make inconsistent predictions – we get either [F-R] + T or T + [F-R] depending on whether we superimpose the Template on T1-T2, or T2-T3 as illustrated below.

(18)

Alternatively, if the rule expressed as Template applies left to right or in reverse, it would yield either FFR or FRR as output, as shown in (26a) and (b) respectively. Needless to say, if Template were to freely apply whenever applicable, it would result in infinite recursion.

(19) a. T T T (left to right)
F R T
F F R

b. T T T (right to left)
T F R
F R R

What, then, decides how the F-R template is to be matched with which substring? The answer comes as no surprise: the template is superimposed on immediate constituents. This is illustrated in the following contrast:

(20) [*huang-shu*] *lang* "weasel" (lit. squirrel + wolf)
31 53 33
a. [31 13] 33 ok
b. 31 [31 13] *

(21) *yang* [*wei-ba*] "goat's tail"
31 53 33
a. [31 13] 33 *
b. 31 [31 13] ok

The patterns exhibited by the above examples are quite consistent. I found 56 examples instantiating the 27 three-tone combinations (3^3).[10] Of these, 50 are consistent with Affinity, while 6 are exceptions.

That Affinity-sensitive Template matching is not translatable in terms of cyclic rule application can be shown by (22). Changing *wei-ba* [53.33] into [53.13] should in no way prevent Template from turning *yang wei* [31.53] into [31.13], resulting in *[31.13.13], whereas the attested sandhi form is [31 [31.13]]. It can be seen that both cyclic rule application and preemptive Template matching are rooted in the more general principle of Affinity.

(22) *yang* [*wei-ba*]
31 53 33
31 31 13
31 13 13 = *
31 31 13 = attested

Note in passing that although kindred in spirit, Affinity is not reducible to cyclicity as a property of lexical rules. (Classical) lexical phonology derives cyclicity from the interleaving of phonological and morphological processes. Therefore, cyclicity follows from the very nature of lexical rules (Kiparsky 1982b, Mohanan 1982, Cole 1995). But the cyclic mode of application extends to the phrasal level as well, as can be shown in the following examples taken from Standard Mandarin. (23a) demonstrates that tone sandhi is not restricted to the lexical domain, but applies between

[10] After eliminating three examples, where morphological constituency is ambiguous.

the subject *gou* "dog" and the predicate phrase *yao mao* "bit the cat." What then blocks tone sandhi in the same syntactic environment of (23b)? Obviously, the application of tone sandhi on the inner bracketed verb phrase *yao wo* "bit me" takes precedence, and therefore bleeds, the same rule on the larger construction.

(23) a. *gou* [*yao mao*] "the dog bit the cat"
214. 214.55

T3 Sandhi (not applicable)

214.214. 55
|
35. 214. 55 T3 Sandhi

b. *gou* [*yao wo*] "the dog bit me"
214.214.214
|
214. 35. 214 T3 Sandhi

T3 Sandhi (not applicable)

3.2 *Temporal Sequence and rhythmic effect*

It is obvious that when the two general processing principles conflict with each other, Affinity generally prevails over Temporal Sequence in both Huojia and Standard Mandarin. Other dialects of Chinese display a less clearcut preference. **Xuzhou**, also a northern Mandarin dialect, has a virtually identical rule to that of Standard Mandarin: the citation tone [213] changes into a rising tone [24] before another "dipping" tone [213].[11] Our source (Su 1985) observed that for left-branching constructions like (24a), there is only one reading, namely [24.24.213]; for right-branching constructions, exemplified by (24b), on the other hand, there are two alternative readings [24.24.213] and [213.24.213].[12] How do we explain such asymmetry? The most natural explanation seems to be this: when Affinity and Temporal Sequence conspire to dictate a left-to-right direction of rule application, as in the case of left-branching constructions like (24a), only one reading is permissible. On the other hand, in right-branching

[11] Phonetically, Xuzhou [213] and [24] correspond to Standard Mandarin [214] and [35] respectively; however, the Xuzhou tones derive from different historical sources from those of Standard Mandarin.

[12] [213] is fully realized only in the final position; elsewhere [213] is truncated to [21] by a process that has an analog in other Mandarin dialects. The truncated [21] is referred to as "half-third tone" (*ban shang*) in Standard Mandarin. I ignore this detail for expository clarity.

constructions, while Temporal Sequence calls for a uniform left-to-right iteration of the sandhi process, Affinity motivates an opposite directionality. The indeterminate dominance relation between these two conflicting constraints gives rise to the alternative readings (i) and (ii).

(24)	a.	[*lu yin*] *ji*	"cassette recorder"
		213.213.213	base tone
		24. 24. 213	sandhi form
	b.	*kai* [*fei ji*]	"to pilot an airplane"
		213.213.213	
	i.	24. 24.213	
	ii.	213. 24.213	

Actually, **Standard Mandarin** allows two alternative readings for right-branching expressions like (25b) as well – except that (25b-ii) is characterized as allegro speech.

(25)	a.	[*lao hu*] *dan*	"brave"
		214.214.214	
	i.	35. 35. 214	ok
	ii.	214. 35. 214	*
	b.	*zhi* [*lao hu*]	"paper tiger, coward"
		214.214.214	
	i.	214. 35. 214	normal tempo
	ii.	35. 35. 214	allegro speech

The standard account (after Shih 1986, 1997) trades on the notion that internal structures are ignored or "flattened" out in fast, casual speech. In allegro speech, therefore, both (a) and (b) are treated as internally unstructured sequences of syllables. Absent the Affinity constraint in an unstructured string, the default mode of phonological processing is left to right, as demonstrated by the fact that reading (ii) of (a) is unacceptable at any tempo, even though such a reading is in principle derivable by applying the tone sandhi rule right to left to the internally unbracketed string [*lao hu dan*] "brave."

Shih's account of the alternative reading (ii) of (25b) as an unstructured prosodic unit seems to be on the right track. Consider (26). There is no reason to posit any structure in a sequence of the three *five*s. Given an internally unstructured p-word like (26), Temporal Sequence dictates a rightward iteration of tone sandhi, yielding reading (i); reading (ii), derivable only through a leftward parse, is unacceptable.

(26)	*wu wu wu*	"five-five-five" (a brand name)
	214.214.214	
i.	35. 35. 214	ok
ii.	214. 35. 214	*

We conclude that in the absence of a countervening principle like Affinity, the default mode of rule application is left to right, consistent with Temporal Sequence. The only difference between Xuzhou and Standard Mandarin is that in the former case Temporal Sequence often prevails even at normal tempo.

As an alternative to the preceding analysis, one could stipulate that in unstructured strings, which do not otherwise impose a cyclic application, phonological rules directly map a base tone to a corresponding sandhi tone in the appropriate contexts (i.e. a [214] when followed – underlyingly – by another [214]). Although direct mapping works in the particular case of Standard Mandarin, it does not hold true in general. Take the Tianjin P3 /FFF/. Recall that in Tianjin, a F surfaces as L before another F, and as H before a L tone. A monomorphemic word like *yi da li* "Italy" with lexically assigned tone pattern /FFF/ presumably represents an internally unstructured string of syllables. Direct mapping would yield the output LLF in (27a). The attested sandhi form HLF (27b) is derivable only by applying the sandhi rules in a right-to-left direction.

(27)	a.	*yi da li*	"Italy"
		F F F	
		| |	by Direct mapping of Dissimilation (c)[13]
		L L F	= *
	b.	*yi da li*	
		F F F	
		|	by Dissimilation (c)
		F L F	
		|	by Absorption
		H L F	= attested

It is worth emphasizing the point that direct mapping or simultaneous multiple application of a phonological rule is not, generally speaking, a viable alternative to directional rule application.

[13] Stated as (15c) in chapter 3.

4 Derivational economy and structural complexity

In our account of Tianjin, aside from Backtracking, derivational economy plays no overt role in determining the sandhi output. Recall that given an input string like /RRR/, a right-to-left application of the Dissimilation rule (RR → HR) yields a perfectly well-formed output RHR in one single move; the attested output HHR, on the other hand, requires two successive derivational steps, in a left-to-right direction. This means that Temporal Sequence ≫ Derivational Economy. Note in this connection that we considered and rejected the output constraint *Complex (which penalizes a contour or branching tone), as a factor in determining the sandhi outputs of Tianjin trisyllabic expressions (see section 5.3 of chapter 3). In general, neither *Complex nor derivational economy (aside from No-Backtracking / One Step Principle) plays any overt role in our analysis of Tianjin, Changting, Huojia, Xuzhou, or Standard Mandarin. This does not mean that derivational economy and structural complexity never play a role in determining how phonological rules are implemented and how they interact with each other. In fact these two concepts hold the key to the analysis of what appears to be a rather complex set of sandhi facts in **Boshan**.

4.1 Boshan: basic facts

Boshan is a northern Mandarin dialect spoken in Shandong province. As documented in Qian (1993), it has three citation tones: /214, 55, 31/. The low falling tone /31/ remains unchanged for the most part. Both 214 and 55 assume various sandhi forms depending on the shape of the adjacent tone to the right. The citation tone 55 comes from two distinct underlying representations, simply marked here as /55a/ and /55b/.[14] Their underlying contrast surfaces in the context of another /55/: /55a.55/ shows up as [53.55], while /55b.55/ turns into [214.55]. /214, 55a, 55b/ and in some cases /31/ become 24 before /31/. These observations are summarized in table 4.3, restated as rules in (28), and instantiated by the examples of (29).

(28) a. 214 → 55 / ___ 214
b. 55a → 53 / ___ 55
c. 55b → 214 / ___ 55
d. 214, 55, (31) → 24 / ___ 31

[14] Generally speaking, /55a/ derives from Middle Chinese lower register *ping* category, while /55b/ is the modern reflex of Middle Chinese *shang* tone.

Table 4.3.

1st tone	2nd tone		
	214	55	31
214	55.T	T.T	24.T
55a	T.T	53.T	24.T
55b		214.T	
31	T.T	T.T	(24.T)

T = base tone, unchanged
a, b = diacritics indicating different historical sources of the citation tone /55/ tones are separated by a dot "."

(29) Examples illustrating rules (28a–d)

a. *chun fen* "the spring equinox"
214.214 base form
55. T[15] sandhi form

b. *qi ma* "horse-riding"
55a.55
53. T

c. *tan bai* "to confess, be candid"
55b.55
214. T

d. *xiang xia* "countryside"
214.31
24. T

e. *cheng shi* "city"
55.31
24. T

f. *ban ye* "midnight"
31.31
24. T

g. *dui xiang* "target"
31.31
31. T (no change)

How do the rules of (28) work together to produce the desired sandhi forms of longer strings of tones? Given four underlying tonal categories,

[15] T indicates unaltered base tone.

Table 4.4.

	Input	Output	Example	
P1	55a.214.214	55.55.214	*[nan guan] jie*	"Nanguan street"
P2	55b.214.214	55.55.214	*gui [chui deng]*	"dirty tricks"
P3	55a.55.31	55.24.31	*[chang jing] lu*	"giraffe"
P4	55a.55a.55	55.53.55	*[tai tou] wen*	"wrinkles on one's forehead"
P5	55a.55b.55	55.214.55	*chang [guo ren]*	"peanut"
P6	55b.55a.55	55.53.55	*[bi liang] gu*	"septum"
P7	55b.55b.55	55.214.55	*[guan li] yuan*	"person in charge"
P8	214.214.214	55.55.214	*[shou yin] ji*	"radio"
P9	214.214.31	55.24.31	*[chuan yi] jing*	"full-length mirror"

there are sixty-four distinct three-tone combinations. Naturally, many of these combinations yield the same outputs regardless of the direction in which the rules apply. However, we can identify nine tonal sequences that are direction-sensitive, that is to say, that give rise to different output forms depending on whether the sandhi rules apply left to right or in the opposite direction.[16] These nine direction-sensitive patterns, numbered P1–9, are listed in table 4.4. Our task is to ascertain the mode of rule implementation that guarantees the correct sandhi forms.

[16] Excluded are three types of tonal combinations: (a) /T.31.31/ or /31.31.T/, (b) /214.55b.55/, and (c) /55b. 55.31/. (a) is excluded because sandhi change on the /31.31/ substring is optional. (b) has two possible outputs (i) and (ii), but I could not ascertain which of the two sandhi forms is the correct one for lack of appropriate instantiation in the source (Qian 1993).

```
        ⇨                          ⇦
(i)  214.55b.55  (n/a)      (ii)  214.55b.55
                                      |
     214.55b.55                   214.214.55
        |                            |
     214.214.55                   55.214.55
```

Finally, (c) is discarded because the attested output is not derivable by applying the rules in either direction:

```
        ⇨                  ⇦
(iii)  55b.55.31     (iv)  55b.55.31
        |                    |
     214.55.31            55b.24.31   (no applicable rule)
        |
     214.24.31
```

For reasons that are not clear to me, the attested output [214.55 # 31] (e.g. *zhi jia gai* "fingernail") consistently blocks tone sandhi between the last two tones (indicated by "#").

4.2 *Complex ≫ Derivational Economy

Temporal Sequence fails to consistently make the correct predictions for table 4.4. As can be seen more perspicuously in the tableau in (30), while P1–2 and P8–9 undergo the tone sandhi rules stated as (28) (equivalently table 4.3) operating from left to right, P3–7 must be scanned in the opposite direction. Nor is there an obvious correlation between morphosyntactic constituency and the direction of rule application. For instance, the example of table 4.4 instantiating P2 /55b.214.214/, namely *gui [chui deng]* "dirty tricks," displays a right-branching structure, but its output [55b.55.214] (P2a) is derivable only by applying the relevant rules in a left-to-right sweep, contrary to what one might expect, if cyclicity based on structural affinity were the overriding principle. Clearly Temporal Sequence and Affinity are low-ranked constraints, dominated by some other higher imperatives.

(30)

						*Complex	Derivational Economy	Temporal Sequence
P1	☞	a	⇨	55a.214.214	(n/a)	*	*	
				55a.214.214 | 55a.55. 214	= Q8			
		b	⇦	55a.214.214 | 55a. 55. 214 | 53. 55. 214		*	**	*
P2	☞	a	⇨	55b.214.214	(n/a)	*	*	
				55b.214.214 | 55b. 55. 214	= Q8			
		b	⇦	55b.214.214 | 55b. 55.214 | 214. 55.214		**	**	*

(30) cont'd

					*Complex	Derivational Economy	Temporal Sequence
P3	a	⇨	55a.55.31 53. 55.31 53. 24.31			**	
☞	b	⇦	55a.55.31 55a.24.31	(n/a) = Q5		*	*
P4	a	⇨	55a.55a.55 53. 55a.55 53. 53. 55			**	
☞	b	⇦	55a.55a.55 55a. 53. 55	(n/a) = Q11		*	*
P5	a	⇨	55a.55b.55 53. 55b.55 53. 214.55		*	**	
☞	b	⇦	55a.55b.55 55a.214. T	(n/a) = Q2	*	*	*
P6	a	⇨	55b.55a.55 214. 55a.55 214.53. 55		*	**	

(30) cont'd

						*Complex	Derivational Economy	Temporal Sequence
P6	☞	b	⇦	55b.55a.55 55b.53. 55	(n/a) = Q11		*	*
P7		a	⇨	55b.55b.55 214.55b.55 214. 214. T		**	**	
	☞	b	⇦	55b.55b.55 55b.214.55	(n/a) = Q2	*	*	*
P8	☞	a	⇨	214.214.214 55.214.214 55. 55. 214	= Q8	*	**	
		b.	⇦	214.214.214 214.55. 214	(n/a)	**	*	*
P9	☞	a	⇨	214.214.31 55.214.31 55. 24. T	= Q5		**	
		b	⇦	214.214.31 214. 24.31	(n/a)	*	*	*

a,b = diacritics indicating different historical sources of the citation tone [55]
Q8, 11 etc. = surface tonal patterns n.8, 11 etc. listed in Qian (1993:27ff.)
TT = (underlined) = current window spanning a two-tone substring
n/a = not applicable to the current window

What then determines the mode of rule application? Take the two alternative derivations of P3 in (30). I will refer to the rules stated in (28) simply as rules (a, b, c, d). Scanning the input from left to right, the sandhi rules of (28) would turn P3a /55a.55.31/ first to [53.55.31] by rule (c), then [53.24.31] via rule (d), in two successive derivational steps. In contrast, the right-to-left application of sandhi rules involves one single step: turning P3b /55b.55.31/ into [55b.24.31] by rule (d) destroys the requisite sandhi context for the application of rule (c), thereby rendering the next step of the operation unnecessary. It is the latter, shorter of the two derivations that produces the surface tonal pattern [55.24.31] (= P3b) for the word *[chang jing] lu* "giraffe" (lit. "long-necked camel"). This choice seems eminently sensible, if we think of phonological rules as embodying a set of well-formedness conditions to be met. Now, of the two alternative ways for the underlying string /55b.55.31/ (= P3) to satisfy the conditions, clearly the leftward scan involves the "least effort," measured in terms of derivational steps. It represents the most "efficient" way of meeting all the conditions stated as (28). This principle of Derivational Economy asserts its dominance over Temporal Sequence in forcing a *right-to-left* direction of rule application on patterns P3 to P7, with the desired results. We measure derivational economy by simply counting the number of steps, as indicated by an asterisk in the column labeled "Derivational Economy" in the tableau in (30).

Needless to say, Economy can converge with Temporal Sequence in dictating a left-to-right scan; this is the case of P1 and P2. There remain only two cases, where length of derivation makes the wrong prediction, namely P8 and P9. In both of these instances, the right-to-left application achieves the goal with the greatest economy of effort. What then takes priority over Derivational Economy? Notice that the winning candidates [55.55.214] (P8a) and [55.24.31] (P9a) carry the fewest instances of [214], a highly complex tone that is both falling and rising. The extra derivational steps are compensated for by the decrease in the complexity of the tonal patterns. Let us refer to this target as *Complex (no complex tone).[17] In (30), each occurrence of [214] (= MLH) is assessed one unit cost (*) in the *Complex column. By imposing the ranking of (31) on the competing constraints, we guarantee the desired winning candidates, as demonstrated in (30).[18]

(31) *Complex ≫ Derivational Economy ≫ Temporal Sequence

[17] A species of *Structure (No structure). Cf. Prince and Smolensky (1993, e.g. p. 25).

[18] (31) adequately account for our core database, which consists of the 19 surface tonal patterns listed by Qian (1993). In addition, however, the author also provides a fairly

5 Concluding remarks

It is worth stressing that what appeared at first to be a fairly complex set of facts across a number of dialects discussed both in this and the preceding chapters turned out to be reducible to being handled by a handful of guiding principles that regulate the "traffic" of phonological rules. This non-trivial result is made possible only by exploiting and extending in some unconventional ways some of the seminal insights of Optimality Theory. In particular, we need to entertain the notion of constraints on derivation as well as constraints on the output. This, in turn, entails ascribing a non-trivial role to derivation or GEN.

Some of the constraints we have found to be crucial are strictly monostratal output constraints, such as WFC (including OCP of various sorts) and *Complex; others are relational, namely faithfulness conditions (Max, Dep, Linearity, and featural Identity). Both of these types of constraints are recognized in standard OT. More controversial are constraints on derivation, including Temporal Sequence, No-Backtracking, One Step Principle, and Derivational Economy. We had argued at length that these derivational constraints cannot be recast in declarative terms, either as output constraints or faithfulness conditions of OT, or as mapping relations modulo two-level rules. In some cases there is a direct translation between derivational economy and faithfulness. Take Boshan for instance. One can simply substitute Faithfulness for Economy in the tableau in (30), since each derivational step entails one degree of deviation from the input. However, as pointed out in section 5.5 of the preceding chapter, the one-to-one equivalence between faithfulness and derivational economy does not hold in general.

By the same token, the sandhi facts are not amenable to an analysis in which we simply stipulate the order of rule application – for the simple reason that many of the cases involve (multiple applications of) one single rule. Nor can we appeal to simultaneous rule application or specify the

extensive body of phonetic transcriptions of running texts and other elicited forms. I have found in the extended corpus such examples as *tuo la ji* "tractor" /214.214.214/ → [214.55.214] (p. 119), side by side with *shou yin ji* "radio" /214.214.214/ → [55.55.214] cited above as Q8 (= P8, p. 28). This suggests some degree of ranking indeterminacy between Economy and *Complex. I take the core database to represent the norm, and interpret the extended corpus as (secondary) variants. Of the core database, there remain two exceptional examples: *pu ji ben* "popular edition" and *cao xie di* "the sole of straw sandals" are underlyingly /55b.55a.55/, and both surface as [214.53.55], side by side with *bi liang gu* "septum (of nose)," which has exactly the same base tones, but surfaces as [55.53.55] (P6), consistent with our analysis.

directionality of iteration on individual rules, for instance, rule X applies from left to right, but rule Y from right to left, etc. Instead, tone sandhi rules are unordered[19] and unspecified with respect to directionality. How a sandhi rule applies singly (but iteratively) or interacts with some other rules (e.g. in a feeding, bleeding, counterfeeding, counterbleeding order) is entirely determined by a set of general, rule-independent principles.

The picture that emerges from the facts presented in the last two chapters is this: the grammar disallows certain tonal concatenations, and specifies their corresponding well-formed alternants. How these input/output conditions are to be satisfied is determined by constraints, some of which are presumably rooted (at least in part) in general principles of language processing.

[19] With the sole exception of the "Preemptive clause" in Tianjin, which stipulates that Dissimilation must precede Tonal Absorption. See chapter 3, section 3.2.

5 *From base tones to sandhi forms: a constraint-based analysis*

Tonal alternations in many well-documented dialects of Chinese seem to instantiate essentially "fortuitous" and "arbitrary substitutions" of one tone by another,[1] a phenomenon that bears little resemblance to the familiar processes of assimilation, dissimilation, or simplification (equivalently, spreading, delinking etc.). Whatever phonetic and/or functional motivations such alternations once had must have faded into the nebulous past. At first blush, the facts of **New Chongming** presented *en bloc* in the Appendix look rather daunting, and appear to reinforce this general impression. I will show that upon closer inspection, the complex set of facts not only yields to a straightforward *description* in terms of standard rewrite rules (see Chen 1991c), but in fact lends itself to a reasonably *explanatory* treatment in a constraint-based approach. The principal appeal of a constraint-based approach lies in its insistence that what often are stated as arbitrary language-specific rules can be best explained in terms of, and in effect reduced to, general constraints on optimal outputs. Instead of simply formulating a set of transformational rules to crank out the correct sandhi forms, I will attempt to render an account of the facts by appealing to a set of general principles. The basic idea is this: while the rich underlying tone system of New Chongming potentially gives rise to a large repertoire of sandhi patterns when tone-carrying syllables combine into lexical compounds, the "tonotactics" of the language rejects all but a tiny fraction of the logically possible tonal combinations. In order to conform to the tonotactics – that is, the general principles governing admissible tonal sequences – the lexically specified tonal strings must undergo modification, often drastically. What specific changes must take place is dictated by the

[1] The quotes are taken from Anderson (1978:157f.), who commented on the baffling patterns of tonal alternation in Xiamen (a.k.a. Amoy). In a similar vein, Schuh (1978:249) characterized Wang's (1967) analysis of the Xiamen "sandhi circle" using paired variables as "a formal tour de force" that merely gives the illusion of capturing a true linguistic generalization. It turns out that the theoretical interest of Xiamen tone sandhi lies not in its phonetic manifestation, but in the syntax–phonology interface, more about which in chapter 10.

principle of economy: make the least "costly" changes to satisfy the imperatives of well-formedness conditions. Thus, suppose word formation strings together into a disyllabic compound two morphemes carrying lexically specified tone A and tone B, and suppose further that the sequence A-B constitutes an unacceptable tonal pattern, one could change A to C, or B to D, or else delete one or the other tone. In many cases, altering the tonal properties (register, melodic contour) turns out to be more "costly" than simply getting rid of one of the two abutting tones. Moreover, tonal deletion does not target just any tone at random; rather, it tends to eliminate the lowest ranked tonal category on a scale of "saliency" defined in terms of complexity (level vs. contour tones), register (high vs. low), and prosodic strength (right- vs. left-prominent). Once we identify the well-formed targets, and know the price tag attached to each of the logically possible moves, we can determine the best course of action and predict the optimal outcome.

Just as the constraint-based framework gives us the necessary conceptual tools and descriptive apparatus to deal with the complex tonal alternations of New Chongming, the facts we are about to present raise interesting issues about OT (Prince and Smolensky 1993, McCarthy and Prince 1993a, 1995a, b), by far the most radical and the most explicitly articulated and empirically tested model in the family of constraint-based approaches (cf. Singh 1987, Paradis 1988, Goldsmith 1990, 1993b, Chen 1991c).[2] In particular, New Chongming affords us empirical data relevant to such issues as opacity and constraint ranking. In OT, most of the work is done by constraints, which "are evaluated *only* with respect to outputs" (Prince and Smolensky 1993:205; my emphasis). On the other hand, the phonological literature abounds in examples of opaque situations, where strictly speaking constraints or significant linguistic generalizations hold true not over the phonetic output, but at some other level of representation. Opacity presents non-trivial problems for an output-driven model (see McCarthy 1995a and Cole and Kisseberth 1995), which will be discussed in the light of New Chongming data (sections 4–5).

At the very heart of OT lies the "Harmonic Ordering of Forms." That is, for a lexical input, a whole set of possible phonological parses is evaluated. "The correct output is the candidate whose complete structure best satisfies the constraint hierarchy. And 'best satisfies' can be

[2] The body of work inspired by constraint-based phonology has grown too large to enumerate; for a sample, see Prince and Smolensky (1993:2) and periodically updated bibliography available through the (electronic) Rutgers Optimality Archive.

recursively defined by descending the hierarchy, discarding all but the best possibilities according to each constraint before moving on to consider lower-ranked constraints" (Prince and Smolensky 1993:19). In this recursive testing of constraint satisfaction, the violation of a higher ranked constraint dooms an output candidate, regardless of how it fares with regard to the lower ranked constraints. While this modus operandi in accordance with what is referred to as Strict Domination principle yields the desired empirical results and makes interesting predictions in the vast majority of cases investigated so far, New Chongming seems to present a limiting case where candidates are evaluated according to the *sum total* pattern of constraint satisfaction. In other words, it may turn out to be the case that constraints are not only ranked, but variably weighted, so that each violation contributes to the overall harmonic status of an output candidate.[3] This issue is discussed in section 6.

The facts of New Chongming are rather complex and the issues involved quite subtle. Some readers may want to get the general idea from sections 1–3, and skip on a first pass the more technical issues raised in sections 4–6.

1 Background

Before we proceed any further, a few words about New Chongming are in order. Chongming is a northern Wu dialect of Chinese, spoken on the island by the same name in the Yangzi delta, about 30 miles north of metropolitan Shanghai. There are two major subdialects: Old Chongming and New Chongming. The former, spoken by older speakers (roughly speaking, aged 40+), has been reported extensively in H-Y. Zhang (1979–80) and Chen and Zhang (1997); the latter, spoken by the younger generation – in our case, by a college student in his early twenties – is the main focus of our attention.[4] New Chongming has an underlying eight-tone system, cross-classified by syllable type, register, and melodic shape, as given in (1). As used here, the term "even tones" refers to monotones or, if you will, single-note melodies, while "oblique tones" stands for contour or branching melodies consisting of several tone segments. Contour tones may have either a rising [MH, LM] or a falling [HM, ML] pitch

[3] See Smolensky's comments on this subject posted over the OT-net (November 18, 1995).

[4] The New Chongming data were collected in two separate fieldtrips, one in the fall of 1986, the other in the summer of 1990. Pan Wuyun, Qian Nairong, Zhang Hongming, and Chen Zhongmin helped me in various aspects of elicitation and transcription. Their invaluable assistance is gratefully acknowledged.

movement. Note in passing that "checked" syllables (CVq, where -q is an obstruent coda) may only carry an even tone; "smooth" syllables (= CVN), on the other hand, may carry an even or an oblique tone.[5]

(1)

	even		oblique	
	smooth	checked	falling	rising
high register	H	Hq	HM	MH
low register	L	Lq	ML	LM

H, M, L = high, mid, low
Hq, Lq represent H and L associated with a checked syllable (CVq).

Restricting ourselves for the moment to disyllabic lexical tone sandhi, given the eight base tones, there are sixty-four (8 × 8) combinatorial possibilities in a lexical compound. In actual fact, only a small number of two-tone sequences are attested; in most other cases, one or the other tone is deleted. The full range of correspondences between underlying tonal sequences and their sandhi forms is laid out in the Appendix. Instead of plunging into the thicket of the complex correspondences between the base tones and the sandhi forms all at once, I shall adopt a divide-and-conquer strategy, by breaking up the larger problem into smaller ones. First of all, we discard a handful (five, to be exact) of sporadic exceptions, marked with an asterisk (*) in the Appendix.[6] Next, by omitting certain details from the broad phonetic transcription of the Appendix, we reduce the sandhi patterns of all disyllabic compounds to just three basic types, listed in (2):

(2) Permissible sandhi patterns in New Chongming

basic types	subtypes	
	T = even	T = oblique
o.T	o.H	o.MH
T.o	H.o, L.o	MH.o
T.T	H.H, L.H	HM.MH (H.M)

T = any tone
o = zero tone (no tone)

[5] Recall that a "smooth" (a.k.a. "slack" or "legato") syllable is CV, CVG, or CVN.
[6] For instance, *kao-shi* "examination" has a perfectly well-formed underlying sequence /HM-MH/; it should therefore surface unchanged, exactly like *xiao-qi* "stingy" /HM-MH/ = [HM-MH]. There is nothing in our account that is capable of deriving the attested sandhi form [H-HM] for *kao-shi*.

For instance, there are two subtypes of [o.T], a sandhi pattern consisting of an atonic syllable (symbolized by "o") plus a tonic or tone-bearing syllable (T). If the T in question is an even tone, then the pattern shows up as [o.H]; if, on the other hand, the surviving T belongs to the oblique category, the sandhi form is phonetically realized as [o.MH]. As for the [T.o] pattern, there are three subtypes: [MH.o] if T = oblique, [H.o], and [L.o], if T = even. In other words, the even tone maintains the H vs. L contrast in the initial position, but not in the final position. Oblique tone, on the other hand, shows up uniformly as [MH] – except when followed by another [MH], in which case it emerges as [HM]. We will elaborate on these and other observations in the sections that follow.

For the sake of clarity, I will make a few simplifying assumptions and expository decisions. First, [o] (lower-case "o") symbolizes the "zero" tone. The zero tones, or more precisely the syllables which lose their lexically associated base tones in sandhi contexts, acquire default pitch values: an indeterminate, mid-level pitch in initial position, and a short, slightly falling pitch in final position. I will assume that the toneless syllables uniformly take on a default L, and attribute the positionally dependent variation to the downward trend in pitch across a phrase, so that the phrase-initial L is somewhat higher than the final L. This cross-linguistic phenomenon of downtrend is given the term "declination" (cf. Liberman and Pierrehumbert 1984 and literature cited therein; with reference to Chinese in particular, see Shih 1987, 1988). In the interest of expository clarity, I will ignore these "low-level" phonetic details in the rest of this chapter. Second, the high-even tone H surfaces in two allotones: [HM] when associated with a "smooth" syllable in final position; [H] elsewhere. Thus, [L.H] in (2) is actually pronounced as [L.HM], when the second syllable is "slack," as in an example like *nian qing* "youth" (see (5), P6 below). This reflects the phonetic effect of an intonational boundary L. This intonational effect does not show up in [L.Hq] (where "q" indicates a checked syllable), simply because checked syllables unexceptionably bar contour tones. Third, the low-level tone [L] when not followed by another tone is phonetically realized as low-rising [LM]. Thus [L.Hq] remains as is, but [L.o] is pronounced with a distinct rising pitch: [LM.o]. The reason for this may well be a functional one, namely to maintain the distinction between an underlying L occupying the only tonic position and a default L assumed by atonic syllables. I will take up the question of H/HM and L/LM in section 5.1. Finally, I regard the parenthesized pattern [H.M] in (2) as a register-dependent variant of [HM.MH], and

defer discussion of its distribution until section 4.3. I have omitted from chart (2) certain phonetic details.

2 Parallel constraint satisfaction

2.1 Restrictions on sandhi patterns

It is obvious from (2) that New Chongming imposes stringent restrictions on well-formed tonal sequences. These restrictions or constraints play a central role in jointly determining the sandhi forms of disyllabic compounds. In particular:

(i) Recall that tones fall into two major categories: E ("evens," or level tones) and O ("obliques," or contour tones). New Chongming bars a sequence of E-O or O-E, although it permits E-E and O-O concatenations. Intuitively speaking, what is at play here is relative tonal salience or stability: evens are dominant over obliques (rising or falling); hence, given two abutting tones of *unequal* saliency, it is the weaker (oblique) of the two that is forced out. Thus E-O → T.o, and O-E → o.T.

(ii) No disyllabic compound may begin with a tone-bearing checked syllable. Presumably this restriction reflects the joint effect of two converging forces: first, checked syllables (CVq) are poor tone carriers, since the glottal stop (-q) abruptly cuts short the vocalic nucleus capable of sustaining a distinctive pitch; second, as we shall see in the subsequent discussion, New Chongming is right-prominent, so that, other things being equal, [σ-σ] exhibits a [weak-strong] prosodic structure. It is not surprising that a poor tone carrier in a prosodically weak (initial) position should lose its distinctive tone. Constraint (ii) takes precedence over principle (i). Thus /Hq.MH/ → [o.MH], in which the surviving tone is the oblique rather than the even, owing to the fact that the even tone Hq is linked to an initial checked syllable.

(iii) The even/oblique dichotomy is absolute. That is to say, while tones may exchange one register for another, or replace a rising by a falling contour, or vice versa, they are segregated into two "water-tight" categories which do not intermingle: neither do even tones turn into a rising or a falling tone, nor do oblique tones flatten out into a level pitch. Put differently in geometric terms, while metathesis (HM ↔ MH, or ML ↔ LM) and featural change affecting [±upper] freely occur, neither inserting (H → MH, HM; or L → LM, ML) nor deleting (MH, HM → H; or LM, ML → L) a tone segment is permitted. This can be verified by examining (2), where all surviving even tones in the middle column emerge as either

H or L, and all oblique tones that escape tone deletion in the right column show up as MH or HM.[7]

(iv) Contrasts in both register and melodic shape are drastically reduced. With respect to melodic shape, even tones do not contrast in this regard to begin with; oblique tones, on the other hand, contrast in shape (rising ≠ falling) underlyingly. Remarkably, all oblique tones surface uniformly as rising [MH] – except when followed by another rising tone, in which case the first rising tone turns into a falling tone yielding the sequence [HM.MH], a familiar OCP effect. Since the distribution of HM is contextually restricted and predictable, there is no longer a phonemic contrast in contour shape between HM and MH.

(v) Finally, the high/low register contrast is lost entirely in oblique tones; it is partially maintained for even tones only in the initial position. The first/last position asymmetry in register neutralization is not uncommon; for instance, it parallels that of Lhasa Tibetan. As reported in Hu (1979; cf. Shih 1986), in this variety of Tibetan, the first syllable of a disyllabic compound maintains its register contrast, but neutralizes its contour to a simple level tone; conversely, the last syllable retains the rising/falling contrast, but merges its tones to a high register. Schematically:

(3) Lhasa Tibetan

	first σ	last σ
Register neutralized to high?	no	yes
Contour neutralized to level?	yes	no

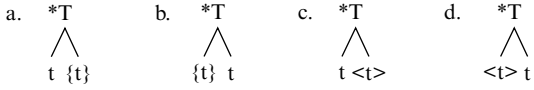
Although the specific details are quite complex, it is in principle reasonably straightforward to devise a set of standard rewrite rules to "transform" the base tones into their corresponding sandhi forms. This I have done in Chen (1991c, 1997). Instead, I will attempt here an OT-style analysis of these facts. The fundamental advantage of this approach is that instead of merely enumerating the generalizations made above as grammatically *inert* well-formedness conditions on the sandhi forms, apart from the transformational rules that generate the output forms, OT makes *active* use of these constraints by assigning them the central task of selecting out

[7] These generalizations are true as far as the patterns of (2) are concerned, abstracting away from the intonational effect of [H] → [HM], and setting aside the register-dependent variant [H-M], which is as yet not fully understood. We will return to these subsets of data in section 5.

of a "teeming pool" of conceivable competing candidates one particular sandhi form that constitutes the optimal parse or phonetic realization of a given input string (see Prince and Smolensky 1993:34, and passim). In order to cash in on the core insights of OT, let us first restate the observations made above as general constraints:

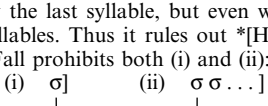
(4)	*E/O	Even-oblique and oblique-even sequences are forbidden.
	E≠O	Do not turn an even tone into an oblique tone, or vice versa.[8]
	*Tq	A checked syllable in weak (initial) position may not carry tone.
	*O↓	No low-registered oblique tone.
	*E↓	No low-registered even tone except in initial position.
	*Fall	No final falling tone.[9]
	OCP	Abutting contour tones may not share identical melodic shapes (rising/falling).
	Prs-E	Parse even tone[10]
	Prs-T	Parse tone
	Ident	Preserve tonal identity (register, tone shape)[11]
	E, O	even, oblique
	T	tone
	Tq	tone linked to a checked syllable
	↑, ↓	high, low register

[8] Since oblique tones branch, evens do not. E≠O has the effect of prohibiting either adding or deleting a terminal tone segment. In other words, E≠O bars any of the following configurations. ((a, b) represents even turning into obliques; (c, d) the other way round.)



{t} = epenthetic tone segment
<t> = unparsed tone segment

[9] *Fall holds on the tonal tier, that is to say, it bars a falling tone not just when it is borne by the last syllable, but even when a falling tone is followed by one or more toneless syllables. Thus it rules out *[HM-o] as well as [o-HM] or [MH-HM]. In other words, *Fall prohibits both (i) and (ii):



where F = falling tone, and "σ . . ." in (ii) stands for an arbitrary string of toneless syllables.

[10] The original Parse X in Prince and Smolensky (1993) is equivalent to Max X in the Correspondence Theory of Faithfulness (McCarthy and Prince 1995b).

[11] This generic Identity constraint will be split up into two, Ident-Reg (preserve register identity) and Ident-C (preserve contour shape identity or, equivalently, preserve the linear sequence of tone segments). See discussion below.

The first seven constraints are surface-true, brooking no violation.[12] In OT terms, they are "undominated." The remaining three constraints are violable, and belong to the family of "faithfulness" conditions which ensure minimal deviations from the underlying tone patterns, consistent with well-formedness conditions.

2.2 Targets of tone deletion

Needless to say, the major task confronting us is the relationship between the base tones and their corresponding sandhi forms. For expository clarity, it is important to bear in mind two major types of processes (to continue to use the processual metaphor): (i) those that delete one or the other of the tones in juxtaposition, but not both;[13] (ii) those that determine the eventual shape of the surviving tones. Let us focus on the first question, namely as to which of the two tones gets deleted, and which is phonetically realized – putting aside for the moment the question regarding the actual tone shapes associated with the surviving tones. The factors that determine tone survival are: the even vs. oblique distinction, and the "smooth" vs. "checked" type (CVq, where -q represents an obstruent, often a glottal stop) of the hosting syllable. The latter distinction is relevant only for the initial syllable. This gives us six classes of two-tone combinations. Abstracting away from tone shape, and ignoring a few sporadic exceptions, the sandhi forms of these six classes are given in table 5.1.

Table 5.1 cross-classifies a two-tone sequence along the even/oblique dichotomy: thus the lower left cell represents the intersection set of tonal strings with an oblique tone linked to the first syllable and an even tone associated with the second syllable. All such underlying O-E strings (89 cases in all) exhibit a sandhi pattern [o.T], that is an atonic syllable followed by a tone-carrying syllable (in this case, the tone belongs to the even category). The shaded cell in table 5.1 represents the O-O combination, which displays a highly complex behavior. We will defer discussion on this subset of cases until section 3. The remaining five tone patterns can be derived quite simply by a combination of *E/O, *Tq and Prs-E, Prs-T, as shown in the familiar tableau format of (5). In (5),

[12] See section 5 for apparent violations of E≠O and *Fall.

[13] Throughout this chapter I will take for granted that a completely toneless prosodic word is ill-formed, even though within a larger prosodic domain words do undergo de-toning. See chapter 6.

Table 5.1.

1st σ		2nd σ			
		even		oblique	
even	checked	o.T	34	o.T	29
	slack	T.T	50	T.o	47[14]
oblique		o.T	89	various sandhi forms, including T.T, T.o, o.T (see table 5.2 and discussion below)	

Numbers indicate type frequency (in a subcorpus of 346 disyllabic compounds)

Es and Eq stand for an even tone associated with a slack or a checked syllable respectively. Patterns 1 and 2 (P1, 2) represent the simplest cases: given that both E-O and O-E patterns are ruled out by *E/O, one alternative is to turn either O to E, or E to O, thereby yielding a well-formed E-E or O-O sequence. This alternative is blocked by the inviolable E ≠ O. Another ploy is to delete either O or E. But since the deletion (underparsing) of evens is more heavily penalized than that of obliques, a generalization encoded in the rank order Prs-E ≫ Prs-T, output candidates [T-o] (P1c) and [o-T] (P2b) (where T = E) are preferred over [T-o] (P2c) and [o-T] (P1b) (where T = O). The winning candidate is marked by the hand sign (☞). But the ordered pair of constraints Prs-E (parse an even tone) ≫ Prs-T (parse tone) is subordinate to *Tq, which bars an initial checked syllable from carrying a tone. Thus in P3b, the winning candidate [o.T] (where T = O) retains an oblique tone, while deleting an even tone, because the competing candidate [T.o] (P3c, where T = Eq) violates the overriding *Tq. Case (P4b) obeys the same imperative of *Tq. Finally, the base form of P5a satisfies all relevant constraints, and therefore is picked as the optimal sandhi form as is, without any adjustment.

[14] Within this cell there are four cases with a o-T [o-MH] pattern. We will return to this interesting class of subregularities in section 6. Additionally, there are two totally unexpected sandhi forms ([HM-MH] and [H-o]) within this same cell. Otherwise, the mapping relation between underlying tone sequences and their sandhi forms (outside of the shaded area) is completely predictable in a subcorpus comprising 346 disyllabic compounds.

(5)

			output		*E/O	*Tq	Prs-E	Prs-T
P1 Es-O		a	T.T		*			
		b	o.T				*	
	☞	c	T.o	= [H.o, L.o]				*
P2 O-E		a	T.T		*			
	☞	b	o.T	= [o.H]				*
		c	T.o				*	
P3 Eq-O		a	T.T		*	*		
	☞	b	o.T	= [o.MH]			*	
		c	T.o		*			*
P4 Eq-E		a	T.T			*		
	☞	b	o.T	= [o.H]			*	
		c	T.o			*	*	
P5 Es-E	☞	a	T.T	= [H.H]				
		b	o.T				*	
		c	T.o				*	

Es, Eq = even tone linked to a slack or checked syllable respectively
P1, P2 . . . = patterns 1, 2 . . .
*E/O, *Tq, etc. stand for the constraints defined in (4)
Hierarchically ranked constraints are separated by solid vertical lines

Actual examples instantiating the various sandhi patterns, such as those of the tableau in (5) will be cited as soon as we can determine the phonetic shape of these abstract patterns T.T, T.o, and o.T, a task we now turn our attention to.

2.3 *Determining tone shape*

Once we ascertain which of the tones deletes and which one remains, determining the actual melodic shapes of the surviving tone or tonal strings (enclosed in square brackets in (5)) turns out to be fairly trivial. The relevant constraints are *O↓, *E↓, E≠O, *Fall, and Ident. The following tableau (6) assumes that one of the three patterns, namely [T.T], [T.o], or [o.T], has already been picked out by the criteria as illustrated in (5).

(6)

input				output	*O↓	*E↓	E≠O	*Fall	Ident
P1 MH-L	= o-T	☞	a	o-H					*
			b	o-L		*			
			c	o-MH			*		*
P2 Hq-ML	= o-T	☞	a	o-MH					**
			b	o-ML	*			*	
			c	o-HM				*	*
			d	o-H			*		**
P3 H-ML	= T-o	☞	a	H-o					
			b	L-o					*
			c	MH-o			*		*
			d	HM-o			*	*	*
P4 L-MH	= T-o	☞	a	L-o					
			b	H-o					*
P5 H-Lq	= T-T	☞	a	H-Hq					*
			b	H-Lq		*			
P6 L-H	= T-T	☞	a	L-H					
			b	H-H					*

The input form /MH-L/ of P1 (pattern 1) is one subtype of O-E, and, therefore, has the [o.T] pattern. Since the surviving T is an even tone, it can surface only as H (candidate a), as dictated by *E↓ (no low-register even tone, except in initial position). Conceivably, P1 may emerge as [o.MH] (candidate c); but such an option is ruled out by E≠O (no switch over between evens and obliques), not because there is anything ill-formed about [o.MH] per se, but because such an output is not the optimal parse given the input form [o.T], where T = E. P2 instantiates the class of Eq-O sequences. The combination of *E/O (prohibiting E-O and O-E sequences) and *Tq (no initial tone-bearing checked syllable) also dictates a [o.T] pattern. The only possible phonetic realization of the surviving oblique tone is [MH], in accordance with *O↓ (no low-register oblique) and *Fall (no final falling), even though [o.MH] (candidate a) entails a double violation of low-ranked Ident (one mark "*" for changing register from low to high, another for turning a falling contour into a rising one). /H.ML/ (P3) and /L.MH/ (P4) constitute subtypes of /Es-O/, both of which yield [T.o] as output. The surviving even tone surfaces as either [H] or [L], depending on the lexically specified register, since, unlike oblique tones,

even tones retain their underlying register contrasts in initial positions. Other conceivable well-formed patterns are weeded out by the faithfulness condition Ident. P5 and P6 represent E-E patterns where both even tones are retained. The optimal outputs are those involving the fewest violations of constraints. Actual examples corresponding to each pattern are given in (7).

(7) Examples corresponding to (6)

		base form	sandhi form	gloss
P1	*xin-xin*	MH-L	o-H	"confidence"
P2	*hei-shi*	Hq-ML	o-MH	"black market"
P3	*jing-li*	H-ML	H-o	"manager"
P4	*chang-shou*	L-MH	L-o	"longevity"
P5	*jing-ju*	H-Lq	H-Hq	"Peking opera"
P6	*nian-qing*	L-H	L-H	"youthful"

In short, the ranking among the relevant constraints is stated in (8). As noted before, the first seven constraints are undominated, and therefore unranked among themselves; they dominate all others, including Prs-E, which in turn dominates Prs-T, as demonstrated by (5), P1–2. Data presented so far do not suffice to establish relative ranking between Prs-T and Ident, both of which need to be unpacked into more specific constraints, as we shall see in the following sections.

(8) {*E/O, *Tq, *O↓, *E↓, E≠O, OCP, *Fall} ≫ Prs-E ≫ {Prs-T, Ident}

3 Constraint ranking

We now turn to the O-O (oblique-oblique) sequences occupying the shaded cell in table 5.1. Table 5.2 "blows up" the shaded cell of table 5.1 in order to bring out more of the fine-grained details.

Oblique tones are cross-classified according to register (high and low) and melodic shape (rising or falling). We therefore get sixteen possible O-O combinations, represented by the sixteen cells in table 5.2. As in the case of other tonal patterns, the crux of the problem lies in ascertaining which underlying O-O sequences map into which of the three basic patterns [T.o], [o.T], or [T.T] (and [T.T′]). If only one oblique tone survives, as in the cases of [T.o] and [o.T], it invariably assumes the rising pitch

Table 5.2. *Oblique-oblique patterns*[15]

		2nd σ			
		falling		rising	
1st σ		Hr: HM	Lr: ML	Hr: MH	Lr: LM
falling	Hr: HM	T.o 6	T.o 3	T.T 4	T.T 6
	Lr: ML	o.T 5	o.T 5	T.T′ 5	T.T′ 3
			T.T′ 1	o.T 2	T.T 2
rising	Hr: MH	T.o 6	T.o 4	o.T 3	T.o 4
				T.o 1	T.T 2
				T.T 2	
				T.T′ 1	
	Lr: LM	o.T 6	o.T 3	o.T 6	o.T 4
			T.o 2	T.T 1	T.o 1

Hr, Lr high, low register
o.T [o.MH]
T.o [MH.o]
T.T [HM.MH]
T.T′ [H.M]

contour, surfacing as [MH.o] and [o.MH] respectively, as dictated by two of the constraints already discussed in section 2, namely *O↓ (no low obliques) and *Fall (no final falling). Where both obliques are retained, they appear as [HM.MH] (annotated as T-T) in the majority of cases. [HM.MH] has a variant [H.M] (coded as T-T′) for lexical compounds whose first syllable is lexically associated with a low-register tone. This generalization is borne out by an inspection of the lightly shaded upper right cells in table 5.2. Notice that the eight instances (out of a total of ten) of [H.M] are derived from the underlying /ML.R/ (where R = rising, that is either MH or LM), in contrast to /HM-R/, which invariably shows up as [HM.MH]. As an expository expedient, in what follows I will regard [H.M] as a register-dependent variant of [HM.MH], putting off discussion of the theoretical issue involved until section 5. In the rest of this section, T-T [HM.MH] subsumes T-T′ [H.M] as a special case.

[15] Omitted from this table are three lexical items that exhibit totally irregular sandhi forms. They are: *kao-shi* "examination" /HM.MH/ → [H.H], *wang-ji* "forget" /LM.LM/ → [o.H], and one of the two readings of *zhi-shi* "instruction, order" /HM.LM/ → [o.H] (the alternative reading is the regular [HM.MH]).

3.1 Oblique-oblique sequences

Under this simplifying stipulation, we can make the following generalization: the only permissible sequence of abutting oblique tones is F-R (falling-rising). We derive this canonical form from the two undominated constraints alluded to in section 2, OCP and *Fall:

(9)

	OCP	*Fall
F-R		
R-R	*	
R-F		*
F-F	*	*

R rising
F falling
OCP abutting contour tones may not share identical melodic shapes (rising/falling)
*Fall no final falling tone (see (4))

Of the four O-O sequences cross-classified according to tone shape, F-R already accords with the canonical template, consequently both tones are retained (lightly shaded upper right quadrant of table 5.2). The minimal modification necessary is the raising of the register of an oblique tone, as required by the constraint *O↓. The remaining three combinations all violate OCP and/or *Fall. In principle, two alternatives present themselves: change the tone shapes to fit the F-R template, or delete one of the two oblique tones. On the whole, the latter alternative is the strategy of first resort. This can be seen in (10):

(10)

base form	sandhi form		
	o-T	T-o	T-T / T-T′
F-R	2	–	20
R-R	13	6	6
R-F	9	12	–
F-F	10	9	1

o.T = o.MH
T.o = MH.o
T.T = HM.MH
T.T′ = H.M

Apart from F-R, most of oblique-oblique patterns surface as either [o.T] or [T.o] (59 out of 66 examples). The question is which of the two oblique tones remains, and which deletes. The overall picture is reasonably clear: for F-F and R-F patterns, the deciding factor is the register of the first tone: if it is high registered (regardless of tone shape), then it is the second falling tone that deletes (the diagonally lined cells of table 5.2); on the other hand, if the first tone carries a low register, then it is this first tone that deletes (the clear cells of table 5.2) in all cases except three. The R-R pattern (the lower right dark-shaded quadrant of table 5.2) behaves less predictably, though by no means randomly. Tonal deletion is sensitive to what we may refer to as tonal saliency. Tonal saliency is jointly determined by register and position: high register tones are more salient than their low counterparts, and a tone occupying a prosodically strong (right) position is more salient than a tone occurring in a weak (left) position. When in conflict, register takes precedence over position; thus a high register tone in a weak position prevails over a low register tone in a strong position. Ignoring sporadic exceptions, we can summarize the correspondences between the lexically assigned tonal categories and the sandhi patterns as follows (parenthesized (T.T) and (T.o) represent minority cases):

(11) Oblique-oblique patterns

base form		sandhi form	
F-R		T.T	
F-F, R-F	T↓.T	o.T	(T.o)
	T↑.T	T.o	
R-R	T↓.T↑	o.T	
	T↓.T↓	o.T	
	T↑.T↑	o.T	(T.T)
	T↑.T↓	T.o	(T.T)

R, F rising, falling
T↑, T↓ high-, low-register tone
T, o tone, zero-tone

3.2 *Well-formed falling-rising patterns*

We now proceed to working out the details of implementation of the intuitive content of our analysis, taking up the unproblematic F-R pattern first

(see (4)). By ranking the undominated OCP, *Fall, *O↓ over the faithfulness condition Ident-Reg (preserve register identity), which says in effect that underlying register must be maintained, we guarantee the optimal output candidate for F-R patterns. Conceivably, an alternative way to satisfy *O↓ is to delete a low-register tone. But we can block this option by ranking Prs-T (parse tone) above Ident-Reg. The ranking hierarchy holding between these two constraints is clearly motivated by P3, where in the majority of cases (five out of seven) sandhi pattern [HM.MH] (candidate b), which entails a violation of Ident-Reg, wins out over [o.MH] (candidate c), which is penalized for failing to comply with Prs-T. P4 shows even more clearly this dominance relation. Compare the two competing candidates [HM.MH] (b) and [o.MH] (c) as output of /ML-LM/. Both entail at least one instance of register raising. After canceling one violation of Ident-Reg for each candidate, the two competing outputs differ in that [o.MH] (c) is assessed one star (*) for violating Prs-T, while [HM.MH] (b) pays the cost of another star for failing Ident-Reg. The latter option consistently wins out. This means that Prs-T ranks higher than Ident-Reg.

(12) F-R patterns

						OCP	*Fall	*O↓	Prs-T	Id-Reg
P1	HM-MH	☞	a	HM-MH	4					
P2	HM-LM		a	HM-LM	–			*		
		☞	b	HM-MH	6					*
			c	HM-o	–		*		*	
P3	ML-MH		a	ML-MH	–			*		
		☞	b	HM-MH[16]	5					*
			c	o-MH	2				*	
P4	ML-LM		a	ML-LM	–			**		
		☞	b	HM-MH[17]	5					**
			c	o-MH	–				*	*

Numbers indicate type frequency

[16] As noted earlier, this output has [H.M] as a register-dependent variant; see section 5.
[17] See preceding footnote.

(13) Examples corresponding to (12)

		base form	sandhi form	gloss
P1	*xiao-qi*	HM.MH	HM.MH	"stingy"
P2	*zhang-gui*	HM.LM	HM.MH	"cashier"
P3	*yan-jing*	ML.MH	HM.MH (→ H.M)[18]	"eye-glasses"
	ye-cai	ML.MH	o.MH	"wild herbs"
P4	*yan-lei*	ML.LM	HM.MH (→ H.M)	"tear" (noun)

3.3 *Tonal saliency and melodic shape*

Aside from F-R, none of the other three patterns – R-R, R-F, F-F – satisfies the well-formedness conditions OCP and *Fall imposed on two-tone sequences. Hence, some adjustment is necessary. As pointed out earlier, there are basically two strategies to bring these offending input patterns in line with the output constraints, namely either to delete one of the two oblique tones (in which case OCP ceases to be relevant), or to change the melodic contour to fit the falling-rising template. Let us consider first the input pattern R-R (the lower right, dark-shaded quadrant of table 5.2). For the most part, tonal deletion is the preferred strategy. That is, instead of turning R-R into F-R, the preferred alternative is either R-o or o-R. This can be seen in the distributional pattern exhibited in (14).

(14) R-R patterns

		o.T	T.o	T.T / T.T′
a.	LM.MH	6	–	(1)
b.	LM.LM	4	(1)	–
c.	MH.MH	3	(1)	3
d.	MH.LM	–	4	2

Figures in parentheses represent sporadic exceptions

The numbers in (14) indicate the type frequency of [o.T], [T.o], or [T.T] as surface manifestations of the four input R-R patterns. Note that out of the twenty-five disyllabic compounds with an underlying R-R string, nineteen delete one or the other tone. This suggests that altering the underlying melodic shape (from rising to falling or vice versa) incurs a more severe penalty than deleting the tone altogether. It is further evident that tonal

[18] [HM.MH] takes on the register-dependent variant [H.M], see section 5.

deletion does not occur at random, but is sensitive to what we may refer to as tonal saliency. Uncontroversially, high tones are more salient than low tones. Another factor that contributes to tonal saliency is prosodic prominence. On the ground that a checked syllable loses its tone only in initial (weak) position by virtue of *Tq (which bars initial CVq from carrying a tone), we inferred that New Chongming is right-dominant. This inference is further strengthened by the observation that, other things being equal, as in the case of /LM.LM/ and /MH.MH/, if tonal deletion applies at all, it is the first syllable that loses its tone, while the last syllable preserves its tonal category. Finally, where two tones vie for saliency, one being high registered, the other being linked to a dominant (right) position, it is the former that wins out, as instantiated in case (d) of (14).[19]

Informally stated, what we see in action is tonal stability or deletion-resistance keyed to some scale of tonal saliency. Recall that an even tone is more salient than an oblique tone, because in order to satisfy *E/O which disallows both E-O and O-E sequences, it is the oblique that deletes. This coupled with the foregoing discussion gives us the following generalizations about tonal saliency:

(15) Tonal Saliency

a. E > O
b. T↑ > T↓
c. Tp > Tw

x > y	x is more salient than y
E, O	even, oblique tone
T↑, T↓	high, low register tone
Tp, Tw	tone in prominent (right), weak (left) position
T	any (other) T

Clauses (a), (b), and (c) of Tonal Saliency (or Saliency for short) are hierarchically ranked. Saliency can be abbreviated as follows:[20]

[19] A similar dominance relation between the inherent salience of high-register tones and the metrically acquired prominence is observed elsewhere, for instance in the northwestern Mandarin dialect of Xining. Xining prohibits two adjacent contour tones. Which of the two abutting contour tones goes is determined precisely by the constraint hierarchy *Tc-Tc (no adjacent contour tones) ≫ Prs-T↑ (parse a high-register tone) ≫ Prs-Tp (parse a prosodically prominent tone on the right). For details, see Zhang and Zhu (1987).

[20] In other words, Saliency is interpreted in the following fashion: if x > y by virtue of clause (a), but y > x on account of clause (b), clause (a) prevails. For instance, a low-register even ranks higher than a high-register oblique in accordance with clause (a); but the pecking order is reversed by virtue of clause (b). In such cases, the relative salience established by (a) prevails. It is only natural to impose on Saliency the general heuristic principle: x > y holds true iff y belongs to the complement set of x. That is to say, E > T↑

(16) Scale of Tonal Saliency: E > T↑ > Tp > T[21]

Since we have already factored out Prs-E as a high-ranking constraint, we can translate the observations made above into the constraints of (17), where we "explode" the generic Prs-T into a hierarchy of more specific constraints. Occasionally we will use the label Prs-Tone to refer to the block of internally ranked constraints comprising Prs-T↑, Prs-Tp, and Prs-T.

(17)

Ident-C	Preserve contour identity: do not change rising to falling or vice versa.[22]
Prs-T↑	Parse high register tones.
Prs-Tp	Parse tones in a prominent (right) position.
Prs-T	Parse any tone.
Prs-Tone	= {Prs-T↑ ≫ Prs-Tp ≫ Prs-T}

The tableau in (18) illustrates how these constraints in conjunction with others already established converge to pick out the optimal candidates.

(18) R-R patterns

						Id-C	Prs-T↑	Prs-Tp	Prs-T	Id-Reg
P1	LM-MH		a	HM-MH	(1)	*				*
		☞	b	o-MH	6				*	
			c	MH-o	–		*	*		*
P2	LM-LM		a	HM-MH	–	*				**
		☞	b	o-MH	4				*	*
			c	MH-o	(1)			*		*
P3	MH-MH	☞	a	HM-MH	3	*				
		☞	b	o-MH	3		*		*	
			c	MH-o	(1)		*	*		
P4	MH-LM	☞	a	HM-MH	2	*				*
			b	o-MH	–		*			*
		☞	c	MH-o	4			*		

Numbers indicate type frequency

is true, unless T↑ happens to be also an instance of E (that is a high-register even tone) – in which case, it goes without saying, clause (a) of Saliency cannot adjudicate between the salience of an E relative to another (high-registered) E. An example illustrating the effect of tonal saliency is P4 of (7) *chang-shou* "longevity" with /L.MH/ as input, where the low-register even L prevails over the high-register oblique MH, as a result, the winning sandhi form is [L.o] rather than [o.MH].

[21] Pike (1948:79), speaking of Mixteco, observes: "High tonemes are more stable than low tonemes; basic high tonemes are seldom changed; basic low tonemes are very frequently changed, eliminated in favor of high tonemes." Cf. Pankratz and Pike (1967), Du (1988).

[22] Equivalently, no metathesis or preserve the linear order of tone segments.

(19) Examples corresponding to (18)

		base form	sandhi form	gloss
P1	*xia-zhi*	LM-MH	o-MH	"summer solstice"
	(*lang-fei*	LM-MH	HM-MH	"to waste"
P2	*ming-ling*	LM-LM	o-MH	"command"
	(*dou-fu*	LM-LM	MH-o	"bean-curd"
P3	*fen-dou*	MH-MH	HM-MH	"struggle"
	tan-qi	MH-MH	o-MH	"to sigh"
P4	*tai-du*	MH-LM	HM-MH	"attitude"
	zha-dan	MH-LM	MH-o	"bomb"

Parenthesized examples represent the minority cases

Here we ignore all candidates that violate the undominated constraints OCP, Fall, and *O↓, which jointly filter out all but three well-formed O-O patterns, namely [MH.o], [o.MH], and [HM-MH]. P1 and P2, both involving a low-register tone in weak position, clearly suggest the ranking Ident-C ≫ Prs-Tone, where Prs-Tone stands for the block of parse-tone constraints: in both cases [o-MH] (candidate b, satisfaction of Prs-C at the expense of Prs-Tone) consistently wins over [HM-MH] (candidate a, vice versa). In other words, OCP is enforced preferably by means of deleting a (low-profile) tone rather than by resorting to changing the tone shape. As for which of the two abutting tones gets deleted, that is a matter of relative tonal saliency. One contributing factor of saliency is prosodic prominence. This is encoded in the ranking Prs-Tp ≫ Prs-T, as shown by patterns 2 and 3, where ceteris paribus (tones of same register, as in /LM.LM/ or /MH.MH/), it is the tone in prosodically weak (first) position that undergoes deletion.[23] Finally, P4 demonstrates that when saliency criteria conflict with one another, namely when a high register tone occupies a weak position while a low register tone occurs in a strong

[23] Given the generic nature of Prs-T, any violation of either Prs-T↑ or Prs-Tp, or for that matter Prs-E, necessarily entails a violation of Prs-T as well. On this interpretation Prs-T need not be ranked with respect to other parse-tone constraints: the latter being additional penalties levied against tone deletion if the tone being deleted happens to be high registered, or prosodically strong etc. Alternatively, the generic Prs-T represents the "Elsewhere" condition, the left-over set not covered by Prs-E, Prs-T↑, and Prs-Tp. On this interpretation, Prs-T means: parse any *low-profile* tone, where a low-profile tone is precisely a tone that does not belong to any of the categories: even, high register, or prosodically prominent. On this interpretation the more specific parse-tone constraints obviously rank higher than the generic Prs-T. The constraint tableaux are constructed with the "Elsewhere" interpretation in mind. This heuristic convention minimizes redundancy and brings out more clearly the intuitive content of constraint ranking.

position, it is the pitch register that determines which of the two tones loses out. In other words, Prs-T↑ ≫ Prs-Tp. Thus, given the input string /MH-LM/, it is [MH.o] (candidate c, satisfaction of Prs-T↑ at the expense of Prs-Tp) that is picked over [o.MH] (candidate b, satisfaction of Prs-Tp at the expense of Prs-T↑). Recall that we had already established the ranking Prs-T ≫ Ident-Reg on the grounds of P3 and P4 in (12).

There remains some indeterminacy of constraint ranking. Examples corresponding to P3 of (18) instantiate a draw between two competing candidates [HM.MH] and [o.MH] (hence two hand signs "☞" in (18)). One could account for the equal split of /MH.MH/ between [HM.MH] (suggesting Prs-T↑ ≫ Ident-C) and [o.MH] (pointing to a reverse dominance relation) by making allowance for an indeterminate ranking between Ident-C and Prs-T↑. The same mechanism of indeterminate ranking between Ident-C and Prs-Tp may be extended to account for P4: given the input /MH.LM/, one can insure an acceptable output by resorting to one of two alternative strategies: (i) changing the contour of the first tone from a rising to a falling contour, yielding [HM.MH] (candidate a);[24] (ii) deleting the second tone, resulting in [MH.o] (candidate c). Strategy (i) keeps intact the second, positionally prominent, tone at the cost of changing the contour of the first tone. In other words, Prs-Tp ≫ Ident-C. On the other hand, strategy (ii) entails preserving the contour of the first tone at the expense of deleting the second, positionally prominent tone. That is to say, Ident-C ≫ Prs-Tp.

In this connection recall that when the choice is between deleting a low-profile tone (thereby incurring a Prs-T violation) and raising its register (in violation of Ident-Reg), the latter alternative prevails. This generalization is borne out by P3–4 of (12). This means that the entire block of parse-tone constraints, namely Prs-Tone, dominates Ident-Reg – in the majority of cases. There is a residue, where the reverse ranking seems to hold. Specifically, the same input string /ML.MH/ (P3 in (12)) emerges as [o.MH] (candidate c, incurring Prs-T violation) in two instances instead of [HM.MH] (candidate b, involving register raising), as expected in the remaining five examples. Despite these residual cases, the relative ranking Prs-T ≫ Ident-Reg is well supported by ten out of the twelve examples of P3–4 of (12).[25]

[24] Simultaneously raising the register of the second tone, on account of *O↓ (no low obliques).

[25] See section 6 for further discussion on ranking indeterminacy.

3.4 Melodic integrity and right prominence

Let us consider the two remaining sets of O-O (oblique-oblique) patterns: R-F and F-F. The pertinent facts are as summarized in table 5.2 and (11). Notice that, aside from one single exception,[26] these patterns surface exclusively as either [o.T] or [T.o]. This suggests that in order to satisfy the undominated constraints OCP and *Fall, the strategy of choice is tonal deletion rather than manipulation of the melodic shape. One obvious way to obtain the desired effect is to rank Ident-C over the generic Prs-Tone, namely the family of constraints that prohibit tonal deletion. As to which of the tones gets deleted, one would expect the decision to be made on the basis of the internal ranking within the parse-tone family of constraints, specifically, Prs-T↑ ≫ Prs-Tp ≫ Prs-T, as previously established. Unfortunately, this most sensible approach does not give us the right results. This conclusion can be demonstrated in tableau form. The tableau in (20) shows that the attested winner output candidate c [MH-o] for an input like /HM-HM/ turns out to be the worst offender of the constraints. Since both input tones are falling, no matter whether only one or both tones survive, all three output candidates [HM.MH], [MH.o], [o.MH] must undergo contour adjustment in order to satisfy OCP and *Fall, thereby violating Ident-C. On that score, all candidates are evenly matched. In addition, both [o.MH] and [MH.o] involve tonal deletion, incurring a violation of Prs-T↑. But worst of all, [MH.o] leaves unparsed the more prominent of the two tones, hence is penalized with an additional star on account of Prs-Tp.

(20)

input			output		Ident-C	Prs-T↑	Prs-Tp	Prs-T	Ident-Reg
HM-HM		a	HM-MH	–	*				
		b	o-MH	–	*	*			
	☞	c	MH-o	6	*	*	*		

Example: *bao-xian* "insurance"

It appears necessary to break down the generic Ident-C into an internally ranked block of more parochial constraints: Ident-Cp ≫ Ident-C. Where called for, I will refer to the family of melody-preserving constraints as Ident-Cont.[27]

[26] *Fan-zui* /ML.ML/ "to commit a crime" has two alternative readings: [o.MH] or [H.M] (= [T.T′]).

[27] Again, I will assume an "Elsewhere" interpretation of the generic Prs-C, as covering the left-over cases not covered by the more specific Prs-Cp, and construct the tableaux accordingly.

(21)	Ident-Cp	Preserve the contour identity of the tone in prominent (right) position
	Ident-C	Preserve the contour identity of any tone
	Ident-Cont	= {Ident-Cp ≫ Ident-C}

The substantive claim here is that a more salient tone tends to be more stable, more resistant to change, whether the change involves deletion or tone shape alteration. It is customary to characterize the prosodic systems of Chinese dialects as right- or left-prominent (cf. Yue and Hashimoto 1986, Qian 1992, Duanmu 1993a, 1994b). Left-prominent systems (prototypically, those of Shanghai and many other northern Wu dialects) tend to preserve the leftmost tone, sometimes spreading it over the entire prosodic domain. Right-dominant systems (notably Mandarin, Min, and some southern Wu dialects) tend to maintain intact the identity of the rightmost tone, while allowing various modifications on the non-final tones in sandhi contexts. Take the case of Tianjin.[28] Tianjin prohibits two adjacent like tones (except two level highs). To satisfy this OCP-based constraint, dissimilation takes place: LL, RR, FF → RL, HR, LF respectively (as usual, L, H, R, F stand for low, high, rising, and falling). Notice that it is the tone on the left (weaker position) and not the tone on the right (stronger position) that undergoes dissimilation. To return to the New Chongming dialect under discussion, the hierarchically ordered pair of constraints Ident-Cp ≫ Ident-C is a reflection of a general trend that manifests itself across many dialects.[29] With this ranking we obtain the desired results for the four ditonic patterns consisting of a high-register oblique tone followed by a falling tone. This is shown in (22):

(22)

						Id-Cp	Id-C	Prs-T↑	Prs-Tp	Prs-T	Id-Reg
P1	HM.HM		a	HM.MH	–	*					
			b	o.MH	–	*		*			
		☞	c	MH.o	6		*	*	*		
P2	MH.HM		a	HM.MH	–	*	*				
			b	o.MH	–	*		*			
		☞	c	MH.o	6			*	*		

[28] For further details, see chapter 3.

[29] In this regard, New Chongming is an anomaly. Geographically located to the far north of the Wu-speaking area, it nevertheless displays a strong right-prominence characteristic of the Mandarin group farther to the north.

(22) cont'd

					Id-Cp	Id-C	Prs-T↑	Prs-Tp	Prs-T	Id-Reg
P3 HM.ML		a	HM.MH	–	*					*
		b	o.MH	–	*		*			*
	☞	c	MH.o	3		*		*		
P4 MH.ML		a	HM.MH	–	*	*				*
		b	o.MH	–	*		*			*
	☞	c	MH.o	4				*		

(23) Examples corresponding to (22)

		base form	sandhi form	gloss
P1	*bao-xian*	HM.HM	MH.o	"insurance"
P2	*kuai-ban*	MH.HM	MH.o	"allegro (rhythm)"
P3	*gai-zao*	HM.ML	MH.o	"reform"
P4	*dui-xiang*	MH.ML	MH.o	"target, objective"

3.5 *Constraint ranking: a summary*

It is useful at this point to sum up the ranking relations holding among the various constraints. In this section we have dealt with the constraints that govern O-O (oblique-oblique) combinations (OCP, Fall, *O↓, Ident-Reg, and the families of constraints Ident-Cont and Prs-Tone), while in section 2 we concerned ourselves with well-formedness conditions on E-O and O-E strings (*E/O, Prs-E, and Prs-T). We need to bring these two sets of constraints together. Prs-E and Ident-Cont are unranked, since they pertain to two disjoint types of tonal combinations: Prs-E is relevant to E-O and O-E strings, while Ident-Cont is involved in O-O sequences barred by OCP and/or *Fall. Prs-E, on the other hand, takes precedence over Prs-T↑ (and transitively over Prs-Tp, etc.). This is demonstrated in (24):

(24) /MH-L/ *tou-ming* "transparent"

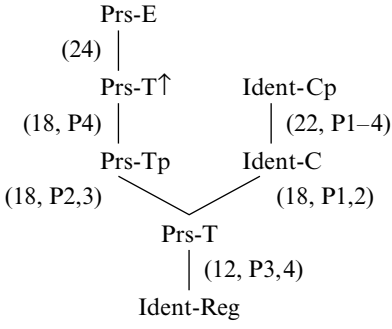
			*E/O	*E↓	Prs-E	Prs-T↑
a		MH-L	*	*		
b		o-L		*		*
c		MH-o			*	
d	☞	o-H				*

Setting aside the undominated constraints, the overall ranking hierarchy we have established is summarized in (25). The vertical shafts indicate the pecking order:

Ci
| = Ci ≫ Cj
Cj

The patterns bearing out the particular dominance relation are indicated in parentheses. There are basically three subsets of constraints: (i) the parse-tone block, reflecting the scale of tonal saliency: Prs-E ≫ Prs-T↑ ≫ Prs-Tp and Prs-T; (ii) the preserve-contour block, consisting of Ident-Cp and Ident-C. The parse-contour constraints are relevant only to oblique tones; hence they cannot be ranked vis-à-vis Prs-E. Furthermore, as noted above, the ranking between Prs-T↑/Tp on the one hand, and Ident-Cp/C on the other is indeterminate; this explains why no line connects this subset of constraints to either Prs-T↑ or Prs-Tp. However, contour identity clearly dominates Prs-T. The last subset (iii) consists of a single constraint, Ident-Reg, and ranks at the bottom of the totem-pole.

(25)



3.6 *Residual problems*

We now come to the last set of oblique-oblique tone patterns. The disyllabic compounds of (26) all have the following input tone pattern: a low-register oblique tone followed by a falling tone; in other words, they differ from those of (20) only in that the tone on the left is low registered. Aside from a small handful of exceptions (three, to be exact), all emerge as [o-MH] (see section 3, table 5.2; the relevant tonal combinations are those occupying the clear cells).

(26)

		base form	sandhi form	gloss
P1	*ye-cao*	ML-HM	o-MH	"weed"
P2	*da-dan*	LM-HM	o-MH	"bold"
P3	*dao-li*	ML-ML	o-MH	"reason"
P4	*yun-dong*	LM-ML	o-MH	"exercise"

The inductive generalization is simply this:

(27) O↓-F
Unparse a low-register oblique tone when followed by a falling one.

For conceptual and technical reasons I will not go into,[30] (27) is unstatable as a bona fide constraint. For the moment I will regard (27) as a "derivational residue," a device of last resort, so to speak.

There is another set of problematic cases. We noted earlier that [HM-MH] has a register-dependent variant [H-M]. [HM-MH] and [H-M] are the only patterns that permit two consecutive (underlyingly) oblique tones. They are annotated as T-T and T-T′ respectively in table 2 (see section 3). Referring back to table 2, one readily notices the following pattern: with a couple of exceptions, two oblique tones may coexist in a disyllabic compound only if the lexically assigned tones form a falling-rising (F-R) pattern – the only well-formed oblique-oblique sequences consistent with OCP and *Fall. The F-R string, then, survives intact, while all other oblique-oblique sequences must rid themselves of one or the other tone. We derive this result essentially by ranking Ident-Cont (= Ident-Cp and Ident-C) over Prs-Tone (= Prs-T↑, Prs-Tp, Prs-T): in other words, in order to satisfy OCP and Fall, it is generally better to delete the offending tone rather than to tamper with its melodic shape. The surviving F-R sequences then split into two subpatterns: [HM-MH], if the first tone has a high register; [H-M] otherwise. The distributional pattern of [HM-MH] and [H-M] is quite predictable: as can be verified against table 5.2. Table 5.3 shows the phonetic details of the upper-right lightly shaded quadrant of table 5.2. The phonetic or functional motivation for the split of F-R into two subpatterns remains mysterious. I will simply refer to whatever process or more general contraints are involved as Melody Simplification, which has the effect of turning a *low*-register falling tone + any rising tone into [H-M].

From the OT perspective, it is unclear why Melody Simplification is restricted to F-R with a low register on the first tone, or why it occurs

[30] I have omitted several long-winded sections that address this issue from an earlier draft.

Table 5.3. *F-R patterns*

1st σ	2nd σ			
	MH		LM	
high register: HM	HM-MH	4	HM-MH	6
low register: ML	**H-M**	**5**	**H-M**	**3**
	o-MH	2	HM-MH	2

at all, since /ML.MH/ or /ML.LM/ yield a perfectly well-formed output [HM.MH] at a minimal cost of violating the low-ranking Ident-Reg. The alternative strategy embodied in Melody Simplification involves a far costlier violation of E≠O. E≠O prohibits turning a level tone into a contour tone, or the other way around (i.e. adding or subtracting a terminal tone node). This prohibition is absolute, except in the narrowly circumscribed class comprising /ML.MH/ and /ML.LM/. It seems obvious that Melody Simplification represents a non-harmonic mapping relation between the input and the output and, so far as I can see, cannot be derived from independently motivated output constraints.

4 Opacity

Recall that New Chongming has an eight-tone system, cross-classified by register (high, low), shape (rising, falling, or level), and syllable type. There are therefore 64 combinatorial possibilities of two tone-carrying syllables. Details aside, the net effect of the various constraints developed in the preceding sections is a drastic reduction of the underlying contrasts. Specifically, the 64 possibilities are reduced to just the eight listed below under column A:

(28)

	A winner candidates	B phonetic realizations
even tones:	H-H	H-**HM**, H-Hq
	L-H	L-**HM**, L-Hq
	H-o	H-o
	L-o	**LM**-o
	o-H	o-**HM**, o-Hq
oblique tones:	HM-MH	HM-MH, **H-M**
	MH-o	MH-o
	o-MH	o-MH

The tone patterns in the left-hand column (A) represent the output candidates picked out by EVAL consistent with the constraints we have established and ranked. However, as we noted at the very outset in section 1, we have omitted certain phonetic details in the interest of expository clarity. The actual phonetic shapes of the tone patterns are indicated in the right-hand column (B), where the phonetic differences are highlighted in boldface. We have already dealt with the register-conditioned split between [HM-MH] and [H-M] in section 3.6. The facts that remain to be accounted for are: /L-o/ surfaces as [LM-o] in actual pronunciation, and /H/ is realized as [HM] in a word-final "smooth" syllable.

4.1 "Counterfeeding" rule relation

To couch the problem in familiar rule-based terms, what we are confronted with is this: an *underlying* oblique tone (in this case /HM/) deletes when occurring in combination with an even tone /H/ by (29a), while a subsequent Allotonic rule such as (29b) creates a surface [H-HM]. In other words, (29b) counterfeeds (29a), as shown in derivation (30). Needless to say, a reversed order of rule application (transparent, feeding order) would obliterate the underlying contrast between /H-H/ and /H-HM/, and incorrectly generate [H-o] as the output for both.

(29) a. Oblique Deletion
oblique tone → ∅ / even tone

b. Allotonic
$H \rightarrow HM \;/\; \underline{\quad}\,]_{word}$
smooth σ

(30)

a. "capital"	b. "watermelon"	
zi-ben	*xi-gua*	
H-HM	H-H	
H-o	–	Oblique Deletion
–	H-HM	Allotonic

In OT terms, the issue becomes the following: output candidate [H-HM] (30b) simultaneously violates *E/O (no E-O or O-E sequences), *Fall (no final falling tone) and E≠O (do not change E into O, or vice versa), all of which are deemed to be undominated in section 2. By all accounts, one would expect any one of these high-ranking constraints to throw out [H-HM] as a serious contender.

A similar situation obtains in the phonetic manifestation of /Ls/ that survive as the only tonic elements in a compound. Take /L-HM/ *tong-ban* "copper coin" for illustration. A sequence of E-O like /L-HM/ is ill-formed, either L or HM must go; since the non-parsing of an even tone is more severely penalized than that of an oblique tone (Prs-E ≫ Prs-T↑), [L-o] is uniquely chosen as the optimal output by the complex of rank-ordered constraints {*E/O, Prs-E ≫ Prs-T′}. Now this [L-o] is actually pronounced as [LM-o] with a distinctly audible rise in pitch on the initial syllable. Let us state this process informally as Tonic Prominence (31). Without it, a lexically assigned L would be indistinguishable from a default L. In functional terms, the purpose of Tonic Prominence is to endow the tonic (i.e. tone-carrying syllable) with a certain perceptual prominence.

(31) Tonic Prominence
L → LM

The problem that arises is this: construed as a strictly output constraint, *O↓ (≫ Ident-Reg) should enforce register raising, an effect that is visible for instance in /HM-LM/ → [HM-MH] *zhang-gui* "manager" (32a). We therefore expect [LM-o] to be realized phonetically as [MH-o], counterfactually. New Chongming maintains a robust contrast between [LM-o] as in *tong-ban* "copper coin" (from /L-HM/, (32b)) and [MH-o] as in *dui-xiang* "target" (corresponding to /MH-ML/). In order to insure the correct results, we need to impose a counterfeeding order Register Raising > Tonic Prominence, as in (32).

(32) a. "manager"	b. "copper coin"	
zhang-gui	*tong-ban*	
HM.LM	L.HM	
–	L.o	Oblique Deletion
HM-MH	–	Register Raising
–	LM-o	Tonic Prominence

In short, Allotonic (H → HM) and Tonic Prominence (L → LM) *counterfeed* Oblique Deletion (to satisfy *E/O) and Register Raising (to avoid *O↓ violation) respectively.

4.2 A representational solution

How are we to account for these opacity effects in a descriptive model that eschews extrinsic rule ordering and makes use exclusively of output

constraints (cf. McCarthy 1995a)? The most promising solution seems to be a representational one.[31]

Consider again the pronunciation [H-HM] as the phonetic parse of /H-H/ in *xi-gua* "watermelon" (30b). It seems eminently plausible to interpret the allotonic variation between [H] and [HM] as an intonational phenomenon. According to Pierrehumbert (1980) and Liberman and Pierrehumbert (1984), the characteristic pitch envelope of an intonation pattern – called "tune" – is decomposed into a sequence of elements (H and L) that are aligned with certain designated elements of the text. These elements include pitch accents and boundary tones. A declarative intonation is characterized, for instance, by a low tone at the end of the phrase. This tonal element is referred to as the boundary L. An interrogative intonation, on the other hand, may be marked with a boundary H or possibly L + H etc. In light of this decompositional view of intonation, a word-final high level tone dips to [HM] by incorporating the boundary L̲ (underlined for clarity) that marks the end of an intonational phrase.[32] We state this process simply as (54). In other words, [HM] is the phonetic realization of the lexically assigned /H/ plus the boundary L̲.

(33) Boundary Low
Parse boundary L̲.

But why does the boundary L̲ manifest itself only in a word-final syllable carrying a H tone? Notice that a disyllabic compound may end in a tonic (toned, T) or an atonic (i.e. toneless, symbolized as "o") syllable. By default, an atonic syllable assumes a low tone. As a consequence, [T.o] = [T.L], on which the attachment of a prepausal boundary L̲ produces no detectable effect. On the other hand, recall that if a word-final syllable carries a tone at all, it must be either a H (if even) or a MH (if oblique).[33] If the word-final syllable carries a H, then the boundary L̲ adjunction produces a falling pitch HM – but only if the syllable is smooth, not if the syllable is checked. Why such a restriction? We may derive this fact in several ways. The simplest way is to assume that slack syllables (CVN,

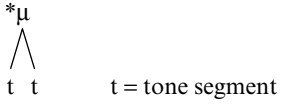
[31] I have discarded other problematic alternatives entertained in an earlier version.

[32] Since all disyllabic compounds we have elicited were pronounced in isolation, bounded by pauses, each elicited sandhi form constitutes a separate intonational phrase.

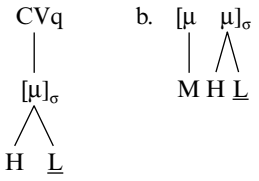
[33] This is so because, as the reader will recall, the register contrast is highly restricted: all oblique tones are phonetically high-registered (by virtue of *O↓), while all even tones are neutralized in the direction of high register in the final position (in accordance with *E↓). In addition, falling tones may not occur word-finally (*Fall). These constraints are listed in section 2, (4).

CVG, and CV [= CV:]) are bimoraic, while checked syllables (CVq) are characteristically short in duration,[34] and therefore monomoraic. On this assumption, the phonetic realization of the boundary $\underline{L}$ is subordinate to constraint (34) that forbids two tone segments riding on a single mora. Constraint (34) rules out a checked syllable carrying a falling tone – namely a [H] plus the boundary $\underline{L}$, as diagrammed in (35a). The same constraint (34) also automatically explains why the boundary $\underline{L}$ has no effect on a word-final syllable carrying an oblique tone MH. The reason is that given the bimoraic limit on syllables, the phonetic realization of the boundary $\underline{L}$ would entail a violation of (34), as shown in (35b).

(34)



(35) a. b.

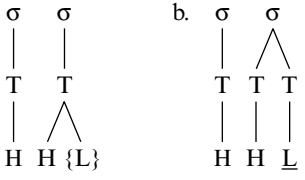


μ = mora
σ = syllable
$\underline{L}$ = boundary L

For expository simplicity, the constraint Boundary Low, simply stated as (33), subsumes the one-to-one relation between mora and tone segment expressed as (34).

With this in mind, we construe the phonetic realization HM of the word-final H not as a change in tone shape (from high level to high falling) or, equivalently, an insertion of a tone segment under the same tone root in the manner of (36a), but rather the adjunction of the boundary $\underline{L}$, arguably an independent root tone, in the fashion of (36b).

(36) a. b.

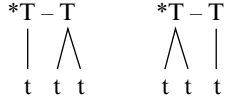


$\underline{L}$ = boundary L
{L} = inserted terminal tone segment L

[34] See Duanmu (1993a) for discussion and experimental evidence.

On this construal, (36b) in no way violates *E/O, stated in (4), which prohibits a heterogeneous mix of even and oblique tones (linear sequence immaterial). *E/O can be spelled out more fully as (37):

(37) *E/O:



E, O = even (level), oblique (contour) tone
T = tone root
t = tone segment

(37) only prohibits (36a), but says nothing about (36b). (36a, b) are configurationally different, though phonetically identical. In other words, *E/O bars a contour tone (36a), but not a tonal cluster (36b) from cooccurring with a level tone.[35] (36a, b) are instantiated by (38a, b) respectively:

(38) a. *zi-ben* "capital" /H-HM/ = [H-o] (= (30a))

b. *xi-gua* "watermelon" /H-H/ → [H-HM] (= (30b))

L = boundary L

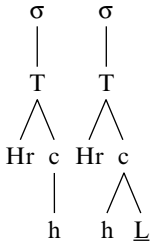
To put it succinctly: the tonotactics of New Chongming only prohibits a falling *tone* in word-final position, but freely permits a prepausal falling *intonation*.

There is one minor detail to take care of. Assuming that the incorporation of the boundary L creates a genuine contour tone, as in (39), we automatically account for the actual phonetic realization of H + L as

[35] For the distinction between contour tones and tone clusters, see Yip (1989).

[HM], since on this assumption the boundary <u>L</u> is construed as the terminal node linked to a high-registered (Hr) tone root.

(39)

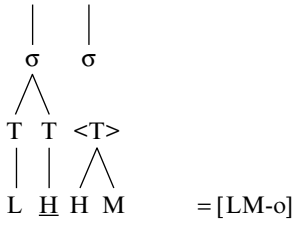


= [H-H<u>M</u>]

Such an explanation is not available if we interpret the boundary <u>L</u> as constituting an independent tone root, as in (38b). However, several alternatives present themselves. We may relegate the realization of the boundary <u>L</u> as a mid-level pitch [. . . M] to some low-level phonetics (for instance, undershooting of the phonetic target), or else impose a constraint to the effect that tautosyllabic tone roots may not disagree in register (cf. Duanmu 1990a). Register agreement will force a high-register interpretation of the boundary <u>L</u>.

Let us now turn to cases like (32b) *tong-ban* "copper coin," underlyingly /L.HM/. The rank-ordered set of constraints {*E/O ≫ Prs-E ≫ Prs-T} picks [L.o] as the output, which is phonetically realized as [LM.o] in accordance with Tonic Prominence, informally stated as (31). Ostensibly, the final phonetic output [LM.o] stands in violation of the undominated *O↓, which prohibits low contour tones. Again in this case intonation holds the key to this puzzle. As we noted earlier in section 4.1, the functional motivation behind Tonic Prominence is to guarantee a certain perceptual prominence to the tonic, i.e. tone-carrying, syllable. Without Tonic Prominence, the tonal element in [L.o] would be indistinguishable from the default L. As a result, [L.o] would be pronounced with a low monotone. We can implement this idea by making use of the notion of pitch accent. As pointed out above, the intonation curve is decomposed into one or more pitch accents <u>H</u> flanked by boundary tones. Pitch accents are associated with stressed or otherwise prominent syllables. Arguably in the prosodic system of tone languages, the tonic (i.e. syllable carrying a lexically assigned tone) represents the prominent element. The phonetic realization of [L-o] as [LM-o] reflects the superimposition of the intonational pitch accent <u>H</u> (underlined) on the tonic syllable *tong* "copper."

(40) *tong-ban* "copper coin," from /L-HM/ (= (32b))

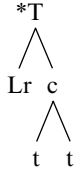


= [LM-o]

H = pitch accent
<T> = unparsed tone

The pronunciation of L + pitch accent H as [LM] is presumably due to the same tautosyllabic like-register constraint that turns H + boundary L into [HM] noted earlier.[36] Details aside, the important point here is that the tonal property associated with *tong* "copper" in (40) is configurationally distinct from that of (41) ruled inadmissible by *O↓. Constraint *O↓ is silent on a structure like (40).

(41) *O↓: *T



Our representational solution to the opacity problem crucially depends on the fundamental distinction between tonal clusters ($[T + T]_{\sigma}$) and true contour tones ($[t + t]_{T}$), that is assumed by virtually all students of Chinese tonology (including Wang 1967, Bao 1990a, Chen 1991b, 1992a), and motivated more explicitly in Yip (1989). It argues against the sequential treatment of Duanmu (1994a), which precludes the possibility of such a structural contrast, as holds between (40) and (41). The force of the argument is strengthened by the observation that tonal saliency determines in large part which of the three sandhi patterns [T.o], [o.T] or [T.T] a disyllabic compound will assume. Tonal saliency in turn is determined first and foremost by the structural complexity of the tonal geometry: simple, nonbranching tones (evens) are more stable than complex, branching tones (obliques), as encoded in the scale E > T↑ > Tp > T (see section 3.3).

[36] The interaction between tone and intonation is no doubt far more complex than the simplistic picture presented here. Chao (1968:39ff.) once compared syllabic tone and sentence intonation with "small ripples riding on large waves." I sidestep the issue of how to model the F_o as the algebraic sum of the two kinds of waves or target levels.

Note parenthetically that pitch accent H shows no overt effect elsewhere. This is easy to understand. Apart from [L.o] (= [LM.o]) we have the following disyllabic tone patterns (ignoring boundary L phenomena): (i) [H.o, o.H, MH.o, o.MH], (ii) [H.H, L.H, HM.MH]. In the case of (i) with one single tonic, by docking to the right of H or MH the pitch accent H produces no audible effect. As for (ii) with two tonics, recall that New Chongming is right-prominent. This means that we expect the pitch accent to fall on the rightmost tonic, again without overt phonetic effect. In particular, we do not expect [L.H] to become [L+H.H], where the pitch accent would create a rising [LM] on the first syllable. Instead, the expected [L.H+H] is indistinguishable from [L.H]. A fortiori, we do not expect [HM.MH] to turn into [HM+H.MH], a pattern that is at odds with the mora–tone constraint noted above as well as the right-prominent metrical organization.

5 Competing strategies

There remain further residual problems. Recall that an input string /MH.MH/ gives rise to two competing candidates [HM.MH] and [o.MH] in a draw, so to speak. For instance, both *fen-dou* "struggle" and *tan-qi* "to sigh" share the same input /MH.MH/, but the former shows up as [HM.MH], while the latter [o.MH]. This is documented in the tableau in (18), P3. One could account for the even split by making allowance for an indeterminate ranking between Ident-Cont (preserve contour identity) and Prs-Tone: Prs-Tone ≫ Ident-Cont in the case of [HM-MH], and Ident-Cont ≫ Prs-Tone in the case of [o-MH].[37] The same mechanism may be extended to account for (18), P4: corresponding to /MH-LM/ we get [HM-MH] in two cases (suggesting Prs-Tone ≫ Ident-Cont), as well as [MH-o] (suggesting Ident-Cont ≫ Prs-Tone) in four instances. While ranking indeterminacy is a logical possibility compatible with OT, it does not offer a completely satisfactory account of the facts. When we take into consideration other oblique-oblique patterns, the ranking between the two blocks of contour-identity and tone-parsing constraints, namely Ident-Cont ≫ Prs-Tone, looks quite robust. For instance, F-F could in principle be transformed into a well-formed R-F by altering the contour at the expense of Ident-Cont. Instead, in all cases but one, well-formedness

[37] Ident-Cont includes both Ident-Cp (preserve the contour of a prominent tone) and Ident-C (preserve any tonal contour).

is achieved by tonal deletion which gives rise to [o.T] or [T.o], Prs-Tone notwithstanding (see (10) in section 3.1).

5.1 Strict domination vs. numerical weighting

More important is the fact that ranking indeterminacy occurs in a non-random fashion. We can adduce a number of systematic subregularities to give substance to this observation. First, the undominated constraint *E/O bars the coexistence of an even tone with an oblique tone; therefore, one of the two adjacent tones must go. In the overwhelming majority of cases, the tone that loses out is invariably the oblique tone. This justifies the ranking: *E/O ≫ Prs-E ≫ Prs-T↑. However, there is one systematic class of exceptions, namely /L.MH/. In accordance with the constraint hierarchy, we expect the winner candidate to be [L.o] (phonetically [LM.o], see section 4.1). This turns out to be true only in one instance, namely *tong-zhi* /L.MH/ [L.o] "comrade." In the majority of cases, the attested sandhi forms are actually [o.MH], suggesting a ranking reversal: Prs-T↓ ≫ Prs-E. What is remarkable is that this exceptional ranking reversal occurs in this particular tonal combination, and nowhere else. That is, while /L.MH/ shows up as [o.MH], all other combinations of /L-o/ – that is /L.HM, L.ML, L.LM/ – emerge as [L-o], bearing out the ranking Prs-E ≫ Prs-Tone. Why? Contrast P1 with P2–4 in the following tableau:[38]

(42)

					Prs-E	Id-Cp	Prs-T↑	Prs-Tp	Id-Reg
P1			/L.MH/						
	a	$	L.o	1			*	*	
	b	☞	o.MH	4	*				
P2			/L.HM/						
	a	☞	L.o	7			*	*	
	b		o.MH	0	*	*			
P3			/L.LM/						
	a	☞	L.o	4				*	
	b		o.MH	0	*				*
P4			/L.ML/						
	a	☞	L.o	8				*	
	b		o.MH	0	*	*			*

Numbers indicate type frequency
P1–P4 = patterns 1–4

[38] I omit Prs-T (preserve any T) from the tableau, since (redundantly) all candidates violate this generic Prs-T. Prs-T is therefore neutral among this group of competing candidates.

(43) Examples corresponding to (42)
P1 *mi-xin* "superstition"
P2 *cun-kuan* "deposit"
P3 *chang-shou* "longevity"
P4 *niu-nai* "milk"

Given the Strict Domination hypothesis, as long as the (a) candidates of P1–4 satisfy the dominant Prs-E, they can flout all the other lower ranked constraints with impunity; conversely, the (b) candidates are doomed by Prs-E, no matter how well they behave relative to all the other norms of conduct. Strict Domination therefore wrongly picks P1a to be the winner. What went wrong? Intuitively speaking, what is going on is this: the normal "repair strategy" (Singh 1987, Paradis 1988) to fix up an ill-formed E-O sequence is to delete the less salient oblique tone. This is what occurs in P2–4. What then sets P1 aside from the other cases? In Chen (1991c) I attributed an analogous anomaly to "derivational economy." In derivational terms, the four cases P1–4 are analyzed as follows:

(44)

"superstition"	"deposit"	"longevity"	"milk"	
P1 *mi-xin*	P2 *cun-kuan*	P3 *chang-shou*	P4 *niu-nai*	
L.MH	L.HM	L.LM	L.ML	
o.MH	o.HM	o.LM	o.ML	E-Deletion
–	o.MH	–	o.MH	Melodic Inversion
–	–	o.MH	o.HM	Register Raising

It can be seen that in the case of P1 the deletion of the even tone L (E-Deletion), instead of the oblique as expected, produces a tone pattern [o-MH] that is well-formed without further modification. This contrasts with each of the other three cases, where a similar strategy leaves behind an "intermediate" form, which requires further "fixing." The repair measures may involve Melodic Inversion (P2), Register Raising (P3), or both (P4). It is this derivational simplicity[39] of P1 in (44) (= P1b of (42)) that prevails over the alternative P1a of (67) which would involve the deletion of an otherwise quite salient tone namely /MH/ which is both high registered and prosodically prominent (in strong, final position).

[39] Otherwise stated as the Harmonic Principle, which says to the effect: "Apply rules minimally to maximally satisfy phonotactic conditions" (see Smolensky 1986, Goldsmith 1990, 1993b; cf. Chen 1991c:17, and 1997).

In OT terms, what this means is that in some limiting cases, instead of checking each candidate against a (partially) rank-ordered set of constraints, discarding candidates as one descends on the hierarchy in the manner prescribed by Strict Domination, EVAL balances the sum of variably weighted violations of each candidate against that of another, and selects the candidate with the lowest total number of violations. As an illustration, suppose of the five constraints involved in (42), the highest ranked Prs-E and Ident-Cp have the weighted value of 4, while Prs-T↑, Prs-Tp, and Ident-Reg are assigned the values of 3, 2, and 1, in a descending order according to the ranking hierarchy established in (25). Calculated this way, the (a) candidates in P2–4 of (42) consistently score better than the (b) candidates. For instance, a total penalty of 5 is assessed against P2a (violation of Prs-T↑ + violation of Prs-Tp = 3 + 2 = 5), compared with 8 against P2b (Prs-E + Ident-Cp = 4 + 4 = 8). On the other hand, it is P1b that fares slightly better than its competitor P1a, since P1b carries a "debit" of 4 (for violating Prs-E), while P1a incurs a total of 5 marks in penalties (= Prs-T↑ + Prs-Tp = 3 + 2 = 5). Let us refer to this method of calculating constraint violation as "Weighted Aggregate." This modus operandi of EVAL is no doubt conceptually more cumbersome and computationally more costly,[40] but may at least suggest a possible source of explanation for why candidate P1b turns out to be the unexpected winner in four out of five cases, while candidate (a) consistently wins out in all the other patterns P2–4.

A number of observations lend support to the evaluation procedure described above. Contrast P1 of (42), repeated here as P1 in (45) below, with P2 of the same tableau. The only difference between the two cases is that while both fail to parse an even tone, P2b runs the extra cost of failing to parse an even tone that is inherently salient on account of its high register (highlighted in the shaded cell). However, under the Strict Domination hypothesis, this difference between P1 and P2 is absolutely

[40] For one thing, in order for EVAL to work, one has to assign a (possibly distinct) value or weight to each constraint, which surely will not be a trivial matter. Goldsmith and Larson (1990) and Goldsmith (1993b) develop a "learning apparatus" that can be trained to perform an analogous task, namely calculating the sonority coefficients for each phonetic feature. For example, based on syllabification patterns of Berber (Dell and Elmedlaoui 1985), laterals (liquids) with a coefficient of .56 are more likely to function as the syllabic nuclei (sonority peaks) than stridents with a coefficient of .19.

There is some on-going debate regarding absolute ranking (strict domination) vs. numerical weighting of constraints. See Joseph Stemberger (November 17, 1995) and Paul Smolensky's (November 18, 1995) comments on this subject disseminated over the OT-net.

immaterial. In both cases candidate (b) is out of the running as soon as it runs afoul of Prs-E; the additional violation of Prs-T↑ is of no consequence. We cannot therefore explain why candidate (a) is picked in the case of P2, but candidate (b) wins out in the case of P1.

(45)

				Prs-E	Id-Cp	Prs-T↑	Prs-Tp	Id-Reg
P1		/L.MH/						
a		L.o	1			*	*	
b	☞	o.MH	4	*				
P2		/H.MH/						
a	☞	H.o	7			*	*	
b		o.MH	0	*		*		

(46) Examples corresponding to (45)
P1 *mi xin* "superstition" (= P1 of (42))
P2 *che piao* "bus ticket"

On the other hand, if EVAL were to compare the two patterns of (45) according to Weighted Aggregate in the manner suggested above, then the tally would be as given in (47). EVAL would then pick the right winners.

(47)

		candidate	penalty score
P1 of (45)		a	5
	☞	b	4
P2 of (45)	☞	a	5
		b	7

There is another piece of evidence in support of numerical weighting. Compare P1 and P2 of (48). In both cases, candidate (a) wins out, because (b) is, in principle, eliminated by the constraint O↓-F, which stipulates that in a O↓-F sequence, initial O↓ should be unparsed (see section 3.6). This leaves unexplained why P1 is split between (a) and (b), while P2 displays a winner-takes-all (5 to 0) pattern. The difference lies in the pattern of Ident-C violation (boxes marked with vertical lines). But the standard procedure is incapable of taking cognizance of this difference: Strict Domination puts a blinker on EVAL, making EVAL blind to whatever difference there may be in the entire shaded regions of the tableaux.

(48)

					O↓-F	Id-Cp	Id-C	Prs-Tp	Prs-Reg
P1			/LM.ML/						
	a	☞	o-MH	3		*			*
	b	☞	MH-o	2	*			*	*
P2			/ML.ML/						
	a	☞	o-MH	5		*			*
	b		MH-o	0	*		*	*	*

(49) Examples corresponding to (48)
P1 a. *yun-dong* "exercise"
b. *da-yu* "downpour"
P2 a. *dao-li* "reason, motivation"
b. no example

To bring out more clearly the difference between the two scenarios portrayed by P1 and P2 of (48), let us restate the contrast in derivational terms as (50) and (51) respectively. Both /LM-ML/ and /ML-ML/ violate a number of well-formedness conditions, specifically a sequence of low-register oblique followed by a falling tone is barred by *O↓ and *Fall. One or the other tone must go. T-Deletion (deletion of a tone in metrically weak position) is the preferred strategy, since a prosodically stronger (rightmost) syllable is more resistant to tone loss (equivalently Prs-Tp ≫ Prs-T). This favors (50a) over (50b). On the other hand, T-Deletion leaves behind a falling tone (barred by *Fall). Hence the adopted strategy entails both shape inversion and register raising (to satisfy *O↓ and *Fall). An alternative course of action (50b) is to delete the last tone /ML/. This option pays the higher cost of Tp-Deletion (compared to the course of least resistance, T-Deletion). Counterbalanced against this extra cost stands the economy of derivation, since the output of Tp-Deletion only calls for a single step of Register Raising. There is, in other words, a close match between the two alternative derivations of (50a, b) in terms of cost. Hence we see the nearly even split of [o-MH] vs. [MH-o] sandhi forms corresponding to /LM-ML/.

(50)	a. "exercise"	b. "downpour"	
	yun-dong	*da-yu*	
	LM-ML	LM-ML	
	o-ML	–	T-Deletion
	–	LM-o	Tp-Deletion
	o-LM	–	Melodic Inversion
	o-MH	MH-o	Register Raising

a = P1a of (49)
b = P1b of (49)

T-Deletion	Delete a tone in metrically weak position
Tp-Deletion	Delete a tone in metrically prominent position

The situation is quite different in the case of (51), the derivational equivalent of P2 of (48). Here the more costly action of Tp-Deletion is not compensated for by a saving in derivational length. In the absence of such a tradeoff, the preferred "repair" strategy is decidedly that of (51a). This explains why we find no instantiated case of (51b): the only attested sandhi form corresponding to /ML-ML/ is [o-MH] (51a).

	"reason"			
(51) a.	dao-li	b.	??	
	ML-ML		ML-ML	
	o-ML		–	T-Deletion
	–		ML-o	Tp-Deletion
	o-LM		LM-o	Melodic Inversion
	o-MH		MH-o	Register Raising

a = P2a of (48)

Strict Domination has been deeply embedded in OT,[41] and strongly supported by a wide range of data. Ultimately, how EVAL calculates the degree of constraint satisfaction, or whether it requires recourse to the "full-blown power of numerical optimization" (Prince and Smolensky 1993:198) is an empirical question that must be decided in the light of carefully controlled and in-depth analyses of limiting cases like New Chongming.

5.2 *Saliency dispersion*

There is a second sense in which apparent ranking indeterminacy is systematic rather than random. It is noteworthy that ranking indeterminacy between Ident-Cont and Prs-Tone (the families of constraints on shape deforming and tone deletion) is observed mostly in those cases where the two adjacent rising tones contrast only minimally in terms of salience. Let us refer to the difference in saliency as "Saliency Dispersion" (by analogy with "Sonority Dispersion").[42] This seems to suggest that tonal

[41] Prince and Smolensky (1993:198): "In Optimality Theory, constraints are ranked, not weighted; harmonic evaluation involves the abstract algebra of order relations rather than numerical adjudication between quantities."

[42] The concepts of minimum sonority distance and sonority dispersion have proven useful in determining patterns of syllabification. For instance, a language may require that onset consonants must display a rising sonority, with a minimum distance of two on a four point scale: stop–fricative–nasal–liquid. In such a hypothetical language, onset clusters *kn-, fl-, tr-* are permissible, but *pf-, sn-, mr* are not. Cf. Steriade (1982), Clements (1990b).

deletion is preferred over contour adjustment (which is the norm) – except when Saliency Dispersion is so small that it is "difficult" to single out which of the two tones is the less salient as a target for deletion. In such a case both tones are retained. Hence, corresponding to /MH.MH/ we have both [HM-MH] and [o-MH] as attested sandhi forms, since in this case we have two tokens of the same tones /MH/, which differ in prominence solely because of their position (see (18), P3). The difference in saliency is even smaller in the case of /MH-LM/ ((18), P4), where the high register prominence of /MH/ and the prosodic strength of /LM/ "cancel" each other out. It is mainly in these cases of minimal Saliency Dispersion that we see ranking indeterminacy as reflected in a split between [HM-MH] patterns (Prs-Tone ≫ Ident-Cont) and [MH.o] (Ident-Cont ≫ Prs-Tone).

Interestingly enough, the converse is also true: tonal deletion is favored when Saliency Dispersion is significant, so that one of the two tones is significantly less salient than the other, and therefore easily targeted for deletion. In this connection recall that when the choice is between deleting a low-profile tone (thereby incurring a Prs-T violation) and raising its register (in violation of Ident-Reg), the latter alternative prevails, as can be seen in section 3, P3–4 of (12), repeated below:

(12) F-R patterns

						OCP	*Fall	*O↓	Prs-T	Id-Reg
P1	HM-MH	☞	a	HM-MH	4					
P2	HM-LM		a	HM-LM	–			*		
		☞	b	HM-MH	6					*
			c	HM-o	–		*		*	
P3	ML-MH		a	ML-MH	–			*		
		☞	b	HM-MH[43]	5					*
			c	o-MH	2				*	
P4	ML-LM		a	ML-LM	–			**		
		☞	b	HM-MH[44]	5					**
			c	o-MH	–				*	*

This means that Prs-T ≫ Ident-Reg – in the majority of cases (ten out of twelve examples). This account leaves a residue, where the reverse ranking seems to hold. Specifically, the same input string /ML-MH/ emerges as [o-MH] (P3c, incurring Prs-T violation) in two examples instead of [HM-MH] (P3b), as expected in the remaining five. Why does apparent

[43] As noted earlier, this output has [H.M] as a register-dependent variant; see section 5.
[44] See preceding footnote.

ranking indeterminacy appear just in those cases like /ML-MH/ (P3)? The reason appears to be that the two tones /ML-/ and /-MH/ are located at the opposite poles of the scale of tonal saliency: ML is both low and prosodically weak, while MH is both high and prosodically strong. Intuitively speaking, what is at play here is the principle that tonal deletion is most likely to take place when one of the two abutting tones is most easily identifiable as the target. In other words, the expected sandhi of F-R is [T.T] or [T.T′]; but where one tone is distinctly less salient than the other, it undergoes deletion, giving rise to [o.T] in a minority of cases (2 out of 22 examples in (12)).

To capture the intuitive content of the foregoing discussion, suppose we assign a value of 1 to a tone if it (i) is high, or (ii) occupies the right-dominant position. If a tone is both high and right-dominant, then it has the cumulative value of 2, and so forth. We then calculate Saliency Dispersion by noting the difference between the two values: Saliency Dispersion = $\Delta(T_1, T_2)$. If our speculation is on the right track, we expect exceptional T-T patterns (from R-R) to concentrate in those cases where Saliency Dispersion is the smallest; conversely, we expect exceptional o-T patterns (from F-R) to concentrate in those cases where Saliency Dispersion is the greatest. Our expectation seems to be borne out, as seen in the following charts.

(52) R-R patterns

	Saliency Dispersion	expected: o.T	T.o	exceptional: T.T, T.T′
LM-MH	2 Δ(0,2)	6	–	1
LM-LM	1 Δ(0,1)	4	–	–
MH-MH	1 Δ(1,2)	3	–	3
MH-LM	0 Δ(1,1)	–	4	2

(53) F-R patterns

	Saliency Dispersion	expected: T.T, T.T′	exceptional: o.T
ML-MH	2 Δ(0,2)	5	2
ML-LM	1 Δ(0,1)	5	–
HM-MH	1 Δ(1,2)	4	–
HM-LM	0 Δ(1,1)	6	–

T.T HM.MH
T.T′ H.M
o.T o.MH
T.o MH.o

In (53), given the perfectly well-formed F-R (abstracting away from register), no tonal deletion is expected, and yet [o.T] occurs, just where Saliency Dispersion is greatest (= 2). Conversely, in (52) R-R is ill-formed, and therefore tonal deletion is expected; instead we find [T.T], most of which occur where Saliency Dispersion is the smallest (equal to or less than 1).

Appendix Sandhi forms of disyllabic compounds (New Chongming dialect)

	2nd σ					
	even T		oblique T			
	smooth σ	checked σ	falling		rising	
1st σ	H, L	Hq Lq	HM	ML	MH	LM
H	H-HM 13	H-Hq 13	H-o 27			
L	L-HM 12	L-Hq 12	LM-o 7	LM-o 8	o-MH 4 LM-o 1	LM-o 4 HM-MH* 1 H-o* 1
Hq, Lq	o-HM 17	o-Hq 17	o-MH 29			
HM	o-HM 45	o-Hq 44	MH-o 6	MH-o 3	HM-MH 4 H-HM* 1	HM-MH 6 o-HM* 1
ML			o-MH 5	MH-o 5 H-M 1	H-M 5 o-MH 2	H-M 3 HM-MH 2
MH			MH-o 6	MH-o 4	o-MH 3 HM-MH 2 H-M 1 MH-o 1	MH-o 4 HM-MH 2
LM			o-MH 6	o-MH 3 MH-o 2	o-MH 6 HM-MH 1	o-MH 4 MH-o 1 o-HM* 1

Numbers indicate type frequency (total subcorpus consists of 346 items).
Starred forms (*) represent exceptions unaccounted for in the analysis.
"o" represents a toneless syllable.
Underlying tonal categories of the first syllable are indicated on the leftmost column, while the base tones of the second syllable are arrayed across the top. Sandhi forms are found in the intersecting cells. Sometimes several base tone combinations converge on a single sandhi form; thus a string consisting of a /H/ followed by any oblique tone /HM, ML, MH, LM/ shows up as [H-o]. Conversely, one underlying tonal string may give rise to several sandhi forms; thus /LM-LM/ has three attested surface phonetic readings [o-MH, MH-o, o-HM].

6 *From tone to accent*

So far we have looked at various types of tone sandhi processes and their interaction with each other. The effects we have examined so far are mostly local in nature, and consist of modifying in one way or another the melodic shape of tones in concatenation. But tone sandhi can have a more global effect on the phonological system as a whole. Goldsmith (1984:48f.) painted the classic scenario where a single tone rule can bring about an overall accentual reinterpretation of an erstwhile full-fledged tonal system. Proto-Bantu fully exploited the paradigmatic H vs. L contrast, so that a disyllabic nominal stem would have all the four logical possibilities: HH, HL, LH, and LL. At some point in the history of Bantu, the well-known Meeussen's rule – or rather its diachronic antecedent – turned a H into a L, when preceded by another H:

(1) Meeussen's Rule
$H \rightarrow L \ / \ H \ \_\_\_\_$

This rule has the effect of reducing the four possibilities to a three-way contrast: HL, LH, or LL, which lends itself to an accentual reinterpretation, according to which a (disyllabic) nominal stem is either unaccented (o.o), or else carries an accent on the first (*.o) or the second syllable (o.*), as depicted below:

(2)

Proto-Bantu	Pre-Tonga	
H.H H.L	H.L	*.o
L.H	L.H	o.*
L.L	L.L	o.o

* = pitch accent

It doesn't matter that the "asterisks" can be interpreted as prelinked Hs (cf. Pulleyblank 1986, Inkelas and Zec 1988, Blevins 1993), the point is that what started out as a paradigm case of tonal contrasts has now evolved into a purely syntagmatic system, where the only relevant lexical information is the location of the accent or prelinked H. Unaccented syllables take the default L. This is, of course, an idealized and grossly oversimplified picture. In actuality, how an accentual system like the one hypothesized for Pre-Tonga is tonologically interpreted depends on the tonal grammar of individual languages. For the historical evolution of the Proto-Bantu tonal system and the tone-to-accent continuum, see contributions to Clements and Goldsmith (1984), Hulst and Smith (1988), and Odden (1985, 1995). The simplified scenario sketched above suffices to illustrate the point we are making.

In this chapter I propose to investigate the tone-to-accent evolution as instantiated in some Chinese dialects. I will start with a simple case, exemplified by Shanghai, where a tonal system has been reduced to an accentual system, at least for one type of sandhi forms (section 1). I then examine in detail the more complex and intriguing case of New Chongming (sections 2–8), which we had already investigated at length with respect to disyllabic compounds in chapter 5. In this dialect, as I will show, tonal characteristics (register, melodic contour) determine the location of the accentual prominence; while the syntagmatic contrast of prominence renders largely redundant the paradigmatic contrasts of tonal categories. On one hand, various other tone/accent reduction rules create a situation where only one tone occurs within a (possibly unbounded) foot; on the other hand, a sweeping Leveling rule neutralizes the sole surviving tone uniformly to a simple H. The net result is that, in connected speech, the only remaining contrast is that between tonic (tone-bearing) and atonic (toneless), or accented and non-accented.

1 Shanghai: an aborted accentual system?

Shanghai has two types of sandhi forms in polysyllabic compounds. The first, type A, is derivable straightforwardly via tone deletion and spread. Tone deletion eliminates all but the leftmost tone; tone spread simply re-aligns the remaining tone segments with the syllables in a left-to-right, one-to-one fashion. This two-step derivation is illustrated by examples like the following:

Table 6.1. *Shanghai*

citation tone	sandhi forms		
	Type A	Type B	Type A ~ B
1 HL.T	*çiã.tã* "altercation" H. L	= Type A	
2 MH.T	*tçin.tsï* "scenery" M. H	*sø.su* "arithmetic" H. L	*ts'ã.p'i* "record" M. H ~ H. L
3 LH.T	*dᴇ.huã* "yolk" L. H	*mᴀ.hu* "careless" H. L	*lɔ.sï* "teacher" L. H ~ H. L
4 Hq.T	*piq.tçi* "tip of pen" M. H		
5 Lq.T	*zoq.k'ᴇ* "to drill open" L.MH		

(3) Type A sandhi forms

a.	*çiã tã*	"altercation"
	HL.MH	base tones
	HL. o	tone deletion
	H. L	tone spread
b.	*dᴇ huã*	"yolk"
	LH. HL	base tones
	LH. o	tone deletion
	L. H	tone spread

Type A is the dominant pattern and abundantly described in the literature (see chapter 7, section 3). There is another subset of sandhi forms, referred to as type B in Xu et al. (1988). Type B forms are not derivable in the usual fashion, and therefore seldom mentioned outside of the primary sources.[1] In addition, there are individual lexical items that fluctuate between type A and type B readings. This state of affairs is summarized in table 6.1.

How do we account for type B sandhi forms? Recall that type A readings are derived by tone deletion and spread. Notice further that the tone that survives and subsequently spreads over the entire disyllabic

[1] With the notable exception of Chang (1992:217f.).

compound is invariably the tone lexically associated with the first syllable. The prevailing analysis (also the one advocated here) sees a direct link between tonal stability and accentual prominence or metrical strength. In other words, Shanghai represents a left-prominent system. Therefore, it is the first, dominant syllable that determines the melodic shape of the compound as a whole. For details and arguments, the reader is referred to chapter 7, section 3. At this point I would like to suggest that type B merely represents a more radical option of encoding left prominence by means of tonal inflection. In a subset of lexical items – labeled here as type B – this strong–weak accentual pattern simply translates as H+L, regardless of the tonal categories lexically associated with the constituent syllables.

The plausibility of such an analysis comes from what at first appears to be an anomaly, namely the fact that type B is restricted to compounds with a smooth syllable (i.e. CVN)[2] in the word-initial position. The phonetics of the smooth vs. checked (i.e. CVq) syllables gives us a clue. Based on measurements provided by Zee and Maddieson (1980:54, fig.4), the former is more than twice the length of the latter (307 msec compared to 125 msec). It is therefore not unreasonable to take CVN and CVq syllables as heavy and light respectively. The distributional restriction on type B lexical entries is easy to account for if we express the obvious generalization as follows:

(4) Type B sandhi forms (Shanghai)

a.	Leftmost:	Assign stress on the leftmost σ
b.	WSP:	Avoid stressing an (initial) light σ

By further stipulating that WSP (Weight-to-Stress Prominence, Prince 1990) – or equivalently Peak-Prominence (Prince and Smolensky 1993:39) – dominates Leftmost, what we are saying is that Shanghai is basically left-prominent, subject to WSP, which disfavors prominence on a light syllable. In other words, Type B forms fall into the two accentual patterns (a, b) of table 6.2. The constraint hierarchy WSP ≫ Leftmost determines the accentual pattern of column 1. Accent is phonetically realized as a H tone (column 2). Posttonic syllables receive a default [l], while pretonic syllables get a [m] or a [l] tone, depending on the onset of the first syllable (column 3): [m] if the onset is voiceless, otherwise [l]. Examples are given in column 4.

[2] Recall that CVN abbreviates CV and CVN. See chapter 1, section 2.1.

Table 6.2. *Type B sandhi forms*

	1 accent	2 pitch-accent	3 default tone	4 examples
a. CVq + σ	(. x) σ σ	(. x) σ σ H	i. (. x) pVq. σ m H	*piq.tçi* m H "tip of pen"
			ii. (. x) bVq. σ l H	*zoq.k'ɛ* l H "to drill open"
b. CVN + σ	(x .) σ σ	(x .) σ σ H	(x .) σ σ H l	*sø.su* H l "arithmetic"

p, b = voiceless, voiced onset
m, l = default M and L
-q = glottal stop

What is remarkable in this account is the rather startling observation that for type B sandhi forms, tonal categories are redundant altogether. Their tonal patterns are entirely predictable from the syllable types and voicing contrast, mediated by the constraints or preference rules on accentuation stated as Leftmost and WSP. In other words, as far as the type B sublexicon is concerned, Shanghai is a pure pitch-accent language, with perfectly predictable placement of tonally realized accents.

The prominence-based analysis along the lines sketched above is further supported by a comparison between this straightforward pitch-accent account and the conventional tone-deletion-cum-reassociation approach. Consider two of the examples cited in table 6.2.

(5) Conventional tone-deletion-cum-spread analysis

a.	*piq.tçi*	"tip of pen"
	Hq.HL	citation tones
	/MHq.HL/	underlying tones
	Mq.H	sandhi tones
b.	*zoq.k'ɛ*	"to drill open"
	LMq.HL	citation tones
	/LHq.HL/	underlying tones
	Lq. H	sandhi tones

Table 6.3. *Disyllabic patterns with an initial CVq*

	1	2	3	4	5
	citation tone	underlying tone	Deletion	Spread	examples
a	Hq.T	MHq.T	MHq.o	Mq.H	*piq.tçi* Mq.H "tip of pen"
b	LMq.T	LHq.T	LHq.o	Lq.H	*zoq.k'ɛ* Lq. H "to drill open"

In order to derive the surface tones of these two examples, the conventional analysis must posit putative underlying representations that depart from the actually occurring citation tones. Specifically, *piq* Hq "pen" and *zoq* LMq "to drill" must have /MHq/ and /LHq/ respectively as their abstract underlying representations. In essence, the abstract underlying representations are motivated by (i) the presence of a M tone on the first syllable in (5a), and (ii) the appearance of a H tone on the second syllable in (5b). In a deletion-cum-spread analysis, both tone segments M and H must have their source in the underlying representations. Hence the need to posit underlying pitch contours that never occur as citation tones. In other words, the analyst is forced to depart from the observable (column 1, table 6.3), and resort to a "theoretical construct" (column 2). Deletion (column 3) and Spread (column 4) eventually produce the sandhi forms of column 5. Needless to say, having posited /MHq/ and /LHq/ as the underlying tonal categories for (a) and (b) in table 6.3, we need extra rules to derive from them the actual citation tones: /MHq, LHq/ → [Hq, LMq].

This is basically the analysis – with minor variations – of all current proposals including Zee and Maddieson (1980), Yip (1980), Wright (1983), Jin (1986, 1995), Z. Zhang (1988), Selkirk and Shen (1990), Duanmu (1993a, 1995), among others.[3] Needless to say, this modus operandi is quite standard in linguistic practice. However, in this particular case, I submit that the abstract underlying representations in column 2 of table 6.3 are nothing but an artifact of the conventional analysis. A better approach is

[3] Yip (1980:312) has the opposite problem: having posited /M/ and /Mq/ instead of /MH/ and /MHq/ as the base form for tone categories B and D, she requires a H-insertion rule to derive the H tone observable in the three- and four-syllable compounds with an initial syllable carrying one of these two tones.

to discard the abstract underlying tones altogether. The H on the second syllable of *zoq.k'ɛ* "to drill open" is the phonetic realization of the pitch-accent. By the same token, the "extra" M in the first syllable of *piq.tçi* "tip of pen" arises out of default pitch assignment. There is no need for underlying tones as distinct from citation tones – in fact, there is no need for any tonal specification at all for polysyllabic compounds with a checked syllable in the initial position.

A final note on Shanghai: our primary source (Xu et al. 1988:147) observes that type B forms are in the minority, favored mostly by older speakers. This suggests a scenario in which the left prominence of the prosodic structure in Shanghai manifests itself in two ways: (i) a milder form of tonal stability and spread; (ii) a more radical and reductionist form of weight-sensitive accentual patterns, tonally realized as [H.o] and [o.H]. The former has prevailed, while the latter has aborted, and survived only in some individually marked lexical items.

While the accentual subpatterns in Shanghai look like a remnant of the past, a similar tone-to-accent development is taking place and gaining momentum in another Wu dialect, none other than the New Chongming variety we have already encountered in chapter 5.

2 New Chongming: an emergent accentual system

With respect to New Chongming, as soon as one goes beyond the narrowly defined scope of disyllabic compounds, one makes a rather startling discovery, namely that this northern Wu dialect has already evolved from a tonal to an essentially accentual system, exhibiting a syndrome of an unmistakably accent-like nature. Specifically, New Chongming displays the following tendencies in varying degrees:

1. **Culminativity**: trisyllabic and longer rhythmic units have exactly one tonic or tone-bearing syllable. This one-per-morpheme restriction is one of the diagnostic criteria for distinguishing accentual from tonal systems (cf. Hyman 1978, McCawley 1978, Hyman and Byarushengo 1984, Odden 1985, 1995, Beckman 1986:44f.).

2. **Leveling**: underlying tonal contrasts of register and contour are neutralized to a simple high-level tone (H) by a process I will call Leveling. Culminativity and Leveling jointly describe a prototypical accentual system which exploits exclusively syntagmatic pitch contrasts. In such a system, linguistic expressions contrast solely in the location of tone, not in terms of either register or melodic shape. Thus, even though New Chongming

has eight underlying tones, instead of the expected 512 combinatorial possibilities (= 8^3), a trisyllabic compound may carry only one of the three tonal patterns: [H.o.o], [o.H.o], or [o.o.H] (where o = atonic or toneless).

3. **Edgemostness**: the tonic syllabic tends to occur at the leftmost or the rightmost margins. Thus New Chongming polysyllabic expressions are overwhelmingly head-terminal [H.o.o] or [o.o.H], except when prosodic weight dictates otherwise (see next paragraph). Edgemostness or head-terminal foot construction is recognized as a manifestation of the demarcative function of accent/stress (cf. Hyman 1977, Halle and Vergnaud 1987).

4. **Prosodic weight**: since there is only one tone (uniformly H) per unit, the question that arises is: "where does it fall?" The location of the tone or "pitch accent" is determined by a number of factors, including notably prosodic weight. In New Chongming, other things being equal, given the morphosyntactic structures [[A B] X] and [X [A B]], it is the structurally more complex and prosodically "heavier" constituent, namely [A B] that typically carries the pitch accent. This weight-dependent prominence is another diagnostic of accentual systems, reminiscent of the well-known stress rule of English to the effect that given nodes $N_1 + N_2$, N_2 is "strong" if and only if it branches, as in (*cóm.pen*)(*sàte*) vs. (*còm.pen*)(*sá.tion*), where parentheses enclose foot-sized units (cf. Liberman and Prince 1977, Hayes 1982, 1995). In section 5 we will bring forth evidence in support of extending the Weight-to-Prominence principle (Prince 1990, cf. Prince and Smolensky 1993) beyond mora count.

5. **Tonic clash**: tone and tonelessness do not contrast merely as [voiced] vs. [voiceless] or [round] vs. [spread]; instead, tonicity (i.e. tone-bearingness) carries with it an inherent accentual prominence. In many cases, general rules of tonic prominence predict output forms like [(T)(T.o)] or [(o.T)(T)], with two abutting tonics, i.e. two syllables bearing distinctive tones (indicated as T).[4] When such a configuration arises, the tonic clash is alleviated either by eliminating one of the two adjacent T's or by shifting one T away from the other. Thus, [(o.T)(T)] → either [o.o(T)] or [(o.T)o] (by de-accenting), or [(T.o)(T)] (by accent shift). De-accenting is functionally equivalent to Destressing in English (cf. (*à*)(*brâ.ca*)(*dá.bra*) → (*à.bra.ca*)(*dá.bra*)), while accent shift is analogous to Iambic Reversal

[4] Where appropriate I will adopt the following convention: (. . .) and [. . .] circumscribe feet and higher prosodic constituents (p-word, phrase) respectively. When called for, tones and syllables are separated by a dot.

(cf. *Ja.pán* → *Jà.pa.nése*, or *Tennessèe* → *Tènnessee Wílliams*)[5] (cf. Liberman and Prince 1977, 1983, Hayes 1984, Selkirk 1984).

6. **Focus** and semantic contentfulness are signaled by means of tonic accent, while the tones of backgrounded elements are optionally suppressed in connected speech.

In the sections to follow we will examine in turn each of these accentual properties of New Chongming. But before we go into the specific details let us first look at the overall picture.

2.1 The overall picture: one tone per unit

The symptoms enumerated above manifest themselves in varying degrees across construction types in New Chongming. Overall, we see a tendency toward the "one tone per unit" target in two senses: (i) only one tone is allowed for each rhythmic unit; and (ii) every tone has one single phonetic realization, namely H. This one-tone-per-unit principle has firmly established itself in trisyllabic lexical compounds, so that a trisyllabic word may have just one of three possible tonal patterns: [H.o.o], [o.H.o], and [o.o.H], illustrated below:

(6)	a.	*[xin zang] bing*	"heart disease"
		H ML LM	base form
		(H o o)	sandhi form
	b.	*[za huo] dian*	"grocery store"
		Lq.MH MH	
		(o H o)	
	c.	*chu [feng tou]*	"to seek the limelight"
		Hq H L	
		(o o H)	

H marks the tonic (tone-carrying) syllable; the symbol "o" symbolizes atonicity (or absence of tone). The pitch values of otherwise atonic syllables are predictably mid level in pretonic position, low and slightly falling post-tonically. These positionally determined default pitch values clearly reflect the well-known intonational phenomenon of "downtrend": as the overall pitch range declines over the intonational phrase, the default L in the beginning of the utterance is phonetically higher than the same default L toward the end (see Pierrehumbert 1980, Liberman and Pierrehumbert 1984). We annotate the pre- and post-tonic default tones with the lower-case [m]

[5] Or its mirror image in German: *Háupt-Vórlesung* → *Háupt-Vorlésung* "keynote address."

and [l], to distinguish them from phonologically specified M and L. If we use an asterisk to signal the prominence of the tonic syllable, then (6) can be represented as (7).

(7) a. *
xin zang bing "heart disease"
(H l l)

b. *
za huo dian "grocery store"
(m H l)

c. *
chu feng tou "to seek the limelight"
(m m H)

The prosodic subsystem depicted here is identical in all essential aspects to the canonically stress-accent system of English. In an article intriguingly entitled "English as a tone language", Goldsmith (1981) derives the tonal patterns of the examples below by positing a single "melody", basically the intonation pattern M-H-L. H is associated with the prominent (accented) syllable, and M and L are linked to the pre- and post-tonic syllables respectively.

(8) a. *
Washington
(M) H L

b. *
Montana
M H L

c. *
Tennessee
M H L

(8a, b, c) correspond exactly to (7a, b, c), down to low-level phonetic details. Notice that in a dactylic pattern like *Wáshington*, the initial MH is simplified to H; on the other hand, the final HL on the final syllable of an anapestic word like *Tennessèe* maintains its falling contour. We see the same asymmetry in New Chongming. The first syllable of *xin-zang-bing* "heart disease" (7a) carries a simple H; but the last syllable of *chu feng*

tou "to seek the limelight" (7c) is phonetically realized with a perceptibly falling pitch HM, which we transcribe here simply as H – explicitly interpreting the phonetic HM as H + boundary L (see chapter 5, section 5.3).

The second set of processes has to do with what we may call "concatenation." Slow, deliberate speech tends to be made up of short, staccato tonic chunks, each containing a distinctive tonic peak. Normal, fluid speech, on the other hand, links two or more such units together into one single continuous stretch. As a corollary of this fusing of units of speech planning, one or other of the tonic elements loses its prominence and becomes a weak, subordinate part of the larger structure. One phonetic effect of this familiar stress subordination manifests itself in the de-toning of the weaker of the two prosodic constituents, so that the less prominent syllables lose their tones entirely and assume the default pitch values. This process is complete within the minimal utterance unit, the (stress-)foot. In larger domains, the prosodically weaker constituents may or may not retain their tones.

2.2 *Lexical vs. postlexical sandhi forms*

One final note before we delve into the particulars. In this chapter we will draw examples from lexical compounds as well as from larger, phrase-like expressions. New Chongming, like many other Wu dialects, distinguishes two types of tone sandhi: lexical and postlexical. Lexical tone sandhi applies within the word;[6] postlexical tone sandhi applies to certain types of constructions beyond lexical compounds such as: number + measure (e.g. two + pails [of water]), verb + pronoun (visit + her), verb + resultative complement (drink + intoxicated = get drunk), verb + directional complement (run + away), various types of reduplication (walk + walk = take a walk), etc.[7] As can be seen from this partial list, postlexical structures are in some sense intermediate between syntactically opaque, often idiosyncratic and fossilized word-size expressions and free, open-ended, productive phrasal constructions. Generally speaking, the sandhi forms of longer expressions are built on the disyllabic base. We have already examined in considerable detail the lexical sandhi forms of disyllabic compounds in chapter 5. As for postlexical tone sandhi, the specific sandhi

[6] More accurately, lexical tone sandhi applies within the prosodic word (p-word). The p-word may be larger or smaller than the morphological word. Cf. Chen and Zhang (1997).

[7] The sandhi forms created by postlexical tone sandhi correspond roughly to what is often referred to as *zhuan-yong shi,* literally "specialized forms," in the nomenclature widely adopted in China. Contraposed to *zhuan-yong shi* is *guang-yong shi* or "general-purpose form," which is produced by what we call lexical tone sandhi.

Table 6.4. *Postlexical sandhi forms (number + measure constructions)*

first σ		second σ			
		postlexical sandhi form		lexical sandhi form	
		even	oblique	even	oblique
even	H	H.H	H.o	(same as postlexical forms)	
	L	L.H	L.o		
	Hq, Lq	o.H	o.MH		
oblique	HM	HM.MH		o.H	MH.o, HM.MH
	ML	H.M		o.H	H.M
	MH	H.H		o.H	MH.o, o.MH, HM.MH
	LM	o.MH		o.H	o.MH

Hq, Lq = H and L associated with checked syllables
Even tones = H, L, Hq, Lq
Oblique tones = MH, LM, HM, ML

form is often construction-specific. Thus, some dialects may have one form for verb + complement constructions (resultative or directional etc.), another for number + measure (or classifier), and so forth (see chapter 2, section 5.2). The details are beyond the scope of this chapter, and not particularly relevant for the main theme we are focusing on. For expository simplicity, I will illustrate the distinction between lexical and postlexical tone sandhi with just sandhi forms of number + measure constructions. Table 6.4 summarizes the disyllabic postlexical sandhi forms corresponding to the underlying tonal categories of the individual items making up the expression.

As usual, disyllabic postlexical constructions are cross-classified according to the underlying tonal categories of the contributing syllables, first syllables down the left column, second syllables across the top. Sandhi forms are found in the cells where the columns and rows intersect. Lexical sandhi forms are given to the right for comparison.[8] It can be seen that postlexical and lexical sandhi forms are identical for disyllabic expressions of which the first syllable carries an even tone, as shown in these examples:

[8] Omitting sporadic exceptions and specific conditions. For details on disyllabic lexical sandhi forms, see chapter 5.

(9)

	phrasal constructions				lexical compounds	
a.	*san ping* H L H H	"three bottles" base tone sandhi form	=	a′.	*hua ping* H L H H	"vase" base tone sandhi form
b.	*qi li* Hq.ML o. MH	"seven miles"	=	b′.	*hei shi* Hq. ML o. MH	"black market"
c.	*liou dui* Lq HM o. MH	"six pairs"	=	c′.	*li shi* Lq. HM o. MH	"history"

Lexical and postlexical sandhi forms differ, however, when the first syllable carries an oblique tone. While oblique + tone (O + T) sequences yield a variety of lexical sandhi forms depending on the identity of the second tone, the corresponding postlexical forms converge on one single sandhi form regardless of the second tone. Thus, the number *jiu* /HM/ "nine" uniquely determines the tonal pattern [HM.MH] of the disyllabic number + measure expression, regardless of what tone is underlyingly associated with the measure word. The difference between lexical and postlexical sandhi forms is best illustrated by the (near-) minimal pairs listed below:[9]

(10)

	base tone	phrasal sandhi form	lexical sandhi form
a.	HM.H	HM.MH *jiu jin* "nine catties"	o.H *jiu jing* "alcohol"
b.	HM.HM	HM.MH *jiu wan* "nine bowls"	MH.o *jiu wan* "wine glasses"
c.	MH.H	H.H *ban tian* "half a way"	o.H *bian tian* "weather change"
d.	MH.HM	H.H *ban wan* "half a bowl"	MH.o *ban dao* "peninsula"
e.	MH-Hq	H.Hq *si bai* "four hundred"	o.Hq *si bo* "fourth uncle"

[9] The dividing line between lexical and phrasal constructions is notoriously fuzzy. We shall have more to say about this issue in chapter 9.

f.	ML.H	H.M *wu jin* "five catties"	o.H *wu jing* "the five classics"[10]
g.	LM.L	o.MH *nian tai* "twenty pieces (of equipment)"	L.H *yan tai* "inkwell"

Since chapter 5 has already dealt with the disyllabic sandhi forms in excruciating detail, in this chapter I will take for granted the monosyllabic and disyllabic sandhi forms, and will not elaborate on the rules that generate them or constraints that select them from among other conceivable alternatives. As will become apparent, the sandhi forms of larger units are almost invariably derivative or extensions of the disyllabic base.

We now proceed to detailing the various accent-like properties of a tonal language like New Chongming, beginning with culminativity.

3 Culminative accent

Recall from chapter 5 that New Chongming has an eight-tone system, and that disyllabic compounds may have any one of the three basic sandhi patterns: [T.o], [o.T], or [T.T]. Furthermore, the tones in these patterns may contrast in terms of even vs. oblique, that is level vs. contour, as in [H.o] vs. [MH.o] and [H.H] vs. [HM.MH]; they can also contrast in high vs. low register, as in [H.o] vs. [L.o] and [H.H] vs. [L.H]. In this sense, tonal distribution is unrestricted, relatively speaking. As soon as one examines longer expressions and more complex structures, one is struck by the drastic reduction of tonotactic possibilities. Each rhythmic unit has exactly one tonic, that is to say, one prominent syllable, marked by its tone-bearing ability. This pattern emerges most clearly in trisyllabic compounds, but extends to phrasal constructions as well.

3.1 One-tone-per-unit constraint

Trisyllabic compounds may carry any one of the three tone patterns of (11); no other sandhi form is admissible as well-formed.

(11) Well-formed trisyllabic patterns
H.o.o
o.H.o
o.o.H

[10] The "five classics" refer to the five canonical books that used to form the core of traditional learning in China.

The patterns of (12) are barred because they contain two or more tones, while those of (13) are ruled out because the only admissible tone of a polysyllabic construction is H – if we disregard the boundary L, an intonational phenomenon.[11]

(12) *H.H.H
*L.H.H
*H.H.o
*o.HM.MH
*MH.o.H
*H.o.MH

(13) *MH.o.o
*o.L.o
*o.o.HM

Let us use the term "culminativity" to refer to the one-tone-per-unit constraint (cf. Hyman 1978, McCawley 1978, Hyman and Byarushengo 1984, Odden 1985, 1995, Goldsmith 1987). Postponing some precisions to be made later on, what we are saying in essence is that a polysyllabic word is organized around a single prominence peak. Only the prominent syllable is accentuated by a distinctive tone, while the recessive syllables are tonally suppressed, reduced to a non-distinct mid or low pitch, transcribed as [o] in our notation. Notice further that the accented syllable surfaces with a uniform H, regardless of the lexically specified tonal category it is originally associated with. We refer to this neutralization of underlying tonal contrasts as Leveling. We may capture these generalizations as follows.

(14) Within a trisyllabic or longer foot, the following conditions must be satisfied:
- **Culminativity**: One and only one tonic, i.e. tone-carrying syllable.
- **Tonicity**: Only accented syllables may carry tone; only tonics may be accented.
- **Leveling**: H is the only permissible tone.

It should be emphasized that none of these conditions hold in disyllabic or monosyllabic units. Thus both [HM.MH] and [H.H] are well-formed, despite their violation of Culminativity. Furthermore, assuming New Chongming to be a right-prominent language, other things being equal, the accent is on the rightmost tonic in [T.T]. This means that the syllable on the left is unaccented but tonic, in violation of Tonicity. Finally, [o.MH] and [MH] contrast with [o.H] and [H], which means that Leveling

[11] [o-o-H] of (11) is phonetically [o-o-HM]; see chapter 5, section 5 for discussion on boundary L and other intonational phenomena.

is inoperative in shorter rhythmic units. I attribute this restriction of tonal distribution to the contextual redundancy of tonal distinctions in polysyllabic strings (see below).

But why are the effects of culminativity and leveling most robust in trisyllabic or longer rhythmic units? To understand this one needs to realize that the forces that lead from a full-fledged tonal system to an inchoate accentual system are none other than those operating on connected speech in general. Specifically, connected speech is governed by two sets of fairly general processes that apply at all levels, both lexical and phrasal. The first is the tendency to limit the occurrence of tonic prominence to one per prosodic unit – a characteristic we have been referring to as Culminativity. The second involves a general process of "leveling" and various kinds of phonetic rules of "smoothing" and tonal coarticulation. Leveling is partial neutralization. With particular reference to New Chongming, by Leveling I mean the effects of register neutralization and/or simplification of a contour tone to a simple high-level tone. The motivation behind Leveling is likely to be functional in nature. All Chinese dialects make do with a small repertoire of permissible syllables.[12] As a consequence, monosyllabic words fully exploit tonal contrasts to minimize potential ambiguity. However, as syllables string together into longer expressions, context renders many of the tonal distinctions redundant.[13] Thus, 64 possible two-tone combinations (= 8^2, given the eight-tone system of New Chongming) are pared down to nine in disyllabic compounds (see chapter 5); by the same token, the 512 (= 8^3) logically possible three-tone sequences are even more drastically reduced to just three in lexical compounds: [H.o.o], [o.H.o], and [o.o.H]. Culminativity and Leveling reflect this inverse relation between length of expression and the informational load of tonal distinctions. It is worth noting that while Leveling is obligatory only within a stress-foot (trisyllabic or longer), it applies with fair regularity elsewhere, as we shall see.

[12] According to W. Chen (1953), the size of Chinese segmental syllabary (i.e. disregarding tones) goes from about 420 in Beijing Mandarin to 720 in Cantonese, compared to more than 100,000 *possible* (not necessarily attested) syllable types in English (figure based on Liberman, Juang, and Soong 1985). When tonal distinctions are fully exploited, the inventory in Mandarin Chinese increases to 1,390, which is less than 4 tones × 420 syllable types, since not all possible tone/syllable combinations actually occur.

[13] By one estimate based on a relatively small corpus of about 30,000 lexical entries in Mandarin Chinese, if we eliminate all tonal distinctions, we find 1,781 *monosyllabic* homophonous entries out of about 3,000 (that is a whopping 59%). The degree of tone-free homonymy decreases dramatically in disyllabic compounds (3,075 out of about 27,000 or just over 11%). See W. Chen (1953).

The remaining question is how to map the 512 (= 8^3) possible three-tone combinations onto one of the only three permissible patterns of (11). We need to be able to identify one of the three syllables as the most prominent in which the construction "culminates." Only this culminating syllable retains its tone, albeit uniformly neutralized to H; all others lose their lexically associated tones entirely. The toneless or atonic syllables are transcribed as [o] in our notation, and assume an intonationally determined default pitch (see section 1.1 above and section 9.1 for further phonetic details). The rule that singles out the most prominent syllable turns out to be surprisingly simple. As a first approximation, I will state the Accent rule informally as follows:

(15) **Accent Rule**

a. H.T.T
   *

b. T.T.E
      *

c. T.T.T
     *

* = accent
E = "even" tone, and includes: {H, Hq, L, Lq}

The Accent rule is interpreted as a set of ordered statements: (a) assign accent (*) to an initial syllable that carries H; (b) *otherwise*, assign accent to the final syllable carrying an even (i.e. level) tone; (c) *otherwise*, accent falls on the medial syllable. Cases (a, b, c) are illustrated by the examples in (16), (17), and (18), respectively. (As usual, square brackets indicate morphological constituency; "o" indicates a toneless syllable, and stress-feet are enclosed in parentheses (. . .).)

(16) H.T.T
     *

a.	[*gong cheng*] *shi*	"engineer"
	H L H	base tone
	(H o o)	sandhi form
b.	*qing* [*gong ye*]	"light industry"
	H H Lq	
	(H o o)	
c.	[*gao she*] *pao*	"anti-aircraft artillery"
	H LM MH	
	(H o o)	

(17) T.T.E
*

a. [*shui xian*] *hua* "narcissus flower"
HM H H
o o H

b. *chu* [*feng tou*] "be in the limelight"
Hq H L
(o o H)

c. *nao* [*chong xue*] "encephalemia"
ML H Hq
(o o Hq)

d. [*mai guo*] *zei* "traitor"
LM Hq Lq
(o o Hq)

(18) T.T.T
*

a. [*jin hua*] *lun* "evolutionism"
MH MH LM
(o H o)

b. *lao* [*qian bei*] "the elders"
ML L MH
(o H o)

Each of the examples in (16) carries a H tone on the initial syllable, and is therefore initially accented by virtue of clause (a) of the Accent rule. The examples of (17) end in an even tone (H, Hq, L, or Lq), and are end-stressed, in accordance with clause (b). The elsewhere cases of (c) are medially accented, by default as it were.

Statistically speaking, the Accent rule informally stated above comes close to a surprisingly accurate summary of the correspondences between the underlying tonal categories and their sandhi forms as far as trisyllabic compounds are concerned. This is borne out by the figures of table 6.5, which shows the distribution of the 108 trisyllabic compounds on our elicitation list.[14] We will return to some of the apparent exceptions in section 5.

[14] After discarding three examples which, unaccountably, split up into two prosodic units. One such example is [*yin yue*] *jia* "musician," underlyingly /H.Lq.H/, phonetically realized as [(H.Hq)(H)] instead of the expected (H.o.o).

Table 6.5. *Trisyllabic compounds (New Chongming)*

base tone	sandhi form		
	H.o.o	o.o.H	o.H.o
H.T.T	**20**	–	–
T.T.E	5	**42**	–
T.T.T	5	2	**34**

E = even tone: {H, L, Hq, Lq}
o = null, zero tone
Figures indicate type frequency

3.2 *Binarity condition*

Culminativity is not restricted to trisyllabic compounds, but extends to longer rhythmic units as well. In general, given *n* syllables, there are exactly *n* logical possibilities. The examples below illustrate the point.

(19) a. [[*gong-yi*]-*pin*]-*chang* "handicraft factory"
H LM HM HM base tone
(H o o o) sandhi form

b. [[*huang-ya*]-*cai*]-*tang* "bean-sprout soup"
L L MH H
(o H o o)

c. *hong*-[[*pu-tao*]-*jiu*] "red grape wine"
L L L HM
(o o H o)

d. *tie*-[[*luo-si*]-*ding*] "iron screw"
Hq L H H
(o o o H)

In contrast to (19), other quadrisyllabic and longer compounds may have one or more tonics. Often the same lexical word has two alternative readings, differing in terms of whether the prosodically weaker foot undergoes further (optional) stress reduction and concomitant tone loss, as illustrated in these examples:

(20) four sea make home
[*si hai*] [*wei jia*] "make one's home wherever one is"
MH.HM.L.H base tones
i. (MH.o) (L.H) alternative sandhi forms
ii. [(MH.o) o.o]

(21) fire on-top add oil
[*huo shang*] [*jia you*] "to incite, instigate"[15]
HM.ML H. L base tones
i. (MH.o) (H. H) alternative sandhi forms
ii. [o o (H. H)][16]

(. . .) = foot
[. . .] = p-word

I assume that (optional) stress reduction removes the foot structure, with the unfooted syllables being adjoined directly to the p-word. In other words, (20-ii) has the rhythmic [(MH.o) o o]. This accounts for the presence of MH, which is allowed only in citation forms or disyllabic compounds. If the defooted syllables were rebracketed with the foot on the left, the result would be an ill-formed *(MH.o.o.o), since all tonal distinctions are obliterated in polysyllabic feet, where the accented syllable uniformly carries a H (see Leveling, discussed in section 8 below).

The reason behind the contrasting behavior of (20)–(21) vs. (19) is quite simple. It is rooted in the Binarity condition, which can be stated quite simply as:

(22) **Foot Binarity**
A foot must be at least disyllabic.

Assuming the prosodic hierarchy, where the p-word dominates the stress-foot (cf. Selkirk 1980, 1981a, Nespor and Vogel 1986, Hayes 1989a), we derive the "minimal word" effect (cf. Prince 1980, McCarthy-Prince 1986, 1993a, Itô 1990, Yip 1991, Prince and Smolensky 1993). The Binarity condition is quite robust across a wide variety of Chinese dialects: a monosyllabic word or sublexical unit (which we will call a "**singleton**") is generally prosodically dependent, and must therefore adjoin a neighboring constituent to form a minimal rhythmic unit, the binary foot.[17] Contrast the behavior of *lang gou* "wolfhound" with that of *gou* "dog" in (23) in Beijing Mandarin.

[15] Literally, to add fuel to the flames.
[16] [o.o (H.H)] is further reduced to [o.o (o.H)], see section 8.
[17] Duanmu (1993a) argues that Chinese dialects fall into two types, M-languages (Mandarin-type) and S-languages (Shanghai-type): the former have bimoraic syllables, while the latter only have monomoraic syllables. To the extent that monosyllabic expressions display characteristic prosodic dependency in M- as well as S-languages, the relevant unit of count is syllable, not mora, for the purposes of ascertaining the minimality effect.

(23) a. wolfhound bite cat

[*lang gou*]	[*yao mao*]			"the wolfhound bit the cat"
2	3	3	1	base tone
(2	3)	(3	1)	footing
(2	3)	(3	1)	tone sandhi

b. dog bite cat

gou	[*yao mao*]		"the dog bit the cat"
3	3	1	
(3	3	1)	footing
(**<u>2</u>**	3	1)	attested
(3)	(3	1)	*

c. dog than cat big

gou	[[bi mao]	da]		"the dog is bigger than the cat"
3	3	1	4	
(3	3	1	4)	footing
(**<u>2</u>**	3	1	4)	attested
(3)	(3	1	4)	*

1, 2, 3, 4 = tonal categories, phonetically [H, MH, L, HM] respectively
* = ungrammatical
<u>2</u> = T2 derived from T3 via T3 Sandhi
foot = (. . .)

(24) T3 Sandhi (Mandarin)
T3 → T2 / ___ T3
condition: within a foot

In the case of (23a), the disyllabic expression *lang gou* "wolfhound" stands on its own as an independent binary foot / p-word, the minimal domain of obligatory tone sandhi. Therefore, tone sandhi rule (24)[18] leaves the base tones in (23a) unchanged. Now, *gou* "dog" in (23b) stands in exactly the same syntactic relation to the VP *yao mao* "bit the cat" as *lang gou*. But the monosyllabic *gou* cannot stand as an independent prosodic unit, but must join *yao mao* into a ternary foot, with the automatic consequence that tone sandhi obligatorily applies, turning a string of 3-3-1 tones into <u>2</u>-3-1. This obligatory stray monosyllable adjunction can in principle create a "super-foot" of indefinite length, as illustrated in (23c), where *da* "big" must join *bi mao* "than cat" into a ternary foot; in turn, *gou* "dog" must attach itself to *bi mao da* to form a quadrisyllabic foot, and hence falls within the obligatory tone sandhi domain.[19]

[18] For details of Mandarin tone sandhi, see chapter 9.
[19] In chapter 9 I will argue that the obligatory domain of tone sandhi in Beijing Mandarin is not the foot, but the "minimal rhythmic unit" (MRU). I will ignore this point here.

Assuming a prosodic structure that mirrors the morphological bracketing, it follows that in New Chongming, in quadrisyllabic words of the form [σ [σ σ σ]] and [[σ σ σ] σ] the monosyllabic constituent cannot stand alone, but must form a single minimal rhythmic unit together with the trisyllabic base. Therefore the entire rhythmic unit, in this case the quadrisyllabic expression, must satisfy the Culminativity condition. The question is: Where does the single tonic peak occur – on the trisyllabic base or the monosyllabic addition? It turns out that the answer is clear-cut: the singleton never carries tone, while the trisyllabic constituent always retains its prominence profile, so that the H always falls on one of the three possible positions within the trisyllabic base. Consequently, none of the tonal patterns of (25) is attested; instead, only those of (26) are well-formed. In other words, the polysyllabic base always retains whatever prominence profile it exhibits when occurring as a trisyllabic unit, while the monosyllabic addition never carries a tone nor is it accented. This generalization falls out from the general Weight-to-Accent principle to be discussed in section 5.

(25)	*[[H.o.o] T]	*[T [H.o.o]]
	*[[o.H.o] T]	*[T [o.H.o]]
	*[[o.o.H] T]	*[T [o.o.H]]
(26)	[[H.o.o] o]	[o [H.o.o]]
	[[o.H.o] o]	[o [o.H.o]]
	[[o.o.H] o]	[o [o.o.H]]

[. . .] = morphosyntactic constituent

Suffice it for our present purpose to illustrate one of the patterns of (26) with an example:

(27)		*xin* [[*lu yin*] *dai*]	"new cassette tape"[20]
		H Lq.H.MH	
	a.	(H) (o H o)	*(violates Foot Binarity)
	b.	(o. o H o)	attested

Note that the second constituent of (27a) exhibits an "amphibrach" pattern consistent with the Accent rule formulated earlier; therefore the sum of the two immediate constituents [(H) + (o.H.o)] is perfectly well-formed – except for the Binarity condition. Binarity, Culminativity, and the Weight-to-Accent principle jointly single out (27b) as the only possible reading.

[20] Adjective + noun constructions are treated as compounds by most students of Chinese. For diagnostics of wordhood, see Packard (1994, forthcoming), especially chapters contributed by John Dai and San Duanmu.

The "lopsided" compounds which display the internal structure of [σ σ σ + σ] or [σ + σ σ σ] (triplet + singleton in any linear order) like that of (27) are in the minority. The overwhelming majority of quadrisyllabic and longer compounds exhibit the "balanced" internal structure of a sequence of disyllabic pairs [σσ + σσ] , exemplified in (20)–(21) or a triplet plus a pair in any order [σσ] + [σσσ] / [σσσ + σσ] (see below). Unlike the monosyllabic singletons, which fall below the minimality threshold, a disyllabic pair (a fortiori a triplet or longer constituent) can form an independent foot.

3.3 Long compounds

In the corpus I have collected, all the pentasyllabic or longer lexical compounds form two or more feet, since all of them exhibit the internal structure of strings of two- and three-syllable constituents. In principle, such long compounds should comprise two or more rhythmic units, each with its own tonic syllable, as in:

(28)	a.	[*shang-hai*] [[*jing ju*] *tuan*]	"Shanghai opera troupe"[21]
		ML. HM H. Lq. L	base tone
		(o H) (H o o)	sandhi form
	b.	China people republic	
		[*zhong hua*] [*ren min*] [[*gong he*] *guo*]	"the People's Republic of China"
		H. L L. L LM.L.Hq	
		(H. H) (L. H) (o o Hq)	

However, it is often the case that they contain only one tonic peak. This is so because even though Culminativity is a well-formedness condition that is enforced consistently only at the foot level, it often prevails even in larger domains, depending on a number of factors, including speech tempo, focus, semantic transparency, and usage-based conventions. Some examples are given below.

(29)	a.	[*zui xiao*] [[*gong bei*] *shu*]	"the least common multiple"
		MH.HM H.ML.MH	
		(MH. o) (H. o. o)	
	→	[(MH. o) o o o]	Defooting
	b.	[*a.er.ba.ni.ya*]	"Albania"
		H.ML.H.L.MH	
		(H.H) (H. o. o)	
	→	[o o (H o o)]	Defooting

[21] The expected reading for *shanghai* in (28a) is [o.MH] rather than [o.H]. The attested sandhi form shows the effect of Leveling, discussed more fully in section 8.

(30)	[[*gong yi*] *pin*] [*zhan lan*]		"handicraft exhibition"
	H.LM.HM HM.LM		
a.	(H o o) (MH.o)		alternative sandhi forms
b.	[(H o o) o. o]		Optional Defooting

Notice that both (29a) and (29b) have only one tonic element, while (30) may optionally have one or two. As can be readily seen, the rhythmic groupings are by and large isomorphic with the morphosyntactic constituency. In the absence of a structural determinant, as in foreign names like *a.er.ba.ni.ya* "Albania" (29b), rhythmic grouping is binary from left to right. The last monosyllable in odd number syllable compounds joins the last binary foot to form a ternary unit.

3.4 Postlexical structures

Culminativity, or its inverse, tonal reduction of prosodically less prominent constituents, goes beyond lexical compounds. As noted above, the foot binarity condition holds true of Chinese dialects in general; however there are certain dialect-specific variations. In Beijing Mandarin, foot binarity is enforced regardless of morphosyntactic status. Thus *gou* in (23b, c) must join the syllables to the right into a single rhythmic unit despite the deep syntactic cut between the subject and the predicate. On the other hand, in most Wu dialects, including New Chongming, the foot binarity condition is restricted to lexical compounds. That is to say, monosyllabic *sublexical units* are strictly forbidden as independent minimal rhythmic units, though monosyllabic *syntactic words* freely occur. To put it another way, footing in New Chongming is word-bound: separate morphosyntactic words do not in principle combine into a single foot.[22] As a consequence, a monosyllabic word constitutes, by necessity a degenerate foot, despite Foot Binarity.[23] The word-bound footing is illustrated by the following examples:

(31)	a.	*si* [*mian-pen*]	"four washbowls"
		MH LM L	base tone
		(MH) (o H)	sandhi form

[22] An analog is syllabification: it is word-bound in English, thus we have *Brit.rail* (for British Railway) vs. *be.tray*. In contrast, syllabification in French cuts across word boundaries. As a consequence, we have [*trã.tã*] for *trente ans* "thirty years" and [*bɔ̃ .na.mi*] for *bon ami* "good friend," and so forth.

[23] This means that the word-boundedness constraint outranks foot binarity.

b. sweep floor go
[*sao di*] *qu* "go sweep the floor"
MH.LM.MH
(MH o) (H)[24]

Phrasal expressions (a) and (b) of (31) form two separate feet each, with *si* and *qu* constituting a monosyllabic foot. In contrast to phrasal constructions, such a bipartite rhythmic organization is impossible with lexical compounds, as shown below:

(32) *fu* [*zhu xi*] "vice-chair"
MH.HM.Lq
a. (MH)(o.Hq) *
b. (o o. Hq) attested

(33) [*ji nian*] *hui* "memorial service"
MH LM LM
a. (MH.o)(MH) *
b. (H o o) attested[25]

(33) is tonally and structurally identical to (31b) in all relevant respects, except that instead of a phrase, it is and behaves like a single word or lexical unit. As a consequence, the two-tonic reading [(MH.o)(MH)] is disallowed (= (33a)). The only attested reading culminates in a single tonic (H.o o) (= (33b)), consistent with Foot Binarity.

It is worth noting that tonal reduction of prosodically less salient constituents is making inroads into phrasal constructions. Thus side by side with (34a) we have (34b).

(34) a. work one day
zuo [*yi tian*] "do one day's work"
MH.Hq.H
(H) (o H)

b. write one CL
xie [*yi shou*] "write one piece" (of poem or song)
HM.Hq.HM
(MH) (o.MH) expected sandhi form
→ (o o H) attested sandhi form

[24] The expected sandhi form of *qu* "go" [MH] is leveled to a simple [H]. The same leveling effect is discernible in many examples to follow. See section 8 for discussion on Leveling.

[25] Consistent with the Accent Rule, under the elsewhere clause (c), we expect *ji nian hui* "memorial service" with the tonal string MH.LM.LM to be paroxytonic (o.H.o), with the accent on the penult. Instead, what we have is a proparoxytonic pattern (H.o.o). The Accent Rule will be fine-tuned below in section 5.

In (34a), the verb *zuo* "work" constitutes a separate foot of its own, given the word-bounded nature of foot formation. On the other hand, in (34b) if *xie* "write" and the object NP *yi shou* "one piece" were to form separate feet, the expected sandhi form would be [(MH)(o.MH)], consistent with the disyllabic sandhi rules developed in chapter 5. Instead, the attested sandhi form indicates that monosyllabic verb *xie* is tonally reduced, and refooted with the binary foot on the right, yielding a trisyllabic foot, with an oxytonic pattern.

Tonal reduction affects only prosodically weak feet. How do we ascertain the relative prominence between two feet, or more generally between two rhythmic constituents? This problem is quite complicated in connected speech, much of which depends on intonation and phrase-level prominence that is only partially predictable (see section 4.2). With this caveat in mind, we can nevertheless isolate two principal determining factors, namely relative tonal saliency and prosodic weight, to be made more explicit in the immediately following sections.

4 Saliency and Edgemostness

4.1 Trisyllabic compounds

The Accent rule informally stated as (15) in section 3.1 comes remarkably close to a descriptively adequate statement: as shown by table 6.3, it makes the correct prediction for nearly 90% of the subcorpus (96 out of 108 trisyllabic compounds). But what is the intuitive content of the Accent rule? First, Accent singles out the initial H. That is hardly surprising, since H ranks highest on the scale of tonal saliency established in section 3 of chapter 5.[26] All the other options rank lower as the ideal docking site for the accent. On the other hand, why does Accent accord the final E a special status? Recall that E (even tone) includes any of the following: H, L, Hq, and Lq; this means that any word-final even tone, regardless of register and host syllable type, attracts accent – unless preempted by an initial H, in accordance with the Accent rule, clause (a) – as shown in the examples of (17a–d). In contrast, neither L nor Hq or Lq in the initial position attracts accent, as illustrated in these examples:

[26] Du (1988) found that her Taiwanese speaking subjects perceive a high-toned syllable as "louder" than a low-toned syllable. Pankratz and Pike (1967:293f.) propose a set of similar tone-sensitive stress assignment rules for Ayutla Mixtec.

(35)	a.	[*wen gong*] *tuan* L H L (o o H)	"song and dance ensemble" base tone sandhi form
	b.	*hei* [*tou fa*] Hq L Hq (o o Hq)	"black hair"
	c.	[*lü dou*] *ya* Lq. LM. L (o. o H)	"green bean sprout"
	d.	[*shuo fu*] *li* Hq Lq Lq (o o Hq)	"persuasive power"
	e.	[*gong cheng*] *shi* H L H (H o o)	"engineer"

The contrast between (a, b) and (e) is particularly striking. (e) instantiates a case of H.T.H sequence; the leftmost H wins out as the prominence peak. On the other hand, both (a) L.T.L and (b) Hq.T.Hq exemplify an E.T.E pattern; but it is the *rightmost* E that attracts accent. Why such an asymmetry? Actually, the asymmetry is only apparent. Recall that the tonotactics of New Chongming precludes an initial checked syllable from carrying a tone. We labeled this constraint as *Tq in chapter 5. This explains why accent does not fall on an *initial* Hq. On the other hand, nothing prevents a *final* checked syllable from bearing tone. Furthermore, register contrast of even tones is neutralized in favor of the high anywhere except in the initial position. This means that any final E is redundantly high, consistent with a constraint referred to as *E↓ which prohibits low-level tones except in initial positions (see chapter 5 for details). By subsuming constraints *Tq and *E↓, the picture that emerges is this: (i) the high-level tone H (and Hq where permitted, i.e. in the final position) attracts accent; (ii) accents tend to be "demarcative," i.e. signaling the beginning or the end of a phonological unit (Hyman 1977). The overwhelming majority of stress systems are "head-terminal" – i.e. with the stress on the first or the last syllable of a rhythmic unit (cf. Halle and Vergnaud 1987, Hayes 1995). With reference to Chinese, Yan and Lin (1988) observe that trisyllabic expressions overwhelmingly favor the [S W S] pattern over [W S W].[27] "Amphibrach" patterns, [o.H.o] with a

[27] Goldsmith (1988:91) makes a similar observation regarding the prominence of the peripheral vowels over the internal ones.

medial accent, represent the residue, where no H appears at the margins. We can express these generalizations in terms of alignment constraints:[28]

(36) AlignL Align accent with the leftmost H.
AlignR Align accent with the rightmost H.

Alignment is subordinate to Culminativity and other constraints established in chapter 5, which I repeat below:

(37) Culminativity One tonic per trisyllabic or longer foot.
*Tq Initial checked syllable may not carry a tone.
*E↓ No low-register even tone except in initial position.

The ranking {Culminativity, *Tq, *E↓} ≫ AlignL ≫ AlignR gives us the desired effects. Consider the simplest cases of (38) and (39).

(38) /H.L.H/ [*gong-cheng*]-*shi* "engineer"

			Culm	*Tq	*E↓	AlignL	AlignR
a	☞	H.o.o					*
b		o.o.H				*	
c		H.o.H	*				
d		o.H.o				*	*

(39) /Hq.H.L/ *chu*-[*feng-tou*] "be in the limelight"

			Culm	*Tq	*E↓	AlignL	AlignR
a		Hq.o.o		*			*
b		o.o.L			*	*	
c	☞	o.o.H				*	
d		o.H.o				*	*
e		Hq.o.H	*	*			
f		o.H.H	*				

In the former case (38), since there are two Hs, the accent could in principle be aligned with either one. The fact that candidate (a) prevails over (b) indicates that AlignL dominates AlignR. Candidate (c) satisfies both AlignL and AlignR, but runs afoul of the overriding constraint Culminativity. Finally, output (d) violates both AlignL and AlignR, and is therefore rejected in favor of alternative (a). Subsumed in the tableaux are Tonicity and Leveling. Tonicity, stated earlier in section 3.1, says that only accented syllables may carry a tone, while Leveling, also formulated there stipulates

[28] For a general theory of "alignment," see McCarthy and Prince (1993b).

that the only permissible tone is H. This means that once we have located the accent, the tonal envelope of the rhythmic unit is entirely predictable.

In the latter case of (39), AlignL picks out output (a), which however is preempted by the undominated *Tq barring initial checked syllables from bearing tone. On the other hand, AlignR favors (b, c); (b) is eliminated by *E↓. Candidates (d, e, f) violate either alignment or culminativity conditions, leaving (c) as the winning candidate.

The reader can easily construct similar tableaux to verify that the hierarchy of constraints unerringly places the accent in the other examples of (16)–(18) and (35).

4.2 *Phrases and long compounds*

Culminativity and Edgemostness, expressed as alignment, extend beyond the trisyllabic compounds. As we noted earlier in section 3.3, postlexical structures are following the footsteps of polysyllabic compounds. Thus (40a) and (40b) coexist side by side, reproduced here for convenience.

(40) a. work one day
zuo [*yi tian*] "do one day's work"
MH Hq H
[(H) (o H)]

b. write one CL
xie [*yi shou*] "write a piece" (of poem or song)
ML ML MH
[o (o H)]

The pattern of tonal reduction is the by-product of the same principles at work in trisyllabic compounds. Tonal saliency encoded in AlignL/R goes a long way toward explaining the patterns of tonal reduction. For instance, the 46 instances of verb + number-measure constructions exemplified in (40a, b) above show the following distribution:

(41) verb + number-measure

sandhi forms	base forms	
	H + [XX]	others
a. [o (XX)]	1	31
b. [(X) o.o]	9	5

(XX), (X) = disyllabic and monosyllabic feet[29]
o = atonic

[29] (XX) is neutral between (T.o), (o.T) and (T.T). (X) represents, of course, simply (T).

Culminativity and AlignL pick (b) as the optimal output; elsewhere it is the heavy constituent, namely the disyllabic substring, that is prominent, and therefore retains the tonic peak (see section 5 on prosodic weight). The effect of AlignL/R is evident in other constructions like number-number + noun, illustrated by (42).

(42) a. five six CL
[*wu liu*] *zhi* "five or six (items)"
ML Lq Hq base tone
[o o (Hq)] by AlignR

b. three four day
[*san si*] *tian* "three or four days"
H MH H
[(H. o) o] by AlignL

c. two three night
[*liang san*] *ye* "two or three nights"
ML H LM
[(o H) o] by Weight (see section 5 below)

Culminativity and Edgemostness (AlignL/R) are not confined to trisyllabic expressions. Longer compounds with internal structure of [[σ σ][σ σ]] may form two separate feet. The weaker foot often, though not obligatorily, undergoes tonal reduction. We have seen examples of this optional tonal reduction in section 3.3. In this respect New Chongming resembles Shanghai (cf. Duanmu 1993a), to be discussed more fully in chapter 7. (43) has two alternative readings (a) and (b) in Shanghai.

(43) a. x
(x .) (x .)
[*qo yã*] [*tço.zɣ*] "professor Ou-yang"
(HL.o) (MH.o) tone deletion
(H. L) (M. H) tone spread

b. x
(x .) (x .)
[*qo yã*] [*tço.zɣ*] "professor Ou-yang"

(x . . .) optional stress reduction
→ [*qu yã*] [*tço.zɣ*]
(HL.o o o) tone deletion
(H. L o o) tone spread

Reading (b) is derived through the optional rule of stress reduction, which eliminates the less prominent of the two feet.

However, there is one fundamental difference between Shanghai and New Chongming. In the former, rhythmic structures are *uniformly* left-prominent at the word and compound level. This means that only non-initial constituents may undergo stress reduction and concomitant tone loss, consistent with the Tonicity constraint stated in (14). The picture in New Chongming is far less predictable. On the assumption of prominence-sensitive tonal reduction, we have the following contrasting prominence configurations for (44) and (45):

```
(44)        x
           (x   .)   (.  x)
           [si hai] [wei jia]          "make one's home wherever one is"
           four sea make home
             MH.HM.L.H                 base tone
      a.   [(MH.o) (L.H)]              alternative sandhi forms
      b.   [(MH.o)  o. o]

(45)                          x
             (x    .)      (.   x)
           [huo shang] [jia you]       "add fuel to fire; to fan the flame"
           fire  on-top  add fuel
           HM.ML       H.L             base tones
      a.   [(MH.o)     (H.H)]          no stress reduction
      b.     [o   o    (H.H)][30]      via stress reduction
```

Given the tone-prominence connection expressed as Tonicity, within a prosodic unit that is trisyllabic or longer only syllables in stressed position may carry tone. This means that the rhythmic unit (MH.o) must be left-headed in (44) and (45). What about two-tonic units, as instantiated by (L.H) and (H.H) in the same aforementioned examples? Recall from chapter 5 that ceteris paribus, the last syllable of a disyllabic compound tends to retain its tonal identity rather than the first. We expressed this positional prominence as Prs-Tp (parse the prominent tone Tp, i.e. the tone in the rightmost position) and Prs-Cp (preserve the prominent melodic contour Cp, i.e. the tone in final position – see chapter 5, section 3.3). For this reason I will assume that a two-tonic foot is right-headed. The problem with (44) and (45) concerns the higher level prominence: at the p-word level, (44) is left-prominent, but (45) is right-prominent. In terms of the p-word level right- or left-headedness, the set of 188 [[σ σ] [σ σ]] lexical compounds in our database shows the distribution seen in table 6.6.

[30] [o.o (H.H)] is further reduced to [o.o (o.H)], see below.

Table 6.6.

Type			
A.	[(XX) o.o]		41
A′.	[(XX) o.o]	or [(XX) (XX)]	24
B.	[o.o (XX)]		34
B′.	[o.o (XX)]	or [(XX) (XX)]	10
C.		[(XX) (XX)]	79
Total			188

(. . .) foot unit
(XX) tonic pair = (T.o), (o.T), or (T.T)
o atonic syllable

In this table I use (XX) to symbolize a **tonic pair**, i.e. a disyllabic foot containing one or more tonic elements. In other words, (XX) is neutral between (T.o), (o.T), and (T.T). Type A and B represent the left- and right-prominent patterns respectively. Both A and B have a subtype A′ and B′ with an alternative reading [(XX)(XX)]. Type C only has the latter reading. The tonic pair carries whatever sandhi pattern is appropriate – (T.o), (o.T), or (T.T) – in accordance with the principles developed in chapter 5. The toneless syllables take a default L, as described in chapter 5, section 1. We can illustrate each type with an example.

(46) ancient Han language[31]

A. [*gu dai*] [*han yu*] "Ancient Chinese"
HM.LM.MH.HM base tone
[(MH.o) o.o] sandhi form(s)

A′. four sea make home
[*si hai*] [*wei jia*] "make one's home wherever one is"
MH.HM.L.H
[(MH.o) (o.H)]
[(MH.o) o. o]

B. water subside rock emerge
[*shui luo*] [*shi chu*] "get to the bottom of the matter"
HM Lq Lq.Hq
[o o (o.Hq)]

[31] "Han" refers to the ethnic Chinese; while "Chinese" includes other ethnic minorities within the political confines of China.

Table 6.7.

	A and A′ [(XX) o.o] and [(XX) o.o] or [(XX)(XX)]	B and B′ [o.o (XX)] and [o.o (XX)] or [(XX)(XX)]
a. E-Pair + E-Pair	4	**20**
b. O-Pair + E-Pair	11	**17**
c. E-Pair + O-Pair	**33**	4
d. O-Pair + O-Pair	**17**	3

(XX) tonic pair = (T.o), (o.T) or (T.T)
o atonic syllable
E-pair even-toned pair, containing an even tone {H, L, Hq, Lq}
O-pair oblique-toned pair, containing an oblique tone {MH, HM}

B′. mess seven eight disaster
[*luan qi*] [*ba zao*] "in a total mess"[32]
LM.Hq Hq.H
[(o.Hq) (o.H)]
[o. o (o.H)]

C. uncommon strange
[*xi qi*] [*gu guai*] "weird"
H.L HM.MH
[(H.H)(HM.MH)]

Putting aside the C type, where no tonal reduction occurs – hence giving no clue to the relative prominence at the compound or p-word level – we have two basic types: the left-prominent A/A′ and the right-prominent B/B′. What then conditions the split of quadrisyllabic compounds along this division line? Unfortunately, hard and fast rules are not to be found; however, certain tendencies can be discerned from the distribution summarized in table 6.7. By and large, the tonal category of the second pair determines where the main stress falls: if the second pair carries an even tone, it attracts the primary stress (cases (a) and (b)), otherwise the word-level prominence peak falls on the left (cases (c) and (d)). In other words, given prosodic constituents [$C_1 + C_2$], C_2 is metrically strong if and only if C_2 carries an even tone. Table 6.7 once more bears out the inherent saliency of even tones.

Examples (46A, B, B′) cited above all conform to this tendency. For instance, the right foot of both (B) and (B′) carries an even tone – Hq and

[32] An idiom analogous to the English expression "at sixes and sevens."

H respectively – both (B) and (B′) are right-prominent, so that tonal reduction affects the recessive foot on the left. By the same token, since the second foot of (A) is oblique-toned, the p-word peaks on the left, leaving the right foot vulnerable to tonal reduction. On the other hand, (46A′) constitutes an exception: *wei jia* "make home" is an even-toned pair (L.H), nevertheless, the primary accent falls on the left. The indeterminacy of the right vs. left prominence at the p-word level and beyond becomes obvious when one observes that occasionally the same lexical or idiomatic expression may have alternative realizations as exemplified by the following doublet (repeated for convenience):

(47) [*huo shang*] [*jia you*] "add fuel to fire; to fan the flame"
HM.ML H. L
a. [(MH.o) (H.H)]
b. [o o (H.H)]
c. [(H. o) o o][33]

In addition to the two alternative sandhi forms (a) and (b) cited earlier, this example has a third reading (c) suggesting a left-prominent structure.

The joint effect of Edgemostness and Culminativity shows up in other interesting ways. Recall that a disyllabic foot may have any one of the three tonal patterns (T.o), (o.T), or (T.T). When one or the other foot is prosodically demoted with concomitant tonal reduction, we expect to find any one of the patterns of (48).

(48) A. [(T.o) o.o] [o.o (T.o)]
B. [(o.T) o.o] [o.o (o.T)]
C. [(T.T) o.o] [o.o (T.T)]

C′ [(T.o) o.o] [o.o (o.T)]

Surprisingly, pattern C is systematically absent in our corpus. Instead, what we find is that all but the leftmost or the rightmost tone disappear. So, corresponding to the expected C, we find C′ instead. The following examples illustrate this point.

(49) a. thousand odd hundred strange
[*qian qi*] [*bai guai*] "all sorts of strange things"
H L Hq MH base tone
[(H.H) (o.MH)] sandhi form with two rhythmic units
[(H.H) o o] stress/tonal reduction; expected sandhi form
[(H.o) o o] attested

[33] The expected form is (MH.o)(o.o), which is further reduced to (H.o)(o.o) by Leveling. See section 8.

b. blood mouth spray people
[*xue kou*] [*pen ren*] "to venomously slander someone"
Hq HM H L
[(o MH) (H H)]
[o o (H H)]
[o o (o H)]

c. fire on-top add oil
[*huo shang*] [*jia you*] "to incite, instigate" (= (21))
HM.ML H. L
[(MH.o) (H.H)]
[o o (H.H)]
[o o (o H)]

These examples demonstrate two strong tendencies: (i) not only each foot, but each p-word has one and only one tonic element (= Culminativity); and (ii) the tonic element is located either at the left or the right edge (= Edgemostness).

5 Prosodic weight and recursive constraint satisfaction

The Accent placement rule restated as alignment is repeated here for convenience:

(50) AlignL Align accent with the leftmost H.
AlignR Align accent with the rightmost H.

What if a trisyllabic compound neither starts nor ends in a H? In the majority of cases, the accent falls on the medial syllable. Conceivably we may cover the leftover cases by a rule such as AlignM:

(51) AlignM Align the accent with any tone in the medial position.

There are several reasons to reject this analysis. Firstly, a priori, rhythmic units are overwhelmingly head-terminal (i.e. left- or right-prominent), rarely head-internal (Halle and Vergnaud 1987).[34] At the word-level, main stress typically falls on the leftmost or the rightmost foot, depending on the directionality of foot building, but rarely on a word-medial foot (cf. Hammond 1985).[35] One would therefore expect accentual organization to

[34] According to Halle and Vergnaud's (1987:10–11) typology, a foot may be binary, ternary, or unbounded. Binary and unbounded feet are head-terminal only; ternary feet are exclusively head-internal. If we eliminate ternary and unbounded feet from metrical primitives (following Prince 1985; cf. Prince and Smolensky 1993:38, Hayes 1995), then all metrical units are by definition head-terminal. However, see Halle and Idsardi (1995) for discussion on foot typology.

[35] Cf. Kager (1995:375) for caveats and counterexamples.

mirror this universal tendency. This also accords with the notion that accent/stress serves the demarcative function of signaling the margins of a prosodic unit (cf. Trubetzkoy 1969, Hyman 1977, Beckman 1986). One predicts therefore that, ceteris paribus, accent occurs peripherally rather than medially. The ceteris paribus proviso could, for instance, allow the saliency principle ("accent the most salient tone") to override the head-terminal principle. This means that the only tonal configurations that call for a medial accent would be /A.B.C/, where the medial tone B is more salient than both A and C – assuming the saliency scale (see chapter 5, section 3.3 for details):[36]

(52) Tonal saliency

E > Hr > T

x > y = x is more salient than y
E = even tone {H, L, Hq, Lq}
Hr = high-registered tone
T = other tones

One such configuration is /O.E.O/ (where O = oblique, or contour tones). Indeed, all instances of underlying /O.E.O/ configurations emerge as [o.H.o], with a medial accent, as in:

(53) a. [*qi-che*] *piao* "bus ticket"
MH H MH
o H o

b. [*hai-nan*] *dao* "Hainan island (place name)"
HM L HM
o H o

But no matter how we might fine-tune or modify the saliency scale, there is no way to predict the head-internal pattern of these examples:

(54) a. [*za-huo*] *dian* "grocery store"
Lq MH MH
o H o

b. [*zhao-xiang*] *guan* "photo studio"
MH MH HM
o H o

[36] Tone-dependent stressability hierarchy has been proposed in various forms by Pankratz and Pike (1967), Hoa (1983), Du (1988), Meredith (1990).

Constraint *Tq (initial checked syllables may not carry tone) eliminates the first syllable in (54a) as a potential host for the accent, since the accented syllable must carry the H tone, consistent with Tonicity. This leaves two identical tones [o.MH.MH]. In accordance with the ceteris paribus clause, we would expect the head-terminal principle to place the accent on the final syllable, contrary to the attested reading [o.H.o]. (54b) creates a similar problem. For the sake of argument, let us suppose that the constraint *Fall (no final falling tone, see section 2.1 of chapter 5) precludes [HM] from surfacing on the final syllable. That leaves us with the same sequence of like tones /MH.MH.o/. The same head-terminal principle would again wrongly locate accent on the initial syllable.

Secondly, an accent placement rule that subsumes AlignM suggests that polysyllabic accentuation is governed by a set of principles that are by and large distinct and unrelated to those determining the sandhi forms of disyllabic compounds.

Finally, AlignM cannot be extended to compounds longer than three syllables, since it underdetermines which of the several medial syllables is to be accented.

5.1 Weight-to-Accent

We can obviate all the problems noted above by recourse to a generalized principle of prosodic weight. The basic idea is that syllables are arrayed on a weighted scale of light, heavy, and superheavy, and a heavier syllable is more likely to attract accent than a "lighter" one. Thus a language may stress the rightmost (or leftmost) heaviest syllable. Abstracting away from extrametricality and specified stress domains (e.g. stress may not fall farther away than three syllables from the edge), this is basically the stress rule attested in such disparate languages as Cairene Arabic (McCarthy 1979), Pirahã (Everett 1988), and Hindi (Hayes 1995). We can extend this Weight-to-Stress Principle (WSP; Prince 1990) to a higher level of metrical organization. Just as syllabic weight ranges from one to three moras, feet are ranked from monosyllabic to binary to *n*-ary units in an ascending order of prosodic weight. A binary/*n*-ary foot is more likely to attract word-level stress than a degenerate, monosyllabic foot. One manifestation of this extended principle is the well-known "unmarked labeling convention," which says to the effect that given nodes N_1 and N_2, N_2 is strong if and only if it branches (i.e. is binary, Liberman and Prince 1977; cf. Hayes 1982). This accounts for the alternating stress

between (*cóm.pen*)(*sàte*) and (*còm.pen*)(*sá.tion*). In the latter case, the rightmost foot is disyllabic or branching, therefore it attracts the main stress. In the former case, on the other hand, the rightmost foot is monosyllabic or non-branching; as a consequence, the main stress docks on the foot farther to the left. This branching or weight-sensitive prominence holds on a higher prosodic level as well. Recall that Liberman and Prince (1977) pointed out the contrast between [[*law degree*] [*LANGUAGE requirement*]], with main stress on the "heavy" branching constituent on the right and [[*LAW degree*] *requirement*] *changes*], with primary accent on the left.

There is some evidence in support of extending WSP to supra-syllabic constituents in Chinese. Years ago, Lü (1963) discovered a robust but curious asymmetry in **Mandarin**: modifier + nominal head constructions (Mod + N) tend to have the internal structure [σ σ + σ], while verb + object NP expressions (V + Obj) predominantly display the mirror image structure [σ + σ σ]. In other words:

(55)

	$[\text{Mod} + \text{N}]_{\text{N}'}$	$[\text{V} + \text{Obj}]_{\text{V}'}$
[σ σ + σ]	✔	
[σ + σ σ]		✔

Here the symbol ✔ stands for "the norm" or "the typical". Typical and atypical examples follow:

(56) Mod + N

typical:	animal science	
	[*dong wu*] *xue*	"zoology"
atypical:	hand organ	
	shou [*feng qin*]	"accordion"

(57) V + Obj

typical:	write essay	
	xie [*wen zhang*]	"to write an essay"
atypical:	waste money	
	[*zao ta*] *qian*	"to waste money"

This asymmetry is so strong that speakers superimpose the two rhythmic templates [σ σ + σ] and [σ + σ σ] on elliptic expressions, and interpret them accordingly. Thus, a quadrisyllabic expression [*fu yin*][*wen jian*] is often shortened, but with different meanings:

(58) duplicate document
[*fu yin*] [*wen jian*]
→ a. [*fu yin*] *jian* "a duplicate document" (Mod + N)
→ b. *yin* [*wen jian*] "to copy a document" (V + Obj)

Where dual interpretations are inappropriate or not available, ellipsis is strictly governed by the principle implicit in the asymmetry. Thus only the (a) versions of the truncated forms are felicitous in the following examples, taken from W.Wu (1986) and Duanmu and Lu (1990), where more can be found:

(59) skill worker
[*ji shu*] [*gong ren*] "skilled labor"
→ a. [*ji shu*] *gong* ok
b. *ji* [*gong ren*] *

(60) repair road
[*xiu jian*] [*ma lu*] "to repair roads"
→ a. *xiu* [*ma lu*] ok
b. [*xiu jian*] *lu* *

How do we explain this asymmetry? Duanmu and Lu (1990) make a very insightful suggestion. They propose that Beijing Mandarin is governed by what they call **Non-Head Stress**. Given $[XP, Y]_{Y'}$ (linear order irrelevant), stress falls on the non-head XP, regardless of whether XP functions as an argument (e.g. object NP), or an adjunct (e.g. a prenominal modifier). Add to this the preference rule that says in effect that stress falls preferably on a longer constituent (in terms of syllable count). We may restate this eurhythmic principle as a variant of WSP (Prince 1990) or Peak-Prominence (Prince and Smolensky 1993):

(61) **Weight-to-Accent**
Accent(x) } Accent(y) if x > y.
x } y x is more harmonic than y
x > y x is heavier than y (in mora, or syllable count)

The observed asymmetry follows directly from the Non-Head Stress and Weight-to-Accent principles, as demonstrated in the familiar tableau form. In the following tableau the accented constituent is capitalized:

(62)

					NHS	Weight
1	[*ji shu*] [*gong ren*]	☞	a	[*JI SHU*] *gong*		
	"skilled labor"		b	[*ji shu*] *GONG*	*	*
			c	*JI* [*gong ren*]		*
			d	*ji* [*GONG REN*]	*	
2	[*xiu jian*] [*ma lu*]		a	[*XIU JIAN*] *lu*	*	
			b	[*xiu jian*] *LU*		*
			c	*XIU* [*ma lu*]	*	*
		☞	d	*xiu* [*MA LU*]		

NSH = Non-Head Stress
Weight = Weight-to-Accent principle
Small caps indicate stress

It can be seen that only (1a) and (2d) are consistent with both Non-Head Stress and Weight-to-Accent. All other truncated forms violate one or both of these constraints.

The preceding account crucially relies on the generalized Weight-to-Accent principle, where prosodic weight not only calibrates syllables in terms of mora count, but also measures larger constituents in terms of number of syllables.[37]

Now let us return to the New Chongming case. Having motivated the Weight-to-Accent Principle, we can immediately explain why the medial syllable in an example like (6b) [*za huo*] *dian* "grocery store" attracts accent: since AlignL and AlignR leave the accent placement underdetermined, Weight-to-accent places the accent on the left-branching constituent, namely on [Lq-MH]. Since an initial checked syllable (Lq) is barred from hosting a tone (by virtue of *Tq), while an accented syllable must carry a tone (enforced by Tonicity), the medial syllable wins out as the accent site.

The Weight-to-Accent Account we have given of (54a) cannot extend to (54b) [*zhao xiang*] *guan* "photo studio" in a straightforward manner. The crucial difference is that while *Tq (no initial CVq may carry tone) automatically rules out [Hq.o.o] as a phonetic realization of (54a), nothing precludes [H.o.o] as the phonetic output of (54b). To make the matter worse, given the choice between [H.o.o] and [o.H.o] for (54b), the head-terminal principle that is the intuitive underpinning of AlignL and AlignR would prefer the former over the latter.

[37] A similar idea can be found in Tsay (1994).

The most plausible account is to assume that output constraints are satisfied recursively – an assumption that is in principle compatible with OT (cf. Goodman 1994, Kenstowicz 1995) and consistent with the interleaving of phonology and morphology that lies at the heart of Lexical Phonology (Kiparsky 1982b, c, Mohanan 1982, inter alia). Notice that *zhao xiang* /MH.MH/ "photo" of (54b) occurs as an independent word with the tonal pattern [o.MH]. Under the assumption of cyclic constraint satisfaction, *zhao xiang* emerges as [o.MH] before it joins *guan* /HM/ "studio" into a larger unit [*zhao-xiang*] *guan* "photo studio." Consequently, the initial syllable is automatically precluded from being accented by virtue of Tonicity. In other words, the relevant input relative to which the output candidates are evaluated is not /MH.MH.HM/ but /[o.MH] + HM/:

(63)

[*zhao xiang*] *guan*		"photo studio"
MH.MH.HM		base tone
[o.MH] + HM		relevant input

			Tonicity	AlignL/R	Weight
a	☞	o.H.o			
b		ó.o.o	*		
c		o.o.H			*

ó = accented but toneless syllable

Candidate (b), with an accented but atonic initial syllable (symbolized as "ó") is patently ill-formed, being in violation of the Tonicity constraint (only tonics may be stressed), while candidate (c) runs afoul of the Weight-to-Accent constraint. The remaining candidate (a) wins. Example (54a) [*za-huo*]-*dian* "grocery store" lends itself to the same treatment. It would have as input on the next cycle the representation /[o.MH] + MH/. The same constraints Tonicity and Weight-to-Accent would pick the correct output in exactly the same manner as shown in (64).

Recursive constraint satisfaction is consistent with all the examples cited above. Take initially accented compounds. The common denominator here is that all of them have the pattern /[T.o] + T/ as the input at the trisyllabic level, and in each case alignment underdetermines the accent placement. Consequently, the lower ranked Weight-to-Accent assumes the decisive role. This is illustrated below.[38]

[38] For discussion on recursive constraint satisfaction from the perspective of OT, see Kenstowicz (1996) and Duanmu (1997).

(64) [*zong tong*] *fu* "presidential palace"
HM.HM.HM base tone
[MH.o] + HM relevant input

			Tonicity	AlignL/R	Weight
a	☞	H.o.o			
b		o.ó.o	*		
c		o.o.H			*

5.2 *Superheavy constituents*

We need to fine-tune the Weight-to-Accent Principle. Examples like these clearly demonstrate that AlignL/R dominates Weight-to-Accent. In both cases AlignL/R assigns the accent to the lighter monosyllabic singleton rather than the heavier disyllabic base.[39]

(65) *qing* [*gong ye*] "light industry"
/H [H.Lq]/ base tone
/H + [H.Hq]/ relevant input

			AlignL	AlignR	Weight-to-Accent
a	☞	H.o.o		*	*
b		o.o.Hq	*		
c		o.H.o	*	*	

(66) [*shui xian*] *hua* "narcissus flower"
/[HM.H] H/ base tone
/[o.H] + H/ relevant input

			AlignL	AlignR	Weight-to-Accent
a	☞	o.o.H			*
b		o.H.o		*	

Now consider these longer expressions:

(67) a. *xin* [[*lu-yin*] *dai*] "new cassette tape"
H Lq H MH base tone
(o H) disyllabic sandhi form
(o H o) trisyllabic sandhi form
(o o H o) quadrisyllabic sandhi form

[39] The disyllabic base need not be the morphosyntactic head: thus *shui-xian* "narcissus" functions as the adnominal modifier of the head *hua* "flower."

b. [[*huang ya*] *cai*] *tang* "bean sprout soup"
L L MH H
(L H)
(o H o)
(o H o o)

For convenience, I indicate the intermediate sandhi forms corresponding to the disyllabic and trisyllabic bases, as well as the ultimate output of the long compounds. Take the intermediate form (H + o.H.o) of (a), which consists of the monosyllabic modifier *xin* "new" and the trisyllabic base *lu yin dai* "cassette tape." Assuming the constraint hierarchy AlignL/R ≫ Weight-to-Accent, AlignL should enforce an initial accent, since *xin* does carry a H. Instead, the accent actually falls on the H of the superheavy trisyllabic base, yielding a paroxytonic sandhi form. By the same token we expect AlignR to accent the final H-toned syllable *tang* "soup" of (b). Instead, what we actually see is a proparoxytonic pattern.

The generalization that emerges is this: if the choice is between accenting a light monosyllable at the word-edge by virtue of AlignL/R or a *heavy* (i.e. disyllabic) constituent in accordance with Weight-to-Accent, the former constraint dominates. On the other hand, if the alternative is between accenting a peripheral singleton or a *superheavy* (i.e. trisyllabic or longer) constituent, the ranking order is reversed. The standard solution to this kind of apparent ranking paradox is to explode the global Weight-to-Accent into more parochial and fine-grained constraints. One way is to revise Weight-to-Accent as follows:

(68) **Weight-to-Accent** (exploded version)
a. **Heavy** Accent a heavy constituent (disyllabic)
b. **Superheavy** Accent a superheavy constituent (trisyllabic or longer)

By interposing AlignL/R between Superheavy and Heavy, we get the desired results. This is illustrated by the examples given below in the usual tableau format:

(69) *xin* [[*lu-yin*] *dai*] "new cassette tape"
H.Lq.H.MH base tone
[H + (o.H.o)] relevant input

		Culm	Super	AlignL	AlignR	Heavy
a	(H.o.H.o)	*				
b	(H.o.o.o)		*			
c ☞	(o.o.H.o)			*		

Super = Superheavy

(70) [[*huang ya*] *cai*] *tang* "bean sprout soup"
L. L. MH. H base tone
[(o. H. o)+H] relevant input

			Culm	Super	AlignL	AlignR	Heavy
a		(o.H.o.H)	*				
b	☞	(o.H.o.o)				*	
c		(o.o.o.H)		*			

In the tableaux above, we ignore candidates that violate other higher ranking constraints,[40] and assume the intermediate representations [H + (o.H.o)] and [(o.H.o) + H] to be the relevant inputs, consistent with recursive constraint satisfaction. The ranking paradox disappears, since we can hold the constraint hierarchy constant and still pick the correct outputs in those cases which motivated the Align ≫ Weight-to-Accent dominance relation. The relevant examples are repeated below:

(71) *qing* [*gong ye*] "light industry"
H. H. Lq base tone
[H + (H. Hq)] relevant input

			Culm	Super	AlignL	AlignR	Heavy
a		(H.H.Hq)	*				
b	☞	(H.o.o)				*	*
c		(o.H.o)			*	*	
d		(o.o.Hq)			*		

(72) [*shui xian*] *hua* "narcissus flower"
HM.H.H base tone
[(o.H) + H] relevant input

			Culm	Super	AlignL	AlignR	Heavy
a		(o.H.H)	*				
b		(o.H.o)				*	
c	☞	(o.o.H)					*

A fortiori, when neither AlignL nor AlignR is relevant, Superheavy determines the location of the accent, again assuming a cyclic mode of constraint satisfaction. This is illustrated below:

[40] For instance, we discard [(H)(o.H.o)] as an output candidate for *xin lu yin dai* "new cassette tape" since it violates Foot Binarity.

(73)	[*wei* [*sheng wu*]] *xi*	"department of microbiology"
	L H Lq LM	base tone
	(H Hq)	disyllabic sandhi form
	(o o Hq)	trisyllabic sandhi form
	(o o Hq o)	quadrisyllabic sandhi form

Or, in tableau form:

(74) [*wei* [*sheng wu*]] *xi* "department of microbiology"
L H Lq LM base tone
[(o.o.Hq)+LM] relevant input

			Culm	Super	AlignL	AlignR	Heavy
a		(o.o.Hq.H)	*				
b		(o.o.o.H)		*			
c	☞	(o.o.Hq.o)					

The robustness of the Weight-to-Accent constraint we formulated as Superheavy is borne out by the fact that the superheavy constituent retains the accent in all but two instances out of a total of 42 quadrisyllabic compounds with the internal structure [σσσ + σ] or [σ + σσσ], that is a triplet plus a singleton, linear order irrelevant.[41]

5.3 *Residue: competing analyses*

There is a small handful of cases that resist the analysis proposed here. Consider the following two examples:

(75) a.	[*zhan lan*] *hui*	"exhibition"
	HM.ML LM	base tone
	(MH.o)	disyllabic sandhi form
	(H o o)	expected
	(o H o)	attested
b.	*bai* [*mu er*]	"white mushroom"
	Lq Lq.ML	base tone
	(o. MH)	disyllabic sandhi form
	(o o H)	expected
	(o H o)	attested

In neither case is AlignL/R relevant, since neither of the word-edge monosyllables carries a H tone. Weight-to-Accent, therefore, predicts that the

[41] Recorded and transcribed in the summer of 1990. The two exceptions are: *xin* [[*jin-hua*] *lun*] "new theory of evolution" /H.MH.MH.LM/ = [H + (o H o)] = (H o o o) instead of the expected (o o H o), and *xin* [[*lu-yin*] *dai*] "new cassette tape," which has an alternate reading (H o o o). See further discussion in section 7 on semantically determined prominence.

Table 6.8. *Alternative analyses of trisyllabic compounds (figures refer to the number of examples; total = 108)*

input	intermediate forms		sandhi forms predicted by:	
			Align/Weight	Accent
a. H.T.T		21	H.o.o	H.o.o
b. T.T.H		46	o.o.H	o.o.H
c. T.T.T	i. [o.T] T	21	o.H.o	
	T [T.o]	6		
	[T.T] T	5		o.H.o
	ii. [T.o] T	7	H.o.o	
	iii. T [o.T]	2	o.o.H	
	T [T.T]	0		

accent remains on the toned syllable of the disyllabic, heavy constituent – that is on the leftmost syllable in (a), the rightmost in (b). Surprisingly, the accent, and therefore the H tone, falls on the medial in both cases. Significantly, both examples constitute the leftover case (c) of the Accent rule (stated as (15) in section 3.1).

To understand this anomaly, let us juxtapose two alternative analyses: (A) the aforementioned Accent rule; and (B) AlignL/R coupled with Weight-to-Accent. Assuming that the subcorpus of 108 trisyllabic compounds is representative of the kind of input to the language learner, s/he can infer two alternative grammars (A) and (B), which are basically equivalent in terms of their empirical coverage. We can show why this is so more clearly in table 6.8.

Of the 108 items under scrutiny, 67 have the input form of either /H.T.T/ or /T.T.E/. Since low-register evens are not permitted except in initial position (by virtue of the constraint labeled as *E↓; see section 4.1), /T.T.E/ is equivalent to /T.T.H/. In these cases – (a) and (b) of table 6.8 – both Accent and Align/Weight make exactly the same predictions: the accent falls on the first H, or else on the last H. The two grammars begin to diverge in the leftover cases of (c). The Accent hypothesis lumps all the leftover cases into a single default category, and uniformly assigns prominence to the medial syllable. The Align/Weight hypothesis, on the other hand, breaks down the residue into a number of subcases, depending on the intermediate forms listed here as (i–iii). Since the examples of class

(c) contain no initial or final H, accent placement is uniquely determined by Weight-to-Accent, which excludes the light singleton from bearing the accent. This means that accent stays on whatever syllable carries a tone (= tonic) within the disyllabic base. Where both syllables of the base are tonic, the syllable on the right is the more prominent one.[42] This means that Align/Weight predicts the accent to be medial – exactly like the Accent hypothesis. So far, therefore, the two grammars are descriptively identical. The remaining nine items of subcases (c-ii) and (c-iii) (in the shaded areas of table 6.8) constitute the only crucial instances that are capable of discriminating between the two hypotheses. Here the empirical evidence is ambiguous between the two grammars: the nine items are virtually evenly split between the two alternative accounts, as seen below.

(76)

		correct output predicted by			
		Align/Weight		Accent	
c-ii	[T.o] T	[H.o.o]	4	[o.H.o]	3
c-iii	T [o.T]	[o.o.H]	0	[o.H.o]	2

Figures indicate numbers of examples

Examples (75a, b) cited above instantiate subcases (c-ii) and (c-iii) respectively. In both instances Accent makes the correct prediction. On the other hand, four instances of (c-ii) are accountable only by Align/Weight, one of which is given below.

(77) [*jin kou*] *huo* "imported goods"
MH.HM MH base tone
(MH.o) disyllabic form
(H o o) trisyllabic form

The alternative Accent rule-based account would treat (77) as one of the elsewhere or default cases (c) of table 6.8, hence wrongly assigning accent to the medial syllable. Other examples, cited earlier, also favor the Align/Weight analysis over the Accent hypothesis:

(78) a. [*zong tong*] *fu* "presidential palace"
HM HM HM base tone
(MH. o) disyllabic form
(H o o) trisyllabic form

[42] Recall that other things being equal, the tone on the right is more resistant to either (i) deletion, by virtue of Prs-Tp (parse tone in right-prominent position) ≫ Prs-T (parse any tone) or (ii) melodic alteration, consistent with Prs-Cp (parse contour of right-prominent tone) ≫ Prs-C (parse contour). See chapter 5 for details.

b. [*qi xiang*] *tai* "weather station, meteorological observatory"
MH ML HM
(MH. o)
(H o o)

c. [*chuan ran*] *bing* "contagious disease"
L ML LM
(LM o)
(H o o)

When faced with choosing between two grammars of nearly equal empirical coverage, the language learner naturally prefers the Accent analysis, which is by all accounts far simpler than Align/Weight. What then motivates a more complex Align/Weight analysis? There are basically two types of reasons for the more complex alternative. One is a priori and theoretical. As pointed out in the opening paragraphs of section 5, the Accent analysis implies that the default, ergo unmarked, accentual pattern is "amphibrach" (o.H.o), at odds with the overwhelmingly head-terminal tendency of rhythmic organization. The other is empirical: as just pointed out, the four examples given above as (77) and (78) are compatible only with the Align/Weight hypothesis. But more significantly, Accent has nothing to say about quadrisyllabic or longer compounds such as "department of microbiology" cited earlier:

(79) [*wei* [*sheng wu*]] *xi* "department of microbiology"

L	H	Lq	LM	base tone
	(H	Hq)		disyllabic form
(o	o	Hq)		trisyllabic form
(o	o	Hq	o)	quadrisyllabic form

In cases such as this, neither the first syllable *wei*, nor the final syllable *xi* carries a H tone. Accent rule, which subsumes AlignM (align the accent with the medial tone) as the unmarked elsewhere condition, is incapable of deciding between the two medial syllables *sheng* and *wu* as the accent site. By contrast, the Align/Weight analysis straightforwardly predicts a paroxytonic pattern (see (74)). From the analyst's point of view, the generalizability of Align/Weight to long compounds is decisive in tipping the balance in favor of the Align/Weight hypothesis. However, from the language-learner's perspective, it is certainly conceivable to devise a "coping" grammar that has one set of rules for trisyllabic compounds (the Accent rule), another for long compounds, which basically says: adjoin monosyllabic constituents as the weak/unaccented member to the trisyllabic or longer base. The fact that these sets of principles seem unrelated to

each other may or may not be apparent or, for that matter, of paramount importance, to the language learner.

Finally, there is an intriguing piece of argument for the Align/Weight analysis. Notice that the Accent rule represents a direct mapping from a string of base tones to trisyllabic forms. Consider the following.

(80)	[*ma xi*] *tuan*	"circus troupe"
	ML.LM. L	base tone
	(H.M)	disyllabic form
	(H o o)	trisyllabic form

In terms of the underlying tonal string /ML.LM.L/, it belongs to the default class of /T.T.T/. A direct mapping hypothesis therefore wrongly predicts [o.H.o] as the output. By contrast, Align/Weight takes the disyllabic sandhi form plus the monosyllabic addition as the relevant input, and determines the ultimate output. Since the disyllabic base *ma xi* "circus" has an unimpeachable underlying falling-rising sequence /ML.LM/, both tones are retained, and surface with a register-dependent [H.M] surface form (see discussion in chapter 5, section 5.6). At this point, the relevant input is /[H.M] + L/. Consequently, AlignL places the accent on the leftmost H-toned syllable *ma*. What is at stake here is the fact that the H tone on *ma* is derived, not underlying. This provides a highly suggestive piece of evidence in support of the cyclic/recursive mode of constraint satisfaction assumed in the Align/Weight analysis.

In short, the choice confronting the language learner is a difficult one: while simplicity favors the Accent hypothesis, generalizability and a priori principles favor the Align/Weight analysis. Chen (1994b) speculates that it is this delicate balance that results in two coexistent and competing grammars, a notion that extends and conceptually enriches Wang's (1969) original hypothesis about competing sound changes as a cause of residue.[43]

6 Tonic clash

Recall that footing is word-bound, a constraint that outranks Foot Binarity. This means that monosyllabic feet are possible. Thus we routinely encounter phrasal constructions with tonal strings like [(T)(o.T)] or [(T.o)(T)]:

[43] The Align/Weight analysis differs in certain details from that of the corresponding so-called 2-Base analysis of Chen (1994b). Hence, the figures of table 5 differ from those of table 1 in the aforementioned article.

(81)	a.	*si* [*mian-pen*]	"four washbasins"
		MH LM L	base tone
		[(MH) (o H)]	sandhi form
	b.	sweep floor go	
		[*sao di*] *qu*	"go sweep the floor"
		MH.LM.MH	
		[(MH o) (H)][44]	

(. . .) = foot
[. . .] = p-word (or higher prosodic unit)

Remarkably, the mirror images [(T)(T.o)] and [(o.T)(T)] are systematically absent. Where we expect such tonal sequences with two abutting tonics to arise by virtue of general principles established in the preceding sections, often we find instead one of the abutting tones deleted or moved away from the other. In other words:

(82)	Expected		Attested			
	[(T)(T.o)]	→	[o (T.o)]	or	[(T) o.o]	Tone Deletion
			[(T)(o.T)]			Tone Shift
	[(o.T)(T)]	→	[(o.T) o]	or	[o.o (T)]	Tone Deletion
			[(T.o)(T)]			Tone Shift

The significance of these facts is that if tone-bearingness is purely paradigmatic in the same sense that other phonological features like voicing or rounding are, then there is no particular reason to expect such a systematic gap. As Goldsmith (1982:57f.) pointed out, Meeussen's rule (which demotes one of the two abutting Hs to L in many Bantu languages) would be unexpected if it is conceived as a purely tonal process. On the other hand, if tonicity or tone-bearingness is seen as entailing by its very nature some phonological prominence, then such a prohibition on adjacent Hs or tonics is analogous to the avoidance of stress clash. Seen in this light, the systematic avoidance of two adjacent tonics in New Chongming is another telltale sign of an emergent accentual system.

In this section we will examine in turn the two alternative strategies for resolving tonic clash: de-accentuation and accent shift.

6.1 De-accentuation and tone deletion

Trisyllabic compounds must form a single foot headed by one and only one tonic syllable; culminativity, therefore, per force precludes the possibility of tonic or accent clash. Longer compounds and phrasal constructions, on the other hand, are not so constrained, as they may form two or more

[44] MH → H, by optional Leveling, see section 8.

rhythmic units, each with a tonic peak. First, let us look at the phrasal constructions of (81a, b) cited above. *Si* "four" forms a separate p-word, and therefore a separate foot, by itself. Likewise, *qu* "go" constitutes an independent foot, consistent with the prosodic hierarchy. In other words, word-bound foot formation licenses monosyllabic feet, overriding the Foot Binarity condition.[45] Logically speaking, therefore, we expect to find examples like the following:

(83) a. *san* [*sheng shi*] "three provinces and municipalities"
H HM ML
(MH o)
[(H) (MH o)] *

b. [*liang san*] *nian* "two or three years"
ML H L
(o H)
[(o H) (H)] *

(84) a. *nian* [*xiao wan*] "twenty small bowls"
LM HM HM
(MH o)
[(MH) (MH o)] *

b. [*liang san*] *wan* "two or three bowls"
ML H HM
(o H)
[(o H) (MH)] *

(83a) is straightforward: it merely strings together the base tone H of *san* "three" together with the sandhi form of *sheng shi* "provinces and municipalities." Likewise (83b) represents the sum of the two constituents, except that L is raised to H, consistent with the constraint *E↓ (no low even tones, except in initial position). By the same token, (84a, b) consist of the disyllabic sandhi form plus the tone contributed by the singleton – subject to the relevant constraints, in particular *O↓ (no low obliques anywhere) and Fall (no final falling tone, see chapter 5). Surprisingly, none of the outputs of (83)–(84) is acceptable (marked by an asterisk "*"). Instead, the attested readings are given in (83′) and (84′), respectively.

(83′) a. [(H) o.o]
b. [o.o (H)]

(84′) a. [o (MH.o)][46]
b. [(o.H) o]

[45] Which bars monosyllabic feet consisting of sublexical units, see section 3.2.
[46] Optionally leveled to [o (H.o)], see section 8 on Leveling.

To frame the problem in broader terms, by limiting ourselves to two possible disyllabic sandhi forms [T.o] and [o.T],[47] we expect to find these four logical possibilities, when a singleton T is added to the disyllabic base:

(85)

Left-branching constructions	Right-branching constructions
[(T) (o.T)]	[(T.o) (T)]
[(T) (T.o)]	[(o.T) (T)]

Instead, three-syllable phrasal constructions show the distribution of table 6.9. Three-syllable-long phrasal constructions are cross-classified according to construction type (columns A, B, C), immediate constituency (left-branching: rows a–e; right-branching: f–j), and sandhi tone patterns (T.o.o, T.o.T, etc.).

It can be seen that, with a small handful of exceptions (5 out of 98), neither *[(o.T)(T)] nor *[(T)(T.o)] occurs in our sample.[48] It is quite transparent that in each case the starred forms contain two abutting accented/tonic syllables, the classic configuration of an accent/stress clash, penalized by a constraint I will refer to as *Clash (for No Clash).

(86) ***Clash**
No abutting accented/tonic syllables.

To avoid such a dysrhythmic pattern, one strategy is to de-accent (therefore de-tone, by virtue of Tonicity, see section 3.1) one of the abutting syllables. The target of this prominence reduction is determined by the same principle we have alluded to, namely AlignL/R ≫ Weight-to-Accent (accent a heavy constituent).

[47] For expository simplicity, I have ignored [T.T] as a possible disyllabic sandhi form. One example of such a form is the following, which instantiates tonal deletion under tonic clash. In addition, the surviving initial tone undergoes Leveling to become [H], see section 8.

photograph on/in	
[*xiao zhao*] *lang*	"in the picture"
HM.MH LM	base tone
[(HM.MH) (MH)]	expected
[(H o) (MH)]	attested

[48] My recordings and field notes contain approximately 500 trisyllabic postlexical expressions of various construction types. Of these very few display either the [(o.T)(T)] pattern or its mirror image [(T)(T.o)].

Table 6.9.

sandhi forms			construction types		
			A	B	C
			VX + D	NuNu + N	Nu + CN
a.		[(T.o)(T)]	5	4	
b.	*	[(o.T)(T)]	1	0	
c.		[o.o (T)]	5	5	
d.		[(T.o) o]	6	5	
e.		[(o.T) o]	11	4	
f.	*	[(T)(T.o)]			4
g.		[(T)(o.T)]			17
h.		[(T) o.o]			9
i.		[o (T.o)]			15
j.		[o (o.T)]			7

Figures indicate type frequency

T, o	tonic, atonic syllable
VX + D	verb + object or reduplicated verb + directional complement
NuNu + N	number expressions + noun
Nu + CN	number + compound noun
*	tonic clash

Exs.:	VX + D	[*sao di*] *qu*	"go sweep the floor"
	NuNu + N	[*liang san*] *nian*	"two or three years"
	Nu + CN	*si* [*mian pen*]	"four washbasins"

To illustrate, consider the examples alluded to earlier:

(87) *san* [*sheng shi*] "three provinces and municipalities"
/H.HM.ML/ = /H + (MH.o)/

			*Clash	AlignL/R	Weight	Parse-T
a		[(H)(MH.o)]	*			
b	☞	[(H) o.o]			*	*
c		[o (MH.o)]		*		*

(88) [*liang san*] *wan* "two or three bowls"
/ML.H.HM/ = /(o.H) + HM/

			*Clash	AlignL/R	Weight	Parse-T
a		[(o.H)(MH)][49]	*			
b		[o.o (MH)]			*	*
c	☞	[(o.H) o]				*

In (87), the H-toned singleton *san* attracts accent by virtue of AlignL, and therefore retains its tone, while the disyllabic base is de-accented; in (88), it is the MH-toned singleton that undergoes tonal reduction, since AlignL is irrelevant here. The general constraints already motivated in earlier sections pick out the winning candidates of (87) and (88). Both of these cases contrast with (89), where the input already satisfies *Clash as is; consequently, no de-accentuation takes place.

(89) *si* [*mian pen*] "four washbasins"
/MH.LM.L/ = /MH + (o.H)/

			*Clash	AlignL/R	Weight	Parse-T
a	☞	[(MH)(o.H)]				
b		[o (o.H)]				*
c		[(MH) o.o]			*	*

6.2 *Accent shift*

Even more remarkable is the fact New Chongming resorts to accent shift to resolve tonic clash. De-accenting is equivalent to beat-deletion in metrical phonology (cf. Prince 1983, Selkirk 1984). An alternative way of resolving a stress clash is accent shift, or beat-movement, in a manner reminiscent of the Rhythm Rule of English (cf. (*Ja.pán*) but (*Jà.pa*)(*nése*), or (*con.dénse*) but (*còn.den*)(*sá.tion*)). Consider the examples of (90).

(90) a. [*yi liang*] *tian* "one or two days"
Hq.ML.H base tone
[(o.MH)(H)] expected
[(Hq.o) (H)] attested

b. [*ba jiu*] *ye* "eight or nine nights"
Hq.HM.LM
[(o.MH) (MH)] expected
[(Hq.o) (MH)] attested

[49] /HM/ → [MH], to satisfsy *Fall (no falling tone, except before a rising tone).

c.	[*chi shui*] *qu*	"go drink water" (lit. [drink water] + go)
	Lq.HM.MH	
	[(o.MH) (MH)]	expected
	[(Hq.o) (MH)]	attested

There is no question that (o.MH) is the expected sandhi form for the disyllabic constituent enclosed between square brackets, as demonstrated by the disyllabic compounds below.[50]

(91)	a.	*hei shi*	"black market"
		Hq.ML	base tones
		(o.MH)	sandhi form
	b.	*se cai*	"color"
		Hq.HM	
		(o.MH)	
	c.	*fa kuan*	"fine, penalty"
		Lq.HM	
		(o.MH)	

This means that in the case of *yi liang tian* "one or two days" (90a), we expect two separate rhythmic units [(o.MH)(H)] with two abutting tonics. One way of resolving the tonic clash is to de-accentuate the medial syllable, thereby yielding [o.o (H)] consistent with AlignR ≫ Weight-to-Accent. Instead, what we find is a leftward accent shift. As a consequence, *yi* "one" is now accented, and therefore capable of carrying tone /Hq/. In the case of *ba jiu ye* "eight or nine nights," the expected intermediate form is /(o.MH) + MH/. Tonic clash resolution via de-accenting would yield [(o.MH) o], eventually [(o.H) o] by Leveling (see section 8), consistent with Weight-to-Accent (when not outranked by AlignL/R). Instead, again what we see is that New Chongming resorts to accent shift to eliminate the offending sequence of two adjacent tonics. *Chi shui qu* "go drink water" parallels *ba jiu ye* in all relevant respects.

We have used trisyllabic phrasal constructions to illustrate the effect of accent shift, but the phenomenon is not limited to utterances of any particular length. One striking example is the following:

(92)	one CL fresh troops		
	[*yi zhi*]	[*sheng li jun*]	"a new force"
	Hq.H	H.Lq.H	
	[(o.H)	(H.o.o)]	expected
	[(o.H)	(o.o.H)]	attested

[50] For tone sandhi purposes, number + number and certain verb + object expressions are treated like lexical compounds. This holds true of most Wu dialects I have examined in any depth.

Consistent with AlignL ≫ AlignR, we expect *sheng li jun* to be initially stressed or accented. Stress clash moves the accent from the initial position to the next best position, namely the final position, in conformity to AlignR that encodes the universal head-terminal rhythmic organization.

Two further remarks are worth making. First, tonic clash avoidance is so strong an imperative as to override the otherwise undominated constraint *Tq, which prohibits a CVq in initial position from carrying a tone (see chapter 5, section 2.1). The initial syllables of (90), transcribed in the Mandarin-based "*pinyin*" as *yi*, *ba*, *chi* actually end in a glottal stop in the New Chongming pronunciation; in other words, they exhibit the checked syllable structure CVq, as indicated by the symbol -q after the tone (Hq, Lq). Nevertheless, as accented syllables, they do carry tone. This means that at the phrase level, *Clash outranks *Tq. Second, notice that in each of the cases of (90) what is shifted is not the tone MH, but accentual prominence; otherwise, the outcome of tonal displacement would have been [(MH.o)(T)] instead of the attested [(Hq.o)(T)] (where T = either H or MH). In other words, the underlying mechanism alters not tonal material per se, but the prominence relation among potential tone-carriers. This further reinforces our contention that what we are witnessing in New Chongming is an emergent, but already robust accentual system.

Potential tonic clash can arise also in lexical constructions comprising two or more feet. Unlike trisyllabic compounds, longer compounds with the internal structure [σσ + σσ] comprise two feet, with optional de-accenting and concomitant tone loss on one of the less salient rhythmic units (see section 4.2). Thus the idiom (93) has two possible readings:

(93)		fire on-top add oil	
		[*huo shang*] [*jia you*]	"add fuel to fire; to inflame"
		HM. ML H. L	base tone
	a.	[(MH. o) (H. H)]	alternative sandhi forms
	b.	[o o (H. H)]	

What is at stake here is the fact that long compounds of certain structural configurations potentially give rise to tonic clash. When such an eventuality arises, we see the same mechanisms of clash resolution at work in long compounds, as illustrated by these examples:

(94)	a.	big scare small surprise	
		[*da jing*] [*xiao guai*]	"to make a fuss"
		LM.H HM.MH	base tone
		[(o. H) (HM.MH)]	expected sandhi form
		[(o. H) (o.MH)]	attested sandhi form

b. wind blow grass move
[*feng chui*] [*cao dong*] "sign of disturbance"
H. H HM. ML
[(H. H)(MH. o)] expected
[(H. H)(o. MH)] attested

c. one side Prt story
[*yi mian*] [*zhi ci*] "biased view"
Hq.LM H.L
[(o.MH) (H.H)] expected
[(Hq. o) (H.H)] attested

Since the input tone sequence of the second rhythmic unit in (a) is perfectly well-formed, the expected phonetic realization ought to be [(o.H)(HM.MH)]. Surprisingly, the actual recorded pronunciation was [(o.H)(o.MH)], suggesting a de-accenting of *xiao* "small." We see the effect of the alternative strategy of accent shift in example (b). Here /HM.ML/ should yield (MH.o) as output, consistent with the principle of Saliency (parsing of a high register prevails over the parsing of other tones, ceteris paribus). Instead, the accent shifts to the right, alleviating the tonic clash at the juncture where the two feet meet. (c) is the mirror image of (b): the expected sandhi form for *yi mian* "one-sided, biased" is (o.MH) (the overriding constraint here being *Tq, which bars a checked syllable from bearing tone); consequently, the whole expression should have the tonal representation [(o.MH)(H.H)]. Instead, we see an accent shift to the left – which explains why the tonic falls on the first rather than the second position.

There is a notable difference between trisyllabic phrasal constructions and long compounds: tonic clash is avoided with fairly high degree of consistency in the former, but tolerated in the latter. Thus, side by side with (94) we find long compounds like those of (95a, b, c), where neither accent deletion nor accent shift comes into play to alleviate tonic clash. The optionality of clash resolution is particularly evident in (95d), where both readings are possible. This difference in clash-tolerance probably reflects the degree to which monosyllables are more closely tied to a neighboring rhythmic unit.[51]

(95) a. hollow dumpling
[*kong xin*] [*tang yuan*] "dumpling without stuffing"
H. H H. L base tone
[(H. H) (H. H)] sandhi form

[51] We return to the prosodic dependency of monosyllables in chapter 9.

b. decide mind
[*da ding*] [*zhu yi*] "make up one's mind"
HM.LM.HM.MH
[(HM.MH) (HM.MH)]

c. clearly know purposefully violate
[*ming zhi*] [*gu fan*] "deliberate breach of law"
L. H MH.ML
[(L. H)(MH.o)]

d. wide open vision
[*da kai*] [*yan jie*] "a real eye-opener"
LM.H HM.MH
[(o.H) (HM.MH)]
[(o.H) (o. MH)]

6.3 *Diagnostic value*

The diagnostic value of tonic clash avoidance is self-evident. We can understand the distributional facts sketched in the preceding sections only in accentual terms. In a purely tonal language, the juxtaposition of tonal elements in an example like *chi shui qu* "go to drink water" [(o.MH)(MH)] (= 90c) is altogether unremarkable, one of the hallmarks of a full-blown tone language being the ability for each potential tone-bearing unit (in this case the syllable) to carry a tone that stands in a paradigmatic relation to another tone. The tendency to avoid such a tonal sequence is totally unexplained. On the other hand, when we recognize in New Chongming an emergent accentual system, we begin to appreciate the highly marked nature of a sequence like [(o.T)(T)] and its mirror image [(T)(T.o)], and the fact that they are barred by a syntagmatic constraint which is quintessentially accentual in nature.

Needless to say, not all tonal sequences may freely occur, even in tonal languages. For instance, one may interpret Meeussen's Law as a manifestation of OCP, which disallows abutting H.Hs.[52] The OCP effect is evident in Chinese dialects as well (see chapter 3 on Tianjin). However it is not possible to reduce tonic clash avoidance to OCP effect. This is clearly demonstrated by (84b = 88), repeated below. Here the two adjacent tones are H and MH. There is no way to interpret the deletion of MH as a purely tonal phenomenon.

[52] One such example is Makua. For relevant facts, see Cheng and Kisseberth (1979–1981), conveniently summarized in Kenstowicz (1994:359ff.).

(84b) two three bowl

[*liang san*] *wan*	"two or three bowls"
ML H HM	base tone
[(o H) (MH)]	expected sandhi form
[(o H) o]	attested

7 Semantically determined prominence

Up to this point, we have proceeded as if relative tonic prominence could be computed automatically on phonological grounds alone, such as tonal Saliency, Edgemostness, Weight-to-Accent, and clash-triggered accent shift. So far we have not reckoned with such contributing factors as intonational meaning, in particular with the informational content and focus structure of utterances.[53] Not surprisingly, when semantics and discourse conditions warrant it, the speaker can bring any constituent of a sentence under focus or emphasis. This focus-related tonic prominence can outweigh all other factors. How these non-phonological factors impinge on tonal behavior is still poorly understood (see chapter 11 on intonation in Wenzhou). For our present purpose, I only wish to make two observations. First, contrastive focus is highlighted by tonic accent and, correspondingly, the backgrounded elements are tonally suppressed. This is illustrated below.

(96) a. third month in

[*san yue*] *li*	"in March"
H Lq ML	base tone
(H Hq)	disyllabic sandhi form
[(H Hq) o]	expected
[(H o) o]	attested

b. fourth month in

[*si yue*] *li*	"in April"
MH.Lq.ML	
(o Hq)	
[(o Hq) o]	
[(H o) o]	

[53] We leave out entirely the question of "expressive meaning" related to the speaker's mood, attitude, and certain aspects of illocutionary force of the utterance. Suffice it to note that unlike English, which encodes such expressive meanings by means of a repertoire of intonational contours, Chinese (Cantonese in particular) tends to use sentence particles for the same purpose. For review and references of relevant literature, see Ladd (1980), Bolinger (1982), and Selkirk (1984, esp. ch. 5).

c. fifth month in
[*wu yue*] *li* "in May"
ML.Lq ML
(o Hq)
[(o Hq) o]
[(H o) o]

In isolation, *san yue* "March" shows up as [H.Hq], while *si yue* "April" and *wu yue* "May" are pronounced with the tone pattern [o.Hq], consistent with the principles governing disyllabic lexical tone sandhi developed in chapter 5. Not surprisingly, the locative enclitic *li* "in, at" is stressless and atonal. What was unexpected is the fact that the tonic falls uniformly on the first syllable. I know of no regular process whereby the tonic accent shifts from the second to the first position in such configurations. The mystery dissipates when one points out that these three examples occurred in succession on a list the informant was asked to read out loud. Obviously, the only syllables that bear contrastive information are the ordinal numbers *san* "third,", *si* "fourth," and *wu* "fifth," which set the three examples apart from each other. Quite appropriately, the informant emphasized this contrastive element by assigning to it the tonic peak and simultaneously demoting all the other elements.

In all likelihood, contrastiveness is responsible for other apparent exceptions as well. Recall that Superheavy (accent a superheavy constituent) dominates AlignL/R (accent the leftmost or rightmost H). This means that in vying for accentual dominance, a H-toned singleton always loses out to a superheavy (trisyllabic or longer) constituent; as a consequence the tonic always falls on the superheavy base in long compounds with the internal structure [σ + σσσ] or [σσσ + σ]. However, there are two exceptions in our subcorpus consisting of 42 entries, one of which is given below:

(97) new evolution theory

	xin	[[*jin*	*hua*]	*lun*]	"new theory of evolution"
	H	MH.MH.LM			base tone
a.		(o	H	o)	trisyllabic sandhi form
b.	(o	o	H	o)	expected quadrisyllabic sandhi form
c.	(H	o	o	o)	attested quadrisyllabic sandhi form

The expected sandhi form is (b), consistent with Superheavy ≫ AlignL/R; instead, the attested reading is (c). I attribute this anomaly to focus-induced accent shift, which serves to emphasize the new vs. the old theories of evolution. This account seems plausible especially in view of the variant readings attested in the only other apparent exception (98).

(98) new record tape
xin [[*lu yin*] *dai*] "new cassette"
H Lq H MH
a. (o H o) trisyllabic sandhi form
b. (o o H o) attested quadrisyllabic sandhi form
c. (H o o o) attested quadrisyllabic sandhi form

The second point worth making is that accent – hence distinctive tone – is used to encode degrees of semantic contentfulness. Notice the following contrasts:

(99) a. scold two sentences
ma [*liang ju*] "utter two sentences in scolding"
LM ML MH base tone
[o (o H)] sandhi form

b. go two times
qu [*liang tang*] "go a couple of times"
MH ML MH
[(H) o o]

The number *liang* "two" can be interpreted either as "two" or, loosely speaking, as "a couple of, several." This semantic contrast between the so-called *shishu* "real number" and the *xushu* "empty number" readings is encoded by means of tonic accent: in the former reading, the accent is on the NP; in the latter reading, the accent is on the verb. Similarly, the word *ji* in example (100) functions either as an "interrogative quantifier" meaning "how many," or as an "indeterminate quantifier" meaning "a few, several."[54] In the former function, the NP *ji tian* is contentful in the sense that it is used to solicit specific information, and is accented, with a full-fledged tonal pattern [HM.MH] (by postlexical tone sandhi, see table 6.4 in section 2.2). In the latter usage, *ji tian* only imparts unspecified information regarding the number of days of imprisonment; accordingly, it is de-accented, and tonally reduced to [o.o].

(100) lock several/how-many day
guan [*ji tian*]
H HM.H
a. [(H) (HM.MH)] "lock up someone for how many days?"
b. [(H) o o] "lock up someone for a couple of days"

[54] The terms "interrogative" and "indeterminate" translate the terms *xunwen shumu* and *buding shumu* used by Chinese grammarians (cf. Lü et al. 1980).

This kind of accenting and de-accenting keyed to meaning is presumably language-universal, if conventionalized and grammaticized to varying degrees in different languages. The southern Min dialect of **Xiamen** exhibits virtually identical patterns to New Chongming, as illustrated by the examples of (101) and (102). In Xiamen, all but phrase-final tones change into a corresponding sandhi tone.[55] Hence we have the pattern [T′ T′ T] in reading (a), where T stands for an unaltered base tone, and T′ for a derived sandhi tone. When the NPs are de-accented and therefore de-toned, the tone associated with the verb becomes phrase-final, and therefore retains its underlying or base form T.

(101) buy two pound
bue [*nng kun*][56]

a.	T′	T′	T	"buy two pounds"
b.	T	o	o	"buy a few pounds"

(102) buy how many/several pound
bue [*kui kun*]

a.	T′	T′	T	"buy how many pounds?"
b.	T	o	o	"buy a few pounds"

T base tone
T′ sandhi tone [accent, level]
o neutral tone

Although the accent in the case of Xiamen and New Chongming has direct tonological consequences, the underlying mechanism of encoding focus and meaningfulness is the same as those employed in more familiar stress/accent systems.

8 Leveling

Tonal Saliency, Edgemostness, the Weight-to-Accent principle, clash-induced stress shift, and tonal deletion, and other non-phonological factors including focus-related prominence, jointly determine the location of the tonic. The tone shapes that can be associated with the tonic are strictly limited. Even in disyllabic compounds, only H, L, MH, and HM may occur on account of various constraints developed in chapter 5, including *O↓ (no low-register oblique) and *Fall (no falling tone, subject

[55] The specific tone shapes are irrelevant for our present purpose. For details, see chapter 10.
[56] *nng* represents a syllable consisting of a dental nasal onset followed by a syllabic velar nasal.

to OCP).[57] The repertoire of possible tones is further limited to exactly one, namely H in trisyllabic compounds. This means that a trisyllabic compound must exhibit one of the only three permissible tonal patterns: [H.o.o], [o.H.o], or [o.o.H], and no other. Now consider longer compounds formed by adding a monosyllabic singleton to a base that is trisyllabic or longer (= Superheavy). The singleton is barred by Foot Binarity from forming a separate rhythmic unit. It therefore adjoins the polysyllabic base into a single unbounded foot. Given Culminativity and the Superheavy constraints, the accent always remains on the superheavy base, with the singleton being prosodically demoted and tonally reduced. This means that long compounds formed by singleton-adjunction must perforce have one of the following patterns:

(103) (o (. . . H . . .)$_\alpha$)
((. . . H . . .)$_\alpha$ o)

(104) (. . . H . . .)$_\alpha$
α = superheavy rhythmic unit (unbounded foot)
". . ." = atonic syllables (incl. null)

The net effect of the aforementioned constraints is this: the only well-formed unbounded foot is one in which the only tonic carries a H, as in (104), where ellipsis ". . ." stands for an arbitrary string of atonic syllables (including null). In other words, the constraint Leveling repeated here as (105) is undominated within an (unbounded) superheavy foot.

(105) **Leveling**
H is the only permissible tone (in a trisyllabic or longer foot)

Beyond the polysyllabic feet, Leveling is less consistently enforced. Recall that compounds with the internal structure [[σσ][σσ]] may form separate feet [(σσ) (σσ)]. Since Leveling is a condition only on trisyllabic or longer feet, H, L, MH, and HM can in principle surface in shorter rhythmic units, as instantiated by the examples of (96) cited in section 6.2. Interestingly enough, cases such as these are distinctly in the minority. In the overwhelming majority of cases, Leveling holds sway not only within the unbounded feet, but reduces every surviving tone to a simple H in any connected speech long enough to provide a sufficient contextual cue. The four- to six-syllable expressions of (107) given below form two-to-three tonic groups in fairly deliberate speech. Where MH or L are expected

[57] As a consequence of OCP ≫ *Fall, HM can occur only before MH.

to occur, they have been leveled to H. The sweeping effect of Leveling is particularly evident in casual speech at faster tempo, as diagnosed by tonal reduction. In these cases virtually all surviving tones emerge simply as H.

(106)a.	thousand change ten-thousand transformation	
	[*qian bian*] [*wan hua*]	"ever changing"
	H MH LM MH	
	[(H o) (o MH)]	expected
	[(H o) (o H)]	attested
b.	[*shang hai*] [*jiao tong*] [*da xue*]	"Shanghai Jiaotong University"
	MH HM H H LM.Lq	
	[(MH o) (H H) (o.Hq)]	expected
	[(H o) (H H) (o.Hq)]	attested
c.	drive trackless tram	
	kai [[*wu gui*] [*dian che*]]	"lawless behavior"
	H L HM MH H	
	[o (L. o)[58] (o H)]	expected
	[o (H o) (o H)][59]	attested
(107)a.	all being gravity	
	[*wan you*] [*yin-li*]	"universal gravity"
	LM.ML ML.Lq	
	[(o MH) o o]	expected
	[(o H) o o]	attested
b.	greatly act-up temper	
	[*da fa*] [*pi-qi*]	"to throw a temper tantrum"
	LM.Hq.L.MH	
	[o o (L o)]	expected
	[o o (H o)]	attested
c.	fire on-top add oil	
	[*huo shang*] [*jia you*]	"to incite, instigate"
	HM.ML H.L	
	[(MH o) o o]	expected (and attested)
	[(H o) o o]	attested (alternative reading)

[58] Recall that when L occurs as the only tonal element in a rhythmic unit, it is phonetically realized as [LM] by virtue of pitch-accent. See chapter 5, section 5.1.

[59] Note in passing the structure-sensitivity of rhythmic structure. (106c) must be analyzed as singleton + superheavy quadrisyllabic constituent, isomorphic to the morphosyntactic bracketing. As a singleton, *kai* yields prominence to the superheavy constituent, and is tonally reduced to [o]. If we parsed (106c) as [σσσ + σσ], the trisyllabic unit (H.L.HM) corresponding to *kai wu-gui* would surface as (H.o.o), consistent with AlignL (accent the leftmost H). One would therefore wrongly predict the reading *[(H.o.o)(o.H)].

The tendency to reduce all tones to a level high is so strong in casual speech that out of 188 tonally reduced quadrisyllabic compounds in our subcorpus, only five carry a tone other than H.[60]

There is evidence that Leveling is making inroads into postlexical or phrasal constructions as well. For instance, Leveling neutralizes the tonal distinction between (108a) and (108b).

(108)a.	burn one burn		
	shao [*yi shao*]		"burn a little"
	H Hq H		
	(H) (o H)		
b.	fry one fry		
	chao [*yi chao*]		"fry a little"
	HM Hq HM		
	(MH) (o MH)		expected
	(H) (o H)		attested, via Leveling

Note that (108a, b) are clearly treated as postlexical structures, otherwise Foot Binarity, which is undominated at the word level, would force the entire three-syllable strings into a single foot, culminating in a single peak (consistent with Culminativity), so that the expected reading of (108a) would be (H.o.o), consistent with AlignL (accent the rightmost H), and that of (108b) would be (o.o.H), in accordance with the Weight-to-Accent principle. In other words, as independent words, the singletons *shao* and *chao* form separate feet of their own. The relevant point here is that *chao* in (108b) should surface with an oblique tone that is phonetically realized as [MH] (consistent with *Fall, which prohibits a falling tone except before a rising tone). Instead, the attested form shows the neutralization of [MH] in the direction of the all-purpose accent marker [H].

Such instances can be multiplied ad infinitum, across a wide variety of construction types. The prevalence of Leveling across postlexical construction types is borne out by our sample: 93 out of 100 instances of NP

[60] These are:

i.	(MH.o)(o.o):	[*si hai*] [*wei jia*]	"make one's home wherever one is"
		[*gu dai*] [*han yu*]	"Ancient Chinese"
		[*ping qi*] [*ping zuo*]	"on an equal footing"
ii.	(o.MH)(o.o):	[*wu si*] [*yun dong*]	"the May Fourth movement"
		[*fu dan*] [*da xue*]	"Fudan University"

In hindsight I strongly suspect that even in these cases MH freely alternates with H. Unfortunately I did not have the foresight to ask my informant whether (H.o)(o.o) and (o.H)(o.o) would have been equally acceptable variants for the examples of (i) and (ii) respectively.

+ locative constructions, 44 out of 49 verb + directional complement, and 33 out of 35 verb + number-measure.[61]

From the foregoing examples it is clear that Leveling converts all surviving tones into a uniform H, obliterating the last trace of paradigmatic contrasts encoded in the underlying tonal categories. The net result is that in a fair majority of cases in connected speech the only remaining contrast is that between tonic and atonic, or accented and non-accented.

[61] There is one major class of exceptions. Number-measure + noun constructions typically form two separate rhythmic units, each retaining its tonal contrasts. For example: [*liang wan*] *fan* "two bowls of rice." Consistent with the rules of postlexical tone sandhi (summarized in table 4, section 2.2), the number-measure expression *liang wan* "two bowls," underlyingly /ML.HM/, assumes the sandhi form [H.M]. The nominal head *fan* "rice" with lexically assigned /LM/ shows up as [MH], consistent with *O↓ (no low-register oblique). As a consequence, /ML.HM + LM/ → [(H.M)(MH)].

7 *Stress-foot as sandhi domain I*

Beginning with this chapter we turn our attention to a leitmotif that underscores all investigations of tone sandhi. We have surveyed the various types of sandhi phenomena, and looked into how sandhi processes interact with each other, but we have not so far directly addressed the issue of the domain of tone sandhi. We have yet to make more precise the principles whereby connected speech breaks down into units within which sandhi rules operate.

The prevailing idea is that speech is hierarchically structured, from the smallest timing units to the utterance as a whole. Each of the prosodic constituents – mora, syllable, foot, word, phrase, etc. – can potentially circumscribe the domain of a phonological rule (cf. Selkirk 1980, 1981a, Nespor and Vogel 1986, Hayes 1989a, inter alia). Of particular importance as operational units of tone sandhi are the metrical foot, the phonological word, phonological phrase, and intonational phrase. In this chapter we consider the metrical foot as the domain of tone sandhi. Subsequent chapters will tackle the larger prosodic units.

There is an emerging consensus among students of Chinese phonology that tone sandhi is intimately related to stress. Kennedy (1953), A. Hashimoto (1980), Yip (1980:163ff.), Wright (1983:156ff.), Shih (1986), Z. Zhang (1988), Chan and Ren (1989), Hsiao (1991), Qian (1992:613ff.), Chang (1992:214ff.), Ao (1992, 1993), Chan (1995), and especially Duanmu (1991, 1992a, 1993a, 1995) all make the explicit claim that tonal stability and the domain of tone association are keyed to the metrical structure, in terms of which they seek to explain a wide range of sandhi phenomena. In particular, many Wu dialects share a common feature: within a stress foot, all unstressed syllables lose their underlying tones, while the stressed syllable extends its lexically associated tone to the rest of the metrical unit.

Paradoxically, despite its pivotal role in determining tonal behavior, stress has proven to be frustratingly elusive acoustically and perceptually. Therefore, before we go any further, we need to address the question of

the phonological status of stress in Chinese. This I will do in the sections 1–2. I will then proceed to use Shanghai as a canonical case study to illustrate how prominence and the stress foot determine sandhi behavior (section 3). Crucially, only the prosodic unit foot – derived by the familiar means of stress assignment, stress clash resolution, and optional stress reduction – can render an insightful account of the tonal pattern in Shanghai.

Shanghai represents the canonical form of a fairly straightforward sandhi system of stress-sensitive tonal deletion and spread. In the next chapter we will see how we may profitably extend the metrical approach to the analysis of more complicated sandhi systems.

1 The phonological status of stress in Chinese

Crucial to our circumscription of the prosodic domain of tone sandhi is the metrical unit "foot," which entails the notion of stress, the linguistically significant prominence around which the metrical unit is organized. But the phonological status of stress in Chinese is controversial, to say the least. In this section I will first point out the phonetic elusive nature of stress with particular reference to Chinese, and suggest various reasons why it is nonetheless "real."

1.1 The phonetic elusiveness of stress

Take Beijing Mandarin. There are, to be sure, well-known minimal pairs such as

(1)	a.	*ma tou*	214.35	"horse's head"
			214.o	"pier"
	b.	*guo nian*	51.35	"to celebrate New Year"
			51.o	"next year"
	c.	*dong xi*	55.55	"east and west"
			55.o	"things"
	d.	*di dao*	51.51	"tunnel"
			51.o	"genuine"
	e.	*sheng qi*	55.51	"to get angry"
			55.o	"vitality"

"o" = atonic, stressless

where the first member consists of two fully toned syllables [T.T], while in the second, the last syllable is atonic [T.o]. The second syllable in the

[T.o] expressions exhibits a marked degree of reduction in terms of tone shape, duration, and vowel quality. No one would quibble with labeling [T.o] words as trochaic, or assigning them a [SW] (strong-weak) metrical representation. But this does not per se establish the phonological status of stress pattern that is independent of tone: the accentual contrast is redundant, given the primary and independently needed tonal distinctions.[1] Moreover, the level of confidence in the [T.o] vs. [T.T] contrast varies from speaker to speaker. In fact, even dictionaries differ on such matters (see W. Li 1981 and particularly C. Chen 1984 for extensive discussion). For instance, the three sources indicated below prescribe three different normative pronunciations for the doublet *da.yi* "gist, main idea" vs. *da.yi* "careless":

(2) dictionaries:

		Xiandai Hanyu	*Guoyu Ribao*	*Practical English–Chinese*
da.yi	"gist, main idea"	[51.51]	[51.o]	[51.51]
da.yi	"careless"	[51.o]	[51.o]	[51.51]

Even more problematic are the [T.T] expressions, with two tone-carrying (tonic) syllables. W. Li (1981) appears to be the first to argue that Chinese is a free-stress language in the same sense as English or Russian: any tonic syllable can attract stress. His position has received support from Yin (1982) and Hoa (1983). Yin divides the 2,100 frequently occurring disyllabic compounds in his corpus into three prosodic types:

(3)	a.	iambic	[T.T]	(W.S)	67+ %
	b.	trochaic	[T.T]	(S.W)	17+ %
	c.	trochaic	[T.o]	(S.o)	14+ %

On this view, Beijing has a three-way contrast that parallels English *raccoon* (W.S), *proton* (S.W), and *atom* (S.o). Here are some triplets gleaned from Hoa (1983, mostly Appendix III, pp. 225–227):

(4)	a.	*gang bi*	55.214	W.S	"pen"
		cu lu	55.214	S.W	"rude, crude"
		xin shui	55.o	S.o	"salary"

[1] In this sense, Yip (1980:57, 84) maintains that tone determines stress, rather than vice versa.

b.	*nu li*	214.51	W.S	"to strive, be diligent"
	gu li	214.51	S.W	"to encourage"
	ben shi	214.o	S.o	"skillful, resourceful"
c.	*bei guan*	55.55	W.S	"pessimistic"
	le guan	51.55	S.W	"optimistic"
	huang gua	35.o	S.o	"cucumber"

We have already seen minimal pairs showing a [T.T] vs. [T.o] contrast. Minimal pairs of the [W.S] vs. [S.W] sort reported by Yin (1982) and Hoa (1983) include:

(5)	a.	*san bu*	51.51	W.S	"to take a walk"
				S.W	"to propagate"
	b.	*gong ji*	55.55	W.S	"to attack"
				S.W	"cock"
	c.	*gong shi*	55.51	W.S	"offensive" (noun)
				S.W	"business, public affairs"
	d.	*pei ke*	35.51	W.S	"to entertain guests"
				S.W	"someone invited to accompany the guest of honor"

It is fair to say that native speakers' judgments on the relative degrees of stress – that is apart from the tonic vs. atonic distinction – are notoriously elusive, and certainly far less clear-cut than the English analogs such as *súbject* vs. *subjéct* or *Whíte House* vs. *white hóuse*. Y-R. Chao rejects precisely this putative contrast between fully toned trochees and iambs on account of "the difficulty of obtaining agreement among native speakers of Peiping [= Beijing]" (Chao 1968:38). Hoa (1983) who, more than anyone else, has investigated in depth the accentuation of the Beijing dialect, freely admits that "En général, les Pékinois non avertis n'ont pas conscience de l'existence d'accents forts et faibles dans leur parler" (p. 7.).[2] Gao-Shi (1963:38) took a more radical stance, and flatly declared that "Chinese has no word stress."

To further confuse this state of affairs, several, often diametrically opposite, hypotheses have been proposed concerning stress assignment. Here are some of the current proposals regarding accentuation in (Beijing) Mandarin:

[2] "Generally speaking, untutored Beijing speakers are unaware of the strong vs. weak accentual contrast in their speech" (my tr.).

A. Uniformly right-prominent
Chao (1968), Yip (1980),[3] M. Lin (1983), M. Lin et al. (1984),[4] Hashimoto (1980 [1987])

B. Free stress: predominantly iambic; some lexically marked trochees
W. Li (1981), Yin (1982), Hoa (1983)

C. Phrases mostly iambic; root compounds indeterminate
Kratochvil (1964, 1967, 1969, 1974)

D. Basically trochaic; iambs result from Trochaic Reversal, Phrase-final Lengthening, etc. Chang (1992)

E. Non-head stress
Duanmu (1990a,[5] 1992a, 1993a, 1995), Duanmu and Lu (1990)

F. Stressability hierarchy: 51 (>) 55 > 35 > 214[6]
Meredith (1990)

G. Indifferent: no lexical stress
Gao and Shi (1963), T. Lin (1985), Du (1988), Duanmu (1993a)[7]

Setting aside for the moment subjective judgments and hypotheses based on purely formal or phonological considerations, let us briefly summarize the experimental findings of M. Lin et al. (1984) and Kratochvil (1969, 1974). Both tested only disyllabic expressions. M. Lin et al. reported that the majority of subjects (six out of eight) judged the second syllable to be more prominent in 89.8% of the cases, and found duration to correlate

[3] Yip conjectures that Beijing Mandarin is right prominent at the phrase-level, but does not have a distinctive word-level stress.

[4] Acoustic and perceptual data limited to disyllabic expressions.

[5] Based on the author's native Chengdu dialect.

[6] That is, a high-falling tone is more stressable than a high-level tone, which in turn is more prominent than a rising tone, and so forth. A similar tone-dependent stressability (HL > ML > H > others) is documented in Ayutla Mixtec (Pankratz and Pike 1967). Du (1988:202f.) shows that, in Taiwanese, high-register tones are perceived as "louder" than low-register tones.

Meredith (1990) not only based his observations on, but also drew inspiration from, Hoa (1983), who has observed the tendency for high (falling) tones to be perceived as more prominent than low (dipping) tones. Hoa noted the following contrast, among others (p. 81):

i.	*la* [*xiao.ti.qin*]	"play violin"
	55. **214**. 35. 35	tone
	2 x x 1	stress (1, 2, x = primary, secondary, tertiary stress)
ii.	*la* [*da.ti.qin*]	"play violoncello / cello"
	55. **51**. 35. 35	tone
	x 2 x 1	stress

[7] Duanmu (1993a:7) states, with respect to all what he calls "M-languages" – namely Mandarin-like prosodic systems – "with respect to stress, there is . . . no distinction among full syllables [i.e. excluding toneless syllables]. Any full rime can be stressed, and no full rime type is more likely to be stressed than another type."

best with the hearers' judgments (the average durations of the first and second syllables are 291 and 323 msec respectively). Since Lin et al.'s sample consisted exclusively of disyllabic expressions (mostly compounds) elicited in isolation in response to "How do you say this?" – pointing to a flash card with Chinese characters written on it – the longer duration of the second syllable may reflect nothing more than the effect of phrase-final lengthening (cf. Lehiste 1980, Selkirk 1984:297f. and copious references cited therein).

Kratochvil (1969, 1974) took a different approach. His data base consisted of 200 disyllabic expressions excised from continuous speech. He found that the index number most closely corresponding to stress judgments is the product of overall fundamental frequency multiplied by overall amplitude multiplied by time. In other words:

$$V = \frac{F_O \times A \times T}{1{,}000}$$

where V = syllabic volume, Fo = fundamental frequency (in Hz), T = time (in msec), and A = amplitude (measured in units of 1/200 in. of the distance by which the amplitude curve deviates from the baseline).[8] The product of Fo, A, and T is divided by 1,000 in order to obtain manageable numerical expressions. He then calculates the proportion (P) of the syllabic volumes of disyllabic constructions by:

$$P = \pm\frac{V_1}{V_2}$$

where V_1 and V_2 stand for the syllabic volume of the first and the second syllable respectively. Using the P values of ±2.0 as the cut-off points, he set up three accentual categories: Trochaic ($P \geq +2.0$), Iambic ($P \leq -2.0$), and Indeterminate (P = between +2.0 and –2.0). Based on the distribution of the 200 disyllabic expressions in his corpus among these three accentual patterns, Kratochvil concluded: (i) phrases are basically iambic; (ii) expressions consisting of a root morpheme (free or bound) plus a bound morpheme (e.g. *mai.le* "buy + aspect," *da.de* "big + subordinator") are typically trochaic; while (iii) most of the compounds consisting of lexical root morphemes (e.g. *cheng.jiu* "accomplishment," lit. "success + attainment") fall in the indeterminate category.

[8] More precisely, Fo and A represent the averages of measurements sampled at six equidistant points over the duration of the syllable.

Other Chinese dialects are equally elusive in terms of stress patterns. For instance, Selkirk and Shen (1990:315) observe that although left prominence has been commonly assumed for Shanghai (e.g. in Yip 1980:164f.; Wright 1983:156), it is supported neither by native-speaker intuition, nor by such phonetic cues as vowel reduction or shortening.[9] To date, Du (1988) represents the most extensive experimental study on tone and stress in Chinese, with special reference to Taiwanese (closely related to Xiamen). Not surprisingly, she found that the only reliable phonetic cue to perceived loudness or prominence is pitch level, consistent with a long line of experimental studies conducted since Fry (1955, 1958) and Bolinger (1958).[10] Thus, more often than not, her subjects picked high-toned syllables [55, 52] as more prominent than low-toned ones [33, 13, 21] (70–79% vs. 21–35%; see figures on pp. 202 and 221). She concluded "there is no independent stress in Taiwanese [i.e. independent of tone]" (p. 222).

1.2 On justifying metrical structures

Actually, the elusiveness of *phonetic* stress is by no means unique to Chinese. Hayes (1995:5f.), for instance, points out that stress has no consistent physical correlate, manifests itself differently in phonetic terms from language to language, is abstract in nature, and is not readily accessible to consciousness (cf. Beckman 1986 for a detailed and critical survey of the phonetic literature on stress). This does not mean that metrical prominence cannot be diagnosed by phonological means such as intonation, rhythm rule, vowel quality, and the interaction of accent with other segmental rules (e.g. flapping and aspiration, as in *a[D]om* vs. *pro[t^h]on*).

Whether or not it is justified to postulate certain metrical structure, even in the absence of clear-cut and phonetically invariant cues, ultimately turns on what sort of phenomena such (putative) structure is able to explain. Bantu languages illustrate this point well. In these languages, a H tone often comes from one location, for instance, the verb root, but surfaces in quite another – typically landing on the penult (as in Chizigula),

[9] Fig. 6 of Zee and Maddieson (1980:59) shows the audio waves of disyllabic sandhi forms. Discounting the CVN vs. CVq distinction, the two syllables are roughly equal in duration. Xu et al. (1988), by far the most detailed and authoritative description of Shanghai, is noteworthy in its total silence on stress.

[10] Unlike previous studies, which typically treat duration and (peak) amplitude as independent attributes, Beckman (1986:188, fig.7.3) demonstrates that total amplitude, i.e. amplitude integrated over the entire duration of the syllable, is a better predictor of accent (72%) than Fo (63%) at least for monolingual speakers of a stress-accent language like English. Her approach is similar in spirit to Kratochvil (1969, 1974) discussed above. Curiously, Beckman made no mention of Kratochvil's experimental works.

or antepenult (as in Nguni languages; see Downing 1990, Kenstowicz and Kisseberth 1990, Kisseberth 1991, cf. Bickmore 1995 and references cited therein). Since penult and antepenult positions are common stress sites, quite naturally this tonal distribution has led linguists to posit a trochaic foot (possibly with an extrametrical syllable) on the right edge of the phonological word in these languages. The metrical prominence in turn attracts the H, regardless of what this H is originally associated with (see Odden 1988, 1995 for a survey of related phenomena). **Lamba**, a Bantu language spoken in Zambia and described in Bickmore (1995), provides us with a slightly more complex and particularly striking example. To greatly oversimplify matters in the interest of clarity, Lamba verbs exhibit the following morphological structure:

(6) Pref $[_{L1}$ Pref + ROOT-Ext-FV

Pref = prefix
FV = final vowel
Ext = extensions
L1 = Level-1 stem

The verbal root is preceded by two types of prefixes: object and certain tense/aspect markers belong to a Level-1 stem, while the subject, negative, etc. pertain to Level-2 morphology. Verb roots as well as both types of prefixes may or may not contribute a H tone. But crucially, where the H actually shows up may be quite different from where it originates. In the examples that follow, H-contributing roots or prefixes are given in bold characters. Roots are capitalized. Syllables which dock the high tone have a H written underneath them. The opening bracket "[" separates Level-2 prefixes (to the left) from Level-1 prefixes (to the right).[11]

(7) H-root

a.	tu-[la-[**LEEMB**-e H	we-habitual-write-FV "we write" (7a)
b.	tu-[ka-mu-**LEEMB**-el-a H	we-fut-him-write-applicative-FV "we will write to him" (21)
c.	tu-[ka-luku-**LEEMB**-e H H	we-fut-progr-write-FV "we will be writing" (19b)
d.	tu-[ka-luku-mu-**LEEMB**-el-a H H	we-fut-prog-him-write-applicative-FV "we will be writing to him" (25a)
e.	tu-[**LEEMB**-e H	we-write-FV "that we write" (subjunctive) (8d)

[11] Parenthesized numbers refer to examples in Bickmore (1995).

(8) H-prefix

a.	**ta**-tu-[le-KOM-a H	neg-we-progressive-hurt-FV "we are not hurting" (16b)
b.	**ta**-tu-[luku-mu-KOM-a H H	neg-we-progressive-him-hurt-FV "we are not hurting him" (22a)

(9) No H-morpheme

a.	tu-[la-KOM-a	we-habitual-hurt-FV "we hurt" (7a)
b.	tu-[luku-KOM-a	we-progr-hurt-FV "we are hurting" (7e)

It is clear that a H tone may originate either in a root like *LEEMB* "to write," or in a negative prefix like *ta-*. However, the source and the eventual resting place of the H need not coincide. Specifically, the H tone contributed by the root *LEEMB* docks on the Level-1 prefixes occupying odd-numbered positions (cases 7a–d) or – in the absence of such prefixes – on the root-initial syllable, by default (case e). (8a, b) show the same distribution of the Hs, except that the H originates in a prefix rather than the root. Needless to say, when there is no H-toned morpheme, the verbal complex surfaces with a L tone (unmarked) throughout, as in (9a, b).

The highly regular patterns of alternating Hs strongly suggest a rhythmic effect. Specifically, by circumscribing the string of Level-1 prefixes as the domain over which trochees are constructed from left to right,[12] we create stress peaks on the odd-numbered syllables – giving rise to a "rhythmic accent" pattern, to use Hyman's (1989) term. Prosodically prominent syllables attract the H tone if there is one.[13] As noted before, a H tone that is not attached to a stressed syllable docks on the root by default.[14]

[12] Allowance being made for monosyllabic degenerate feet.

[13] Similar alternating patterns of H-toned syllables/moras have been documented in Makua (Cheng and Kisseberth 1979–1981, Kenstowicz and Kisseberth 1979), Shona (Pulleyblank 1986, Sietsema 1989), and Kirundi (Hyman 1989, Goldsmith and Sabimana 1989). For a general survey and appraisal of the tone/accent issue concerning African languages, see Odden (1988, 1995).

[14] Bickmore accomplishes the same thing in a more formal fashion, by placing a right parenthesis on the right of the root-initial syllable à la Halle and Idsardi (1995). In other words, the stress domain of Lamba is:

$$\ldots [_{L\text{-}1} (\sigma \ldots [_{Rt}\ \sigma)] \ldots$$

(10)

```
        (x  .) (x  .)
. . . [L1 (σ σ  σ σ) Root . . .
          \ | /
            H
```

In the Lamba case (as in Bantu languages generally), the justification for metrical structure is chiefly "positional," that is to say, the property of attracting tone to a particular location or locations. In particular, it shows that tone and stress (i.e. metrical organization) are two independent properties that may coexist within the same language and interact in a systematic fashion (cf. Goldsmith 1981, 1988, Hyman 1989, Hyman and Katamba 1992): we need tonal specification to ascertain whether or not there is a H tone somewhere within a verbal complex; in addition, we need stress to determine where exactly this H eventually shows up. In this sense, the motivation for positing metrical structure in Lamba is purely formal, not phonetic. The question of invariant perceptual cues for stress does not arise.

In the same spirit, we must look elsewhere in order to motivate *phonological* stress or metrical prominence. The fundamental insight implicit in virtually all studies on tones in context is that tonal behavior is diagnostic of metrical prominence in that tonal stability is a characteristic of accentual prominence, while tonal modification, neutralization, or complete loss typically affect syllables in metrically recessive positions. It is well known that in many languages such as Fasu (Hyman 1978), Pame and Trique (Hollenbach 1984), full-fledged tonal contrasts are realized only on accented syllables. For instance, Trique is famous for its five-level-tone system. Any one of the five level tones may fall on the final syllable, which bears the main stress. Elsewhere, tone is entirely predictable: a low-mid [2] before high {1, 2}, and mid elsewhere, before {3, 4, 5} (5 = high, 1 = low). In this respect, tone resembles other phonological features, including phonation type, vowel quality, and so forth. Mock (1988:205) reports that the three-way distinction between plain, glottalized, and laryngealized (creaky voice) vowels is lost on unstressed syllables. The reader need not be reminded of the correlation between vowel reduction and stresslessness in English (recall the contrast between *at*[ə]*m* and *prot*[ɔ]*n*). In this connection, Beckman (1986:104) notes that the atonal (i.e. neutral-toned) syllables in Mandarin Chinese "are a precise phonological analogue of reduced syllables in English and define the comparable level in the hierarchy" (i.e. of metrical prominence). Linguists have also long noted the phenomenon of stress-related tonal stability.

For instance, Maddieson (1978a:341) observes that tones associated with stressed syllables are marked and play a dominant role in assimilatory processes in the sense that tone under stress generally remains unchanged, while tones in metrically weak positions tend to assimilate to the more prominent tone.

Taken together, the tone sandhi and other stress-sensitive phenomena we will detail in the rest of this chapter make a strong case for invoking metrical structure in Chinese.

2 Stress-sensitive tonal phenomena

In this section, as a prelude to the main discussion of the stress foot as a prosodic domain for tone sandhi rules, I will review some of the stress-related tonal phenomena in Chinese.

2.1 Loanword phonology

First of all, stress can be reinterpreted as tonal contrast. **Cantonese** best illustrates this point. Consider the Cantonese loanwords of English origin given below (from Cheung 1986; cf. A. Hashimoto 1972:101f.):

(11) Cantonese

a.	*ha:i.fa:i*	H.H	"hi-fi"
b.	*si.ta:m*	L.H	"stamp"
c.	*pou.fei*	M.H	"buffet"
d.	*ok.lön*	H.L	"Oakland"
e.	*p'a.si.wik*	M.H.L	"Pacific"
f-i.	*p'a.t'i*	H.L	"party"
f-ii.	*p'a.t'i*	H.MH	(alternative pronunciation)

It is clear that Cantonese encodes English stress syllables by means of a high tone (a). By contrast, unstressed, pretonic syllables carry a low tone, if epenthetic (b), otherwise a mid tone (c). Post-tonic syllables carry the expected default low tone [L] (d, e, and f-i), or else a rising tone [MH] (f-ii). At first blush, the high-rising tone on the post-tonic syllable may come as a surprise. Notice, however, that there is a regular tone change (see chapter 1, section 4.4.2) that turns a final low to a high rise in Cantonese, as exemplified by:

(12) Tone change

a.	*wɔ:ng.sik*	L.Hq	"yellow"	(lit. yellow + color)
b.	*tʌ:n.wɔ:ng*	M.L	"yolk"	(lit. egg + yellow)
	→	M.MH		

where *wɔ:ng* "yellow" changes from its underlying L in (12a) to a rising MH in (12b). This explains why the post-tonic syllable surfaces with a high-rising tone.[15]

(13)	pár.ty	English
	H	stress coded as H
	H. L	Default L
	H. MH	by regular tone change

Likewise, (Beijing) **Mandarin** treats English stress and the lack thereof as high and low tones. C. Cheng (1968) argues, for example, that the so-called T3 Sandhi is best stated as

(14) L → MH / ___ L

which neatly accounts for the following contrast:

(15)	a.	*hao proféssor*	"good professor"
		L L H L	
	→	MH L H L	
	b.	*hao stúdent*	"good student"
		L H L	

The accentual pattern of *professor* is tonally interpreted as L.H.L; it therefore triggers tone sandhi when juxtaposed to a low-toned syllable like *hao* "good" (15a). *Student*, on the other hand, does not provide the right sandhi environment, being construed as having a H.L pattern.

Tone and stress clearly have their independent status in Chinese and English respectively. They happen to intersect in a relatively peripheral corner of loanword phonology. But can tone and stress coexist in the same language, and interact in a more systematic way? The answer appears to be yes. For evidence we turn to Tangxi and (Beijing) Mandarin.

2.2 *Tangxi: compound vs. phrase stress*

As we pointed out earlier, by far the most important diagnostic of prominence is the retention of tone and the ability to determine the shape of a neighboring tone. By the same token, a syllable in weak position tends to undergo tonal modification, reduction, or loss – modulo vowel merger, reduction, devoicing etc. of stressless or unaccented syllables.[16]

[15] I owe this tone change account to Hashimoto (1972:101f.). For further details on Cantonese loanword phonology, see Cheung (1986), Yip (1990), and Silverman (1992).

[16] Vowel reduction (e.g. English) and partial merger (e.g. Russian) are well known. Vowel devoicing in weak position has been reported in Southern Paiute (Halle and Clements 1985) and Chatino (Kenstowicz and Kisseberth 1979, citing Gleason 1955).

The northern Wu dialect of **Tangxi**[17] instantiates a canonical case of stress-dependent tonal distribution. This Wu dialect, one of the earliest reported in the tonological literature (Kennedy 1953), distinguishes two major types of constructions, A and B, by and large corresponding to lexical and phrasal constructions respectively. Kennedy stated without elaboration that in disyllabic expressions "It is phonetically demonstrable that in A the louder stress is on the first syllable, in B on the second" (Kennedy 1953:372) – very much like compound vs. nuclear stress in English. Crucially, we owe Kennedy the original insight of linking the dichotomy between A and B sandhi tone patterns to the contrast between left and right accentual prominence of the two construction types. Specifically, only stressed syllables retain their original tone; recessive syllables uniformly lose their tonal contrasts. Surviving tones spread *rightward* to span over the entire lexical unit; syllables preceding the nuclear stress carry a low-level tone by default. This analysis, pre-dating the advent of autosegmental and metrical phonology, is exceptional in both simplicity and elegance. Some examples will illustrate the essential aspects of his analysis. In representing stress, I will adopt the following convention: parentheses enclose the prosodic unit (foot, p-word, and so forth); the gridmark (x) indicates the head, while a dot (.) functions as a placeholder for the non-head constituents (cf. Hayes 1995). Thus the one-line notation of (16a) abbreviates the more conventional two-line representation of (16b).[18]

(16) a.
```
(x  . )
 σ  σ
```
b.
```
x
(x  x)
 σ  σ
```

(17)
```
      (x   .)
a.    tseo.sea          "disinfectant" (lit. smelly water)
      LH. HL            base tone
      LH.  o            tone deletion
      L    H            tone association

       (x   .)
b.    kwen.sea          "boiling water"
      HL   HL
      HL    o
      H     L
```

[17] Kennedy (1953) transcribed this place name as *Tangsic*. Tangxi lies about 30 km due north of the provincial capital Hangzhou.
[18] In Kennedy's transcription, "e" stands for schwa.

```
      (x    .    .)
c.    bhao.moa.then          "racecourse"
      LH   M   LH
      LH   o    o
      L    o    H

      (x    .    .    .)
d.    pao.hyie kong. s[19]   "insurance company"
      HL HL HL HL
      HL  o   o   o
      H   o   o   L
```

Stress falls uniformly on the leftmost syllable of the compounds in (17). The rising (LH) or falling (HL) tone melody of the first syllable then extends over the entire compound, which can be as long as four or more syllables. We implement this tone spread by an edge-in association: link the tone segments to the left- and right-most syllables. We then interpolate the pitch curve and superimpose it over the toneless syllables (marked "o") in medial positions. This edge-in tonal association with pitch interpolation is quite common among Wu dialects – documented for instance for Wuxi (Chan and Ren 1989; cf. Yip 1989, 1995) and a variety of Shanghai reported in Sherard (1980).

In contrast to lexical compounds, phrasal constructions exhibit a right-prominent pattern – with automatic consequences for their tonal behavior. Both (18a) and (18b) are construed as verb + object constructions. Tone deletion of stressless syllables applies as usual. However, since tones associate only rightwards, the initial syllables remain toneless, and carry a default pitch level, lower than the mid-level tone M by one half to a whole note, which Kennedy (1953:372) calls the "suspense tone".

```
(18)  (.   x)
a.    yao zoe          "to row a boat"
      LH. LH
       o  LH

      (.     x)
b.    bhao moa         "to race horses"
      LH   HL
       o   HL
```

The structure-dependent right vs. left prominence produces interesting surface tonal contrasts. Here are some examples that have become the staple of tonological literature:

[19] Here "s" stands for a syllabic spirant.

(19)		(. x)	phrase stress
	a.	*qea nyin*	"to love people"
		LH. LH	base tone
		o LH	tone deletion
		(x .)	compound stress
	b.	*qea-nyin*	"sweetheart (loved person)"
		LH. LH	base tone
		LH. o	tone deletion
		L H	tone association
(20)		(. x)	phrase stress
	a.	*tsao vae*	"to fry rice"
		HL LH	base tone
		o LH	tone deletion
		(x .)	compound stress
	b.	*tsao-vae*	"fried rice"
		HL LH	base tone
		HL o	tone deletion
		H L	tone association

2.3 *Mandarin: iambic reversal*

Despite its phonetic elusiveness, the accentual pattern of (Beijing) Mandarin lurks just beneath the surface, and exerts a subtle but detectable influence on the tonal phonology of Mandarin.

Chao (1968:27f.) reports what he describes as "[a] tone sandhi of minor importance," which changes a rising MH to a high-level HH (= H), as illustrated below, where the affected tone segment (M) is underlined, and the vertical shaft indicates the sandhi change in question. This process is sometimes known as **T2 Sandhi** (MH = Tone 2 in Mandarin, see section 4.1 of chapter 1).

(21) a. [*bo luo*] *kuair* "pineapple cubes"

H. <u>M</u>H. HL
↓
H

b. [*tian wen*] *tai* "observatory"

H. <u>M</u>H. MH
↓
H

c. *shei* [*neng pao*] "who can run?"
MH. <u>M</u>H. L
↓
H

The phonetic motivation for this process, stated as (22) by Hyman (1975:94)[20] is quite transparent. Basically, M is raised when flanked by two Hs, a familiar "plateauing" or "bridging" effect attested in many Bantu languages (cf. Carter 1980, Furere and Rialland 1985, Chen 1992a).

(22) T2 Sandhi
H. <u>M</u>H. T
↓
M

What appears at first to be utterly mysterious is the requirement that the raising context (H___H) must be followed by another tone. This is demonstrated by the inapplicability of T2 Sandhi in the disyllabic expressions such (23a, b), where the asterisk marks an unattested sandhi form. Why should the presence of a tone determine whether or not the process applies to a non-adjacent tone segment? In other words, the rule violates the locality condition (cf. Goldsmith 1976a, Clements 1985, Steriade 1986, Pulleyblank 1986, McCarthy 1989, Archangeli and Pulleyblank 1994). The mystery is compounded by the fact that the rule blocks when another syllable precedes the raising context, as illustrated by (23c), due to Z. Zhang (1988:76).

(23) a. *su lian* "the (former) Soviet Union"
H. <u>M</u>H
↓
H*

b. *tong xue* "schoolmate"
MH.<u>M</u>H
↓
H*

c. [*lao jin*] [*chan jiu*] "Lao Jin craves wine"
MLH.H.<u>M</u>H.MLH
↓
H*

[20] Except that Hyman used the tone letters: 35 → 55 / {55, 35} ___ T.

All this begins to make sense when we take rhythmic structure into consideration. Under the assumption that words and phrases are basically right-prominent (cf. Chao 1968, Hoa 1983), strings of two and three tonic[21] syllables are expected to exhibit the following patterns:[22]

```
(24)      (.  x)
     a.    σ  σ
           su.lian                "the Soviet Union"

           x
          (x)  (.      x)
     b.    σ   [σ      σ]
           shei.neng.pao          "who can run?"

                   x          x
          (.  x)  (x)        (x  .)  (x)
     c.  [σ  σ]   σ     →   [σ σ]    σ
           bo.lo.kuair
           "pineapple cube"
```

Cases (a) and (b) are unremarkable. Case (c) entails a classic stress clash, which Standard Mandarin resolves by resorting to what is functionally equivalent to Iambic Reversal in English (Hoa 1983, esp. pp. 42–82; cf. Liberman and Prince 1977, Selkirk 1984, Hayes 1984, inter alia). Under these assumptions, two and three syllable expressions have these prominence patterns, regardless of internal structure:

```
(25)      2 1          1 = strongest
          2 3 1
```

On this point, virtually all students of Chinese phonology are in agreement. Chao (1968:35) asserts that in both trisyllabic compounds and phrases, the strongest stress falls on the last, the secondary accent on the first, and the weakest prominence on the medial syllable. Hoa (1983:73) states categorically that trisyllabic expressions, whether right- or left-branching, "ne peuvent jamais recevoir un sommet accentual sur la deuxième syllabe, quel que soit le contexte accentuel" [can never receive an accentual peak on the second syllable, regardless of the accentual

[21] I.e. discounting neutral toned syllables.

[22] If we equate stress with tonicity, i.e. tone-bearingness, each of the tone-bearing syllables like *su* "Soviet" and *lian* "Union" would constitute a separate foot. Therefore, strictly speaking, *su.lian* would have the following presentation:

```
(.    x)
(x)  (x)
su.  lian
```

contrast]. Assuming this stress pattern, we can account for the facts by restricting the T2 Sandhi rule to apply only to metrically weak positions – which is exactly the suggestion made in Yip (1980:291).[23] This metrical condition immediately explains why T2 Sandhi never applies to the last tone, given the right-prominent rhythmic pattern of Mandarin Chinese. Hence the obligatory presence of an extra tone to the right of the T2 Sandhi rule.

But, what about the failure of T2 Sandhi in the case of (23c)? As shown in (26), the syllable *chan* "to crave" does occupy a metrically weak position, hence should be able to undergo the raising rule, but doesn't, at least at normal tempo, as reported in Zhang (1988:76). The most reasonable answer is to say that the tone-raising rule is restricted to the minimal rhythmic unit (MRU) at normal tempo.[24] There is a strong tendency across Chinese dialects to group syllables into units of two and three, a minimal word effect, where the binarity condition is defined over the syllables (cf. McCarthy and Prince 1986). This means that while each of the trisyllabic expressions of (21) constitutes a single MRU, the quadrisyllabic string of (26) forms two MRUs. As a result, the requisite tonal string H.MH straddles two separate prosodic units.

(26)
```
                    x
(.    x) (.     x)
[lao jin] [chan jiu]        "Old Jin craves wine"
L.  H. MH. LH
        ↓
        H*
```

Let us consider one last case. We expect (27) to exhibit the same rhythmic pattern as (26), with the second syllable in a metrically strong position. Curiously, T2 Sandhi applies with perfectly natural output. Why?

(27)
```
                    x
(.    x)  (.    x)
[qing hua] [da xue]         "Qinghua University"
H.  MH  HL.MH
    ↓
    H
```

[23] Although Yip has in mind the allegro rule that turns all non-final rising tones into high level tones.

[24] I defer discussion on the nature of MRU until chapter 9.

The answer lies in an insightful observation made by Hoa (1983:79). According to her, "Plus le syntagme est lexicalisé, plus il est usuel, plus le schéma aura tendance à avoir les deux accents les plus forts éloignés l'un de l'autre, formant une sorte d'arc accentuel" [the more the expression becomes lexicalized and conventional, the more the rhythmic pattern tends to place the two strongest accents far apart from each other, forming a sort of accentual arc], as diagrammed below:

(28) "Arc accentuel"

```
                x
x               x
x ( . . . x . . . ) x
```

This penchant for putting a maximal distance between the two highest rhythmic peaks is kindred in spirit to one of the eurhythmic principles proposed by Hayes (1984:52):

(29) Phrasal rule
"A grid is more eurhythmic if its second highest level bears two marks, spaced as far apart as possible."

To return to the Chinese example at hand, a highly lexicalized expression like (27) undergoes Iambic Reversal not on account of stress-clash, but in order to achieve the optimal accentual contour. This accounts nicely for the application of T2 Sandhi in (30) but not in a phrasal construction of low frequency like (26).

(30)

```
                    x
 (x      .)  (.    x)
[qing hua] [da xue]        "Qinghua University"
 H.  MH  HL.MH
     ↓
     H
```

Crucial to our account of T2 Sandhi is the postulation of Iambic Reversal, either triggered by stress clash or motivated by the eurhythmic target of a cup-shaped accentual contour SWS.[25] What evidence do we have that Iambic Reversal indeed occurs? First of all, unlike in the case of disyllabic expressions, there is intersubjective agreement on the prominence profile of polysyllabic constructions. Both Y-R. Chao (1968) and Hoa (1983), to date the most extensive studies of Beijing Mandarin

[25] Yip (1980) suggests there is no word-level stress, only phrase-level stress: [2 (3) 1].

accentuation, assert that greater prominence falls on the peripheral syllables of polysyllabic expressions, regardless of internal structure, as shown below, where 1 = primary accent.

(31) a. 2 3 1 (Chao, Hoa)
b. i. 2 3 3 1 (Chao)
ii. 2 3 3 1 ~ 3 2 3 1 (Hoa)
1, 2, 3 = primary, secondary and tertiary stress

Assuming a uniformly right-prominent prosodic system, this intersubjectively verified accentual pattern is unaccountable for without Iambic Reversal:

(32)
```
        x              x
(. x) (x)   →   (x .) (x)
 3 2   1          23   1
```

The intersubjective consensus is corroborated by cross-dialectal evidence. In an experimental study of Taiwanese, Du (1988) asked his subjects to pick the syllables that sounded louder or more prominent. The results (p. 199) are given below: 90% of the judgments favor prominence on the initial or the final syllable, or both.[26]

(33) Taiwanese (Du 1988:199)

	percentage of responses	percentage subtotal
x o o	44	
o o x	38	90
x o x	8	
o x o	10	10

x = louder

Beyond Chinese, Iambic Reversal has well-known analogs in English; and a similar pattern has been documented for French as well (Hoa 79). Note in passing that peripheral prominence is consistent with the "delimitative function" of stress: accentual prominence serves to demarcate the endpoints of a prosodic unit (Garde 1968, Trubetzkoy 1969, Hyman 1977, Beckman 1986).

[26] Quadrisyllabic expressions exhibit more scattered accentual patterns, presumably owing to the tendency of splitting 4-σ strings into bisyllabic units.

Finally, there is a subtler argument. Recall that in (24c) the normally iambic nominal *bo.luo* must turn into a trochee via Iambic Reversal in order to undergo T2 Sandhi. We can show that this has indeed taken place by deductive reasoning. Iambic Reversal in grid terms is implemented by a *horizontal* beat movement. In other words, to alleviate stress clash, Iambic Reversal can move the offending beat of (34) (marked in boldface) not just to any position, for instance to position 2 – with or without the added gridmark (x) to comply with the "Continuous Column Condition" (Hayes 1995:34) – but laterally to the next column of the same height (position 1).

```
(34)                x                   x
                  **x** x               **x**  x
            x     x  x          x (x) x  x
    a.      x x   x  x   →  b.  x  x  x  x   = *
            1 2   3  4          1  2  3  4

                                            x
                                **x**       x
                                x      x x
                         →  c.  x  x  x  x   = ok
                                1  2  3  4
```

With this conventionally accepted premise in mind, we can demonstrate that Iambic Reversal must apply to expressions like [*bo luo*] *kuair* "pineapple cubes" (24b) by embedding it in a larger construction like (35).

```
(35)                        x
                      **x**  x
                 x    x     x
              x  x    x     x   x
    a.  [[*bo luo*] *kuair*] [*guan tou*]        "a can of pineapple cubes"
        pineapple cube can

                            x
                 **x**      x
                 x    x     x
              x  x    x     x   x
→   b.  [[*bo luo*] *kuair*] [*guan tou*]        = *
```

Looking at the grid from the top down, the stress clash holds between *kuair* and *guan*. Horizontal beat movement would land the offending x on *luo*, which yields an unacceptable reading (35b).

Instead, the sole acceptable reading (36c), is derivable only by first moving the beat on *bo* horizontally to *luo*, in the process creating the canonical [2-3-1] subgrid on *bo luo kuair* as in (36b), then "fixing" the

stress clash on the pentasyllabic compound as a whole by moving the beat on *kuair* in parallel fashion to *bo*.

(36) a. = (35a)

```
                    x
            x       x
    x       x       x
    x   x   x       x   x
```

→ b. [[*bo luo*] *kuair*] [*guan tou*]

```
                    x
    x               x
    x       x       x
    x   x   x       x   x
```

→ c. [[*bo luo*] *kuair*] [*guan tou*]

```
H.MH
  ↓
  H
```

To the extent that native speakers can judge the differential degree of acceptability between (35b) and (36c), there is no question about the reality of Iambic Reversal on *bo luo kuair* "pineapple cubes" or *tian wen tai* "observatory" (= 21b). In both cases, the stress pattern is corroborated by the applicability of T2 Sandhi.

3 Shanghai: stress-foot as sandhi domain

In Tangxi, discussed in section 2.2, we see the first intimation of a number of recurrent themes, notably, tone-to-stress attraction and (directional) tone spread or reassociation. In these respects, Shanghai, also a northern Wu dialect, but much more accessible, far better documented and more thoroughly analyzed, represents variations on the same basic themes. More importantly, it argues more strongly for a metrical approach to tone sandhi in the following sense: as far as Tangxi is concerned, we could in principle simply tie tone sandhi rules directly to the morphosyntactic structure, so that only the initial syllable of the rightmost word within a phrase retains and extends its lexical tone rightwards to the entire compound. Such an account, awkward as it is, falls short of the full range of sandhi facts we know about Shanghai. The fundamental reason lies in the non-isomorphism between morphosyntactic and prosodic constituents. Specifically, the prosodic constituents are derivable only by means of prototypical rules of metrification, stress clash resolution, and optional stress reduction, to be made precise below.

In the presentation that follows I will draw on Zee and Maddieson (1980), Jin (1986), Xu, Tang et al. (1988), Selkirk and Shen (1990) and particularly Duanmu (1991, 1992a, 1993a, 1995) for both data and analysis.[27] Shanghai has a five-tone system:

(37)	A	53	HL	*tɔ*	"knife"
	B	34	MH	*tɔ*	"island"
	C	23	LH	*dɔ*	"peach"
	D	5q	Hq	*toq*	"sincere, earnest"
	E	12q	LMq	*doq*	"to read"

Zee and Maddieson (1980) supply detailed phonetic data on both citation tones and the pitch envelopes of polysyllabic sandhi forms.

Various attempts have been made to reduce the Shanghai tonal categories (see chapter 1, section 3). For now, I will sidestep the issues surrounding the inventory of underlying tonal categories, and take the citation tones as given, in order to focus on the interplay between tone and metrical prominence.

3.1 Word/compound stress

At the word level, Shanghai displays a highly regular, left-prominent rhythmic organization, which in turn conditions tone deletion and circumscribes the domain of tone spread (cf. Yip 1980, Wright 1983). It lends further evidence for the link between tonal stability and metrical prominence as determined by the familiar principles of metrical organization that characterize prototypical stress-accent systems. More importantly, the stress-like property of prominence is further demonstrated by processes that operate on formal objects such as the metrical grid, notably the mechanism for resolving stress clash. At the heart of the analysis are the following metrical rules, variations of which have been proposed by Duanmu (1991, 1992a, 19s93a).

(38) a. Morpheme level:
Line 0: trochee, left to right, ignore degenerate foot
Line 1: left-headed, unbounded stress
b. Word/compound level: Assign cyclic left-headed stress
c. Phrase level: Assign cyclic right-headed stress
d. Stress Reduction: Optionally delete Line 1 stress
e. Clash Resolution: Remove the stress column next to a higher column

[27] Cf. Sherard (1972, 1979, 1980), Jin (1995), Hung (1987a:115f.), Z. Zhang (1988:168f.), Xu et al. (1981–1983).

Cyclic stress assignment subsumes Stress Equalization Convention:

(39) Stress Equalization Convention
When two or more constituents are conjoined into a single higher-level constituent, the asterisk columns of the heads of the constituents are equalized by adding asterisks to the lesser column(s). (Halle and Vergnaud 1987:265)[28]

In the absence of morphosyntactic structure, for instance, in polysyllabic proper names of foreign origin, trochees are constructed from left to right, ignoring degenerate feet, consistent with Hayes (1995), who allows degenerate feet only in strong positions.[29] Once the metrical pattern is determined, tonal distribution works pretty much as in Tangxi: only stressed syllables retain their underlying tone segments, which are reassociated one-to-one with the stressless syllables from left to right (rather than edge-in, as in the Tangxi case). Otherwise toneless syllables take on a default low (simply annotated here as "o"). All examples given immediately below are transliterations of foreign names, so that the rhythmic grouping is uninfluenced by the morphological structure. Metrical constituents are enclosed in parentheses.

(40)

	(x .)	
a.	*pa. li*	"Paris"
	(HL.o)	tone deletion
	(H. L)	tone association
	(x . .)	
b.	*ni. k'aq. sõ*	"Nixon"
	(LH. o o)	
	(L. H o)	
	x	
	(x .) (x .)	
c.	*qo. ta. li. ya*	"Australia"
	(MH.o)(LH.o)	
	(M. H)(L. H)	
	x	
	(x .) (x . .)	
d.	*ku. r. pa. tço.fu*	"Gorbachev"[30]
	(HL.o) (HL. o o)	
	(H. L) (H. L o)	

[28] Cf. revised version in Hayes (1995:378).
[29] I will assume for our purposes that the stray monosyllable is adjoined to the left to form a ternary foot.
[30] [r] is syllabic.

Take *ni.k'aq.sõ* "Nixon" (b) for illustration. The left-to-right metrical organization leaves a monosyllable *sõ* unfooted. Only the initial stressed syllable retains its underlying tone /LH/. These tone segments re-associate in a familiar one-to-one, left-to-right fashion, leaving a final toneless syllable (annotated as "o"), which is phonetically realized with a default low pitch.

The accentual properties of the rhythmic organization in Shanghai are demonstrated by the familiar stress clash and its resolution. This is illustrated by the following contrast (from Duanmu 1993a:21, exs. 40a, b):

```
(41)        x                      compound stress
           (x  .)  (x  .)          word stress
     a.    [qo.yã] [tço.zɤ]        "professor Ou-yang"
           (HL.o) (MH.o)           tone deletion
           (H.  L) (M.  H)         tone association

            x
           (x) – (x  .)            stress clash
     b.    [wã] [tço.zɤ]           "professor Wang"
           (LH) (MH.o)

            x                      clash resolution
     →     [wã] [tço.zɤ]
           (LH  o   o)             tone deletion
           (L   H   o)             tone association
```

Crucially, in Shanghai as well as in Tangxi, word/compound stress is cyclically assigned on the leftmost constituent. The uniformly left-prominent rhythmic structure generates a "perfect grid" in the case of example (a), but creates a classic stress clash (indicated by "–"), with two abutting stress columns, in the case of example (b). This stress clash is alleviated by removing the stress column next to a higher one – in exactly the same way the second syllable *bra* is destressed in the English word *àbracadábra*; analogously, the first syllable *ba* of *banána* is accentually demoted.[31]

```
(42)                x
            x       x           x
            x – x . x   .       x   .  .  x   .
     a.     a.bra.ca.da.bra  →  a.bra.ca.da.bra

             x
            x – x .          →     .  x  .
     b.     ba.na.na             ba.na.na
```

[31] By appropriately named Abracadabra and Banana/Bandana rules. In Selkirk's (1984) analysis, the initial syllables receive their prominence either by the Initial Basic Beat or via the euphonic Beat Addition (right to left) rules. See Hayes (1982) for an equivalent tree-based account of the relevant facts of English.

The destressing of *tço* in (41b) renders the constituent *tço.zɤ* "professor" headless, which is therefore rebracketed with *wã* "Wang" to form a ternary foot. The stress alternation exhibited by *tço.zɤ* entails automatic consequences in its tonal behavior. Specifically, the de-stressed *tço* loses its underlying tone, and becomes the target of tone spread.

Certain facts of Shanghai tone pattern that may appear at first to be quite surprising lend further support to the metrical approach outlined above. The examples given below are taken from Duanmu (1995:229, 232):

(43)	a.	*ka. li. foq. ni. ya*	"California"
		HL.LH.MH.LH.LH	base tones
		(HL.o) (MH. o o)	tone deletion
		(H. L) (M. H o)	tone association
	b.	*nø. ka. li. foq. ni. ya*	"Southern California"
		LH.HL.LH.MH.LH.LH	base tones
		(LH. o. o) (MH. o o)	tone deletion
		(L. H. o) (M H o)	tone association

(a) is unremarkable: a left-to-right footing yields the binary + ternary pattern. On the other hand, (b) calls for a ternary + ternary rhythmic organization. Word-initial ternary unit is unexpected, given the left-to-right binary footing. However, under the assumption of cyclic foot formation, clash-induced destressing and rebracketing makes exactly the right prediction, as illustrated below:

(44)	x	word stress
	(x .)(x . .)	foot stress
	ka. li. foq. ni. ia	"California"
	x	compound stress
	(x)(x . . . .)	word stress
	(x)(x .)(x . .)	foot stress
→	*nø. ka. li. foq. ni. ya*	"Southern California"
	x	
	(x . .) (x . .)	Clash resolution
→	*nø. ka. li. foq. ni. ya*	

Once destressing and rebracketing have created the two ternary stress domains, tone deletion and association operate as usual to produce the observed sandhi forms indicated in (43b).[32]

[32] Instead of creating ternary feet (at the expense of enlarging the foot inventory), Kenstowicz (1995:423ff.) leaves *li* and *ya* unparsed: (*nø. ka*) *li.*(*foq. nyi*) *ya*, but makes the same point that foot construction (or its OT equivalent) must be cyclic.

Note that *qo.ta.li.ya* "Australia" has two alternative pronunciations, one with two stress feet or tonal domains, as in (45a), the other with one single tonal domain:

(45)		x	
		(x .) (x .)	
	a.	*qo. ta. li. ya*	"Australia"
		(MH.o)(LH.o)	tone deletion
		(M. H)(L. H)	tone association = (40c)
		(x . . .)	Stress Reduction (optional)
→	b.	*qo. ta. li. ya*	
		(MH.o o o)	tone deletion
		(M. H o o)	tone association

The alternative reading (b) arises from optional Stress Reduction stated as (38d), which has the effect of conflating several stress feet into one. Stress Reduction accounts for alternative readings in compounds as well:

(46)		x	compound stress
		(x .) (x .)	word stress
	a.	[*baq. çoq*] [*kõ. tsï*]	"Snow White" (lit. white snow + princess)
		LH. MH. HL.MH	base tones
		(LH. o) (HL. o)	tone deletion
		(L H) (H L)	tone association
		(x . . .)	stress reduction
→	b.	[*baq. çoq*] [*kõ. tsï*]	
		LH.LH HL. LH	
		(LH o o o)	tone deletion
		(L H o o)	tone association
(47)		x	compound stress
		(x .) (x)	word stress
	a.	[*<u>h</u>õ. moq*] *zã*	"redwood bed"[33]
		LH LH LH	
		(LH o) (LH)	tone deletion
		(L H) (LH)	tone association
		(x . .)	stress reduction
→	b.	[*<u>h</u>õ. moq*] *zã*	
		LH LH LH	
		(LH o o)	tone deletion
		(L H o)	tone association

[33] <u>h</u> = voiced h.

3.2 Phrasal prominence

The tone–stress connection is equally robust at the phrasal level. Here are a few examples taken from Xu et al. (1988):

(48)	x	phrase stress
	(x .) (x)	word stress
a.	*lʌ.lʌ zã*	"to stretch out" (lit. to pull + long)
	HL.HL.LM	base tone
	(H. L) (LM)	tone deletion and tone association
	x	
	(x .) (x)	
b.	*k'uɛ.dɤ du*	"have a good physique" (lit. build, physique + big)
	MH.LM.LM	
	(M. H) (LM)	
	x	
	(x .) (x)	
c.	*dø.dø tsø*	"to turn round and round" (lit. circles + turn)
	LM.LM.MH	
	(L. M) (MH)	
	x	
	(x .) (x .)	
d.	*bong dɤ sɛ fʌq*	"unkempt" (lit. disheveled head + messy hair)
	LH.LH.MH.Hq	
	(L. H)(M. Hq)	

Tone deletion and re-association apply within the stress-feet in the usual fashion (for simplicity, hereafter we collapse the two processes into a single step). Two minor details deserve some comment. Stress reduction is typically restricted to lexical expressions; it does not usually apply to phrases. Thus, while (*h̲õ.moq*)(*zã*) "redwood bed" optionally undergoes stress reduction and coalesces into one stress domain (*h̲õ.moq.zã*), as illustrated above, phrasal constructions (*lʌ.lʌ*)(*zã*) "to stretch out" (V + resultative complement), (*k'uɛ.dɤ*)(*du*) "build + big" (subject + predicate), (*dø.dø*)(*tsø*) "to turn round and round" (adverb + verb), and (*bong dɤ*)(*sɛ fʌq*) "disheveled head + messy hair" (coordinate construction) normally constitute separate rhythmic units. The second point worth making is that in connected speech the non-final tones are further neutralized in the direction of a middle tone. Thus Xu et al. uniformly annotates the second syllable in all the examples of (48) with an undifferentiated M tone. I will regard this as an intonational phenomenon, and omit it from the sandhi form, in accordance with the general practice adopted by Duanmu, Selkirk and Shen, and Jin, among others.

Function words and grammatical particles are inherently stressless (cf. Selkirk 1984).[34] Since the basic rhythmic unit, the foot, is left-prominent, these syllables cliticize to the left as the weak member of an (unbounded) foot. To illustrate, these examples are taken from Selkirk and Shen (1990:321f.) and Jin (1986:44):

(49)	x	phrase stress
	(x .) (x .)	foot/word stress
a.	*zï laq zã̃ he*	"live in Shanghai"
	LH LM LH MH	base tone
	(LH o) (LH o)	tone deletion
	(L H) (L H)	tone association
	x	
	(x . .) (x)	
b.	*tɔ qiq pe zo*	"pour a cup of tea"
	MH MHq HL LH	
	(MH o o) (LH)	
	(M H o) (LH)	
	x	
	(x . . .) (x)	
c.	*paq ngo yiq pəng sï*	"give me a book"
	MH. LH. Hq.MH. HL	
	(MH. o o o) (HL)	
	(M H o o) (HL)	

The preposition *laq* "in, at," pronominal objects *ngo* "me," and the number + classifier expression *qiq pe* "one cup" or *yiq pəng* "one volume" are unstressed in connected speech; the stressless syllables then adjoin the prominent syllable to the left to form a left-headed foot (the only well-formed foot in Shanghai). Both *laq* and *qiq* acquire a high tone via tone association within the domain circumscribed by the stress foot.

Now consider the contrast exemplified by the classic doublet:

(50)	x	compound stress
	(x) (x)	word stress
a.	*tsʼɔ.vɛ*	"fried rice" (compound)
	MH.LH	base tones
	(x .)	clash resolution
→	*tsʼɔ.vɛ*	
	(MH. o)	tone deletion
	(M. H)	tone association

[34] For a detailed discussion of the function words – their classification, membership, and tonal behavior in Shanghai – see Jin (1986:35–66).

	x	phrase stress
	(x) (x)	word stress
b.	*tsʼɔ vɛ*	"to fry rice" (phrase)
	(MH)(LH)	
	(H) (LH)	tone simplification (see below)
	(. x)	
→ b′.	*tsʼɔ vɛ*	= *
	. (x)	
→ b″.	*tsʼɔ vɛ*	= *

In the compound (50a), Clash resolution triggers destressing and tone deletion with concomitant re-association of the surviving tones in the now familiar pattern, resulting in [M.H] as the sandhi form. Being a phrasal construction, (50b) carries the nuclear stress on *vɛ* "rice" to the right by virtue of nuclear stress; hence we have a perfect mirror image of (a). (b) also instantiates a case of stress clash. If Clash resolution were to apply symmetrically, removing in this case the lower grid column on the verb *tsʼɔ* "to fry", it would either produce (b′), if the unstressed syllable is rebracketed to form a right-dominant iamb, or else leave *tsʼɔ* unfooted, as in (b″). The first option violates a fundamental WFC on foot form in Shanghai, which allows exclusively left-headed feet, while the second alternative runs counter to the Exhaustivity condition, which requires that syllables must be parsed into metrical feet (Halle and Vergnaud 1987:15–16). Under these circumstances, argued Duanmu (1993a, 1995), stress clash is tolerated.[35] We will return to the question of asymmetric stress clash in chapter 8, section 2; suffice it for now to note that the verb *tsʼɔ* "to fry" retains its underlying tone /MH/, which undergoes simplification to emerge with [H].

How does tone simplification work? Let us examine additional examples of phrasal construction taken from Xu, Tang et al. (1988).

(51)	x	phrase stress
	(x) (x .)	word stress
a.	*tʼø* [*çiɔ.li*]	"penny-wise" (lit. seek + small gains)
	(HL)(MH.LH)	
	(HL)(M. H)	tone deletion/association
	(H) (M. H)	tone simplification

[35] In OT terms, this means Exhaustivity and Left prominence constraints dominate Clash Avoidance. Another alternative entertained by Duanmu (1995) is to add a mora at the phrase boundary to separate the verb from the object NP in (50b), thereby alleviating stress clash.

```
            x
      (x)   (x)
b.    su    mən          "to lock the door"
      (MH)(LH)
      (H)  (LH)

            x
      (x)   (x)
c.    dzi   i            "to quit smoking"
      (LH)(HL)
      (M)  (HL)

            x
      (x)   (x)
d.    dzüq  tçin         "to dig a well"
      (LMq)(MH)
      (Lq)  (MH)

            x
      (x)   (x  .)
e.    ziq   [ʌ.çi]       "to pick up garbage"
      LMq.LH.MH
      (LMq)(L. H)
      (Lq)  (L. H)
```

It is clear that weakly stressed syllables may not carry contour tones: all contour tones are reduced to a simple, level tone: H, M, and L. Furthermore, the distribution of H, M, and L is entirely predictable: H for high-registered tones; M or L for low-registered tones depending on whether they fall on smooth or checked syllables. In other words, tones in weak position no longer distinguish between rising and falling contours, both being neutralized to a simple level tone. This is summarized in the following table:

(52)

	strong position		weak position	
	smooth σ	checked σ	smooth σ	checked σ
high register	HL		H	Hq
	MH	MHq		
low register	LH	LMq	M	Lq

The phrase-level tone sandhi of Shanghai exemplified an extremely common phenomenon, namely partial tonal neutralization in prosodically weak positions. To summarize, Shanghai distinguishes three degrees of prominence:

(53) a. main stress: may carry a contour tone
b. subordinate stress: may carry only a level tone
c. stressless syllable: toneless (except by tone reassociation or by default)

3.3 Stress-based vs. end-based accounts

The metrical approach sketched above to Shanghai tone sandhi contrasts markedly with the so-called "end-based" alternative advocated by Selkirk (1986) and Selkirk and Shen (1990).[36] The end-based theory holds that prosodic domains can be defined by means of setting basically two parameters: right or left edge of an X^o (word) or X^{max} (maximal phrasal projection). As Duanmu (1992a, 1995) argues convincingly, the metrical analysis has several advantages over the end-based alternative. I will mention only a few. First, the stress-based approach makes an explicit connection between prominence and tone stability, a generalization that would be fortuitous in the "end-based" theory.

Second, the metrical approach offers a better explanation for word-length sensitivity. According to Selkirk and Shen, the left edge of a word (X^o) demarcates the beginning of a prosodic word, which defines the scope of tone deletion and association. This algorithm for deriving prosodic categories should in principle be insensitive to word-length, and is therefore incapable of explaining why a trisyllabic proper name like (*ni.kˈaq.sõ*) "Nixon" forms one single tonal domain, but longer strings like (*qo.ta*)(*li.ya*) "Australia" tend to break up into two or more feet. Generally speaking, the mismatch between lexical words and tonal domains constitutes a problem for the end-based theory. Thus the tonal domain is bigger than the lexical word in (*wã.tço.zɤ*) "Professor Wang", but smaller than the proper name (*ku.r*)(*pa.tço.fu*) "Gorbachev." By marking the left edge of all and only X^o, we wrongly predict *(*wã*)(*tço.zɤ*), and *(*ku.er.pa.tço.fu*).

(54)	x	compound stress
	(x) (x .)	word stress
a.	*wã* [*tço.zɤ*]	"Professor Wang"
	LH.MH.LH	
	(x . .)	Clash resolution
→	*wã* [*tço.sɤ*]	
	(L H. o)	tone deletion/association

[36] Selkirk and Shen (1990:313) entertained the possibility of alternative stress-based analyses, but did not elaborate.

	x	word stress
	(x .) (x . .)	foot stress
b.	*ku. er. pa. tço. fu*	"Gorbachev"
	HL.LH.HL.MH.HL	
	(H. L) (H. L. o)	

Third, a stress-based approach accounts better for structural asymmetry. As pointed out earlier, [*h̲õ.moq*] *zã* "redwood bed" (repeated here) has two alternative readings – in contrast with *h̲õ* [*moq zã*] "red wooden bed", which allows only a one-foot reading.

(55)		x	compound stress
		(x .) (x)	word stress
	a.	[*h̲õ. moq*] *zã*	"redwood bed"
		LH LH LH	
		(L H)(LH)	tone deletion/association
		(x . .)	Stress reduction
→	b.	[*h̲õ. moq*] *zã*	
		LH LH LH	
		(L H o)	

(56)		x	compound stress
		(x) (x .)	word stress
	a.	*h̲õ* [*moq zã*]	"red wooden bed"
		LH. LH. LH	
		(LH)(L. H)	= *
		(x . .)	Clash resolution
→	b.	*h̲õ* [*moq zã*]	
		LH LH LH	
		(L H o)	tone deletion/association

The reading in (55a) is possible, since the monosyllabic *zã* "bed" carries a word-level stress (by virtue of Stress Equalization convention, see above). Since it involves no stress clash, and Stress Reduction is optional, *zã* can retain its status as an independent stress unit. In (56a), on the other hand, uniform compound stress on the left creates a stress clash, which in turn triggers obligatory removal of the offending stress column (in this case on *moq* "wood") with automatic rebracketing. Hence, (56) can form only one single stress/tone domain. There is no mechanism in Selkirk and Shen's analysis to capture this asymmetry.

Finally, the most compelling argument concerns rhythmic effect. H-M. Zhang (1992:266) observed that two monosyllabic phonological words (p-words) may merge into a single one under certain conditions. This is illustrated below.

(57)	(((	phonological phrase (XP, left)
	(((	prosodic word (X°, left)
	kɤ [*ngo* [*lo.pe*]]	
	dog bite boss	"the dog bit the boss"
	(MH)(LH)(LH.MH)	p-word
	(MH. LH)(LH.MH)	p-word merger
	(M H) (L H)	tone deletion, association

Assuming that syntactic words (X°) and phrases (XP) project their corresponding prosodic categories, in accordance with Selkirk and Shen's end-based theory, we expect (57) to have three p-words, hence three tonal domains. However, the first two monosyllabic p-words can optionally merge into one, as demonstrated by the deletion of LH from *ngo* "bite," and the subsequent reassociation of H originally linked to *kɤ* "dog" with the de-toned *ngo* "bite." But what Zhang failed to make explicit is the exact condition under which p-word merger takes place. For instance, Duanmu (1995) points to the contrast between (57) and (58):

(58)		(	((	phonological phrase (XP, left)
		(	((	prosodic word (X°, left)
		[*lo.pe*]	[*ngo kɤ*]	
		boss	bite dog	"the boss bit the dog"
		(LH.MH)	(LH)(MH)	p-words
		(LH.MH)	(LH. MH)	p-word merger
	a.	*(L H)	(L H)	tone deletion, association
	b.	(L H)	(LH)(LH)	

Like (57), (58a) fuses the two monosyllabic p-words *ngo* and *kɤ* into one, but with an unacceptable result: only reading (58b) is attested. Neither Zhang nor Selkirk and Shen have in their account a principled explanation for this contrast, which Duanmu argues persuasively is rooted in the metrical structure of these two constructions:[37]

(59)	x	phrase stress (sentence)
	(x) (x)	phrase stress (NP, VP)
	(x) (x) (x .)	word stress
a.	*kɤ* [*ngo* [*lo.pe*]]	
	dog bite boss	"the dog bit the boss"
	x	
	(x .) (x .)	Clash Resolution
→	*kɤ* [*ngo* [*lo.pe*]]	
	(MH.LH) (LH.MH)	
	(M H) (L H)	tone deletion and association

[37] I modify Duanmu's grids to make them consistent with the convention adopted here.

```
                    x          phrase stress (sentence)
  (x    )    (    x)           phrase stress (NP, VP)
  (x   .)    (x) (x)           word stress
b. [lo. pe]   [ngo kɤ]
  boss       bit   dog         "the boss bit the dog"
  LH.MH.  LH.  LH
  (L    H) (LH)(LH)
```

In (a) stress clash is resolved by deleting the stress column on *ngo* "bite" next to the main stress, in the process creating a well-formed left-prominent foot over *kɤ*] *ngo* "dog bite." By contrast, in (b), *ngo* again carries the offending stress abutting the main prominence on *kɤ* "dog." However, clash resolution by de-stressing *ngo* creates a dilemma: if *ngo* is rebracketed with *kɤ* "dog," it will create an unlawful right-headed foot; on the other hand, if *ngo* is re-metrified together with *lo.pe* "boss" on the left, it will give rise to a non-optimal ternary foot. Under these circumstances, stress clash is permitted.[38]

[38] See chapter 8, section 2.3–2.4 for a different interpretation of a rightward stress clash.

8 *Stress-foot as sandhi domain II*

The Shanghai case we have scrutinized in the preceding chapter represents a prototypical stress-sensitive tonal system in its purity and simplicity, both tonally and metrically. The picture is considerably more complicated in other dialects. Curiously enough, in some cases the tone that survives and spreads is not the one lexically associated with the stressed syllable, as in the case of Shanghai, but some unexpected melodic pattern. The most promising hypothesis is one proposed by Chan and Ren (1989) and (Chan 1995) that involves stress-sensitive tonal substitution followed by stress shift. Two Wu dialects have been analyzed in this light, namely Wuxi (section 1) and Danyang (section 2). The Danyang case is of particular significance in that it clearly demonstrates that stress clash is directionally asymmetric: only a leftward clash (i.e. with the higher grid column on the left) enforces stress removal of the less prominent syllable, while a rightward clash can remain unresolved.

The chapter closes with an extended analysis of Nantong (section 3). Nantong exhibits a more complex, tempo-sensitive rhythmic organization. Furthermore, it is noteworthy for having two sets of domain-specific rules: in addition to the familiar tone deletion and spread within the stress foot, it calls for various tonal modification rules that operate at the edges or within the domain defined by the phonological word.

1 Wuxi: stress shift

The metrical approach to Shanghai demonstrated above promises to provide us with a key to unlock the mystery of other more complex sandhi patterns richly documented in other Wu dialects, for instance Danyang. Before we tackle Danyang, it behooves us to look at a simpler case, that of Wuxi, which turns out to be of great significance in its own right for a number reasons: in particular (i) it further illustrates the intimate connection between prominence and tonal behavior; and (ii) it sheds light on

what appears at first to be an utterly arbitrary and whimsical relationship between base tones and their phonetic forms in sandhi contexts.

In what follows, I rely on Chan and Ren's (1989) insightful paper for both description and analysis of Wuxi (cf. Yip 1989 for fine-tuning). This northern Wu dialect has four basic tonal patterns, labeled A, B, C, and D, as indicated below:

(1) Wuxi tone patterns

	Tone shape	MC
A	LLH	T3, 6
B	HHL	T1, 4, 7q
C	LHH	T5
D	LHL	T2, 8q

T7q, 8q are checked tones
MC = corresponding Middle Chinese tonal categories 1–8
Odd and even numbered tones represent register a (high) and b (low) tones.

Wuxi distinguishes two types of constructions with respect to tonal behavior:

(2) Type I
a. reduplicated verbs
b. verb + resultative or directional complements
c. expressions with number + classifier
d. reduplicated nouns (in baby talk only)

Type II
a. compounds (including regular reduplicated nouns)
b. phrases

Type I correspond to canonically left-prominent constructions independently attested across Chinese dialects, including Beijing Mandarin. For instance, the underlined syllables are all stressless and carry a neutral tone in Beijing: *kan.kan* "take a look" (reduplicated verb), *pa.shang.lai* "climb up hither" (verb + directional complement), *liang.ge ren* "two Cl. people", and *nai.nai* "grandma" (reduplicated noun, baby talk). Not surprisingly, Type I expressions undergo tone deletion and spread, pretty much as in Tangxi (and, with minor modification, Shanghai): only the first – metrically strong – syllable retains its underlying tone and spreads it over the tonal domain. Chan and Ren refer to this sandhi process as Pattern Extension. Specifically, the tone pattern, annotated as a three-tone melody,[1]

[1] Ignoring OCP.

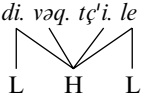
is mapped onto the syllables in an "edge-in" fashion, from opposite ends, with the middle tone segment being associated with all the syllables. This mode of tone mapping is illustrated below:

(3) Pattern Extension

a.	*dɛ.dɛ* LHL.o LH.HL	"to play (e.g. piano)" tone deletion tone spread	(reduplicated verb)
b.	*dɛ.dɛ* LLH.o L.LH	"egg"	(reduplicated noun)
c.	*g∧. ts'in.ts'∧w* HHL. o o H. H. HL	"to do completely"	(do + complete)
d.	*di. vəq. tç'i. le* LHL. o. o. o LH. H. H. HL	"unable to bring up"	(bring + not + up + hither)

Take the last example. Only the leftmost prominent syllable *di* "bring, lift" keeps its underlying tone melody /LHL/. The peripheral Ls are associated with the first and the last syllables respectively. The middle H is linked to each of the syllables in the verb–directional complement construction, creating a rise-high-fall pattern. In other words, the notation of the last line of (3d) is equivalent to (4):

(4)



In sharp contrast to Type I constructions, Type II expressions – the vast majority of lexical as well as phrasal constructions – display a sandhi pattern that appears at first totally random and puzzling. Take the following example:

(5)

a.	*fi* HHL		"to fly"
b.	*fi.* HHL. L.	*tçi* HHL LH	"airplane" (fly + machine) base form sandhi form
c.	*fi* HHL. L.	*tçi.* HHL. L.	*p'î* LHH LH	"air ticket" base form sandhi form

The base tone /HHL/ surfaces in citation form (a). In connected speech, however, polysyllabic compounds take on sandhi forms that bear little

Table 8.1. *Wuxi Pattern Substitution*

1st σ		2nd σ		
		T3, 4, 5, 6	T1, 2	T7, 8
A LLH	T3	C LHH		
	T6	D LHL		
B HHL	T1	A LLH		
	T4	A LLH	D LHL	
	T7q	A LLH	D LHL	B HHL
C LHH	T5	B HHL		
D LHL	T2	B HHL		
	T8q	A LLH	D LHL	

T1–T8 = Middle Chinese tonal categories

recognizable relation to the citation form. In particular, /HHL/ is replaced by /LLH/ as the tonal melody that eventually spreads to the entire compound in the usual edge-in manner. Arbitrary though this paradigmatic substitution may be, it is not random, but systematic. Basically, pattern substitution must take place before pattern extension. Pattern substitution is context-free in some cases (replace pattern X of the initial syllable by Y), and context-sensitive in others (replace pattern X of the initial syllable by Y, if the second syllable carries tone P, and by Z, if the second syllable carries tone Q, etc.). Systematic pattern substitutions are summarized in table 8.1.

Table 8.1 reads as follows: tone patterns of first syllable indicated by the leftmost column are replaced by one or more patterns in the lower-right quadrant, depending on what tone the second syllable carries. Remarkably, what determines the choice of tonal substitutions is not the actual tone shape, but the Middle Chinese tonal categories of the constituent syllables. For instance, pattern A [LLH] is replaced by pattern C [LHH] if pattern A derives from Middle Chinese (MC) T3; if the same pattern A [LLH] comes from MC T6, it is replaced by pattern D [LHL] instead. As noted above, in some cases, one pattern is replaced by two, occasionally three, patterns. The two- or three-way split is conditioned by the tonal properties of the second syllable. But what is interesting is that the second

syllables are partitioned into subsets not definable by their contemporary tone shapes, but in terms of their historical tonal categories. For instance, the set of {T3, 4, 5, 6} defies definition as a natural class in phonetic terms: it comprises LLH, LHH, and HHL (only if HHL comes from MC T4), to the exclusion of LHL and HHL (if HHL derives from ancient T1). In terms of the historical scheme of tonal classification, {T3, 4, 5, 6} belongs to the set of so-called "oblique" tones (*shang* and *qu* in traditional terminology) generally understood to have been contour tones (in contraposition to *ping* and *ru*, hypothetically level tones).

Details aside, Type II constructions undergo one extra step of Pattern Substitution, before tone deletion and spread. The derivation of (5c) proceeds as follows:

(6)	*fi.*	*tçi.*	*p'î*	"air ticket"
	HHL.	HHL.	LHH	base tone
	LLH.	HHL.	LHH	Pattern Substitution (according to table 8.1)
	LLH.	o.	o.	tone deletion
	L	L.	LH	tone spread (edge-in)[2]

The following pair illustrates pattern split conditioned by the second syllable:

(7)	a.	*vu*	*tçi*	"weapon"
		HHL.	LHH	base tone
		LLH.	LLH	Pattern Substitution
		LLH.	o	tone deletion
		L.	LH	tone spread
	b.	*na.*	*dã*	"milk candy"
		HHL.	LHL	base tone
		LHL.	LHL	Pattern Substitution
		LHL.	o	tone deletion
		LH.	HL	tone spread

The point of particular interest is this: as seen in the preceding pair, which pattern is substituted for which is determined by the syllable on the right, which generally keeps its base tone unchanged, while the syllable on the left undergoes tonal modification. This canonical schema of tone sandhi

[2] I ignore one additional detail: the H tone of pattern A [LLH] is linked to the second syllable of a trisyllabic or longer string, if the second syllable is lexically associated with any one of the set of historical tonal categories {T3,4,5,6}. An example would be *ts'iq.tsi.mɛ*, where the middle syllable belongs to T3:

ts'iq.tsi.mɛ	"seven sisters"
LLH. o. o	tone deletion
L. LH. o	tone spread

(8) $T \rightarrow T' \ / \ \_\_\_ \ T$

(where T′ = modified tone) is widespread among Mandarin, Min, and southern Wu dialects, generally regarded as *right-prominent* systems (cf. A. Hashimoto 1987, Duanmu 1991, 1992a, 1993a, 1995). At the point when pattern substitution was in force, tone deletion could not have run its course – otherwise, the tonal categories of the second syllable could not in principle condition the pattern split between (7a) and (7b). On the other hand, once the initial syllable has replaced its tonal pattern, it is the (new) pattern associated with the initial syllable that is retained and extended over the entire string, suggesting a *left-prominent* pattern. Chan and Ren see in this paradoxical state of affairs evidence of a more conservative system in transition from right-to-left prominence, the endpoint of a long evolution already reached by other better known Wu dialects such as Tangxi, Shanghai, and Suzhou, among others.

To recapitulate and make more explicit Chan and Ren's stress-shift theory, here are the hypothetical steps of derivation. Crucially, stress shift must take place between pattern substitution on the one hand, and tone deletion/spread on the other, as illustrated by the last example cited above.

(9)	(. x)	right-prominent
	na. dã	"milk candy"
	HHL.LHL	base tone
	LHL. LHL	Pattern Substitution
	(x .)	left-prominent (by stress shift)
→	*na. dã*	
	LHL. LHL	
	LHL. o	tone deletion
	LH. HL	tone spread

Chan and Ren's insight explains, at least in part, the otherwise puzzling mapping relation between base tones and their sandhi forms. By the same token, to the extent that their analysis crucially depends on assumptions about stress patterns and the historical shift from right-to-left prominence, it argues persuasively for the need to take into consideration prominence patterns in our description of tone sandhi facts.

2 Danyang: asymmetric stress clash

We are now in a position to tackle the more complex and far more interesting case of Danyang. Particularly noteworthy is how tone sandhi

domain in this dialect is determined by *asymmetric* clash resolution, a notion to be made more precise in the following.

First reported in extraordinary detail and analyzed in great depth by Lü (1980), Danyang, a northern Wu dialect, has been the subject of considerable interest (cf. Chen 1991a, 1996; Z. Zhang 1988; Chan 1989, 1991, 1995; Yip 1989, 1995; Duanmu 1991, 1994a; H-M. Zhang 1992). Colloquial Danyang[3] has six citation tones.

(10) Citation tones

	CVN	CVq
level tones	H, M, L	Hq, Mq
contour tone	LH	

CVN = smooth syllable (CV, or CVN)
CVq = checked syllable

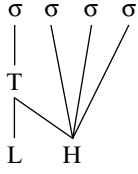
Characteristically, the CVq's only take even (i.e. level) tones. More importantly, Danyang has six word melodies that may be mapped onto polysyllabic compounds and more or less set phrases. Following Chen (1991a, 1996), I will label these word melodies from A to F.

(11) Melodies

		2-σ	3-σ	4-σ
A	H	H.H	H.H.H	H.H.H.H
B	M	M.M	M.M.M	M.M.M.M
C	L	L.L	L.L.L.L	L.L.L.L
D	HL	HL.L	HL.L.L	HL.L.L.L
E	LH	LH.H	LH.H.H	LH.H.H.H
F	HL-LH	HL.LH	HL.HL.LH	HL.HL.HL.LH

Melodies A–E span the entire polysyllabic string by simply extending the rightmost tone segment over the syllables on the right. Thus, melody E extends the H tone rightwards to cover the entire quadrisyllabic string:

(12)

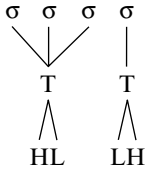


= LH.H.H.H

[3] We will ignore for the most part the literary layer of Danyang.

Melody F requires special treatment. One possibility suggested by Yip (1989) is to stipulate an edge-in association for this particular melody. Rightward spread of the falling contour tone (HL) produces the desired sandhi form.

(13)



= HL.HL.HL.LH

Needless to say, this account is less than satisfactory, on two counts: (i) it calls for a mode of association for melody F that is distinct from the left-to-right directionality for all other melodies; (ii) tone spread targets the tone root (HL) as a whole in (13), but only extends the terminal tone element in (12). I will not go into the details here, but refer the reader to Chan (1995) for a critical comparison of the various solutions to this problem and a promising new proposal.

In the sections that follow, I will first only say a few words about the relationship between the citation tones and the word melodies, then go on to the focal point of our interest, namely the question as to how to circumscribe the domain onto which the word melodies are mapped.

2.1 Deriving word melodies

The relationship between citation tones and word melodies is highly complex, as illustrated by the following examples:

(14)	a.	*hua.sheng*	"peanut"
		M. M	base tone
		HL. L	sandhi form
	b.	*dan.yang*	"Danyang" (place name)
		M. LH	
		H. H	

Again, what holds the key to unraveling the mystery surrounding Danyang tone sandhi is the decomposition of the process into two steps: Pattern Substitution and Tone Deletion/Spread, already instantiated by the related dialect of Wuxi sketched in the preceding section. Specifically, the examples above involve context-sensitive substitutions statable as:

(15) M → HL/___ M
H /___ LH

Note again that at this point pattern substitution presupposes a right-prominent metrical structure, with the tonal categories on the right playing the dominant role of determining the tone shape of the first syllable. However, the roles are reversed after pattern substitution: it is the first syllable that retains and extends its (modified) tone to the syllable(s) on the right. In other words, the oblique and often opaque relation between citation and sandhi forms is mediated by pattern substitution and tone deletion/spread, reflecting a right-to-left prominence stress shift:

(16)	(. x)	(. x)	
	hua.sheng	*dan.yang*	
	M. M	M. LH	
	HL. M	H. LH	Pattern substitution (15)
→	(x .)	(x .)	Stress shift
	hua.sheng	*dan.yang*	
	HL. M	H. LH	
	HL. o	H. o	tone deletion
	HL. L	H. H	tone spread

The rule stated as (15) is only part of a more complex, but discernible pattern of substitutions. For an excellent summary, see Chan (1995). In what follows, I will assume that (i) a particular word melody has been determined in accordance with Pattern Substitution; and (ii) the right-to-left stress shift has taken place. Accordingly, the focus will be exclusively on how a polysyllabic string is cut up into substrings onto which the selected word melody is mapped.

2.2 Determining tonal domains

Danyang tonal domains are determined by exactly the same principles we appealed to in our description of Shanghai. Specifically, each of the word melodies is mapped onto a stress foot; and foot-size units are constructed in accordance with none other than the rules proposed for Shanghai (by Duanmu 1991, 1992a, 1993a), repeated here for convenience:

(17) a. Morpheme level:
Line 0: trochee, left to right, ignore degenerate foot
Line 1: left-headed, unbounded stress
b. Word/compound level: assign cyclic left-headed stress
c. Phrase level: assign cyclic right-headed stress.
d. Stress Reduction: optionally delete line 1 stress
e. Clash Resolution: remove the stress column next to a higher column

It is worth noting that Lü (1980:89) made the point that stress falls on the first syllable of a tonal span, namely the substring associated with a single word melody. Lü's remark lends independent auditory support to the fundamental principle of left-headed metrical organization expressed as (17a, b).

Here are some examples that illustrate how these simple principles combine to predict the attested sandhi forms.

(18) a. x Clash
(x) (x .) → (x . .)
zhen [*hong mu*] *zhen* [*hong mu*] "true red wood"
HL

b. x Reduction
(x .) (x) → (x . .)
[*jin shi*] *yan* [*jin shi*] *yan* "myopia"
(lit. near sight eye)
L LH L
(i) = ok (j) = ok

c. x Reduction
(x) .) (x .) → (x . . .)
[*shang hai*] [*che zhan*] [*shang hai*] [*che zhan*] "Shanghai station"
L H L
(i) = ok (j) = ok

d. x
(x) (x . .) Clash x Reduction
(x) (x .) (x) → (x . .) (x) → (x . . .)
jiu [[*jiao.da*] *che*] *jiu* [[*jiao.da*] *che*] *jiu* [[*jiao.da*] *che*]
"old bicycles"
L M L
(i) = ok (j) = ok

A right-branching expression like (a) entails stress clash, and therefore obligatorily undergoes clash-induced stress deletion, resulting in a single stress foot. As a consequence, the melody HL, as determined by Pattern Substitution, spreads to the end of the word. Optional Stress Reduction accounts for the alternative readings of (b) and (c). Finally, (d) involves both Clash Resolution and Stress Reduction.

Lü (1980:102f.) cites 95 instances of trisyllabic (x .)(x), and 22 of (x . .), suggesting that stress reduction is optional – alternatively, that monosyllabic constituents optionally form degenerate feet. The same author also remarks that the overwhelming majority of quadrisyllabic expressions give rise to two separate feet (x .)(x .) mirroring their bipartite morphosyntactic structure (p. 112).

There is some evidence suggesting that Clash Resolution applies cyclically. Consider the following derivation:

```
(19)        x
            (x) (x    .    .)
            (x) (x)  (x    .)        stress clash
      a.    ye [bai [ ju hua]]       "wild white mums"

            x
            (x    .)  (x    .)       Clash Resolution
→     b.    ye [bai [ ju hua]]
             \/       \/
             M        L              = ?? (unattested)
```

Clash-induced destressing of *bai* "white" yields (b). Since Stress Reduction is optional, we expect (b) to surface as a legitimate sandhi form. However, (b) is unattested in the source. Assuming (b) to be a systematic gap, we can explain its absence by taking clash avoidance to be a well-formedness condition that must be satisfied at every morphological level. In other words, Clash Resolution must apply cyclically, as illustrated below:[4]

```
(20)        x
            (x)  (x    .)           stress clash
      a.    bai [ ju hua]            "white mums"

            (x    .    .)            Clash Resolution
→     b.    bai [ ju hua]

            x
            (x  (x    .    .)       stress clash
→     c.    ye [bai [ ju hua]]       "wild white mums"

            (x    .    .    .)       Clash Resolution
→     d.    ye [bai [ ju hua]]
             \ \ /  /
              \ \/ /
               \ /
                M                    (attested)
```

[4] In OT terms, Kenstowicz (1995) made the similar point that constraints must be satisfied cyclically. See Kenstowicz (1996) for an alternative view in terms of base-identity.

Cyclic Clash Resolution predicts a uniformly right-branching structure like (20) to form a single stress foot, hence carry a single word melody.

2.3 *Directional asymmetry*

The picture becomes somewhat complicated in quadrisyllabic expressions which do not break up neatly into symmetrical halves. Cyclic stress assignment yields (21-i), which entails a clash between the last two syllables. We expect therefore Clash Resolution to de-stress *ji*, the less prominent of the two clashing syllables, yielding (21-j). Instead, (21-i), with three separate stress feet, hence three separate melodies and tonal domains, is the only attested reading.

(21)

```
(x . . .)
(x . .) (x)   Clash      (x          )
(x .) (x) (x)   →        (x . .) (x)
[[tuo la] ji] shou       [[tuo la] ji] shou        "tractor operator"
  \/  |   |                \|/     |
  H   M   H                 H      H
(i) = ok                 (j) = *
```

Examples like this present an obvious problem for the metrical analysis in the spirit of Duanmu (1991, 1992a, 1993a), and prompted H-M. Zhang (1992) to propose, instead, a structure-based account. Specifically, he defines the tonal domain in terms of c-command:

(22) Tonal Domain (H-M. Zhang 1992:192)
α is spread rightward to β iff β is c-commanded by α.

There is some a priori plausibility to Zhang's syntax-based approach. Details aside, what it says is that the melody of α spreads to all elements α is in construction with. C-command simply encodes this semantic relatedness. For instance, *tuo* (21) is in construction with *la*, with which it forms a compound verb meaning "to pull"; on the other hand, *tuo* does not per se enter into direct meaning relation with *ji* "machine." Hence, it makes intuitive sense that the word melody of *tuo* does not extend to *ji*. A fortiori, *ji* is related to *shou* "operator"[5] only indirectly, through the mediation of the larger construction *tuo.la.ji* "tractor." On this account, it is easy to understand why (21) breaks down into three separate sandhi domains.

[5] Lit. hand, as in the English expression "farm hand" and the like.

Unfortunately, Zhang's syntactic account fares no better than Duanmu's metrical analysis. Consider some of the examples given above, repeated below for convenience:

(23) a. x Reduction
(x .) (x) → (x . .)
[*jin shi*] *yan* [*jin shi*] *yan* "myopia"
L LH L
(i) = ok (j) = ok

b. x Reduction
(x .) (x .) → (x . . .)
[*shang hai*] [*che zhan*] [*shang hai*] [*che zhan*] "Shanghai station"
L H L
(i) = ok (j) = ok

c. x
(x) (x . .) Clash x Reduction
(x) (x .) (x) → (x . .) (x) → (x . . .)
jiu [[*jiao.da*] *che*] *jiu* [[*jiao.da*] *che*] *jiu* [[*jiao.da*] *che*]
"old bicycles"
L M L
(i) = ok (j) = ok

The c-command account fails in (a) as well as (b): neither does *jin* "near" c-command *yan* "eye" in (a), nor *shang* c-command *che.zhan* "station" in (b), and yet the L-melody optionally spreads rightwards to the end of both compounds (j-readings). Conversely, in the case of (c), *jiu* "old" does c-command *che* lit. "car", and yet the L-melody fails to spread to the end of the word in one of the two alternative readings, both attested in Lü (1980:109). Ironically, a construction like (c), which Zhang (1992:189) took to be problematic for Duanmu's stress-based analysis, turns out to be a major stumbling block for the c-command account, but quite straightforward in metrical terms: obligatory Clash Resolution predicts reading (c-i), and optional Stress Reduction yields reading (c-j).

Furthermore, a c-command analysis fails by necessity in "flat" structures, such as coordinate constructions and proper names of foreign origin. In each of the following cases, the entire polysyllabic entry is presumably c-commanded by the first syllable, and therefore should form one single tonal domain, counterfactually. From the metrical perspective, the partition into several tonal domains is straightforwardly predictable

from a left-to-right metrification, allowance being made for monosyllabic feet.

(24)

a. (x .) (x)
shang.zhong.xia "top, middle, and bottom"
LH LH

b. (x .) (x)
wei.ni. si "Venice"
HL M

c. (x .)(x .)
mao.li.qiu.si "Mauritius"
LH M

So, we are probably better off sticking with a metrical analysis. But what about example (21)? Why isn't stress clash resolved by the removal of the accent on *ji* in (21-j)? In quadrisyllabic three-plus-one strings, there is a fairly clear-cut contrast between two types of constructions, depending on the internal structure of the trisyllabic element on the left. Example (21), repeated below as (25), instantiates a left-branching first element consisting of [[*tuo la*] *ji*] "tractor." Clash-triggered stress reduction would produce the wrong – at least unattested – sandhi form (25-j); in fact, the rightward clash between *ji* and *shou* is tolerated, with the result that both *ji* and *shou* constitute an independent stress/tone domain (= reading 25-i).

(25) Type A

x			
(x . .) (x)	Clash	x	
(x .) (x) (x)	→	(x . .) (x)	
[[*tuo.la*] *ji*] *shou*		[[*tuo.la*] *ji*] *shou*	"tractor operator"
H M H		H H	
(i) = ok		(j) = * (not attested)	

On the other hand, example (26) illustrates a right-branching first element. In this case, resolution of the leftward clash between the first two syllables, *xiao* "little" and *dong* "east" is obligatory. Failure to de-stress *dong* results in an unacceptable (unattested) reading (26-i).

(26) Type B

```
   x
  (x          )  (x)   Clash      x
  (x)  (x     .)  (x)   →        (x    .     .)  (x)
[xiao [dong.men]] qiao          [xiao [dong.men]] qiao      "Little East-
   |    /\    /\                    \   |   /     /\        Gate Bridge"
   H   HL    LH                      M          LH
(i) = * (not attested)          (j) = ok
```

This *asymmetric* directional stress-clash tolerance is quite systematic. Lü (1980:109) remarks that the "general tendency" is for Type B to form two tonal domains (σσσ) (σ) , while all five examples of Type A reported in the aforementioned article break up into three (σσ) (σ) (σ). Here is a minimal pair: [[*dou.fu*] *gan*] (lit. tofu + dried) and [*hu* [*luo.bo*]] (lit. foreign + radish) instantiate an internal left- and right-branching structure, respectively, embedded in a quadrisyllabic expression. The rightward clash in Type A is left unresolved, while leftward clash obligatorily triggers destressing in Type B.

(27) Type A

```
   x
  (x          )  (x)
  (x    .)  (x)  (x)
[[dou.fu] gan] si          "shredded tofu"
   \   /    |    |
    L      M    M
```

(28) Type B

```
   x
  (x          )  (x)          x
  (x)  (x   .)  (x)   →    (x    .    .)  (x)
[hu [luo.bo]] si         [hu [luo.bo]] si        "shredded carrot"
                            /\    /    |
                           LH          M
```

The correct generalization about the facts, then, is that in Danyang clash resolution is directionally asymmetric, stated as follows:

(29) Asymmetric Clash Resolution
  a. Leftward clash must be resolved by de-stressing
  b. Rightward clash is tolerated

The reader will no doubt object to Asymmetric Clash Resolution as an ad hoc, counter-intuitive artifact of the analysis – at best a merely inductive summary of the observed facts, at worst a desperate attempt to salvage the

stress-based account. One justifiably demands some independent motivation of such a move.

The first thing to note is that this asymmetry is quite systematic in Danyang. Quadrisyllabic expressions can fuse into a single tonal domain (with one single melody), or break up into several, each with its distinctive melody. Lü (1980:88, table 3) lists the following possibilities:

```
(30)        x
     a.   (x . . .)

          x[6]
     b.   (x .) (x .)      (optionally → a, by Stress Reduction)

                  x
     c.   (x . .) (x)

                     x
          x          x
     d.   (x .) (x) (x)

                     x
          x          x
    *e.   (x) (x .) (x)   (obligatorily → c, by Clash Resolution)
```

Lü (1980:89) was careful to note that when the string is cut up into several tonal domains, it is the last syllable that carries the primary accent. The gridmarks of (30) are intended to reflect the impressionistic prominence judgments offered by Lü. Cases (a, b, c) do not involve stress-clash, and are non-problematic. The remaining two patterns (d, e) exhibit very different distributional patterns. Of particular interest here is the systematic absence of pattern (e). The reason is that (e) entails a leftward clash, and obligatorily undergoes Clash Resolution, merging with pattern (c) in the process. Pattern (d), on the other hand, involves a rightward clash, which does not automatically trigger stress reduction. Hence (d) survives as a possible sandhi form.

2.4 *Analogs in Wu and other dialects*

It turns out that this directional asymmetry is quite robust across Chinese dialects. For instance, it has been observed in **Nantong** (to be discussed at length in section 3). According to Ao (1993), the only possible metrification patterns for the quadrisyllabic words in this Wu dialect are:

[6] Lü is silent on whether the first or the third syllable is more prominent. On the basis of examples like (*shang.hai.che.zhan*), I take the primary stress to fall on the first constituent.

(31) Nantong

a. Normal tempo: (x .) (x) (x) (**optionally** → b, by Stress Reduction)

b. Fast tempo: (x . .) (x)

c. Unattested at any tempo: *(x) (x .) (x) (**obligatorily** → b, by Clash Resolution)

Again, pattern (c) is systematically absent, since leftward clash is strictly "verboten," and mandatory resolution is enforced by de-stressing.[7]

Next, recall that **Shanghai** resembles Danyang in the relevant respects. Contrast the following examples.

(32)		x	compound level
		(x) (x .)	word level
	a.	*çiã* [*bi.zɔ*]	"perfumed soap"
		(x . .)	Clash Resolution
		HL.LH.LH	base tone
		(HL. o o)	tone deletion
→	b.	(H L o)	tone association

(33)		x	phrase level
		(x) (x .)	word level
	a.	*tʼϕ* [*çiɔ.li*]	"penny-wise" (lit. covet + small advantages)
		HL.MH.LH	base tone
		(HL)(MH.o)	tone deletion
		(HL)(M. H)	tone association
		(H) (M. H)	(by tone simplification)
		(. x .)	
	b.	*tʼϕ* [*çiɔ.li*]	= *
		. (x .)	
	c.	*tʼϕ* [*çiɔ.li*]	= *

Compound and phrase-stress rules place primary accent on the left in (32), and on the right in (33), giving rise to stress clash in both cases. However, clash triggers stress removal only in (32), not in (33). Why? One hypothesis is that the well-formed foot is uniformly left-prominent. Therefore, by destressing *bi.zɔ* "soap" in (32), we create a left-prominent, albeit ternary,

[7] Left prominence is an overriding constraint also in Nantong, which automatically bars the following: (x) (x .) (x) → * (. x .) (x).

foot. On the other hand, by defooting *tɸ* "to covet" in (33), we either create an ill-formed amphibrach or medially accented foot, as in (33b), or else we leave *tɸ* unfooted, as in (33c), in violation of the Exhaustivity principle, which requires that all syllables must be parsed into feet. Neither of these options is viable. Hence, under this condition, stress clash is tolerated. To couch this solution in OT terms, what we have is a constraint hierarchy:

(34) {Left, Parse} ≫ *Clash ≫ Bounded

Left:	Metrical feet are left-headed
Parse:	Parse syllables into feet
*Clash:	Avoid stress clash
Bounded:	Foot ≤ 2 syllables

Given this constraint hierarchy, we make the correct predictions: in the case of (32), *Clash is satisfied at the cost of creating a ternary foot, in violation of Bounded. (33), on the other hand, cannot satisfy *Clash without incurring heavier penalties, as brought out more perspicuously in the familiar tableau form in (35).

(35)

				Left	Parse	*Clash	Bound (at most 2)
(32)	*çiã* [*bi.zɔ*] "perfumed soap"	a	x (x) (x .)			*	
		b ☞	x (x . .)				*
(33)	*t'ɸ* [*çiɔ.li*] "penny-wise"	a ☞	x (x) (x .)			*	
		b	x (. x .)	*			*
		c	x . (x .)		*		

But this cannot be the whole story. Consider the following contrast first pointed out by H-M. Zhang (1992:266), and discussed earlier in section 3.3:

(36)

	x	phrase stress (sentence)
	(x) (x)	phrase stress (NP, VP)
	(x) (x) (x .)	word stress
a.	*kɤ* [*ngo* [*lo.pe*]]	"the dog bit the boss"

```
                    x
        (x   .)   (x   .)          Clash Resolution
→   b.  kɤ [ngo   [lo.pe]]
        (LH. LH)(LH. LH)           base tone
        (LH.  o) (LH.  o)          tone deletion
        (L     H) (L     H)        tone association

(37)                      x        phrase stress (sentence)
        (x   )    (       x)       phrase stress (NP, VP)
        (x   .)   (x)     (x)      word stress
    a.  [lo.pe]   [ngo  [kɤ]]      "the boss bit the dog"
        boss bite dog
        LH.LH.  LH.  LH            base tone
        (LH. o) (LH) (LH)          tone deletion
        (L   H) (LH) (LH)          tone association

                    x
        (x   .)  .  (x)
→   b.  [lo.pe] [ngo [kɤ]]         = *

                    x
        (x   .    .)  (x)
→   c.  [lo.pe] [ngo [kɤ]]         = *
```

Assuming that lexical and phrasal constructions are left- and right-prominent respectively, consistent with the metrical rules stated in section 2.2, both of the above examples involve stress clash. In the former, (36) clash-triggered defooting and refooting yield the desired sandhi form. In the latter, (37), however, stress clash is left unresolved. Why? Duanmu (1995:255f.) attempts to account for this contrast by arguing that enforcement of bidirectional Clash Resolution would lead to two undesirable consequences, namely either leaving *ngo* "bite" unparsed, in violation of Exhaustivity, as in (37b), or else rebracketing *ngo* with *lo.pe* "boss," thereby creating a less than ideal ternary foot as in (37c). Either way, defooting in no way improves the metrical structure. Under this condition, so argues Duanmu, clash is tolerated.

But this line of reasoning is inconsistent with the constraint hierarchy we have established on the strength of *çiã.bi.zɔ* (ex. 32 in (35)), which unambiguously demonstrates that clash avoidance dominates the boundedness (i.e. at most disyllabic) condition. In other words, creating a ternary foot is a lesser evil than leaving stress clash unresolved. Therefore, given the constraint hierarchy Parse ≫ *Clash ≫ Bounded (34), the expected winner candidate is ex. 37c (marked "$"). The prediction turns out to be wrong. Instead, the attested sandhi form is ex. 37a:

(38)

ex.				Left	Parse	*Clash	Bound (at most 2)
(36)	*kɣ.ngo.lo.pe* "dog bites boss"	a	x (x)(. x) (x)(x)(x .)			*	
		b	☞ x (x .)(x .)				
(37)	*lo.pe.ngo.kɣ* "boss bites dog"	a	☞ x (x) (. x) (x .)(x)(x)			*	
		b	x (x).(x)		*		
		c	x $ (x . .)(x)				*

$ = expected winner candidate
☞ winner candidate

It seems that the only way to resolve the ranking paradox is to split *Clash into two: *Clash-Left and *Clash-Right, with Bounded wedged between the two:

(39) {Left, Parse, *Clash-Left} ≫ Bounded ≫ *Clash-Right

*Clash-Left: Avoid x x (x)

*Clash-Right: Avoid x x (x)

(40)

ex.				Left	Parse	*Clash-L	Bound	*Clash-R
(32)	*çiã.bi.zɔ* "perfumed soap"		a x (x)(x .)			*		
		☞	b x (x . .)				*	
(33)	*tˈɸ.çiɔ.li* "penny-wise"	☞	a x (x)(x .)					*
			b x . (x .)		*			
			c x (x . .)	*				

(40) cont'd

ex.				Left	Parse	*Clash-L	Bound	*Clash-R
(36)	*kɤ.ngo.lo.pe* "dog bites boss"		a x (x)(. x .) (x)(x)(x .)			*		
		☞	b x (x .)(x .)					
(37)	*lo.pe.ngo.kɤ* "boss bites dog"	☞	a x (x) (. x) (x .)(x)(x)					*
			b x (x .) . (x)		*			
			c x (x . .)(x)				*	

This asymmetry is respected across Chinese dialects. Take **Beijing** Mandarin disyllabic expressions. While both [T.T] and [T.o] occur, the third logical possibility [o.T] is systematically absent.[8]

(41) a. T.o *di.dao* "genuine"
b. T.T *di.dao* "tunnel"
c. o.T unattested

T = syllable carrying a full-fledged tone
o = atonic syllable

The southern Min dialect of **Xiamen** displays exactly the same distribution of atonic syllables.

(42) a. T.o *kiã si* "scared to death"
b. T.T *kiã si* "coward, be afraid of death"
c. o.T unattested

One way to account for this cross-dialectal asymmetry is to simply posit two types of structures: left- or right-prominent. In a right-prominent configuration as in (43a), leftward clash enforces destressing, thereby producing a trochaic foot. The completely stressless syllable loses its tone, yielding a [T.o] pattern like (41a) and (42a). On the other hand, a right-prominent rhythmic pattern like (41b) and (42b) involves only a rightward clash, hence both syllables remain stressed, therefore capable of carrying a tone.

[8] Kratochvil (1974) and Wu (1985) make similar observations.

(43)
a. x
(x)(x) → (x .) = T.o

b. x
(x)(x) = T.T

3 Nantong: stress-foot and p-word

Nantong, the last Wu dialect we will examine in connection with metrical structure, displays an astonishingly rich array of sandhi patterns. It invites a close scrutiny in as much as its complex sandhi facts serve as a probe into how speech is prosodically structured. It has two sets of tone sandhi rules, keyed to two distinct prosodic domains, the metrical foot and the phonological word (p-word). Of particular interest are the now familiar rules of Tone Deletion and Tone Spread. They are foot-internal, and therefore dependent on and, by the same token, diagnostic of the metrical foot as a prosodic unit. Nantong exhibits a highly intricate tempo – as well as structure-sensitive rhythmic organization. As we shall see, the underlying principles of rhythmic organization are best expressed in OT-style constraints.

There is a second set of sandhi rules, including various types of tone raising, dissimilation, and default tone assignment. These rules apply at the p-word level, often targeting word-initial or -final positions.

The city of Nantong is located 180 km northwest of Shanghai across the Yangzi river. Geographically, it finds itself wedged between the vast Mandarin landmass to the north and the Wu-speaking Yangzi delta to the south. For facts and analysis of Nantong, I draw liberally on a recent dissertation by Benjamin X. Ao (1993), the first detailed report on this fascinating dialect. I have restated most of Ao's rules. In some cases, the reformulation consists in omitting some technical details that do not bear on the issue at hand. In other cases the revision reflects a different way of looking at the facts and recasts the analysis from a different perspective. However, all examples are cited from Ao (1993) or confirmed by the aforementioned author.[9] All examples are cited in the Pinyin notation, unless otherwise indicated.[10]

[9] Numbers to the right of examples cited in the text refer to the *loci* in Ao (1993). I am indebted to Benjamin Ao for providing me with additional crucial data and for extended discussion on Nantong through email.

[10] I have chosen Pinyin over the IPA symbols because, in addition to tone sandhi, Nantong has a number of segmental rules. IPA transcription incorporating these segmental alternations would unnecessarily complicate our presentation.

I will first formulate the relevant rules and show how they operate (sections 3.1–3.2), deferring until sections 3.3–3.6 the question of how to circumscribe the prosodic domains.

3.1 Foot-level vs. word-level rules

Nantong has a five-tone system: H, L, MH, LM, and HM. There are two sets of domain-sensitive sandhi processes, foot-level and word-level rules. Within a foot (which Ao (1993) symbolizes as ϕ), all non-initial syllables lose their underlying tones, and the initial rising tone is then stretched, accordion-wise, over the entire span of the metrical foot. Otherwise, toneless syllables assume the default value to be specified later. Deletion and Spread are stated below.

(44) Tone Deletion

$T^n \rightarrow \varnothing \ / \ (T \ \_\_\_ \ )_\phi$

(45) Tone Spread

a. Obligatory

$(\sigma \ \sigma^n)_\phi$

{MH}
LM

b. Optional

$(\sigma \ \sigma^n)_\phi$

H

ϕ = foot
σ^n = any string of toneless syllables
ǂ = delinking

Spread (a) has the effect of relinking the terminal tone segment of a rising tone (LM or MH), one to one, left to right, with the second tone segment being multiply linked, if the foot is more than two syllables in length. Since all but the foot-initial tones are lost via Deletion, Spread targets only atonic syllables. Thus, Spread (a) turns (MH.o) and (LM.o) into (M.H) and (L.M) respectively. As usual, the foot is enclosed between parentheses (. . .), and the zero "o" stands as a place-holder for a toneless syllable. Spread (b), which is optional, simply extends the H rightwards to all the syllables within a foot.

The example given below illustrates how Deletion and Spread work together in a straightforward manner.

(46) "man-made lake" = (171)

a. [*ren gong*] *hu*

(MH.L) (MH)	base tones
MH.o MH	Deletion
M. H MH	Spread (a)

	"pardoned him"	= (159)
b.	[*rao le*] *ta*	(pardon + asp + him)
	(MH.o.o)	(*le*, *ta* = toneless enclitics)
	M. H. H	Spread (a)
	"gave"	
c.	*xo tə*[11]	(give + asp) = (161)
	(H. o)	(*tə* = inherently toneless)
	(H. H)	Spread (b)
	"gave"	
d.	*xo tə*	
	(H. o)	
	(H. l)	Default (see below)

Note that Spread can extend over several syllables which are either inherently toneless or lose their tones through Deletion; this is shown in example (b). We will defer to section 3.4–3.6 the question of how the foot is constituted. Optional Spread (b) applies to example (c), yielding (H.H) as output; the alternative reading (H.l) is derived through a Default rule, to be discussed immediately below.

Tone Spread is limited to high-level or rising tones. Hence the falling tone HM does not spread to the diminutive suffix -*er* in (47a, b), which, therefore, remains toneless. Instead, -*er* assumes a default tone, low [l] in p-word-final position (47a), mid [m] p-word-internally (47b). In the examples that follow, the p-word is enclosed in [. . .].

(47)	"rabbit"	= 167a
a.	*tu er*	
	[HM.o]	
	–	Spread
	HM. l	Default
	"rabbit fur"	= 167a
b.	*tu er pi*	
	[(HM.o) (MH)]	
	–	Spread
	HM. m. MH	Default

m, l = default mid and low tones

Recall, too, that Spread is foot-bound, the initial toneless syllables are unaffected. (48a, b) show that the default value is [m] not only in medial, but also in initial positions.

[11] In IPA, not Pinyin.

(48) "on the table" (lit. on + table + Loc) = (157)

a. *zai zhuo shang*

[o. (HM. o)]

– – Spread

m HM. l Default

"(someone) is reading" (lit. at + read + book) = (157)

b. *la-xê*[12] *kan shu*

[o. o. (MH) (L)]

– – Spread

m. m. MH L Default

m. m. MH ML other (see below)

We can, therefore, state the Default rule as follows:

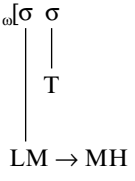
(49) (Tonal) Default
Default tone is [l] in p-word-final position; [m] elsewhere.

What is going on, of course, is that Nantong has a left-prominent foot and that the tones occupying a weak position within the foot are neutralized. The toneless syllables either assimilate to the tone of the prominent syllable (by Spread) or takes on some unmarked pitch value (by Default). Of particular importance from our point of view is the observation that while Deletion and Spread are foot-bound, Default is a word-level rule, the value of a default tone being determined by its position relative to the p-word.

In order to derive the actual sandhi forms of the examples, we need to briefly describe some of the other sandhi processes operating at the p-word level. A set of rules has the effect of raising the low-register tones, /L/ and /LM/. Since /L/ and /LM/ are raised, respectively, to M and MH in overlapping, but not identical environments, it is not possible to collapse the rules into a single schema. It is therefore necessary to state them separately, as follows:[13]

(50) L-Raise $L \rightarrow M \ / \ [\ldots \ \_\_\_ \ T \ldots]_{\omega}$

(51) LM-Raise



[12] *la.xê* marks the Progressive aspect; it has no equivalent in Mandarin.

[13] L-Raise, LM-Raise, and Reg-Dissimilation correspond, respectively, to Ao's Register Deletion (181), Initial Register Deletion (186), and Penultimate Register Dissimilation (190). Ao breaks down register raising into two steps: register deletion, followed by a Register Default rule (166), which fills in the value [–upper] in p-word-final position, [+upper] elsewhere.

(52) Reg-Dissimilation LM → MH / ___ Lr]$_\omega$
ω = p-word
Lr = low-register tone

L-Raise raises all non-final /L/ to M. /LM/, on the other hand, is raised to MH in two contexts: in word-initial position, when followed immediately by another tone-bearing syllable (= LM-Raise); and in p-word-final position by dissimilation to a final low-register tone (= Reg-Dissimilation).[14]

Crucially ordered before L-Raise and LM-Raise is what Ao (p. 118) refers to as Initial Decontour, which I paraphrase as follows:

(53) Initial Decontour
LM → L / $_\omega$[__ {MH, H}

This has the effect of merging /LM/ with /L/ in the specified contexts.

Finally, we need to mention a few other late phonetic rules. There is a general M-insertion rule that turns L and LM into ML and MLM respectively.[15] The derived MLM is simplified to ML except in p-word-final position.

(54) M-Insertion
L, LM → ML, MLM

(55) MLM-Simplification
MLM → ML / __ T

By and large, high-register tones H, MH, HM remain unchanged, except for one minor rule that "decontours" MH into M before MH and H. Note parenthetically that checked syllables also assume high level through a process which Ao calls Obstruent-Induced Decontour and Register Deletion. I will set these details aside.

3.2 *Effects of sandhi rules*

We are now ready to see how these rules produce the attested sandhi patterns of Nantong. Consider the multifarious sandhi forms of just one

[14] L-Raise and Reg-Dissimilation operate exclusively on the tonal tier. Consequently, they are unaffected by intervening toneless syllables. In contrast, LM-Raise requires a sequence of two tones, LM and T, which must be adjacent both on the tonal tier and on the syllable tier; an intervening toneless syllable would therefore block LM-Raise. This explains the two-tier statement of LM-Raise.

[15] Notice that the L that turns into ML cannot be the default L. This can be handled easily by ordering the rule which supplies the default L (and M) last. In a non-derivational – intuitively more explanatory – account, M-insertion functions to make an accented L more salient by creating a falling contour. The accented L is the underlying /L/ occurring in f-initial, metrically strong, position. The default L occurs only in w-final, prosodically weak context.

single morpheme *dian* "electric, electricity," which happens to be highly productive in creating compound words. All examples given in (56) are uttered at normal speed.

(56) "electricity" = (178)

a. [(*dian*)]

LM	
MLM	M-Insertion

"telephone" (lit. electric + speech); = (184)

b. [(*dian*)(*hua*)]

LM LM	
MH –	LM-Raise
– MLM	M-Insertion

"telephone operator" = (179)

c. [(*dian.hua*)(*yuan*)]

LM LM MH	
LM o MH	Deletion
– – –	LM-Raise
L M MH	Spread
ML M MH	M-Insertion

"wire" (lit. electric + thread) = (188)

d. [(*dian*)(*xian*)]

LM. MH	
L MH	Decontour
M MH	L-Raise

"television" (lit. electric + vision) = (178)

e. [(*dian*)(*shi*)]

LM. HM	
– –	Decontour
MH. HM	LM-Raise

"old movies" (lit. old + electric + shadow)

f. [(*lao.dian*)(*ying*)]

H LM H	
H o H	Deletion
H H H	Spread

"power plant"

g. [(*fa.dian*)(*chang*)] (lit. generate + electricity + factory)

HM LM H	
HM o H	Deletion
– – –	Spread
HM M H	Default

In citation form (a), the underlying low-rising tone of the lexical root *dian* "electric, electricity" acquires a complex falling-rising contour via M-Insertion. In disyllabic expressions, each syllable constitutes a separate foot as in (b). In this case, /LM/ is raised to MH by LM-Raise. In trisyllabic compounds, on the other hand, the first two syllables are grouped into a single foot. As a consequence, Deletion, Spread, and finally M-Insertion apply to generate the actual output in (c). The effect of Decontour is illustrated by (d). Decontour fails in (e) because LM is followed by HM rather than MH or H, as required by the rule. Finally, in both (f) and (g), *dian* occurs in a foot-final, weak position, and therefore loses its underlying tone by Deletion. In the former case, *dian* takes on the tone of the preceding syllable via the optional H Spread (b); in the latter, since HM does not spread, the word-internal *dian* assumes the value of a mid-level tone by Default.

What comes under our focus of interest is the observation that all these sandhi rules are domain-sensitive – in other words, in addition to the appropriate tonal sequences, the rules discriminate between domain-initial, internal, or final sandhi contexts. Deletion and Spread are both foot-bound, i.e. limited only to the domain of the metrical foot. Default assigns a pitch value to otherwise toneless syllables depending on the latter's position relative to the p-word: initial, final, or medial. The three register-raising rules all operate on a tonal string circumscribed by reference to the beginning (LM-Raise), end (Reg-Dissimilation), or medial position (L-Raise) of a p-word. To highlight this point, contrast (56e) with (57a) and (57b).[16]

(57) produce electricity factory

a.	[(*fa*)(*dian*)(*chang*)]	"power plant"
	HM LM H	
	– – –	Decontour
	– – –	LM-Raise
	– – –	Reg-Dissimilation
	HM MLM H	M-Insertion
	HM ML H	MLM-Simplification

produce electricity machine

b.	[(*fa*)(*dian*)(*ji*)]	"generator"
	HM LM L	
	– – –	Decontour
	– – –	LM-Raise
	HM MH L	Reg-Dissimilation
	HM MH ML	M-Insert

[16] I am very grateful to Benjamin Ao (p.c.) for supplying the critical examples of (5a, 7b).

The examples of (57a, b) are uttered at a slow tempo, so that each syllable constitutes a separate foot. We shall have considerably more to say about tempo in the next section. For the moment note that in contrast to (56e), *dian* occupies a medial position in (57a). Hence it is not subject to either of the p-word-initial rules of Decontour and LM-Raise. Reg-Dissimilation doesn't apply either, since the last syllable carries a high register tone. Instead, *dian* retains its underlying low register tone /LM/, and undergoes M-Insertion and MLM-Simplification to emerge as [ML] on the surface. Likewise, in (57b) the medial *dian* foregoes Decontour and LM-Raise, but is subject to Reg-Dissimilation, and ends up as [MH].

3.3 Rhythmic effects on tone sandhi

At various points in the foregoing section, I alluded to tempo. In this section I will describe in greater detail the rhythmic effects on tone sandhi. Nantong organizes a string of syllables into prosodic units of feet and p-words in a highly tempo-sensitive manner. Ao (1993:131ff.) distinguishes three tempos: slow, normal, and fast. Putting aside underlyingly toneless syllables for the moment, (58) summarizes the rhythmic patterns of one to five syllable expressions at three different speeds.

(58)

	slow	normal	fast
σ	[(σ)]	same	same
σ σ	[(σ)(σ)]	same	same
σ σ σ	[(σ)(σ)(σ)]	[(σ σ)(σ)]	same
σ σ σ σ	[(σ)(σ)][(σ)(σ)]	[(σ σ)(σ)(σ)]	[(σ σ σ)(σ)]
σ σ σ σ σ	[(σ)(σ)][(σ)(σ)(σ)]	[(σ σ)(σ σ)(σ)]	[(σ σ σ σ)(σ)]

(. . .) = foot
[. . .] = p-word

Putting off until section 3.4 the question of how we derive the prosodic structures of (58), let us examine the far-reaching effects the various tempo-sensitive structures may have on tone sandhi. This point can be made most clearly and strikingly with an example. *Er*, underlyingly /LM/, means "two." However, a five-digit number "2-2-2-2-2" can be read off at increasingly faster speed, with dramatically different tonal patterns, as illustrated in (59a, b, c). Without assuming the rhythmic organization of (58) and the domain-sensitive tone sandhi rules established in sections 3.1–3.2, we cannot even begin to render anything that approaches a reasonable account of the facts set before us.

(59) er- er- er- er- er "2 2 2 2 2"
/LM.LM.LM.LM.LM/

a. slow
[(LM)(LM)][(LM)(LM)(LM)]

MH. LM	MH. LM. LM	LM-Raise
MH. LM	MH. MH. LM	Reg-Dissimilation
MH. MLM.	MH.MH.MLM	M-Insertion
MH. ML.	MH. MH. MLM	MLM-Simplification = (208)

b. normal
[(LM.LM)(LM.LM)(LM)]

LM.o	LM.o	LM	Deletion
–	–	–	LM-Raise
LM.o	MH.o	LM	Reg-Dissimilation
L. M	M .H	LM	Spread
ML.M	M. H	MLM	M-Insertion = (224)

c. fast
[(LM.LM.LM.LM)(LM)]

LM. o	o	o	LM	Deletion
–	–		–	LM-Raise
MH. o	o	o	LM	Reg-Dissimilation
M. H	H	H	LM	Spread
M. H	H	H	MLM	M-Insertion = (229)

In each of the examples cited above, we give a stepwise derivation. How each of the relevant rules operates in each case by now should be self-explanatory.

Needless to say, the same principles apply beyond numerical expressions. To cite just one more example: consider the three renditions of the quadrisyllabic proper name for Bolivia, *bo.li.wei.ya* in Chinese, with the underlying tonal sequence /L.LM.MH.L/.

(60) *bo.li.wei.ya* "Bolivia"
/L.LM.MH.L/

a. slow
[(L)(LM)][(MH)(L)]

M LM	MH L	L-Raise
M MLM	MH ML	M-Insertion
M ML	MH ML	MLM-Simplification = (214)

b. normal
[(L.LM)(MH)(L)]

L. o	MH	L	Deletion
M. o	MH	L	L-Raise
M. o	MH	ML	M-Insertion
M. m.	MH	ML	Default = (224)

c. fast
[(L.LM.MH)(L)]

L.	o	o	L	Deletion
M.	o	o	L	L-Raise
M.	o	o	ML	M-Insertion
M.	m	m	ML	Default

Examples can be multiplied ad infinitum (see Ao 1993, chapter 6). What is worth emphasizing is that the domains which define the sandhi context are obviously not morphosyntactic constituents, but are strictly prosodic in nature. It was in order to show this point that we have chosen to illustrate the rhythmic effect on tone sandhi with morphosyntactically unstructured strings of syllables such as multi-digit numbers and foreign place names. To further corroborate this point, consider the following example. The fact that the initial H extends to the penultimate syllable shows *gei-le ta ben* to constitute a single foot, the domain of Spread. *Gei-le ta ben*, which we gloss as "give + perfective aspect + him + classifier," is patently a non-constituent in terms of morphosyntax.

	give-Perf him Cl. book	
(61)	*gei-le ta ben shu*	"gave him a book"
	H o o o L	
	(H o o o) (L)	Foot construction (see (62b) below)
	[(H o o o)][(L)]	Unbounded p-word[17]
	H H H H L	Spread[18]
	H H H H ML	M-Insertion = (161a)

It goes without saying that one should not conclude from the selected examples that rhythmic organization is syntax-blind, as will become clear in section 3.6.

3.4 Rhythmic organization

In the preceding sections we took for granted the prosodic structures stipulated at the outset of section 3.3. We now turn to the question of what principles lie behind such rhythmic organization. For our purposes, our primary focus will be on foot construction. According to Ao (1993:138),

[17] P-word is bounded by the morphological word, i.e. while a morphological word can break up into several p-words, separate morphological words may not form a single p-word.

[18] High tones spread *optionally* (see Ao 1993:106ff.). Thus, an alternative reading of (61) is [H.m.m.m.ML], where the medial syllables assume the Default [m], instead of assimilating to the initial H.

we first generate the slow tempo patterns through the construction rules given in (62) below, then derive the normal and fast speech patterns through foot- and word-adjunction.

(62) Foot and word construction
- a. Place a unary foot on every underlyingly toned syllable
- b. Link every underlyingly toneless syllable to the first foot on its left
- c. Build right-dominant binary phonological words from left to right
- d. Build a unary phonological word on the remaining foot or adjoin it to the preceding phonological word if possible

Toneless syllables aside, (63) illustrates how the procedural steps spelled out above yield the pentasyllabic patterns spoken in a slow, deliberate manner:

(63) Slow tempo

σ σ σ σ σ	
(σ)(σ)(σ)(σ)	(a)
[(σ)(σ)][(σ)(σ)](σ)	(c)
[(σ)(σ)][(σ)(σ)][(σ)]	(d) alternative (i)
[(σ)(σ)][(σ)(σ)(σ)]	(d) alternative (ii)

(. . .) = foot
[. . .] = p-word

In contrast to the "staccato" rhythm of slow speech, punctuated as it were by a succession of monosyllabic feet, normal speech tends to group syllables into trochees, or binary left-prominent feet. There is an exception clause to this rule: the last two syllables do not join into a single foot (see further discussion below). Ao (p. 143) states this process as follows:

(64) Normal Tempo Defooting

ϕ = foot

Defooting demotes a non-final syllable and cliticizes it to the footed syllable on the left. Defooting is intended to apply iteratively from left to right.

At normal tempo, an unbounded right-dominant phonological word is constructed (which we call simply Unbound). (65) illustrates how Defooting and Unbound work on two- to five-syllable strings.

(65) Normal tempo

σ σ	σ σ σ	σ σ σ σ	σ σ σ σ σ	
(σ)(σ)	(σ)(σ)(σ)	(σ)(σ)(σ)(σ)	(σ)(σ)(σ)(σ)(σ)	Foot Construct (62a)
–	(σ σ)(σ)	(σ σ)(σ)(σ)	(σ σ)(σ σ)(σ)	Defooting
[(σ)(σ)]	[(σ σ)(σ)]	[(σ σ)(σ)(σ)]	[(σ σ)(σ σ)(σ)]	P-word

As stated, Defooting requires at least three syllables; it therefore blocks in disyllabic strings. As for quadrisyllabic compounds, Defooting leaves two syllables at the end, which fail to undergo a second round of Defooting for exactly the same reason. Trisyllabic expressions undergo Defooting as expected. The longer, pentasyllabic sequences are unremarkable in that Defooting merely applies twice, iteratively from left to right.

Fast speech differs from normal speech only in that at accelerated tempo foot binarity no longer holds, so that iterative foot adjunction creates an unbounded foot, subject to the same exception clause forbidding a binary foot in p-word-final position. This gives us [(*bo.li.wei*)(*ya*)] "Bolivia" in fast speech (= (60c)).

(66) Fast tempo

bo. li. wei. ya	"Bolivia"
(L.LM.MH)(L)	base form
(L. o o) (L)	tone deletion
(M. o o) (L)	L-Raise
(M. o o) (ML)	M-Insertion
(M. m. m) (ML)	Default

At this point a few words about underlyingly toneless syllables are in order. Affixes and clitics, including quite an assortment of grammatical particles and function words such as aspect markers, classifiers, pronouns, etc., are typically toneless underlyingly. By virtue of (62b), therefore, they all attach themselves to a tone-bearing syllable on the left to form a (unbounded) foot. As a consequence, the foot-bound rule Spread extends the H tone to both of the toneless syllables on the right in an example like *rao-le ta* (gloss: pardon + perfective + him) repeated here:

(67)

[*rao le*] *ta*	"pardoned him" = (159)
(MH.o.o)	Foot Construction (62b)
M. H. H	Spread

3.5 Rhythmic organization: an OT perspective

While Ao's rules work flawlessly and yield the right results, they conceal some of the motivating principles. For instance, the non-final condition

on foot-adjunction is a curious one. Defooting contributes to fluent speech in that it joins syllables into ideally binary units, thereby creating a more natural rhythmic flow instead of a string of separate monosyllabic feet. Remarkably, Defooting, repeated below for convenience but in a slightly different format, requires an extra syllable to the right, effectively barring two final syllables from joining together into a foot. Why should this be so?

(68) Normal Tempo Defooting

(x)(x)(x) → (x .)(x)

(. . .) = foot
x = head

Defooting is reminiscent of the T2 Sandhi rule in Beijing discussed in section 2.2 of chapter 7: T2 Sandhi turns M into H in the environment H.__H – provided there is another tonic syllable to the right. Obviously, we have not yet discovered the explanatory principle behind an inductive generalization stated above as Defooting.

We begin to understand what is going on by noting that on the one hand, the foot in Nantong is uniformly left-headed, while the p-word is right-prominent, that is to say, tends to place the main stress on the rightmost syllable. This being the case, Defooting a final syllable has the effect of moving the main stress to the left, as shown below

(69)
```
          x                 x
. . . (x)(x)]ω   →   . . . (x  .)]ω
      σ  σ                σ σ
```

ω = p-word

since assigning a word-level (line 2) prominence on anything but the head of the foot (line 1) would create an ill-formed grid:

(70)
```
          x
. . . (x  .)]ω
      σ σ
```

Defooting, in other words, automatically disqualifies the final syllable as the head of the p-word. The only condition under which we find a non-final word-level stress is where the final syllable or syllables are inherently or otherwise toneless. For instance, in a construction like *rao le ta* "pardoned him", both the perfective marker *le* and the object pronoun *ta* behave as toneless enclitics, and therefore incapable of functioning as the

head of a foot or, a fortiori, of a word. The metrical structure of (71b) is therefore ill formed, and only (71a) is acceptable, even though the word stress does not fall on the absolute final position.

(71)

```
        x
       (x   .  .)
a.     rao le ta        "pardoned him" (gloss: pardon + perfective + him)
       (MH.o.o)[19]     = (159)

              x
       (x    .)(x)
b.     rao le ta
       (MH.o)(o)        = ill formed
```

Within limits set by metrical well-formedness, a polysyllabic string is organized, from left to right, ideally into binary feet.

What we have then is another classic example of conflicting constraints, which lends itself to fairly straightforward OT-style treatment. We can simulate the directionality of foot construction by some version of "alignment" (see McCarthy and Prince 1993b; cf. Mester and Padgett 1993, Crowhurst and Hewitt 1995, and discussion in section 5.2 of chapter 3). The evaluation of Alignment satisfaction and assessment of penalties for violation thereof are somewhat complicated. For our purposes, a shorthand alternative will do. Since binary footing proceeds from left to right, the first two syllables always form a foot.[20] We can therefore mimic the left-to-right directionality of foot formation by simply prohibiting word-initial monosyllabic feet.[21]

The relevant constraints are:

(72) Rhythmic constraints

a. Tonicity: only tonic syllables may be stressed; only stressed syllables may carry tone.
b. Left: Build left-prominent feet.
c. Right: Assign word-level stress to the rightmost syllable.
d. LRF (Left-to-right footing): No initial degenerate feet.
e. Bounded: The foot must bounded (at most disyllabic).
f. Binary: The foot must be binary (at least disyllabic).

The tableau in (73) illustrates exactly how these general principles combine to predict the optimal rhythmic organization for strings of three to

[19] By the late phonetic rule Spread, (MH.o.o) is realized as (M.H.H).

[20] In this sense, metrification is syntax-blind, so to speak (see Ao, pp. 150f.). We will see in the next section how morphosyntax constrains rhythmic structure.

[21] More precisely, monosyllabic feet can occur only (next to each other) at the end of a word. I will ignore this fine point here.

five syllables at *normal tempo*, as given at the outset of section 3.3. For expository simplicity, I will focus exclusively on the normal tempo, deferring discussion of fast and slow speech until section 3.7. To factor out morphosyntactic interference in metrification, I have chosen examples that have a flat (coordinate) structure or no structure at all, as in the case of foreign names, and strings of numerical expressions. We will deal with structure-dependent rhythm in the next section.

(73) Normal tempo

ex.		Left	Right	LRF	Bound (≤ 2)	Binary (≥ 2)
1. *shang.zhong.xia* "top, middle, and bottom" = (234)	a x ☞ (x .) (x)					*
	b x (x . .)		*		*	
	c x (. . x)	*			*	
2. *bo.li.wei.ya* "Bolivia" = (229)	a x (x .) (x .)		*			
	b x (. x) (. x)	*				
	c x (x . .) (x)				*	
	d x ☞ (x .) (x) (x)					**
	e x (x) (x .) (x)			*		**
3. *er.er.er.er.er* "2-2-2-2-2" = (229)	a x ☞ (x .) (x .) (x)					*
	b x (x . .) (x .)		*		*	
	c x (x . .) (x) (x)				*	**
	d x (x) (x .) (x) (x)			*		***

The tableau in (73) is interpreted in the usual fashion. Take a quadrisyllabic input like *bo.li.wei.ya* "Bolivia" for illustration. The highly ranked Left and Right automatically rule out candidates ex.2a and ex.2b. The fact that candidate 2d wins over 2c means that Bounded ranks higher than Binary on the constraint hierarchy. In other words, if it is not possible to break down a string into ideally binary units, it is better to leave monosyllabic feet alone, rather than to fuse them into ternary or longer metrical units. The optimal rhythmic structure of *Bolivia* manifests itself tonally as follows:

(74)	*bo. li. wei. ya*	"Bolivia"
	(M. m)(MH)(ML)	normal tempo
		= (60b)

3.6 *Structure-dependent rhythm*

The OT approach exemplified above sheds light on another otherwise puzzling aspect of the rhythmic organization of Nantong. Ao (1993:150ff.) flatly states that morphological structure is irrelevant to the rhythmic organization of trisyllabic compounds. Consider these examples:

(75)		grape sugar	
	a.	[*pu tao*] *tang*	"glucose" = (234)
		(MH.MH)(MH)	
		(MH. o) (MH)	Deletion
		(M H) (MH)	Spread
		red ten character[22]	
	b.	*hong* [*shi zi*]	"red cross" = (234)
		(MH. LM)(LM)	
		(MH. o) (LM)	Deletion
		(M. H) (MLM)	Spread
	c.	*sheng yun diao*	"onset, rhyme, tone" = (234)
		(L. LM)(HM)	
		(M. LM)(HM)	L-Raise
		(M. o) (HM)	Deletion
		(M. m) (HM)	Default

They demonstrate that regardless of their internal structure (left-, right-branching, or flat), the sandhi patterns of compounds exhibit a uniformly [(xy) (z)] prosodic organization. Take (75b), for instance. If its rhythmic structure were to mirror its morphological constituency, it would have a derivation leading to an unacceptable phonetic output:

[22] The Chinese word for "cross" *shi-zi* literally means "the character for ten," which in Chinese orthography happens to have the shape ✚.

(76)	*hong* [*shi zi*]	"the Red Cross"
	(MH)(LM.LM)	
	(MH)(LM.o)	Deletion
	(MH)(L. M)	Spread
	*(MH)(ML.M)	M-Insertion

Curiously, however, when a right-branching structure like *hong* [*shi zi*] is embedded into a larger structure, as in (77), instead of the expected [(σ σ)(σ)(σ)], as in the case of internally unstructured expressions like *bo.li.wei.ya* "Bolivia," this example enforces a trisyllabic foot, which otherwise occurs only in fast tempo speech. In this respect (77) contrasts with a bare [σ [σ σ]] structure like (75b) on the one hand, and with a flat or unstructured quadrisyllabic string like *Bolivia* (repeated below) on the other.

(77)	[*hong* [*shi zi*]] *hui*	"the Red Cross (Society)"[23] = (236)
	(MH.LM.LM)(LM)	
	(MH. o o) (LM)	Deletion
	(M. H. H) (LM)	Spread
	(M. H. H) (LM)	M-Insertion
(78)	[*bo. li. wei. ya*]	"Bolivia" = (224)
	(L.LM)(MH)(L)	
	(L. o) (MH)(L)	Deletion
	(M. o) (MH)(L)	L-Raise
	(M. m) (MH)(ML)	M-Insertion, Default

One cannot simply say without qualification that immediate constituents (IC) must belong to the same foot, for this is patently false, witness (75b), where *shi* "ten" and *zi* "character" are split between two metrical feet. Confronted with this paradoxical situation, Ao was forced to add an awkward stipulation, which reads as follows:

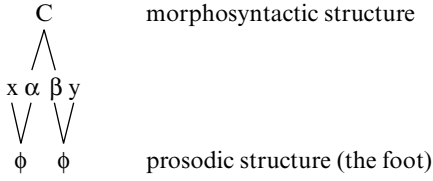
(79) Morpheme Integrity constraint
A foot dominating α and β dominates every γ that c-commands β and does not c-command α, provided that γ precedes a footable δ. (Ao 1993:154)

Here α, β, γ, δ refer to a linear sequence of syllables. Apart from the complicated stipulations on c-command relations, Morpheme Integrity constraint is saddled with a seemingly unrelated appendage concerning a footable δ. Clearly this constraint has failed to capture the essence of what is truly going on.

[23] *Hui* literally means "society" or "association."

A careful examination of the fairly large corpus of data presented in Ao reveals that Nantong normal speech obeys the same principle of "syntactic scansion" that is well attested in Chinese poetry (see Chen 1979, 1980, 1984, inter alia). Poetic syntax prohibits what I called "foot-straddling" in Chen (1979:411). Foot straddling occurs when the ICs α and β of a constituent C stand astride two metrical feet (ϕ), as pictured below:

(80)



(Here x and y represent arbitrary strings, including ∅.) Metrico-syntactic mismatches of this type are virtually unknown in the regulated verse of Chinese (see Chen, op.cit.). I propose No Straddling as an independently motivated principle.

(81) No Straddling
Immediate constituents must be footmates.

The simply worded constraint prohibits precisely a metrico-syntactic mismatch like the one diagrammed in (80). Note however that No Straddling in no way implies the converse, namely that foot-mates must be ICs: it does not prevent non-constituents like *fu* "deputy" and *zong* "general" in the following example from forming a single foot:

(82) deputy general commander
fu [*zong* [*si ling*]] "deputy commander-in-chief"
(x .) (x)(x)

We can bring out the effect of No Straddling interacting with other constraints most perspicuously in the customary tableau form. By correctly ranking these constraints, we can predict the attested sandhi forms. In particular, let us see how a constant constraint ranking unerringly picks out the attested winner candidates across a wide variety of construction types. Of particular interest in the tableau in (83) is the way [*hong* [*shi zi*]] *hui* "the Red Cross (Society)" contrasts with (i) *hong* [*shi zi*] "the Red Cross" ex.2; (ii) an unstructured quadrisyllabic proper name like *bo.li.wei.ya* "Bolivia" ex.3; and (iii) a structurally isomorphic expression like *xu.pə.pə.ti* "very irritable" with an atonic syllable -*ti* in the final position ex.4. The disyllabic compound *dian shi* "television" ex.5 is added to illustrate the dominance of metrical well-formedness (Left and Right) over the structural cohesion expressed as No Straddling.

(83)

		Tonic	Left	Right	NStr	LRF	Bnd	Bin
1 [*hong* [*shi zi*]] *hui* "the Red Cross" = (236)	a x (x)(x .)(x)					*		**
	b x (x .)(x .)			*	*			
	c x (x .)(x)(x)				*			**
	d x (x)(x)(x)(x)				*	*		****
	e x ☞ (x . .)(x)						*	*
2 *hong* [*shi zi*] "the Red Cross" = (234)	a x (x)(x .)			*		*		*
	b x (x . .)			*			*	
	c x (x)(x)(x)				*	*		***
	d x ☞ (x .)(x)				*			*
3 *bo.li.wei.ya* "Bolivia" = (224)	a x (x .)(x .)			*				
	b x (. x)(. x)		*					
	c x (x . .)(x)						*	*
	d x ☞ (x .)(x)(x)							**
4 *xu.pə.pə.ti* "very irritable" = (257)	a x ☞ (x .)(x .)			*				
	b x (x .)(x)(x)	*			*			**
	c x (x . .)(x)	*					*	*
5 *dian.shi* "television" = (178)	a x (x .)			*				
	b x (. x)		*					
	c x ☞ (x)(x)				*	*		**

LRF = Left-to-right footing: no initial monosyllabic foot
Bnd = Bounded (at most 2)
Bin = Binary (at least 2)

Let us quickly run through the examples of the tableau in (83). For ex.1, *hong shi zi hui* "the Red Cross (Society)," the footing of candidate (a) is isomorphic to its morphological structure, but runs afoul of the Left-to-Right Footing (LRF) constraint. Candidate (b) displays the perfect alternating rhythm, but at the cost of shifting the word prominence leftward. Candidate (c) satisfies the Right prominence constraint, but incurs the No Straddling violation, which is not ameliorated by simply building a degenerate foot over each monosyllable (candidate (d)). The winning candidate (e) satisfies the higher ranking constraints of Right prominence, No Straddling and Left-to-Right Footing at the expense of the lower ranked Bounded and Binary.

However, No Straddling does not go unchallenged. The winning candidate (d) of ex.2 *hong shi zi* "the Red Cross" stands in flagrant violation of No Straddling, with the ICs *shi* and *zi* torn asunder between two metrical feet. All efforts in the interest of conformity with the No Straddling constraint (candidates 2a, b) are thwarted by the overriding imperative of Right prominence.

We have already dealt with the case of *bo.li.wei.ya* (ex.3). The important contrast worth noting is that it is internally unstructured, therefore not subject to No Straddling. This explains the two different rhythmic organizations attested for two equally quadrisyllabic expressions:

(84) a. (*bo.li*)(*wei*)(*ya*) "Bolivia"
b. (*hong.shi.zi*)(*hui*) "the Red Cross (Society)"

Ex.4 *xu.pə.pə.ti* "very irritable" instantiates a productive process in many dialects of Chinese which turns a plain adjective into a "vivid" descriptive expression by adding XX-*de*, where XX is a string of reduplicated syllables (onomatopoetic in some cases) and -*de* is a toneless suffix. Examples in Mandarin include: *hei* "dark" → *hei-qi-qi-de* "pitch-dark," *ying* "hard" → *ying-bang-bang-de* "stiff and inflexible," and so forth. A similar morphological process occurs in Nantong, where the toneless -*ti* corresponds to the Mandarin suffix -*de*, and is inherently toneless. Morphologically speaking, an expressive adjective (A) of the form *A-X-X-ti* has the internal structure [A-[X-X]]-*ti*.[24]

(85) a. [*xu* [*pə pə*]] *ti* "very irritable"
b. [*hong* [*shi zi*]] *hui* "the Red Cross (Society)"

[24] [xu.pə.pə.ti] is the phonetic form given in Ao (p. 257), not a Pinyin transcription.

Structurally speaking, (85a) parallels (85b) exactly, except that the final syllable of the former is toneless. Interesting consequences follow. Consistent with the Tonicity constraint stated earlier, the tonelessness of the final syllable -*ti* disqualifies it from occupying a metrically prominent position, either as the head of the foot, or, a fortiori, of the p-word. Tonicity therefore rules out candidate 4b (*xu.pə*)(*pə*)(*ti*) and 4b (*xu.pə.pə*)(*ti*) of (83), which parallel exactly the winning candidates (*bo.li*)(*wei*)(*ya*) (ex.3d) and (*hong.shi.zi*) (*hui*) (ex.1e) of the same tableau respectively. Instead, the only attested reading at the normal tempo is:

(86)	[*xu* [*pə pə*]] *ti*	"very irritable"
	(x .) (x .)	
	H.HM. HM.o	base tones
	(H. o) (HM.o)	Deletion
	(H. H) (HM.o)	Spread
	(H H) (HM. l)	Default

Finally, ex.5 *dian.shi* "television" of (83) shows that the metrical well-formedness expressed as (word-level) Right prominence and (foot-level) Left prominence jointly doom candidates 5a and 5b, and weigh in favor of 5c, which violates every lower ranked constraint it can conceivably breach.

3.7 Constraint ranking

To summarize, the data presented in (83) establish the following ranking relations among the constraints, with crucial examples cited in the right-hand column, where A ≀ B means candidate A is more "harmonic" than B.

We now turn to the rhythmic structure of allegro or fast speech. Instead of (*bo.li*)(*wei*)(*ya*) (ex.3d of (83)), which is the attested reading for "Bolivia" at the normal tempo, it is (*bo.li.wei*)(*ya*), that is ex.3c that prevails when speech tempo accelerates:

(87)	Fast tempo	
	bo. li. wei. ya	"Bolivia"
	(M. m. m) (ML)	= ex. (60c)

This tempo-sensitive variation is handled quite simply by constraint reranking:

(88) Fast tempo:
Binarity ≫ Bounded

Table 8.2. *Constraint ranking for normal tempo*

Tonicity ≫ Right	x (x .)(x .) [*xu* [*pə.pə.*]] *ti*	x (x .)(x) (x) [*xu* [*pə.pə.*]] *ti* **or**
		x (x . .) (x) [*xu* [*pə.pə.*]] *ti*
Left, Right ≫ No Straddling	x x (x) (x) (x .) *dian.shi*	x (. x) *dian.shi* **or**
		dian.shi
Right ≫ LRF, No Straddling	x x (x) (x) (x .) *dian.shi*	
		dian.shi
No Straddling ≫ LRF	x (x . .) (x) (x) (x .) (x) [*hong*[*shi.zi*]] *hui*	x [*hong*[*shi.zi*]] *hui*
No Straddling ≫ Bound	x (x . .) (x) (x .)(x) (x) [*hong*[*shi.zi*]] *hui*	x [*hong*[*shi.zi*]] *hui*
Bound ≫ Binary	x (x .) (x) (x) *bo.li.wei.ya*	x (x . .) (x) *bo.li.wei.ya*

Now, what about the slow tempo rendition of Bolivia, repeated here as (89)?

(89) Slow tempo
bo. li. wei. ya "Bolivia"
(M)(ML)(MH)(ML) = (60a)

Here, the overriding imperative, it seems, is to parse, i.e. phonetically realize, all lexically assigned tones or their sandhi alternants. Recall that Tonicity, repeated below, expresses a bidirectional entailment: stressed ↔ tonic, stressless ↔ atonic.

(90) Tonicity
Only tonic syllables may be stressed; only stressed syllables may carry tone.

This means that if two or more syllables are joined into a single foot, all but the first (stressed) syllable will lose their tones. To derive the slow tempo effect, all we need to do is to posit the following ranking:

(91) Slow tempo:
{Tonicity, Parse-Tone} ≫ Bound ≫ Binarity

This is illustrated below:

(92) Slow tempo

	bo.li.wei.ya "Bolivia"	Tonic	Parse-Tone	Bound (at most 2)	Binarity (at least 2)
a	x (x .)(x)(x) T.o T T		*		**
b	x (x . .)(x) T.o.o T		**	*	*
c	x (x .)(x)(x) T.T T T	*			**
d ☞	x (x)(x)(x)(x) T T T T				****

Candidates (a), (b), and (c) lose out in slow tempo, because footing the first two or three syllables together entails either the elimination of the tones associated with the metrically weak syllables, thereby incurring a Parse-Tone violation (cases a, b), or else linking the tone to a stressless syllable, in contravention of the Tonicity constraint (case c). This leaves candidate (d) as the winner, despite the fact that it increases the number of degenerate feet, in violation of Binarity.

In summary, the richly articulated rhythmic structures of Nantong at various tempos can be seen as a response to a number of maximally simple and general, but often conflicting constraints on how a string of syllables is rhythmically organized into metrical feet. In the immediately following chapter, we examine the next larger prosodic unit, namely the "minimal rhythmic unit" or MRU.

9 *Minimal rhythmic unit as obligatory sandhi domain*

Arguably the best known case of tone sandhi is one that takes place in **Beijing Mandarin** (often referred to as Mandarin, *tout court*). This lingua franca, or *Putonghua* (lit. "the common language"), has a small inventory of four contrastive tones, exemplified by this oft-cited quadruplet:

(1) Mandarin tone system

T1	high level	[55]	*ma*	"mother"
T2	rising	[35]	*ma*	"hemp"
T3	low dipping	[214]	*ma*	"horse"
T4	falling	[51]	*ma*	"to scold"

T1–T4 stand for the four tonal *categories*, whereas the digits enclosed in square brackets represent the pitch values on Y-R. Chao's (1930) five-point scale. We surveyed the various sandhi processes observed in Beijing Mandarin in the introductory chapter (section 4.1). Here I will concentrate exclusively on one of them, namely the so-called T3 Sandhi, by far the most important and most intensively scrutinized of the tone rules. T3 Sandhi basically turns T3 [214] into T2 [35], when followed by another T3. This is stated simply as

(2) **T3 Sandhi** (TS)
T3 → T2 / ___ T3

and illustrated by examples such as:

(3) wild grass

a. *ye-cao* "weed"
3 3
2 3

b. *mai jiu* "buy wine"
3 3
2 3

Henceforth I shall refer to the T3 Sandhi rule stated above simply as TS (for Tone Sandhi). For simplicity, all subsequent examples are annotated

with the single digit tone *categories* (1 to 4) rather than multi-digit tone values (214, 35 etc.). Furthermore, I will often use the symbol ***s*** (for sandhi tone) to indicate a derived T2, as distinct from an underlying T2. This has the advantage of collapsing two lines of derivation into one, so that (a) is equivalent to (b), yet distinct from (c):

(4)	a.	*fen$_1$-chang*	"flour factory"
		s 3	
	b.	*fen$_1$-chang*	"flour factory"
		3 3	base tone
	→	2 3	sandhi tone
	c.	*fen$_2$-chang*	"graveyard"
		2 3	base tone, no change

We can tell *fen$_1$* "flour, powder" with an underlying T3 from *fen$_2$* "grave, tomb" with a lexically assigned T2 by simply putting them in a different context, in which their basic identity comes to the surface. Thus before a T4, both T2 and T3 maintain their underlying tones unchanged:

(5)	a.	*fen$_1$-sui*	"to pulverize"
		3 4	base tone, no change
	b.	*fen$_2$-mu*	"grave"
		2 4	

As we will see, the notion of the word plays a central role in our analysis of Beijing tone sandhi, I will link the sublexical units together with a hyphen, to distinguish words from larger constructions. Thus, we have *ye-cao* "weed" (one lexical compound made up of two words *ye* "wild" and *cao* "grass") vs. *mai jiu* "buy wine" (two words). Where appropriate I will separate word-size units with "#" for clarity and emphasis (e.g. *mai # jiu* "buy wine" vs. *ye-cao* "weed"). We defer until sections 3–4 the somewhat complex issues related to wordhood in Chinese.

The phonological aspects of Beijing tone sandhi are straightforward and, for our purpose, of marginal interest. On the other hand, the question as to the domain or scope that circumscribes the operation of TS is a thorny one that has engaged the attention of many linguists. In the rest of this chapter I will concern myself exclusively with this issue.

Stated in its stark simplicity, the TS rule tells only part of the story. As will become amply apparent, it is not the case that any sequence of two T3's would turn into T2 + T3. Even if it were, the rule gives no clue as to how it might apply to a string of three or more T3's; as a matter of fact

an input string /3-3-3/ shows up as [2-2-3] in some instances, and as [3-2-3] in others. In approaching this enormously complicated problem I will proceed in the following manner: First, I will present the core of an analysis based on the concept of **Minimal rhythmic unit** (MRU), inspired in large part by Shih (1986, 1997 [=1990]), although couched in terms of OT-style constraints, which bring out more perspicuously the conceptual content of such an approach (section 1). I will then argue that MRUs are constructed in a two-pass fashion: first on lexical, then on phrasal constructions (section 2). This raises the question of the wordhood in Chinese and the sandhi behavior of clitics (sections 3–4). Finally, after a summary of the analysis proposed here (section 5), I compare it with other approaches based on a more elaborate prosodic hierarchy (section 6), syntactic juncture (section 7), and meaning relations (8).

1 Minimal rhythmic units

The basic idea is that connected speech is broken up into "**Minimal rhythmic units**" (**MRUs**), within which TS applies obligatorily. What constitutes an MRU, and how are they derived is the crux of the matter at hand. Once we determine the MRUs, TS applies cyclically or, in the absence of morphosyntactically defined subconstituents, iteratively from left to right.

1.1 Minimality and maximality effects

The fundamental insight underlying our description of Beijing TS was captured by Zhou (1964 [1961]:266), who put it quite succinctly: "The dominant drift of contemporary Chinese rhythm is toward disyllabification." This drift toward disyllabicity as the minimal prosodically independent unit manifests itself in a variety of ways. First, most lexical entries in contemporary Chinese are polysyllabic. For instance, 85% of all nouns in a survey of 3,000 high-frequency expressions are disyllabic or longer. In order to meet the disyllabic requirement, various "padding" devices are employed, including otherwise pleonastic expressions, e.g. each of the following words *mei-li* "beautiful," *shu-mu* "tree," *gou-mai* "buy," and countless others is made up of two synonyms, broadly defined (both *mei* and *li* mean roughly "beautiful"). The disyllabicity requirement is so strict that monosyllabic foreign names must be augmented in some way: thus one must refer to "(Bernard) Shaw" not as **Xiao* ([çiaw] in IPA), but as *Xiao-shi* (Shaw + Mr.); likewise, "Bonn" calls for a disyllabic transliteration as *Bo-ang*, even though a shorter *bang* would be a better phonetic

approximation. More importantly, monosyllabic words that do exist in the lexicon cannot occur freely in actual speech. Thus, to the question "What is your (family) name?" one can answer "Ou-yang,"[1] but not "Wang." In order to bring the monosyllabic surname above the disyllabic threshold, one has to repeat (part of) the question: *wo xing wang* "My name is Wang." Some languages enforce this minimal word rule even more rigidly. One case in point is Mixteco, as described in K. Pike (1948:79): "With the possible exception of a few particles that have not been found in isolation, all Mixteco morphemes are basically disyllabic. Every Mixteco morpheme found in isolation is without exception disyllabic, though when included in phrases morphemes frequently become monosyllabic." For further discussion on the minimal word effect see McCarthy and Prince (1986, 1993a, 1994a, 1995a), Itô (1990) and, Prince (1990), Prince and Smolensky (1993), with particular reference to Chinese, Zhou (1961 [1964]), Lü (1963), K. Lee (1976), Yip (1991).

The optimal rhythmic organization is not only at least disyllabic, but ideally disyllabic as well. This "bounded" rhythmic organization manifests itself in the tendency for trisyllabic expressions to shorten, as in *luo-hua-sheng* → *hua-sheng* "peanut," *wai-guo-yu* → *wai-yu* "foreign languages," *niu-nai-lao* → *nai-lao* "cheese," etc. Statistically, disyllabic [σσ] and quadrisyllabic (= [σσ + σσ]) compounds, set phrases, and idioms outnumber by a wide margin trisyllabic [σσ + σ] / [σ + σσ] and quadrisyllabic [σσσ] + σ / [σ + σσσ] expressions (see Lü 1963).

We add to this the left-to-right metrification principle, which we have seen at work in Tianjin (chapter 3), Changting (chapter 4), Shanghai (chapter 7), Danyang, and Nantong (chapter 8). We can state the eurhythmic principles of Beijing as follows:

(6)	Binarity	The MRU is at least disyllabic.
	Boundedness	The MRU is at most disyllabic.
	LtoR	MRUs are constructed from left to right.[2]

Each instance of a subminimal MRU is penalized by one asterisk. Likewise, each syllable in excess of the upper bound of "at most two" is marked by one asterisk under Boundedness. By ranking Binarity over

[1] Most Chinese surnames are monosyllabic; *Ou-yang*, *Si-ma* are some of the common exceptions known as *fu-xing* (lit. double surnames).

[2] LtoR can be restated as AlignL(MRU, IP), that is align the left edge of every MRU with the left edge of an intonation phrase (IP), the distance between the two edges is measured in terms of syllables (cf. McCarthy and Prince 1993b, Mester and Padgett 1993, Crowhurst and Hewitt 1995).

Boundedness, we derive exactly the MRUs observed in those cases where rhythmic organization is unfettered by morphosyntactic structure, as illustrated by a string of two-to-five-syllable-long numerical expressions:

(7)

ex.					Binary (≥2)	Bound (≤2)	LtoR
1	wu.wu		a	(wu)(wu)	**		
	"5-5"	☞	b	(wu.wu)			
2	wu.wu.wu		a	(wu.wu)(wu)	*		
	"5-5-5"		b	(wu)(wu.wu)	*		*
		☞	c	(wu.wu.wu)		*	
3	wu.wu.wu.wu	☞	a	(wu.wu)(wu.wu)			
	"5-5-5-5"		b	(wu.wu.wu.wu)		**	
			c	(wu)(wu.wu.wu)		*	*
4	wu.wu.wu.wu.wu		a	(wu.wu)(wu.wu)(wu)	*		
	"5-5-5-5-5"	☞	b	(wu.wu)(wu.wu.wu)		*	
			c	(wu.wu.wu)(wu.wu)		*	*
			d	(wu)(wu.wu)(wu.wu)	*		*

MRUs are enclosed in (. . .)

The rhythmic structures of these expressions are diagnosed by the way TS operates on them. Notice in the examples that follow, within an MRU, every occurrence of *wu*, which carries an underlying T3, turns into a T2.

(8) a. wu.wu
(s 3)

b. wu.wu.wu
(s s 3)

c. wu.wu.wu.wu
(s 3) (s 3)

d. wu.wu.wu.wu.wu
(s 3) (s s 3)

s = T2 derived from T3

The analysis given here extends to other morphosyntactically unstructured expressions. Thus, (9a) is the only possible reading[3] of *Somalia*, which when transliterated into Chinese consists of four T3 syllables:

[3] I.e. discounting allegro speech and marked intonation breaks.

(9) *suo-ma-li-ya* "Somalia"

a. (s 3) (s 3) ok

b. (3) (s s 3) *

c. (s s 3) (3) *

d. (s 3) (3) (3) *

1.2 Structural congruence

Needless to say, one does not break down an utterance into optimally binary MRUs in total disregard of morphosyntactic structure. Quite on the contrary, the rhythmic organization of Beijing Mandarin is exquisitely sensitive to constituent structure. Let us examine the following quadrisyllabic expressions:

(10) strike-down governor

a. [*da-dao*] [*sheng-zhang*] "down with the governor!"
(s 3) (s 3)

which kind wine good

b. [[*nei zhong*] *jiu*] *hao* "which kind of wine is better?"
(s 3) (s 3)

c. *gou* [*yao* [*xiao-mei*]] "the dog bit Xiaomei"
(s 3) (s 3)

d. [*zhi*-[*lao-hu*]] *pao* "the paper tiger is running"
(3 s s 3)

look-for cowardly devil

e. *zhao* [[*dan-xiao*]-*gui*] "look for the coward"
(3 s s 3)

These exhaust the logically possible tree structures.[4] Case (a) is unremarkable. The rhythmic organization is isomorphic with the syntactic structure. But this symmetry breaks down in the other cases. Why is a bipartite rhythmic structure permitted in (b–c), but barred from (d–e)? The answer lies in a constraint that was originally conceived as a principle of syntactic scansion in classical Chinese verse (see Chen 1979, 1980, 1984), which I restate below:

(11) No Straddling (NoStr)
Immediate constituents must be MRU-mates.

[4] Assuming exclusively binary branching.

We have already encountered this structure-sensitive metrification principle at work in Nantong.[5] Unlike in Nantong, No Straddling is undominated in Beijing Mandarin. This means that, together with Binarity, it enforces an unbounded MRU on (10d) and (10e), overriding the Boundedness constraint, as illustrated in the following tableau:

(12) *zhao* [[*dan-xiao*]-*gui*] "look for the coward"
(3 s s 3)

			NoStr	Binarity (≥2)	Bound (≤2)
	a	(*zhao*)(*dan-xiao*)(*gui*)		**	
	b	(*zhao dan*)(*xiao gui*)	*		
	c	(*zhao dan-xiao*)(*gui*)		*	*
	d	(*zhao*)(*dan-xiao-gui*)		*	*
☞	e	(*zhao dan-xiao-gui*)			**

A syntactically congruent rhythmic structure (a) leaves two subminimal MRUs at the edges, in violation of Binarity. Conversely, the bipartite MRUs of (b) create the optimal rhythm that satisfies both Binarity and Boundedness constraints, but at the expense of No Straddling. Binarity rules out candidates (c) and (d), leaving (e) as the only viable parse of a center-embedded structure, despite the fact that it creates an unbounded MRU. The two asterisks under the column Boundedness are intended to show that the quadrisyllabic MRU of (e) exceeds the "at most two" limit by two syllables.

Notice that the MRU-initial syllable surfaces with an unchanged T3. The reason why this is so will be made clear shortly. Example (10d) [*zhi*-[*lao-hu*]] *pao* "the paper tiger is running" is an exact mirror image of (10e) (= (12)). The same constraints guarantee the same output without further elaboration.

What about (10b) and (10c)? The rhythmic structures are at odds with the syntactic bracketing. As noted before in connection with Nantong, although No Straddling requires that immediate constituents be MRU-mates, it in no way entails the converse, namely that MRU-mates must be immediate constituents in mutual c-command. Therefore, neither (*jiu hao*) "wine good" nor (*gou yao*) "dog bite" violates No Straddling. As a consequence, Boundedness imposes binary MRUs on the strings.

[5] Except that in Nantong, the relevant metrical unit is the stress-foot. See section 1.4 for terminological clarification.

(13) *gou* [*yao* [*xiao-mei*]] "the dog bit Xiaomei"
(s 3) (s 3)

		NoStr	Binarity	Boundedness
	a (*gou*)(*yao*)(*xiao-mei*)		**	
	b (*gou*)(*yao xiao-mei*)		*	*
	c (*gou yao xiao-mei*)			**
☞	d (*gou yao*)(*xiao-mei*)			

No Straddling is nothing but a special case of a more general constraint on finding the best fit between syntactic and rhythmic structures, subject, of course, to the eurhythmic principles of Binarity and Boundedness. Consider the following example:

(14) guess water very hot
cai-xiang [*shui* [*hen re*]] "(I) guess the water is very hot"
1 3 3 3 4 base tone
(1 3) (s 3 4) i. ok
(1 s 3) (3 4) ii. ??

Since No Straddling is undominated, we can ignore output candidates that split up either *cai-xiang* "guess" or *hen re* "very hot." In principle, *shui* could go with either *cai-xiang* or *hen re* to form an MRU. But reading (14-i) is appreciably better than (14-ii). In (14-i), *shui* and *hen* are members of the same MRU, therefore, *shui* undergoes TS, changing to T2. In the alternative rhythmic division, *shui* is separated from *hen*, and therefore does not undergo TS; instead it provides the requisite context that triggers TS on *xiang*. Therefore, in addition to the more specific No Straddling we need a generic syntax/rhythm Congruence constraint, which I state informally in "processual" terms:

(15) Congruence
Group X forms an MRU with its closest morphosyntactic mate

The effect of Congruence is illustrated below:

(16) guess water very hot
cai-xiang [*shui* [*hen re*]] "(I) guess the water is very hot"
(1 3) (s 3 4)

		NoStr	Binarity	Bound	Congr
☞	a (*cai-xiang*)(*shui hen re*)			*	
	b (*cai-xiang shui*)(*hen re*)			*	*
	c (*cai-xiang*)(*shui*)(*hen re*)		*		

Needless to say, Congruence must outweigh the structure-blind Left-to-Right Footing constraint, otherwise the latter would always prevail, in effect nullifying whatever effect Congruence might have. This ranking relation is illustrated below:

(17) that kind wine harmful
[[*nei zhong*] *jiu*] [*you-hai*] "that kind of wine is harmful"
(4 s 3) (3 4)

		NoStr	Binarity	Boundedness	Congruence	LtoR
	a (*nei zhong*)(*jiu you-hai*)			*	*	∅, 2
☞	b (*nei zhong jiu*)(*you-hai*)			*		∅, 3
	c (*nei zhong*)(*jiu*)(*you-hai*)		*			∅, 2, 3
	d (*nei zhong jiu you-hai*)			***		∅

Candidate (b) wins over (a) despite the fact that the second MRU (*you-hai*) lies farther away from the beginning of the IP. Candidate (c) is eliminated by the binarity requirement. Candidate (d) loses out because of the multiple violations of Bound, exceeding the "at most two" upper ceiling by three syllables.

Notice, unlike the undominated No Straddling, the structure-sensitive Congruence must rank below Binarity and Boundedness. This is shown by (10b) presented in a tableau format in (18):

(18) which kind wine good
[[*nei zhong*] *jiu*] *hao* "which kind of wine is better?"
(s 3) (s 3)

		NoStr	Binarity	Bound	Congr
	a (*nei zhong jiu*) (*hao*)		*	*	
	b (*nei zhong jiu hao*)			*	
☞	c (*nei zhong*)(*jiu hao*)				*

Note that the winning candidate (c) satisfies Binarity and Boundedness at the cost of violating Congruence. A comparison between (b) and (c) demonstrates the ranking Boundedness ≫ Congruence.

To summarize, the preliminary ranking among the various constraints applicable to MRUs is as follows (see section 1.5 for fine-tuning):

(19) {No Straddling, Binarity} ≫ Boundedness ≫ Congruence ≫ LtoR

1.3 Cyclic application of TS

Having isolated the MRUs as the domain of obligatory TS, we next address the question of how the TS rule applies to a string of T3's within an MRU. We alluded earlier to the unexpected survival of T3 within an MRU. Recall (10d) of section 1.2, repeated here:

(20)	[*zhi*-[*lao-hu*]] *pao*	"the paper tiger runs"
	3 3 3 3	base tone
	(3 s s 3)	sandhi form

All we need to say is that TS applies cyclically. In fact, as I will show in section 2, MRUs must be constructed cyclically first within the word, then at the phrase level. The derivation of (20) proceeds as follows:

(21)	[*zhi*-[*lao-hu*]] *pao*	"the paper tiger runs"
	3 3 3 3	base tone
	(2 3)	Lexical MRU, TS
	(3 2 3)	Lexical MRU, TS not applicable
	---	---
	(3 2 2 3)	Phrasal MRU, TS = (20)

The application of TS to the inner-bracketed *lao-hu* "tiger" preempts TS applying to the next larger constituent *zhi*-[*lao-hu*] in the usual bleeding relation. Nothing prevents TS from operating on the outermost layer [*zhi*-[*lao-hu*]] *pao*. The same mode of rule application generates the attested sandhi pattern of *zhao* [[*dan-xiao*]-*gui*] "look for the coward" (10e).

Where structure does not determine the order of derivation, TS operates iteratively from left to right. This accounts for

(22)	wu.wu.wu	"5-5-5"
	3 3 3	base tone
	(2 2 3)	MRU, TS = (8b)

Across MRUs TS *optionally* applies, given the appropriate sandhi contexts. This accounts for the two equally acceptable readings in (23):

(23)	head simple	
	[*tou-nao*][*jian-dan*]	"naive, simple-minded"
	2 3 3 1	base tone
	(2 3) (3 1)	i. MRU-internal TS not applicable
	(2 2) (3 1)	ii. TS across MRU

In any case, obligatory intra-MRU tone sandhi takes precedence over the optional inter-MRU application. This point is illustrated below:

(24)

want buy antique		
xiang [*mai* [*gu-dong*]]		"wants to buy antiques"
3 3 3 3		base tone
(2 3) (2 3)		intra-MRU TS
		inter-MRU TS not applicable

Intra-MRU TS on *gu-dong* effectively removes the requisite environment for the reapplication of inter-MRU TS on *mai*.

1.4 The place of MRUs in the prosodic hierarchy

The preceding three subsections encapsulate the core of the analysis proposed here. What I have done so far is to recast and make more explicit the basic ideas that were implicit in a set of foot formation rules proposed by Shih (1986:110):[6]

(25) **Foot Formation Rules**

a. Immediate Constituency: Link immediate constituents into disyllabic feet.
b. Duple Meter: Scanning from left to right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.
c. Super-foot Construction: Join any leftover monosyllable to a neighboring binary foot according to the direction of syntactic branching.

Clause (a) couches the No Straddling constraint in processual terms: all immediate constituents automatically coalesce into a foot. Clause (b) embodies the same principle as a left-to-right, binary rhythmic organization. Clause (c) follows directly from the undominated Binarity condition: since monosyllables are subminimal, they must per force join a neighboring foot. The branching directionality condition on (c) translates our generic metrico-syntactic Congruence constraint.

The remarkably simple and general principles of eurhythmy (Binarity, Boundedness, LtoR) and metrico-syntactic match (No Straddling, Congruence) – or their equivalent formulation as Foot Formation Rules – are capable of handling most of the fundamental facts of Beijing tone sandhi. However, there remains a sizable residue of recalcitrant cases, which will be the subject of the ensuing sections. In particular, I will argue that most of these remaining problems can be resolved by a two-pass MRU construction. That is to say, MRUs are built first over word-size units, then over larger, phrasal constructions. Before we proceed, we need to make a terminological clarification and take care of a few other matters.

[6] Revised in Shih (1990:18). A similar set of foot formation rules has been proposed to relate the syntactic structure of a line to the metrical template of classical Chinese poetry, see Chen (1984).

Following Chen (1979, 1984), Shih (1986, 1997) and subsequent works on tone sandhi have generally used the term "(metrical) foot" to refer to the sandhi domain. This creates something of a terminological confusion. In prosodic morphology, the foot is a sublexical unit that circumscribes, for instance, the scope of reduplication (cf. McCarthy and Prince 1986, 1990, 1993a). In metrical phonology, the foot is a stress unit, organized around a head (stressed syllable or mora). As a consequence, a stressable word is at least coextensive with a stress-foot as in (*bird*), (*e.dit*), (*Wa.shing.ton*); often larger than the foot as in (*Mis.sis*)(*sip.pi*), (*per.so*)(*na.li.ty*). Two stressable words[7] therefore always constitute two separate feet (cf. Liberman and Prince 1977, Hayes 1982, 1989, inter alia). This is also how the term "foot" is used in Yip (1980, 1995), Beattie (1985), and others. Quite reasonably, Yip (1995:490) reserves the term "foot" for "units containing only one toned syllable." In this light, the term "(stress) foot" is entirely appropriate when applied to the Wu-type dialects discussed at length in the preceding chapter. The foot in Shanghai and other Wu dialects is *word-bound*, i.e. the foot is either a sublexical constituent as in (*ø.li*)(*moq.soq*) "to speculate, conjecture" (lit. "groping in the dark"), or coextensive with a lexical compound as in (*tçiɔ.çiã.hiʌq*) "symphony," but not a full phrasal construction. Thus sentence (26) can be parsed only as (a) not as (b), as diagnosed by the tonal distribution.

(26)	boat	again	toward	front	row	Prt	
	zø	*iɤ*	*zɔ*	*zi*	*<u>h</u>iɔ-ləq*		"The boat started to row forward again"
	LH.	HL.	LH.	LH.	LH. o		base tone
a.	(LH)	(HL)	(LH)	(LH)	(LH.o)		ok (from Xu et al. 1988:527)
→	(LH)	(HL)	(LH)	(LH)	(L. H)		via Spread = ok
b.	(LH.	HL)	(LH.	LH)	(LH.o)		
→	(LH.	o)	(LH.	o)	(LH.o)		via Deletion
→	(L.	H)	(L	H)	(L. H)		via Spread = *

Notice that each of the first four syllables forms a separate foot, hence is stressed and retains its own lexical tone.[8] *Ləq*, a sentence particle with an inchoative/perfective meaning, behaves like an inherently toneless enclitic. Cross-word footing would wrongly predict foot-internal tone deletion and spread as in (b).

[7] At least two content words. For the treatment of function words in metrical phonology, see Selkirk (1984).

[8] Ignoring certain postlexical tonal modifications; for details, see chapter 7, section 3.

If we equate a foot with a unit containing only one tone-bearing syllable, as Yip (1995) proposes, then the term "foot" is inappropriate as a characterization of the domain of obligatory tone sandhi in Beijing Mandarin. As stated, TS requires two abutting T3-carrying syllables. Should we perhaps use the term p-word to refer to MRUs, as suggested for instance by Yip (1995:490)? Unfortunately, the term p-word raises problems of its own. P-word has at least two different meanings. In one usage, let's call it p-word$_1$, it is organized around a root or stem. Typically a p-word$_1$ consists of a stem plus surrounding affixes or other specified elements (cf. Nespor and Vogel 1986:109f.). In no case does the p-word$_1$ spill over the limits of a syntactic word. According to another usage (as in Beattie 1985, L. Cheng 1987, Hsiao 1991), the p-word – let's label it p-word$_2$ – is synonymous with "clitic group." A p-word$_2$ or clitic group is usually defined as a content word plus surrounding function words, which behave like prosodically dependent pro- or en-clitics (cf. Nespor and Vogel 1986:145f.). We cannot equate MRU with either p-word or clitic group, simply because the MRU in the Mandarin-type of languages is not restricted to word-size units, organized around a single content word. In principle, MRU can cut across *all and any* syntactic junctures, as clearly seen in the following examples:

(27) a. simplex, monomorphemic words
ma-yi "ant"
(s 3)

b. compound words
guard danger
bao-xian "to insure"
(s 3)

small rain
xiao-yu "drizzle"
(s 3)

external sister
biao-[*zi-mei*] "female cousin"
(s 3 4)

c. phrases
nine foot long
[jiu chi] chang "nine feet long"
(s 3 2)

very early rise
[*hen zao*] *qi* "get up very early"
(s s 3)

write novel
xie [*xiao-shuo*] "write a novel"
(s 3 1)

d. clauses
water very hot
shui [*hen re*] "the water is very hot" (Subject + Predicate)
(s 3 4)

I write book
wo [*xie shu*] "I am writing a book"
(s 3 1)

e. interclausal
want buy car
xiang [*mai che*] "plan to buy a car"
(s 3 1)

he leave good
[*ta zou*] *hao* "it is good that he left"
(1 s 3)

These examples clearly demonstrate that projections of any rank, ranging from single morphemes to biclausal constructions, can coalesce into one single MRU. In this sense, MRU is not "rank-sensitive."[9] Conversely, one cannot equate MRU with a larger-than-word unit such as the phonological phrase, because the MRU can be a sublexical unit as well, as in

(28) *suo.ma.li.ya* "Somalia"
3 3 3 3
(s 3)(s 3)

There is no question about *Somalia* as a single word, in fact an unanalyzable monomorphemic word; on the other hand, we know that (28) comprises two separate MRUs, otherwise, *ma* and *li* would obligatorily undergo TS.

In short, MRUs can be both sublexical and supralexical; as a consequence, they do not fit neatly into the conventional prosodic hierarchy of foot–p-word–clitic-group–phonological phrase (cf. Selkirk 1980, 1981a, 1986, Nespor and Vogel 1986, Hayes 1989). Viewed in this light, the prosodic organization of Beijing Mandarin is of considerable typological significance. Clearly, the MRU stands apart from the conventional prosodic hierarchy; it is basically a device to group *syllables* of a wide variety

[9] Subject to the proviso that the lexical MRU take logical precedence over its phrasal counterpart (see section 2).

of grammatical ranks and status into rhythmic units, as determined largely by constituency or tree configuration. The MRU constitutes a prosodic unit sui generis; it is off-scale, hors-séries, so to speak. For this reason, I have settled on MRU as a neutral term vis-à-vis the conventional prosodic hierarchy.

Note, however, MRUs are not merely artefacts created exclusively to account for the tone sandhi phenomena. The notion of MRU is rooted in the long tradition of Chinese versification. It corresponds to the *poetic* unit, for which the term "foot" was originally used (Chen 1979, 1980, 1984), in keeping with the established conventions. To illustrate, the heptasyllabic lines in regulated verse are traditionally dissected into three "feet" each:[10]

(29) (*chun-can*) (*dao si*) (*si fang jin*)
spring silk-worm until death silk only-then exhaust
"The spring silk-worm keeps spinning silk until it dies (of exhaustion)"

(*la-ju*) (*cheng hui*) (*lei shi gan*)
candle become ash tear only-then dry
"The candle keeps shedding tears until it is reduced to ashes"

Notice that, except for the initial two metrical units, each of the poetic feet is a supralexical entity, made up of phrase-like constructions, for instance: (*dao si*) "until death," (*cheng hui*) "become ashes" can hardly be construed as lexical compounds. In this respect, the Chinese use of the term (poetic) foot is not different from the Western tradition, in which the corresponding metrical unit of verse can be larger or smaller than a word, as illustrated by such lines taken randomly from Kiparsky (1975:599):

(30) a. And art / made tongue- / tied by / author / ity (Son.66)
b. Nothing, / sweet boy. / But yet / like prayers / divine (Son.108)
c. Making / dead wood / more blest / than li / ving lips (Son.128)
d. Simply / I cred / it her / false-speak / ing tongue (Son.138)

/ separates poetic feet

A (poetic) foot may consist of a word (*nothing*, *divine*), of fragments of a word (autho / rity), or span several lexical items (*sweet boy*, *more blest*); the foot, as a unit of poetic scansion may encompass several stress-feet (*dead wood*, *like prayers*), and need not coincide with a constituent, morphosyntactically or phonologically (*I cred* / *it her*).

[10] With apologies to the poet Li Shang Yin (fl. AD 813–858): much of the beauty of his metaphors of undying love is lost in the translation.

1.5 Tempo, intonation, and emphasis

A sketch of Mandarin tone sandhi would not be complete without a word about tempo and intonation. The cyclic mode of rule application predicts a contrasting pair like (31a, b), where the right-branching structure (31b) can only have the sandhi pattern (3 s 3) (= reading (31b-i)).

(31) tiger gall

a. [*lao-hu*]-*dan* "fearless"
(s s 3)

b. *zhi*-[*lao-hu*] "paper tiger"
(3 s 3) i. ok
(s s 3) j. ok, only in allegro speech

This is certainly true in normal, perhaps deliberate, speech. However, in fast or allegro speech, it is possible for TS to apply simultaneously to a right-branching structure (3 [3 3]), as if it were a "flat" or "unstructured" foot, producing (s s 3) as an alternative sandhi pattern.

Furthermore, the parsing of connected speech into MRUs interacts with pragmatics and intonation phrasing in intricate ways. It appears that MRUs are bounded by intonation phrases (IPs). That is to say, an utterance can be cut up into IPs in several ways; within each IP, the syllables are grouped into MRUs, consistent with the various constraints formulated in the preceding sections. Crucially, MRUs may not cut across IP boundaries. This accounts for the alternative readings of (32).

(32) wolfhound bite Xiaoming

lang-gou [*yao xiao-ming*]		"the wolfhound bit Xiaoming"
2 3 3 3 2		base tone
(2 3)%(s 3 2)		i.
(2 s 3)%(3 2)		j.

% = intonation break

I take (32-i) to be the default reading, with the subminimal *yao* "bite" being stray-adjoined to the object NP, pursuant to the Congruence constraint. In addition, (32-j) is a possible reading, assuming an intonation phrasing of the sort indicated there.

Much of the indeterminacy that prompted Shih (1997) to abandon the directional branching condition on clause (c) of her Foot Formation Rules (equivalent to our Congruence constraint on MRU) can be traced to alternative intonational phrasings. As a case in point, (32-j) is only an apparent exception to Congruence. Since (32-j) is cut up into two separate IPs, and since MRU is IP-bound, *yao* "bite" has no place to go but join

lang-gou "wolfhound" on the left in order to satisfy Binarity. In all subsequent discussion I will assume the following inviolable constraint:

(33) **IP-Bound**
MRUs are IP-bound.

One final note: as Z-S. Zhang (1988) and Shih (1997) point out, speakers tend to mark an element under focus or contrast by placing before it an emphatic boundary (!). This boundary, like the IP boundary, becomes a new reference point for foot formation. Thus the usual rhythmic organization yields the unmarked reading (34-i). On the other hand, the emphatic reading of (34-j) presupposes a structure derivable only by a rhythmic parsing that is sensitive to pragmatically determined markers of emphasis and contrast.

(34) only buy stocks not sell stocks[11]
zhi [*mai gu-piao*] *bu* [*mai gu-piao*]
(s 3) (3 4) . . . i.
(3) !(s 3 4) . . . j.

! = emphatic boundary

The fact that the emphatic boundary enforces a subminimal rhythmic unit *zhi* "only," suggests the constraint we refer to as IP-Bound must outweigh Binarity.

The final ranking of the relevant constraints for MRUs is, then, the following:

(35) {No Straddling, IP-Bound} ≫ Binarity ≫ Boundedness ≫ Congruence ≫ LtoR

In order to better investigate the principles of rhythmic organization, in the subsequent sections I will steer clear of these interferences by limiting our discussion whenever possible to constructions uttered at normal tempo, and in one single "neutral" intonation, i.e. without emphasis or internal break.[12]

2. A two-pass MRU formation

The MRU-based account outlined in the preceding section readily accounts for an impressive array of sandhi facts. The model treats surface structures

[11] "Buy" and "sell" in Mandarin are *mai* T3 and T4 respectively.
[12] For further details concerning the interaction between intonation phrasing and foot formation, see Z. Zhang (1988), Shih (1997), and Mok (1993).

essentially as unlabeled trees, over which MRUs are built, consistent with the eurhythmic and metrico-syntactic congruence principles. This overly simple view is clearly in need of refinement. In particular, I will argue that MRUs are constructed first cyclically at the word level, then across the board at the phrase level.

2.1 *Lexical integrity*

Minimally, MRU formation must recognize the category rank "word." Consider the following contrasts. For emphasis I separate words from each other by "#".

(36) I plan buy book
a. *wo* # [*xiang* # [*mai* # *shu*]] "I plan to buy books"
(s 3) (3 1)

become inverse ratio
b. *cheng* # [*fan*-[*bi-li*]] "become inversely proportional"
(2 3) (3 4) i*.
(2 s 3 4) j.

= word boundary

The two examples are geometrically isomorphic, with a uniformly right-branching tree. If MRU formation treats these two examples indiscriminately as unlabeled trees, then we would be at a loss as to how to explain why two pairwise identical trees like (36a) and (36b) behave quite differently with respect to TS. Crucially, the reading (36b-i), which parallels (36a) is totally unacceptable. (36b), with an embedded trisyllabic word must constitute a single unbounded MRU. This pattern is absolutely robust, as corroborated by the examples of (37):

(37) play ghost trick
a. *shua* # *gui*-[*ba-xi*] "play dirty tricks"
(s 3) (3 4) i*.
(3 s 3 4) j.

invite old manager
b. qing # lao-[zhang-gui][13] "invite the old manager"
(s 3) (3 4) i*.
(3 s 3 4) j.

[13] It will be shown that "old manager" actually constitutes a single lexical compound. See section 4.

Mirror images of (36) and (37) are exemplified below.

(38) this kind wine good
a. [[*zhe* # *zhong*] # *jiu*] # *hao* "this kind of wine is good"
(4 3) (s 3)

cock tail wine good
b. [[*ji-wei*]-*jiu*] # *hao* "the cocktail is good"
(1 3) (s 3) i*.
(1 s s 3) j.

The lexical integrity of left-branching constructions is equally robust, as seen below:

(39) luxury item few
a. [*she-chi*]-*pin* # *shao* "there are few luxury items"
(1 3) (s 3) i*.
(1 s s 3) j.

fall-guy[14] stupid
b. [*ti-si*]-*gui* # *sha* "the fall-guy was stupid"
(4 3) (s 3) i*.
(4 s s 3) j.

In all cases, it is not possible to break up the embedded trisyllabic word to form two binary MRUs.

It was facts like these that prompted Shih (1986:136f.) to conclude that her Foot Formation Rules (FFR), which she considers to be postlexical (p. 138), must nevertheless respect the "Lexical Integrity" of words; that is to say, FFR do not break up lexical items in order to use part of a word to form a foot with another syllable. Curiously, she abandoned the notion of Lexical Integrity in favor of a meaning-based filter that rules out foot constructions in violation of lexical integrity (cf. Shih 1990 and p.c.). I will argue against a meaning-based filter in section 8. Instead, the position I am advocating here is this: MRU formation is both lexical and postlexical; that is, it applies both in the lexicon and at the phrase level – in that order, consistent with the commonly accepted assumptions of lexical phonology. Once we accept the dual status of the MRU, there is no special need to either stipulate lexical integrity or posit any intermediate prosodic layer between the MRU and the IP. In other words, the contrast between (36a) and (36b) is straightforwardly derivable in the following fashion:

[14] lit: [[instead-of die] ghost], one who takes the blame instead of somebody else.

(40) "I plan to buy books" "become inversely proportional"

a. *wo # xiang # [mai # shu]* b. *cheng # [fan-[bi-li]]*[15]

3	3	3	1	2	3	3	4	base tone
					(s	3	4)	Lexical MRU, TS
(s	3)	(3	1)	(2	s	3	4)	Phrasal MRU, TS

In (b) the word *fan-bi-li* "inversely proportional" exits the lexicon with a "prefab" ternary MRU, so to speak. Therefore, **Free Element Condition** (Prince 1985), which prevents a previously metrified element from splitting off to form a new unit with some other member, bars *fan* from breaking off to form a binary MRU with *cheng* "to become." As a consequence, the only way the phrase *cheng # fan-bi-li* can satisfy Binarity, is to co-opt the monosyllabic *cheng* into the already structured (*fan-bi-li*) to form a longer unbounded MRU. The situation is quite different in the case of (a). It consists exclusively of monosyllabic words. Consequently, at the lexical level, no legitimate MRUs can be formed. The words, therefore, remain "free elements" for the purposes of grouping into MRUs. A bipartite (σσ)(σσ) prosodic structure is the only one that satisfies both Binarity and Boundedness.

The contrast between (38a) and (38b) can be handled in exactly the same manner.

2.2 *Cyclic MRU formation*

Since MRUs are formed in the lexicon as well as postlexically, it follows that MRU formation works cyclically or, equivalently, the constraints governing the MRUs must be satisfied cyclically (cf. Kenstowicz 1995, 1996). Both (41a) and (41b) constitute a single lexical compound (uninterrupted by an internal word boundary).

(41) exhibit hall director

a. *[[zhan-lan]-guan]-zhang* "director of the exhibition hall"

3	3	3	3	
(s	3)	(s	3)	i*.
(s	s	s	3)	j.

false small broadcast

b. *jia-[xiao-[guang-bo]]* "false rumor"[16]

3	3	3	1	
(s	3)	(3	1)	i*.
(3	s	3	1)	j.

[15] *Bi-li* "proportion" occurs as a free-standing word. *Fan* (lit. anti-, against), on the other hand, is a bound form, and functions as a derivational prefix.

[16] "Rumor" = lit. small + broadcast.

Under the assumption of cyclic MRU formation, Free Element Condition together with the Binarity constraint will force each successively added monosyllable into joining the preexisting MRU to form an increasingly longer rhythmic unit. Thus (41a) is derived as follows:

(42)	[[*zhan-lan*]-*guan*]-*zhang*				"director of exhibition hall"
	3	3	3	3	base tone
	(s	3)			MRU, TS
	(s	s	3)		MRU, TS
	(s	s	s	3)	MRU, TS

(41b), with the mirror-image construction, is derivable straightforwardly in the same manner. Cyclic TS application yields the desired results.

(43)	false small broadcast				
	jia-[*xiao*-[*guang-bo*]]				"false rumor"
	3	3	3	1	base tone
			(3	1)	MRU, TS
		(s	3	1)	MRU, TS
	(3	s	3	1)	MRU, TS not applicable

Theoretically, the MRU that grows in size by annexing one monosyllable at a time can span across as many syllables as allowed by the principles of word formation and phrase-structure constraints. (44) gives a pentasyllabic example.

(44)	fountain		pen[17] good		
	[[[zi-lai]-shui]-bi] # hao				"fountain pens are good"
	(4	2)			Lexical MRU, TS not applicable
	(4	2	3)		Lexical MRU, TS not applicable
	(4	2	s	3)	Lexical MRU, TS
	(4	2	s	s	3) Phrasal MRU, TS

2.3 *Lexical and phrasal TS*

The picture that emerges is this: TS must apply at both lexical and postlexical strata. This conclusion is supported by the following piece of evidence: (27c–e) of section 1.4 clearly show that TS must be postlexical, affecting obviously phrasal constructions such as *xie xiao-shuo* "to write novels" and *shui hen re* "the water is very hot." How do we know that TS operates also at the lexical level? We know this because TS must precede phonological rules that are demonstrably lexical. Consider the following triplet:

[17] "Fountain pen" = lit: [[[self-running] water] pen].

(45) sister sister

a.	*jie-jie*	"older sister"
	3 o	base tone, unchanged
	young sister	
b.	*xiao-jie*	"miss" (unmarried woman)
	3 3	base tone
	2 3	TS
	2 o	Deletion
	big sister	
c.	*da-jie*	"the oldest sister" (an older woman)
	4 3	base tone, unchanged

The kinship term *jie-jie* "sister" is derived through a morphological process of reduplication that copies only the segmental material, without the tone (cf. Packard 1990, Chen 1992b, Sproat and Shih 1993). Therefore, the second *jie* in (45a) is inherently toneless. This explains why the first *jie* does not undergo TS, since there is no second T3 to trigger TS. The same *jie* in (45b) is also toneless phonetically. However, it must start out with an underlying T3, otherwise we cannot explain why *xiao* "young, small" surfaces with a derived T2. The obvious analysis is to posit /3-3/ as the underlying tonal representation for (45b), to which TS and Tone Deletion apply – in that order. Now, there is no question that Tone Deletion, also known as the Neutral Tone rule, belongs to the lexical stratum. Tone Deletion is lexically idiosyncratic. For instance, while it applies to (45b), it does not in (45c). Chao (1968:39) has pointed out that the tonal reduction in weakly stressed syllables is unpredictable on phonological, syntactic, or semantic grounds, and that each lexical entry must be individually marked whether or not to undergo the neutral tone rule. This lexical idiosyncrasy is amply borne out by a recent study by C-Y. Chen (1984). To cite a few examples: *lao-hu* "tiger" [2-o] but *lao-shu* "mouse" [2-3], *da-yi* "careful" [4-o] but *da-yi* "summary, core idea" [4-4], etc. Lexical idiosyncrasy is widely accepted as a hallmark of bona fide lexical rules (cf. Hayes 1990, Kaisse 1990). Since TS must apply before the Neutral Tone rule, witness (45b), TS must apply in the lexicon as well. By the same token, since tone sandhi presupposes MRU, MRUs must be constructed at the lexical level as well.

To summarize, there are three types of evidence for "two-pass" MRU formation, once at the lexical level, and then at the phrasal level: (i) lexical MRU formation precedes and preempts phrasal MRU construction, consistent with the Free Element condition; (ii) lexical MRU formation is

cyclic, while phrasal MRUs are constructed iteratively left to right; (iii) TS, hence MRU formation by implication, must apply before some lexical rules such as the Tonal Deletion or Neutral Tone rules.

That lexical MRU formation and lexical TS are cyclic is made clear above. In contrast, phrasal MRU formation and TS must operate across the board. Compare these two tonally and configurationally identical strings:

(46) false small broadcast

a.	*jia*-[*xiao*-[*guang*-*bo*]]					"false rumor"
	3	3	3	1		base tone
			(3	1)		Lexical MRU, TS not applicable
		(s	3	1)		Lexical MRU, TS
	(3	s	3	1)		Lexical MRU, TS not applicable
	want		write	novel		
b.	*xiang* # [*xie* # *xiao-shuo*]					"plan to write a novel"
	3		3	3	1	base tone
				(3	1)	Lexical MRU, TS not applicable
	(s		3)	(3	1)	Phrasal MRU, TS i = ok
	(s		s)	(3	1)	Cross-MRU TS optional j = ok

If a ternary MRU were constructed on the complement VP *xie xiao-shuo* in (46b), the prediction would be a reading that is identical tonally speaking to (46a). Instead, the most natural reading of (46b) is a two-MRU sequence (s.3)(3.1), although (s.s)(3.1) is also possible via optional inter-MRU tone sandhi.

3 The syntactic word

In the preceding section we appealed to the notion of the word[18] to circumscribe the domain of cyclic lexical foot formation. Given the crucial use we make of this notion, it is important to make more precise the notion of the word. No one questions the lexical status of compounds like *ji-wei-jiu* "cocktail," *fan-bi-li* "inverse ratio," *gui-ba-xi* "dirty tricks," or *zi-lai-shui-bi* "fountain pen" as syntactic words on account of their limited productivity, idiomaticity, and so forth. In other cases, the lexical status of certain expressions is far less clear-cut. I will examine two of

[18] I will use the term "word" tout court to refer to the syntactic/morphological word, i.e. a lexical item that passes the standard morphosyntactic tests for wordhood. Where a distinction is relevant, I will use the term "p-word" (phonological word) to refer to a phonologically cohesive word-like unit.

them: the modifier + noun and verb + resultative complement constructions. Phonologically they behave like integral lexical items. Interestingly, independent syntactic tests bear out their status as words rather than phrasal constructions.

3.1 Modifier + noun

Side by side with the more or less idiomatic expressions like *ti-si-gui* "the fall guy" (lit. "a devil that dies instead of someone else"), we also have a host of regular and transparent [M N] (= modifier + noun) constructions like *xiao-[ru-ge]* "small + squab = small squab," or *[pu-tao]-jiu* "grape + wine = grape wine," etc. How we decide on their wordhood has immediate consequences for their sandhi behavior. Consider the following contrast (from Hsiao 1991):

(47)
a. want roast squab
xiang # [*kao* # *ru-ge*] "want to roast squabs"
(s 3) (3 2)

b. roast small squab
kao # [*xiao* *ru-ge*] "roast small squabs"
(s 3) (3 2) i*.
(3 s 3 2) j.

The first sentence breaks into two binary MRUs. The second, however, must form a single MRU. If we parse *xiao* + *ru-ge* "small squabs" as a phrase consisting of two syntactic words, on a par with *kao* + *ru-ge* "to roast squabs," then we have no way of accounting for the contrast pointed out by Hsiao. On the other hand, if we construe *xiao* + *ru-ge* as a lexical compound, then the lexical MRU constructed on (*xiao-ru-ge*) would prevent *kao* from joining *xiao* into a separate MRU at the postlexical level, consistent with the Free Element Condition. As a matter of fact, there is independent evidence that points to [M N] as lexical rather than phrasal constructions.

Shih (1986) and Sproat and Shih (1991) argue that Mandarin Chinese has two types of [M N] constructions: (i) direct modification [M N] as in *jiu shu* "old book," *ci zhuan* "ceramic tile," *xiao ru-ge* "small squab," *mu-tou fang-zi* "wooden house" (lit. wood + house), *zui-di gong-bei-shu* "smallest common multiple," where the modifier is juxtaposed with the modified; (ii) indirect modification [M-de N], where the modifier is linked to the head by the subordinator *de*, which also marks possessives and relative clauses (cf. C. Huang 1987). Examples of the second type are: *hong de hua* "red flower," *hen xiong de lang-gou* "very ferocious wolfhound,"

dai yan-jing de nan-hai-zi "a boy that wears eye-glasses" (gloss: [wear eyeglasses DE] boy). They further note that [M N], but not [M-de N], is subject to certain (universal) *Adjective Order Restrictions*. Thus the strict sequencing SIZE > QUALITY > SHAPE (X > Y means X precedes Y) imposed on [M N] – but not on [M-de N] – constructions is illustrated by examples (48)–(49).

(48) a. *xiao lü hua-ping* (SIZE > QUALITY)
"small green vase"

b. **lü xiao hua-ping*
*"green small vase"

c. *xiao de lü de hua-ping*
small DE green DE vase
"a green vase that is small"

d. *lü de xiao de hua-ping*
"a small vase that is green"

(49) a. *hao yuan pan-zi* (QUALITY > SHAPE)
"nice round plate"

b. **yuan hao pan-zi*
*"round nice plate"

c. *hao de yuan de pan-zi*
nice DE round DE plate
"a round plate that is nice"

d. *yuan de hao de pan-zi*
"a nice plate that is round"

Now, there is no question about [M-de N] as phrases, since the M-de could be a simple A, an AP, or a full-fledged relative clause.[19] In contrast, there are good reasons to think that [M N] constructions are syntactic words or nominal compounds.

First, Zhu (1956; cf. Duanmu 1993b) have already pointed out that [M N] constructions in Chinese are not fully productive, but often represent more or less frozen collocations with many unpredictable gaps. Thus all the (b) forms in (50)–(52) are odd in contrast with the (a) forms. There is nothing semantically incongruous about the [M N] combinations in (b), as shown by their English translations. Furthermore, they all turn into perfectly normal expressions when the subordinator *de* is added, as seen in (53).

[19] In fact Sproat and Shih (1991) regard all M-de constructions as (reduced) relative clauses.

(50)	a.	*gao shan*	"tall mountain"
		gao lou	"tall building"
	b.	**gao shu*	"tall tree"
		**gao ren*	"tall person"
(51)	a.	*zang yi-fu*	"dirty clothes"
	b.	**zang dian-ding-pao*	"dirty bulbs"
(52)	a.	*hua-ji dian-ying*	"funny movie"
	b.	**hua-ji tong-xue*	"funny classmate"
(53)	a.	*gao de shu*	"tall tree"
	b.	*gao de ren*	"tall person"
	c.	*zang de dian-ding-pao*	"dirty bulbs"
	d.	*hua-ji de tong-xue*	"funny classmate"

Secondly, [M N] is often non-compositional in meaning. Thus the DE-less *hei-ban* of (54a) is interpreted non-compositionally, which accounts for the non-contradiction of (54b). The [M-de N] phrase, on the other hand, is semantically ill-formed.

(54)	a.	*hei-ban*	"blackboard"
	b.	*bai de hei-ban*	"white blackboard"
	c.	**bai de* [*hei de ban*]	"white black board"

Finally, a number of restrictions on [M N] constructions can be explained if we construe them as word-level, not phrase-level, projections. On this construal it follows that neither M nor N can be a phrasal category, but must both be word-level constituents, for otherwise we would have the anomaly of X^o-internal phrases: $[M\ NP]_{N^o}$ or $[XP\ N]_{N^o}$. This prediction is borne out. According to Zhu (1956) and Sproat and Shih (1991), direct modifiers – that is DE-less M – only modify word-level projections of N, whereas modifiers with -DE modify higher projections (Sproat and Shih, p. 572). This accounts for the ungrammaticality of (55b), since the DE-less *xin* "new" modifies a phrasal constituent *gui de da-yi* "expensive overcoat." If this is true then the grammaticality of (55c) identifies the inner [M N] [*fen-hong da-yi*] "pink overcoat" as a word-level projection N^o: otherwise it could not be further directly modified by the DE-less *xin* "new."

(55) a. *gui de* [*xin da-yi*]
expensive DE new overcoat
"expensive new overcoat"

b. **xin* [*gui de da-yi*]
new expensive DE overcoat
"new expensive overcoat"

c. *xin* [*fen-hong da-yi*]
"new pink overcoat"

Just as there are limitations on the nominal head in [M N] constructions, there are stringent restrictions on the modifier as well: only words, not phrases, can function as direct prenominal modifiers. Thus, when the adjective *xin* (in 56a) is itself expanded into an AP, it can no longer directly modify *da-yi*, as shown in (56b). AP, on the other hand, can enter into an [M-de N] construction such as (56c). This observation follows from the assumption that [M N]s are word-level projections, since X^o-level constructions do not generally allow an internal constituent of a higher rank.[20]

(56) a. *xin da-yi*
"new overcoat"

b. *[*bi zhe jian xin*] *da-yi*
than this CL new overcoat
"the overcoat that is newer than this one"

c. [*bi zhe jian xin de*] *da-yi*
than this-CL new DE overcoat
"the overcoat that is newer than this one"

By analyzing [M N]s as word-level projections, cyclic MRU formation and TS straightforwardly account for the contrastive pair (47a, b). Lexical FFR first builds a ternary MRU foot over (*xiao-ru-ge*) "small squab" thereby preempting a binary MRU (*kao xiao*) "roast + small." The derivation of (47b) is given as below:

(57) roast small squab
kao [*xiao*-[*ru-ge*]]

3	3	3	2	base tone
		(3	2)	Lexical MRU
	(s	3	2)	Lexical MRU, TS
(3	s	3	2)	Postlexical MRU, TS = not applicable

[20] Duanmu (1993b:16, fn.4) cites a few isolated counterexamples such as *zui gao ji* "most high level = the highest level," *zui da xian-du* "most large limit = the top limit." As he observed, these expressions are not productive, and all involve the word *zui* "most." One could add that *zui* being a bound form, it could in principle be analyzed as an affix like the superlative *-est* in English.

3.2 Complex predicates

As a further confirmation of the approach we have taken, the cyclic effect of lexical MRU formation can be discerned in verbal expressions as well. Consider examples like (58).

(58) dog noisy wake:up Xiaoming
gou $_{VP}$[*chao* $_{S}$[*xing xiao-ming*]] "the dog woke up Xiaoming by making noise"
(s 3) (s 3 2) i. *
(3 s 3) (3 2) j. ok

On one account (cf. Hashimoto 1971), the action verb *chao* "make noise" takes a clausal complement like [*xing xiao-ming*], in which *xing* is the verb of the complement clause, which means something like "so that Xiao-ming wakes up." Such an analysis is at odds with its tone sandhi behavior. Given a structure like (58), we would predict (58-i) to be the most congruent MRU. The reason is that, on this construal, *xing* is a clause-mate of *xiao-ming* and should therefore be rhythmically integrated with the latter, consistent with the Congruence constraint, for the same reason we parse (59) as (σσσ) (σσ), rather than (σσ) (σσσ).

(59) guess water very cold
cai-xiang $_{S}$[*shui* [*hen* *leng*]] "(I) guess the water is very cold"
(1 3) (3 s 3) i. ok (= (14))
(1 s 3) (s 3) j. ??

This makes the wrong prediction for (58), where *xing* and *xiao-ming* demonstrably belong to two separate MRUs. In order to guarantee the correct MRU parsing of (58-j), we must construe *chao-xing* as a complex predicate, consisting of a verb + resultative complement:

(60) dog noisy-wake:up
gou $_{VP}$[$_{V}$[*chao-xing*] $_{NP}$[*xiao-ming*]]
3 3 3 3 2 base tone
(s 3) (3 2) Lexical MRU, TS

(3 s 3) (3 2) Postlexical MRU,
TS not applicable / optional
= (58-j)

Tone sandhi behavior is consistent with the syntactic analysis of (60) motivated on independent grounds. Both *chao* "be noisy" and *xing* "to wake up" are intransitive verbs, but the verbal complex *chao-xing*,

is a two-place predicate, licensing an NP like *xiao-ming,* which functions as the object of the morphologically complex verb *chao-xing* rather than simply as the subject of *xing*, as suggested by the passive construction (61).

(61) *xiao-ming bei gou chao-xing le*
Xiao-ming Pass dog noisy-wake:up Asp
"Xiaoming has been awoken by the dog making noise"

Chao-xing then behaves as a complex predicate. It is not crucial for our purposes whether the verbal complex *chao xing* is derived by means of lexical operations (Thompson 1973, Bresnan 1982, Di-Sciullo and Williams 1987), or by syntactic transformations through verb-raising, head-to-head movement, or clause union from a biclausal source like (62) (cf. Hashimoto 1971, Baker 1988).

(62) *gou chao-de* [*xiao-ming xing le*]
dog noisy DE Xiaoming wake:up Asp
"the dog made so much noise that Xiaoming woke up"

In either case, the most widely accepted assumption is that "all predicate forming operations . . . yield complex morphological objects, i.e. morpho-phonologically integrated wordforms" (Ackerman and Webelhuth 1992:5; cf. references cited therein).[21]

There is independent evidence that, regardless of their syntactic or lexical origin, the end result *chao-xing* and similar verbal complexes clearly exhibit word-like behavior. The unitary wordhood of such expressions can be diagnosed by standard tests such as affixation, phrase-structure conditions, collocational restrictions, and so forth. Without belaboring the point, let us briefly review some of these tests. Aspect markers like *-zhe* (progressive), *-le* (perfective), and *-guo* (indefinite past, i.e. "sometime in the past") are word-level affixes rather than phrase-level clitics. This can be seen in (63a, b).

(63) a. *da-guo wang-qiu*
play-Asp tennis
"have played tennis (in the past)"

b. **da wang-qiu guo*

[21] Assuming that changes in valence and argument structure are generally diagnostic of lexical operations, the process involved in the creating of expressions like *chao xing* is clearly lexical.

If *chao xing* indeed constitute separate verbs in series, we would expect such aspectual markers to attach themselves to either one of the constituent verbs. This wrongly predicts the grammaticality of (64a). Instead, the only grammatical position for the aspect marker is as shown in (64b), suggesting that the host to which the aspect marker is affixed is the entire morphologically complex word/stem *chao-xing*.[22]

(64) a. **chao-guo xing xiao-ming*
noisy-Asp wake:up Xiao-ming
"woke up Xiaoming (some time in the past) by making noise"

b. *chao-xing-guo xiao-ming*

Another diagnostic of the wordhood of *chao-xing* derives from a phrase-structure condition of Chinese, which stipulates that a verb may be followed by at most one constituent (J. Huang 1982). This accounts for the contrast between (65) and (66):

(65) a. *dan-xin ta di-di*
carry-heart he younger-brother
"be worried about his younger brother"

b. *fu-ze zhe-jian shi*
bear-responsibility this-CL matter
"take responsibility for this matter"

(66) a. **bo pi ju-zi*
peel skin orange
"peel an orange"

b. **kai wan-xiao tong-xue*
make fun schoolmate
"tease a schoolmate"

In each case the verbal complexes *dan-xin* (lit. carry heart/mind), *fu-ze* (bear responsibility), *bo pi* (peel skin), and *kai wan-xiao* (make fun) have the internal structure of verb + object; but whereas the first two (65a, b) may take an "outer" object, the last two (66a, b) may not. This means that the internal verb + object structures of *dan-xin* and *fu-ze* are syntactically opaque as far as the phrase-structure condition is concerned.

[22] In order to rule out (64a), we cannot simply stipulate that aspect markers can be attached only to the last of a series of verbs. This would wrongly predict the following sentence to be ungrammatical:
ta ceng-jing xiang-guo [*qing bao-mu*]
s/he once consider-Asp hire babysitter
"he once considered hiring a babysitter"

In this respect *chao-xing* parallels *dan-xin* and *fu-ze* in that *-xing* is not construed as a postverbal constituent, but as part of the verb itself, so that it licenses a postverbal complement, in this case *Xiaoming.*

Finally, the wordhood of *chao-xing* can be tested by collocational criteria as well. For instance the degree adverb *hen* "very" can modify only adjectives or stative verbs. Thus (67a) is perfectly fine. But the fact that (67b) is ungrammatical (though not semantically incongruous) suggests that the collocational restriction no longer holds between *hen* and *chao* as a separate stative verb from *xing*; instead *chao-xing* is treated as a single action/causative verb, and can no longer cooccur with degree adverbs like *hen*; cf. (67c).

(67) a. *ke-tang / hai-zi hen chao*
classroom / child very noisy
"the classroom / the child is very noisy"

b. **hai-zi hen chao-xing xiao-ming*
child very noisy wake Xiaoming
"the child woke up Xiaoming by making a lot of noise"

c. **ta hen da wang-qiu*
"he plays tennis a lot"

It is for reasons like these that verb + resultative complexes, including those in the so-called *potential mode* (e.g. *jiang-de-qing-chu* lit. explain-can-clear, "explain clearly"; *kan-bu-dao* lit. chop-not-fall, "cannot chop down"), are generally analyzed as complex predicates having the status of unitary syntactic words instead of phrasal constructions (cf. Chao 1968, Thompson 1973, Packard 1990, Dai 1992, Sproat and Shih 1992, Duanmu 1993b).

Strictly speaking, in order to insure the proper rhythmic parsing of (60), we need to say only that [*chao xing*] are immediate constituents, though not necessarily pieces of a single (syntactic) word. The lexical effect of verb + resultative constructions is more clearly demonstrated by (68).

(68) repair fine watch
[*xiu-li*]-*hao* # *biao* "repair the watch so that it works fine"[23]
(1 s s 3) i
(1 3) (s 3) j*

[23] Needless to say, there is another reading of (68),
xiu-li [*hao biao*]
"to repair a fine watch"
which is not the one intended here. More importantly, *hao* "fine" in this context is not the aspectual marker of completion as in

Correct reading (i) can be generated only by a MRU formation that cycles on the lexical structure *xiu-li-hao*, as shown in (69).

(69)

repair fine watch		
[*xiu-li*]-*hao* # *biao*		
1 3 3 3		base tone
(1 3)		Lexical MRU
(1 s 3)		Lexical MRU, TS

(1 s s 3)		Postlexical MRU, TS = (68-i)

This ternary MRU built over *xiu-li-hao* preempts the formation of a binary MRU over *hao* and *biao* as in (70).

(70)

repair fine watch		
[[*xiu-li*]-*hao* # *biao*		
1 3 3 3		base tone
(1 3)		Lexical MRU
---		
*(1 3) (s 3)		Postlexical MRU, TS = (68-j)

The joint lexical effects of verb + resultative and modifier + noun compounds are evident in (71). It is only by taking *gao-xing-si* (lit. make-happy to death) and *lao-zhang-gui* (old manager) to be lexical items that we can correctly predict reading (71-j).

(71)

happy die old manager		
[*gao-xing*]-*si* # *lao*-[*zhang-gui*]		"(something) made the old manager extremely happy"
1 4 3 3 3 4		
(1 4) (s 3) (3 4)		i*.
(1 4 3) (s 3 4)		j.

If *si* "to die" and *lao* "old" were treated not as indivisible parts of larger lexical compounds, but as free agents, so to speak, nothing would prevent them from joining each other into a binary MRU. Recall that while No Straddling bars immediate constituents from splitting up into two MRUs,

mai-hao cai
buy-Asp grocery
"have finished grocery shopping"

Instead, the -*hao* in (68) corresponds to a sentence like:

biao xiu-li de [*hao bu hao*]
watch repair DE fine not fine
"Is the watch (working) well as a result of the repair?"

Such a construction is not available with -*hao* as aspect marker, cf:

**cai mai de* [*hao bu hao*]
grocery buy DE Asp not Asp

it in no way implies the converse, namely that MRU-mates must be immediate constituents. Further more, we know that the Boundedness constraint (no more than disyllabic) outranks the generic Congruence on account of examples like (72), where *xie* "write" forms a separate bounded MRU with a more distant *xiang* "want" rather than joining with its closest syntactic mate *xiao-shuo* "novel" into an unbounded rhythmic unit.

(72) want write novel
xiang # [*xie* # *xiao-shuo*] "plan to write a novel"
(s 3) (3 1)

Given this Boundedness ≫ Congruence ranking, we would predict that by the same logic *si* and *lao* would form a binary MRU, under the assumption that *si* and *lao* remain free elements as they enter into phrasal composition. The fact that the correct reading (71-j) presupposes two ternary MRUs argues for *gao-xing-si* and *lao-zhang-gui* as lexical units, each with its own "prefab" MRUs before they meet at the phrase level.

There is a bonus to the analysis proposed here. The "Duple Meter" clause (b) of Shih's (1986) Foot Formation Rules (see section 1.4) imposed an extra condition to the effect that Duple Meter is blocked if adjacent syllables branch in opposite directions structurally speaking.[24] This proviso was designed to block Duple Meter precisely in cases like (71). Subsequently Shih (1997 and p.c.) sought to appeal to some sort of *sense unit* condition to eliminate a semantically incoherent prosodic unit like (*si lao*) "die old" in (71), and discarded the opposite branching condition as a purely ad hoc stipulation. Unfortunately, as we shall see in section 9, we cannot count on a meaning-based filter to reliably rule out unacceptable prosodic organizations. In any event, parsing connected speech into MRUs at both lexical and phrasal levels automatically takes care of this kind of construction (as well as others; see the next section), and renders the opposite branching condition superfluous.

4 The phonological word

4.1 Clitic group

In all the cases discussed so far, the domain of cyclic MRU formation coincides with the syntactic word. We will now consider some additional facts.

[24] I plead guilty of the same shortcomings, since I appealed to basically the same FFR in Chen (1984).

(73) "as for famous paintings, there are (quite a few)
inside the presidential palace"
famous painting president palace inside have
ming hua [[[*zong-tong*]-*fu*]$_{NP}$ *li*]$_{PP}$ *you*

(s	s	s	s	3)	i ok
(s	3)	(s	s	3)	j *
(s	s	3)	(s	3)	k *

Zong-tong "president" is a free-standing word, consisting of two bound forms. *Fu* "palace, mansion," however is a bound form; it can occur only as part of a compound. The locative apposition *li* "inside" is also a bound morpheme, but with a difference: whereas *fu* is a root morpheme that combines only with other morphemes or words to create a word-level complex, *li* "inside" is a bona fide syntactic word – albeit phonologically bound – in construction with phrasal constituents. Thus, *fu* cannot combine with a phrase, as demonstrated by the contrast between (74a) and (74b); but *li* can, as illustrated by (74c):

(74) a. *[*na ge zong-tong*]$_{NP}$ fu
that CL president palace
"that president's palace"

b. *na ge* [[*zong-tong*]-*fu*]
that CL president palace
"that presidential palace"

c. [*na jian xue-xiao*]$_{NP}$ *li*
that CL school inside
"inside that school"

To label [NP *li*] as a word-level entity would create the anomaly of a word-internal XP, in contravention of the so-called expandability test of wordhood (cf. Dai 1992, Duanmu 1993b). Accordingly, (73) is bracketed as indicated above. Cyclic rhythmic parsing derives a trisyllabic MRU on the word *zong-tong-fu*. This leaves two dangling syllables *li* and *you*, which we expect to join into a binary MRU by phrasal left-to-right MRU formation. The predicted TS output is the ungrammatical reading (k) for (73), given in (75):

(75) president palace inside have
. . . [[[*zong-tong*]-*fu*]$_{NP}$ *li*] $_{PP}$ *you*

(s	3)				Lexical MRU, TS
(s	s	3)			Lexical MRU, TS
(s	s	3)	(s	3)	Phrasal MRU, TS = (73-k*)

In order to guarantee the only correct sandhi form of (73-i), it is necessary to make use of the notion of *phonological word* (p-word), which for our purposes is defined as the lexical host plus an adjoining clitic.[25] Clitics resemble affixes in being bound morphemes, but differ from the latter by their ability to attach themselves to a phrasal host. For this reason, the English possessive is characterized as a clitic rather than a suffix as in *[the queen of England]*$_{np}$*'s hat* (cf. Klavans 1982 inter alia). By this definition, the entire locative phrase *zong-tong-fu-li* constitutes one single p-word. The phonological dependency of *-li* identifies it as a clitic, something of a tertium quid between a word and an affix: as a word, its distribution is syntactically determined; as an affix (i.e. bound form), it does not occur freely, but must anchor itself in a neighboring host. For this reason, clitics are sometimes referred to as phrasal affixes. On this account, the combination NP-*li* constitutes a phrase (say, prepositional or postpositional phrase); however, this does not prevent NP-*li* from being treated as a phonologically cohesive entity, namely as a phonological word. Hence the cyclic effect of MRU formation operating on the p-word (as well as the syntactic word) correctly insures an unbounded MRU,[26] which in turn guarantees the correct sandhi form in (73-i). This is illustrated in (75′), where there is only one word boundary, signifying the lexical integrity of elements linked by the hyphens.

(75′) president palace inside have
. . . [[*zong-tong*]-*fu*]-*li* # *you*

(s 3)	Lexical MRU, TS
(s s 3)	Lexical MRU, TS
(s s s 3)	Lexical MRU, TS
(s s s s 3)	Phrasal MRU, TS = (73-i)

It is worth emphasizing rhythmic parsing treats p-words, i.e. word + clitic, as single units, on a par with syntactic words (s-words). The need for a prosodic entity such as p-word, also known as clitic group, is well recognized in prosodic phonology (cf. Zwicky 1977, 1985, Kaisse 1985, Nespor and Vogel 1986, Hayes 1989). We can further exploit this notion to explain a number of otherwise puzzling facts. By the same token, its explanatory value for these unrelated facts lends it further empirical support.

[25] In this sense, p-word (i.e. p-word$_2$ mentioned in section 1.4) is synonymous with the clitic group. I adopt this usage to facilitate comparison with other prosodic accounts.
[26] On the cyclicity of lexical MRU formation, see section 2.2.

Consider (76). The normal readings as given in (4) are not derivable from normal MRUs and TS, since the footing of immediate constituents takes precedence over all other means of foot building.

(76) cat than dog small
a. *mao* [[*bi gou*] *xiao*] "the cat is smaller than the dog"
(1 3)(s 3)

I toward north walk
b. *wo* [[*wang bei*] *zou*] "I walked toward the north"
(s 3) (s 3)

To account for these facts, Hsiao (1991) suggests that Foot Formation discriminates between "*lexical*" (content words) and "*functor beats*" (function words). There are a number of empirical as well as conceptual problems with this approach, which I will not go into at this point.[27] Instead, I am inclined to think that we can subsume the notion of functor beats under the well-attested process of cliticization.[28]

A good place to start talking about cliticization is the dichotomy between free and bound morphemes along the lines suggested, for instance, by Chao (1968:143f.). All free morphemes constitute words by definition.[29] Polysyllabic expressions, regardless of their category membership and grammatical function, occur as free forms; they behave therefore like lexical items. Examples of polysyllabic function words are: *ru-guo* "if," *sui-ran* "although," *na-me* "in that case." As for bound morphemes, they belong to three types: (i) *Bare roots* that cannot occur as minimal free forms, like *ya* "duck," *fu* "carry, bear," *li* "beautiful"; these root morphemes must combine with an affix or another root morpheme (free or bound) to form a free-standing word. Thus *ya-zi* (duck + suffix), *fu-ze* (bear + duty = take responsibility, both bound roots), *mei-li* (free root + bound root, both

[27] For instance, assuming "functor beats" for bound forms like *li* "inside" and *ye* "also," Foot Formation would proceed as follows:

ocean inside also pitch-dark
da-hai li ye hei-qi-qi "it is pitch dark also in the ocean"
4 3 3 3 1 1 1 base tone
(4 3) (1 1 1) Footing of "lexical beats"
*(4 3) (s 3) (1 1 1) Footing of "functor beats," TS

After the footing of the two lexical items *da-hai* "ocean" and *hei-qi-qi* "pitch dark," the two remaining functor beats should be footed together, with unacceptable sandhi effects.

[28] Hsiao (1991:90) observes that "in all likelihood, all functors are subject to cliticization." There seems to be no compelling reason to appeal to the notions of *both* functor beats *and* cliticization. However, the validity of the notion of "functor beat" seems well supported by the rhythmic effects exhibited in verse, nursery rhyme etc. For details see Hsiao (1991).

[29] Cf. Bloomfield's (1933:178) definition of the word as "the minimal free form."

meaning "beautiful"). (ii) Bound word-level *affixes*, such as the ordinal number prefix *di-* (*di*-N = "N-th," thus *san* "three," *di-san* "third,") the aspectual suffix *-zhe* (progressive), and the potential/resultative infix *de/bu* (e.g. *jiang-de-qing* "talk-DE-clear = able to explain clearly"; *xie-BU-wan* "write-NOT-finish = cannot finish writing"). (iii) *Clitics*, a tertium quid between words and affixes. Clitics are word-like in that their distribution is syntactically determined, i.e. they occur in construction with some phrasal constituent in a syntactic domain. On the other hand, they are affix-like in that they characteristically lack prosodic independence, and must attach themselves to a neighboring element to form a phonological constituent (p-word). One of the peculiarities of clitics is that they can attach themselves to a phonological host that is distinct from its structural (i.e. phrasal) anchor. This relative freedom of attachment creates precisely the kind of asymmetry between syntax and phonology that is manifest especially in tone sandhi.[30]

4.2 Sandhi behavior of clitics

With this in mind, let us examine the sandhi behavior of clitics, with particular reference to (i) prepositions, (ii) classifiers or measure words, and (iii) object pronouns.

4.2.1 Prepositions

First, reconsider (76a, b). Prepositions like *bi* "than" and *wang* "toward" are bound morphemes. But they are not word-level prefixes, since they can take fully expandable NPs, as illustrated in (77).

(77) *xiao-ming bi* [*na-ge dai yan-jing de hai-zi*] *cong-ming*
Xiaoming than that CL wear eyeglasses DE boy intelligent
"Xiaoming is more intelligent than the boy that wears eyeglasses"

Consistent with the phrase-structure rules of Chinese, prepositions, *bi* and *wang* occupy phrase-initial positions. Phonologically speaking however, they exhibit the kind of dependency characteristic of clitics, and therefore cliticize *leftwards* to form p-words with their phonological hosts *mao* "cat" and *wo* "I." The resulting clitic group, {host=clitic} (where "=" symbolizes the phonological *liaison* that links a clitic to its host), constitutes a single p-word. By extending the cyclic domain of lexical MRU formation to p-words as well as s-words, we automatically guarantee the

[30] Concerning the criteria for discriminating words from clitics, and clitics from affixes, see Zwicky and Pullum (1983), Zwicky (1985), Klavans (1985), Kaisse (1985), Dai (1992).

preemptive rhythmic organization of the clitic groups. This accounts for the asymmetry between the morphosyntactic bracketing and the prosodic organization of (76a, b) reanalyzed here as (78a, b).

(78) cat than dog small

a. *mao* $_{VP}$[$_{PP}$[*bi gou*] *xiao*] "the cat is smaller than the dog"
mao = *bi* $_{VP}$[*gou xiao*] Cliticization
1 3 3 3 base tone
(1 3) Lexical MRU
(1 3) (s 3) Phrasal MRU, TS

I toward north walk

b. *wo* $_{VP}$[$_{PP}$[*wang bei*] *zou*] "I walk toward the north"
wo = *wang* [$_{VP}$ *bei zou*] Cliticization
3 3 3 3 base tone
(s 3) Lexical MRU, TS
(s 3) (s 3) Phrasal MRU, TS

= links a clitic to its host

Notice that cliticization is not limited to or motivated by subminimality. Cliticization creating unbounded MRUs is quite common:

(79) owl[31] than dog small
[[*mao-tou*]-*ying*] [[*bi gou*] *xiao*] "the owl is smaller than the dog"
mao-tou-ying = *bi* Cliticization
(1 2 1 3) (s 3)

4.2.2. Classifiers

Classifiers and measure words also behave as enclitics. This accounts for the well-known (quasi-)minimal pairs (80a, b) and (81a, b).[32]

(80) buy iodine

a. *mai* [$_{VP}$ *dian-jiu*] "buy iodine"
3 3 3 base tone
(s 3) Lexical MRU, TS
(3 s 3) Phrasal MRU, TS not applicable

b. buy some wine
mai [$_{NP}$ *dian jiu*] "buy some wine"
(*mai* = *dian*) *jiu* Cliticization
3 3 3 base tone
(s 3) Lexical MRU, TS
(s s 3) Phrasal MRU, TS

[31] lit. "owl" = [[cat-head] eagle].

[32] A restructuring analysis similar to one advocated here was first proposed by Poteet (1985). It has been alluded to also in Shih (1986, 1997), L. Cheng (1987), and Hsiao (1991), among others.

(81)

	buy umbrella	
a.	*mai* [$_{NP}$ *yu-san*]	"buy umbrella"
	3 3 3	base tone
	(s 3)	Lexical MRU, TS
	(3 s 3)	Phrasal MRU, TS not applicable
	buy CL umbrella	
b.	*mai* [$_{NP}$ *ba san*]	"buy an umbrella"
	(*mai* = *ba*) *san*	Cliticization
	3 3 3	base tone
	(s 3)	Lexical MRU, TS
	(s s 3)	Phrasal MRU, TS

Dian-jiu "iodine" (80a) is a single lexical item. The homophonous *dian jiu* in (80b), on the other hand, is an NP, consisting of a measure word *dian* "some" plus the noun *jiu* "wine." As a measure word, *dian* is a bound morpheme, and cliticizes obligatorily leftwards onto the verb *mai* "buy." This restructured VP is rhythmically organized in the usual fashion, with predictable and correct sandhi outcome. (81a, b) parallel (80a, b) in all relevant respects.

4.2.3 Object pronouns

Finally, the behavior of object pronouns as clitics is demonstrated by an example like (82):

(82)

"the landlord doesn't allow me to keep kittens"	
landlord not allow me keep kitten	
fang-dong bu yun-xu wo [$_{S}$ *PRO yang xiao-mao*]	
(*yun-xu* = *wo*)	Cliticization
2 1 4 3 3 3 3 3 1	base tone
(2 1) (s 3) (3 1)	Lexical MRU, TS
(s s 3)	Lexical MRU, TS
---	---
(2 1) (4 s s 3) (s 3 1)	Phrasal MRU, TS

Unless cliticization takes place to create a preemptive ternary MRU (*yun-xu* = *wo*), nothing could prevent the two unpaired syllables *wo* "I" and *yang* "keep, raise" from joining together into a binary unit (*wo yang*), resulting in a reading like (83), which is highly unusual, if not ungrammatical.

(83)

. . . *bu yun-xu wo* [$_{S}$ *yang xiao-mao*]	
(s 3) (3 1)	Lexical MRU, TS
---	---
(4 s 3) (s 3) (3 1)	Phrasal MRU, TS

Parenthetically, one might attribute the oddity of the rhythmic structure of (83) to some structural properties of such constructions. Conceivably, one might speculate that (83) is prohibited by a constraint against an MRU straddling the boundary between two clauses. But there is nothing against pairing two words on opposite sides of a clause boundary into an MRU. This is demonstrated by (84).

(84) "the landlord doesn't believe that I speak German"

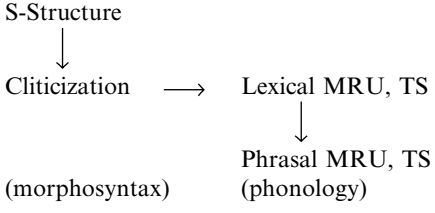
landlord		not believe			I speak German				
fang-dong	*bu*	*xiang-xin*	[$_S$	*wo*	[*jiang*	*de-yu*]]			
		(*xiang-xin* = *wo*)							Cliticization
2	1	4	1	4	3	3	2	3	base tone
(2	1)		(1	4	3)		(2	3)	Lexical MRU, TS not appl.
(2	1)	(4	1	4	3)	(3	2	3)	Phrasal MRU, TS not appl.

Assuming the standard analysis according to which *yun-xu* takes a VP complement, while *xiang-xin* takes a clausal complement, *wo* functions as the matrix object in (82), but as the subordinate subject in (84). If clause boundary presented an unsurmountable barrier to MRU, it would rule out the attested tonal behavior of (84). This is so because TS is obligatory between members of the same MRU. Since TS does not apply between *wo* "I" and *jiang* "speak," they must belong to separate MRUs. In other words, *wo* "I" must be parsed rhythmically together with *xiang-xin* "believe" on the wrong side, so to speak, of the clausal divide, rather than with its clause-mate the predicate *jiang* "speak."

5 Summary

To recapitulate, the core of the analysis I envisage for Mandarin TS consists of the following elements:

(85)



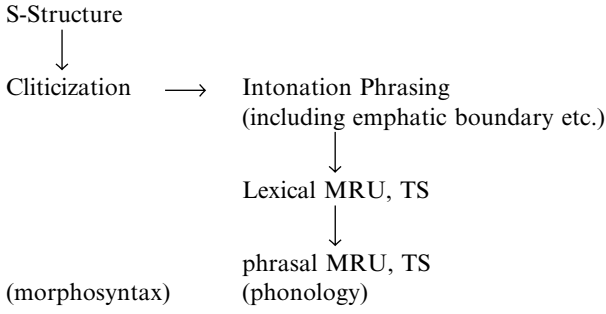
S-structure, rebracketed in certain cases via cliticization, is organized into prosodic units called MRUs. The resulting prosodic structures then serve as input to TS. MRU construction is stratum-ordered: it applies (cyclically) at the lexical level (including p-words), then iteratively across

the board at the phrasal level. Likewise, TS applies first within the MRU, then (optionally) across MRUs.

It is worth highlighting the fact that one single set of constraints governs the parsing of connected speech into one single type of prosodic entity, namely the MRU. In fact, I see little reason to posit anything intermediate between the MRU and intonation phrase, as far as Beijing Mandarin TS is concerned.

There is one aspect we have barely touched on, namely intonation. As mentioned in passing in section 1.5, MRU is IP-bound: that is, an utterance is first broken into IPs, then within IPs syllables are grouped into MRUs. MRUs do not cross IP boundaries. Furthermore, other intonational effects, such as emphasis, also introduce certain breaks which MRUs must respect. Even though the intricate problem of intonational effects lies beyond the scope of this chapter, minimally we need to revise (85) as (86):

(86)



The parsing of connected speech, whether we are talking about word-size units or sentences of unlimited length, is governed by a single set of hierarchically ranked constraints:

(87) {No Straddling, IP-Bound} ≫ Binarity ≫ Boundedness ≫ Congruence ≫ LtoR

How does the present analysis compare with other conceivable accounts? I will consider three other alternatives in the next three sections.

6 The prosodic hierarchy

6.1 Branchingness and restructuring

The first alternative we will consider is one which entails a more articulated prosodic hierarchy. Prosodic phonology (cf. Selkirk 1980, 1981a, 1986,

Nespor and Vogel 1982, 1986, Hayes 1989, inter alia) found its clearest expression as it pertains to Mandarin TS in Beattie (1985) and L. Cheng (1987). Beattie (1985) imported into Mandarin the entire prosodic hierarchy developed in Nespor and Vogel (1982) and elaborated more fully in Nespor and Vogel (1986):[33]

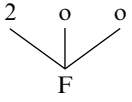
(88) Prosodic hierarchy

U	utterance
I	intonation phrase
P′	a P plus a non-branching P
P	phonological phrase
W′	two non-branching W's
W	word
F	(stress) foot
S	syllable

Foot, in Beattie's terminology (following Yip 1980), is a stress-foot: each stressed (i.e. tone-bearing) syllable constitutes a foot together with toneless syllables, if any, surrounding it. Thus (89a), with only one tone-carrying syllable, constitutes one single F, but (89b) consists of two.

(89) friend Pl "friends"

a. [*peng-you*]-*men*



happy

b. *gao-xing*

1,2,3,4 = tonal categories
o = toneless (neutral tone)
F = stress-foot

The units on the next two (or four) levels up on the prosodic hierarchy, namely W/W′ and P/P′, are defined in a somewhat complex way. Basically, morphosyntactic words count as W; whereas P (= ϕ in Nespor and Vogel's notation) is circumscribed as follows:

[33] Even though P/P' "is the highest prosodic domain on which 3TS is defined" (Beattie 1985:36). Beattie's 3TS is equivalent to our T3 Sandhi.

(90) ϕ Construction
Join into a ϕ any lexical head (X) with all items on its non-recursive side within the maximal projection and with any other non lexical items on the same side (e.g. prepositions, complementizers, conjunctions, copulas . . .). (Nespor and Vogel 1982:228)

Both W and P undergo restructuring. Specifically, two non-branching Ws in mutual c-command join into W′; and a non-branching P joins another P on the left to form a P′ under the conditions stated in (91).

(91) Optional ϕ restructuring
A non-branching ϕ which is the first complement of X on its recursive side loses its label and is joined to the ϕ containing X under a new node labeled ϕ′. (Nespor and Vogel 1982:230)

Given the phrase structure of Chinese, it is hard to tell the recursive from the non-recursive side relative to the lexical head. NPs are one-sided (left), since they are consistently head-final. AP and VP expand recursively in both directions. Take the VP headed by *xiao* "tease" in (92).

(92) [$_{VP}$ *ba* [$_{NP\ S'}$[*hai-xiu de*] *gu-niang*] *xiao* [$_{S'}$ *de ku* [$_{S'}$ *de yan-jing hong le*]][34]
BA shy DE$_1$ girl tease DE$_2$ cry DE$_3$ eye red Prt
"(someone) teased the shy girl so that she cried until her eyes got red"

Both the object NP on the left (*hai-xiu de gu-nian* "the shy girl") and the resultative complement S′ (*de ku de yan-jing hong le* "so that she cried until her eyes got red") on the right of the head of the VP *xiao* "tease" have embedded structures, and can expand recursively.

L. Cheng (1987) eschews the recursive/non-recursive dichotomy, further streamlines the prosodic hierarchy, and seeks to eliminate the need to restructure W′ and P′ by incorporating branchingness directly into the circumscription of W and P (labeled PWd and PPh respectively). Her definition of PWd and PPh incorporates both Selkirk's (1986) end-based theory and Cowper and Rice's (1987) [±branching] as a prosodic parameter:

(93) Mandarin end settings
PWd – $_{word}$[, [branchingness +]
PPh – $_{X}$head[, [branchingness +]

The following example illustrates how this works.

[34] *Ba* is the object marker (or preposition); *de$_1$* is the relative clause marker; and *de$_{2,3}$* is the complementizer that introduces the resultative clauses.

(94) Xiaomei BA pen give Xiaoming
[*xiao-mei*] [[*ba bi*] [*gei xiao-ming*]] "Xiaomei gave the pen to Xiaoming"

3	3	3	3	3	3	2	base tone	
(2	3)	(2	3)	(3)	(3	2)	PWd	= i*
(2	3)	(2	3)	(2	3	2)	PPh	= j
(2	3)	(2	3)	(3)	(3	2)	PPh	= k*

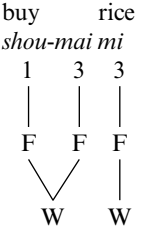
(. . .) = PWd or PPh

Let us grant that at the PWd level, the word-internal structure is visible. Hence *xiao-ming* is internally branching, and therefore a PWd boundary is placed before *xiao-ming*. *Gei* "give" is the head of a branching structure, namely $_{V}$[*gei xiao-ming*], therefore *gei* too demarcates the left edge of a PPh. Crucially, *bi* "pen" is morphologically simple, non-branching. It therefore cannot constitute a free-standing PWd. Instead it joins the function word *ba*, which for our purpose is regarded as a preposition, to form a single-branching PWd. The leftover *Xiao-mei* constitutes another PWd. This procedure cuts up sentence (94) into four PWds. Notice that unlike (*ba bi*), *gei* "give" cannot join *xiao-ming* to form a PWd, for two reasons: (i) two content words cannot be grouped into one PWd; (ii) *xiao-ming* already constitutes a PWd by itself. (i) is a common assumption in prosodic phonology; (ii) is stipulated by L. Cheng (1987:25). We infer from the unacceptability of (94-i*) that TS is obligatory within PPh. The significance of this last remark will become clear in the discussion that follows.

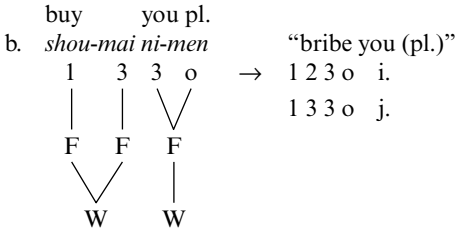
6.2 *Wrong predictions*

There are several problems with the approach sketched above. First of all, Mandarin is clearly a syllable-counting, not a stress (equivalently stress-foot) counting language. This is demonstrated by the following.

(95) buy rice
a. *shou-mai mi* "buy rice"



→ 1 2 3 i.
1 3 3 j*.



If stress-foot conditions TS, then TS should treat (95a) and (95b) alike, since in each case the second W consists of one single F. This prediction is false: TS is obligatory between the two Ws in (95a), but optional in (95b).[35] In fact, there is scant evidence for stress-foot thus defined as an operational unit for Mandarin TS.[36] On the other hand, that syllable-counting MRUs play a crucial role in determining the domain of TS in Beijing Mandarin will become even more apparent as we proceed.

Second, at the heart of the prosodic analysis as made explicit in Beattie (1985) and L. Cheng (1987) is the notion of branchingness. It will become apparent that branchingness as a *syntactic* property makes wrong predictions; as a *phonological* notion, i.e. essentially a syllable-counting device, it obliquely encodes generalizations that are better captured by means of MRU formation. The reason is not hard to see. At the phrase level, word-internal structures are opaque. In other words, word-level projections are regarded as non-branching constituents for the purpose of PPh formation. This is so because otherwise the morphologically complex proper name *xiao-ming* (prefix *xiao* "little" + name *ming*) in (94) would form a separate PPh, *xiao-ming* being the head within its own maximal projection NP. This wrongly generates (94-k*).[37] Therefore, morphologically complex words, i.e. X^o, must be construed as non-branching heads for the purpose of constructing PPhs. Thus interpreted, branchingness as a syntactically defined parameter makes the following prediction: unadorned subject + verb and verb + object constructions always constitute a single PPh, within which TS is obligatory. This prediction is counterfactual. Take the verb + object construction.

[35] In Beattie's (1985) account (p. 34), TS is obligatory within W/W′, but optional across W/W′.

[36] That is not to say that the stress-foot doesn't play a central role in determining tone sandhi domain in other dialects. See the chapters 7–8.

[37] Here we discount, as usual, marked intonation phrasing.

(96) repair watch

a. *xiu-li biao* "repair watches"

(1 3) (3)[38] PWd i*.

(1 2 3) PPh j.

repair bicycle

b. *xiu-li jiao-da-che* "repair bicycles"

(1 3)(3 2 1) PWd i.

(1 2 3 2 1) PPh j.

L. Cheng's (1987) prosodic hierarchy predicts that (96a) and (96b) will behave identically with respect to TS. Syntactically speaking, the polysyllabic noun *jiao-da-che* "bicycle," just as the monosyllabic noun *biao* "watch," constitutes a bare noun, a non-branching constituent. Both, therefore, should merge into a single PPh – hence both obligatorily undergo TS. While this correctly accounts for the unacceptability of (96a-i*) as well as (94i*), it wrongly predicts (96b-i) to be equally unacceptable. Conversely, in view of (96b), one might stipulate that TS is obligatory within the PWd, but optional within the PPh. This makes the opposite, equally wrong, prediction that (96a-i*) and (94i*) are acceptable.

To further highlight the problem with the complex prosodic hierarchy adopted by Beattie (1985) and L. Cheng (1987), let us look at the mirror image of (96). In (96) we keep the verb constant and substitute one object NP with another. In the examples of (97) we do the opposite. Given the directionality of end-setting theory, there can be no difference whatsoever in the prosodic structure of (97a) and (97b). The object NP is identical in both cases. A PPh boundary should occur either in both (97a) and (97b), or in neither one. The internal structure of the verb is irrelevant, since [V NP] always constitutes a branching structure. A left-end setting is, by its very nature, incapable of discriminating between (97a) and (97b).

(97) a. *zhao yan-jing* "look for eyeglasses"

(3) (3 4) PWd i*.

(s 3 4) PPh j.

look:for eyeglasses

b. *xun-zhao yan-jing* "look for eyeglasses"

(2 3) (3 4) PWd i.

(2 s 3 4) PPh j.

[38] Here I ignore the problem of a non-branching PWd like *biao* "watch." Presumably it is for cases like this that L. Cheng (1987:25) qualifies her end-setting parameters by saying: "PWd's in Mandarin must be branching *if possible* [emphasis added]."

Prosodic hierarchy, as instantiated in the works of Beattie and L. Cheng, fails to capture the fundamental fact that identical grammatical structures (interpreted via end-setting), behave differently in terms of TS.

Integrating branchingness into the parameter setting of PWd and PPh was intended to eliminate the need to restructure these prosodic units. Unfortunately it doesn't accomplish what it was intended to do. Given its left-end setting, a right-branching structure like (97) will always have a PPh boundary between the first and the second PWd. This wrongly predicts (97-i*) to be a possible reading, since TS is obligatory within PPh, but optional between PPhs within an IP. In order to prevent (97-i*), one must have recourse to some sort of restructuring, whereby non-branching PPhs coalesce into one. It is for reasons like this that Beattie (1985:46f.) introduced "a constraint blocking the occurrence on the phonetic level of two consecutive non-branching P's [= PPhs]." Only by joining the first two PPhs can we derive reading (98-j).

(98)	dog	want	bite	cat	"the dog wants to bite the cat"
	gou	[*xiang*	[*yao*	*mao*]]	
	(3)	(3)	(3)	(1)	PWd
	(3)	(3)	(3	1)	PPh
	(3)	(3)	(3	1)	TS (n/a) = i*.

	(3	3)	(3	1)	PPh Restructuring
	(2	3)	(3	1)	TS = j.

Clearly the constraint on adjacent non-branching PPhs and consequent PPh coalescence is a round-about way of saying: no adjacent monosyllabic PPh *qua* obligatory domain of TS.

In short, the branchingness in question is either syntactic, in which case it makes counterfactual predictions about (96–98), or else it is a phonological property in syntactic disguise and is better captured by MRU. In terms of MRUs, the rhythmic parsing of sentences (94, 96–98) is totally unremarkable. MRU formation proceeds exactly as described in section 1. In fact, as far as sentences (94, 96–98) are concerned, the input to MRU formation can be parsed simply as unlabeled trees.

6.3 Need for PPh?

Notice that both L. Cheng (1987) and Beattie (1985) posit PPh as the domain of obligatory TS. Hsiao (1991) also postulates PPh as a prosodic unit for a different reason. He points to the contrast illustrated in (99) (Hsiao 1991:57f.):

```
(99)        dog    bite    woman              "the dog bit the woman"
   a.  S[gou VP[yao NP[nü-ren]]]
         3       3       3  2
        (2       3)     (3  2)                FFR        = i.
      ------------------------------------------------------------
         3  %    3       3  2                 IP
        (3) %   (2      (3  2))               FFR        = j.

        keep dog    guard gate                "keep a dog to guard the gate"
   b.  VP[yang gou VP[shou men]]39
         3     3       3     2
        (2     3)     (3     2)               FFR        = i.
      ------------------------------------------------------------
         3 % 3         3     2                IP
        (3)%(2         3     2)               FFR        = j*.
```

% = intonation break
IP = intonational phrasing
FFR = Foot Formation Rules (see section 1.4)

Hsiao (1991) accepts the core idea behind Shih's (1986, 1990) FFR (see section 1.4).[40] FFR predict that the first two syllables in both (99a) and (99b) will join into a binary foot. This correctly predicts reading (i) in both. How do we account for reading (j) of (99a)? Hsiao assumes, as we do (see section 1.5), that FFR (consequently, TS) is IP-bound: that is, FFR may not cut across IP boundaries. By placing an intonation break between *gou* "dog" and *yao* "bite" in (99a), IP-bound FFR has no alternative but to form a monosyllabic (defective) foot, plus a ternary foot. This yields reading (j) of (99a). The question then arises: Why can an intonation break occur between the first and the second syllables in (99a), but not in (99b)?

To solve this problem Hsiao resorts to the notion of PPh. The basic idea is that IPs are built bottom-up, from the prosodic units on the next level down, namely from PPhs. The PPhs are circumscribed along the lines of Nespor and Vogel (1982, 1986), Beattie (1985), and L. Cheng (1987). From this perspective, the pair of sentences given in (99) yields the following rhythmic parses:

[39] I assume, with Hsiao, a flat ternary [V NP VP] structure for the construction under discussion.
[40] With modifications that need not concern us here.

(100) dog bite woman "the dog bit the woman"

a. $_{S}$[*gou* $_{VP}$[*yao* $_{NP}$[*nü-ren*]]]

(3)	(3)	(3 2)	PPh	
(3) %	(3)	(3 2)	IP	
(3) %	(3	3 2)	FFR	
(3) %	(2	3 2)	TS	= (99a-j)

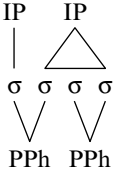
keep dog guard gate "keep a dog to guard the gate"

b. $_{VP}$[*yang* $_{NP}$[*gou*] $_{VP}$[*shou men*]][41]

(3	3)	(3 2)	PPh	
(3 %	3)	(3 2)	IP	
(3) %	(3	3 2)	FFR	
(3) %	(2	3 2)	TS	= (99b-j*)

The end-setting parameters[42] cuts up (100a = 99a) into three PPhs. The second and third PPhs optionally coalesce into one IP. IP-bound FFR and TS produce the desired output. In contrast, the second syllable *gou* "dog" in (100b = 99b) is a non-branching NP; consequently it cannot mark the left edge of a PPh. As a result, there are only two PPhs in (100b). Since IPs are built from PPhs, the only intonation break, if there is one, should coincide with the partition between the two PPhs, namely between *gou* "dog" and *shou men* "guard gate." To introduce an intonation break between *yang* "keep" and *gou* "dog" would cut a PPh asunder, in violation of the *Strict Layer Hypothesis* (cf. Hayes 1989), as diagrammed in (101).

(101)



This bottom-up intonation phrasing correctly eliminates (100b) as a possible prosodic structure and consequently as a possible sandhi form.

Under closer scrutiny, Hsiao's account is open to question. If branchingness is a syntactic configuration, then *nü-ren* "woman," an indivisible syntactic compound, is a non-branching constituent at the phrase level. Hence (99a) must be divided into two PPhs as shown in (102).

[41] I assume, with Hsiao, a flat ternary [V NP VP] structure for the construction under discussion.

[42] Actually, Hsiao's end-setting parameter for PPh is {left, X^{max}, +branching}, not {Xhead}. This makes no difference for the case at hand.

(102)

	dog	bite	woman	
	$_{S}$[*gou* $_{VP}$[*yao* $_{NP}$[*nü-ren*]]]			"the dog bit the woman"
	(3)	(3	3 2)	PPh
	(3)	(3 %	3 2)	IP
	(3	3) %	(3 2)	FFR
	(2	3) %	(3 2)	TS = (99a-i)

This bottom-up intonation phrasing would wrongly rule out (102) as a possible intonation pattern, since on this account the IP break (%) would fall in the middle of a PPh.

If, on the other hand, branchingness is simply an oblique way to refer to syllable count, then it predicts a PPh boundary on the left edge of the branching (that is, nota bene, polysyllabic) noun *mu-ji* "hen" in (103).

(103)

	raise hen	lay egg		
	yang mu-ji $_{VP}$[*sheng dan*]			"raise hens to lay eggs"
	(3)	(3 1)	(1 4)	PPh
	(3) %	(3 1)	(1 4)	IP_1
	(3) %	(3 1)	(1 4)	FFR
	(3) %	(3 1)	(1 4)	TS (na) = i*.
	---	---	---	---
	(3)	(3 1) %	(1 4)	IP_2
	(3)	(3 1) %	(1 4)	FFR
	(2)	(3 1) %	(1 4)	TS = j.

Given the tripartite PPh, both intonational phrasings IP_1 and IP_2 ought to be possible. In fact, IP_1 resulting in reading (i) is unacceptable; only IP_2 with the sandhi pattern (j) is attested. The bottom-up intonation phrasing offers no principled way to rule out IP_1.

I agree with Hsiao that a proper account of (100a, b) does involve IP-bound FFR. But as demonstrated above, it is doubtful that intonation phrasing has anything to do with PPh. Instead, I believe intonation phrasing, together with focus and emphasis, is controlled by semantic and pragmatic factors. In particular, *Sense Unit Condition* (SUC) (Selkirk 1984; cf. Halliday 1967) plays a role in intonation phrasing. To oversimplify the matter somewhat, X and Y coalesce into an IP provided that X and Y stand in some semantic relation, characterizable as head/complement or head/adjunct. SUC straightforwardly rules in (100a) and rules out (100b) as well as (103-i). In (100a) *yao* and *nü-ren* "bite + woman" stand in a subject–predicate relation; hence SUC licenses V–NP as an IP. As for (100b), the reason why its intonation phrasing is ill-formed becomes apparent when we annotate the sentence more fully as (104). (104) makes it clear that *gou* "dog" is the object of *yang* "keep"; it

is not, per se, the subject of *shou men* "guard the gate," even though it is the antecedent that controls the PRO in the purposive adjunct clause S′. Consequently, *gou shou men* is semantically incoherent, and therefore cannot constitute an IP, in accordance with SUC.

(104) keep dog guard gate
$_{VP}$[*yang gou*$_i$ $_{S'}$[*PRO*$_i$ *shou men*]] "keep a dog to guard the gate"
(3) % (s 3 2) IP

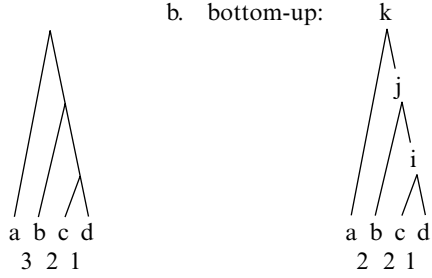
(103-i*) is ungrammatical for exactly the same reason. We have more to say about SUC in section 8.

I conclude there is no basis for positing PPh as a domain of Mandarin TS. In fact, there is no compelling evidence for *any* prosodic entity intermediate between the MRU and the IP, as far as Mandarin TS is concerned.

7 Syntactic juncture

We have seen that the alternative and more elaborate prosodic hierarchy either fails to account for the facts or else mimics MRU by indirection. Let us consider other possible approaches. What if we dispense with prosodic structure altogether and allow TS to directly operate on s-structures? Intuitively speaking, phonological processes are most likely to affect elements that are most closely linked together. Or, to put it another way, a phonological rule is most likely to be blocked by the strongest grammatical junctures that separate distantly related constituents. There are two ways to measure structural distance or junctural strength: top-down or bottom-up. They make somewhat different predictions.

(105) a. top-down: b. bottom-up:



a, b, c, d = constituents
1, 2, 3 = junctural strength (higher number = stronger juncture)

In the top-down model (105a), junctural strength straightforwardly encodes the hierarchy of IC cut. In the bottom-up model (105b), structural

distance is measurable by counting the nodes separating two adjacent constituents. Thus c and d are separated by only one node i; but b is equidistant between a and c, since in each case it is separated by two nodes: b is separated from a by j and k, from c by i and j.[43] C. Cheng (1970, 1973a) represents a classic analysis in the spirit of a direct syntactic conditioning of phonological rules from a top-down perspective. This approach makes demonstrably false predictions, as illustrated by familiar and mundane examples such as these:

(106) dog bite small cat "the dog bit the kitten"
a. *gou* [*yao* [*xiao-mao*]]
(s 3) (3 1)

plan write novel "plan to write a novel"
b. *xiang* [*xie* [*xiao-shuo*]]
(s 3) (3 1)

(. . .) = obligatory domain of TS (= MRU)

In (106a) the highest IC cut is between the subject *gou* "dog" and the predicate *yao xiao-mao* "bit the kitten"; in (106b) it is between the verb *xiang* "plan" and the complement VP *xie xiao-shuo* "write a novel." And yet, TS applies across this juncture, while blocking at the weaker juncture between the second and the third syllable.

The bottom-up model fares no better. It predicts that TS will apply indifferently either between the first and the second, or between the second and the third syllables. Hence, it is not capable of predicting the unmarked, default readings of (106a, b).

Duanmu (1990b) represents a slight variation of a direct syntax-to-phonology mapping, by requiring that TS applies cyclically to morphosyntactically defined constituents. Operating on ever-larger constituents, as illustrated in (107), TS generates the highly marked, possibly unacceptable, tonal pattern [3 s 3 1] as the sandhi output.

(107) plan write novel
$_{VP1}$[*xiang* $_{VP2}$[*xie* $_{NP}$[*xiao-shuo*]] "plan to write a novel"

3	3	3	1	base tone
		(3	1)	NP cycle
	(s	3	1)	VP2 cycle
(3	s	3	1)	VP1 cycle = ?,*

? = highly marked
* = unacceptable

[43] These two ways of computing grammatical relatedness correspond to what Clements (1978) calls syntactic depth (top-down) and syntactic distance (bottom-up).

There are other variations and permutations of the direct syntax-to-phonology mapping. For instance, Liu (1980, as reported in Kaisse 1985:170ff.), proposed a so-called *Edge Condition* to the effect that a language may stipulate that a phonological rule applies between X and Y (or Y and X), provided that X constitutes a non-branching unit lying at the edge of a larger constituent containing X and Y. Thus TS applies at the sandhi site marked "=" in cases (a) and (b), but blocks at "/" in case (c):

(108) Edge condition
a. [X = [Y . . .]]
b. [[. . . Y] = X]
c. [[. . . X] / [Y . . .]]

Edge condition also makes counter-factual predictions – bidirectionally. On the one hand, it predicts that in a uniformly right-branching structure like (106a, b), TS should apply across the board, since each syllable lies at the edge of the containing constituent. This means that, given the base tones of (106a, b), we would expect [s = [s = [3 1]]] as the sandhi form. This is possible, but certainly not the only, or even the most natural, reading of (106a, b). Instead, the unmarked, default reading is [(s-3)(3-1)]. Conversely, Edge condition predicts that the internal sequence of /3-3/ as in (109) would not undergo TS. As we can see in reading (109-j), TS freely applies between the two internal syllables *nao* and *jian*. It doesn't help to argue that both *tou-nao* "mind" and *jian-dan* "simple" are syntactically indivisible, i.e. non-branching, albeit morphologically complex, words; for in this case, one cannot explain the failure of TS in reading (109-i).

(109)	mind		simple		"simple-minded"
	[[*tou-nao*][*jian-dan*]]				
	(s	3)	(3	1)	i.
	(s	s)	(3	1)	j.

While the direct syntax–phonology mapping founders on the most elementary facts alluded to, these facts lend themselves to a straightforward account in the MRU-based model proposed here. In the case of (106), TS fails to apply between *yao* "bite" and *xiao-mao* "kitten," despite the fact that they satisfy the Edge condition (*yao* lying at the edge of the constituent VP); this follows from the fact that TS is not obligatory *across* MRU. Conversely, TS does apply between *tou-nao* "mind" and *jian-dan* "simple" in (109), notwithstanding the fact that the two abutting

T3-bearing syllables (i.e. *-nao* and *jian-*) are internal to the sentential predicate "(being) simple minded"; this follows from the same principle, namely that TS is MRU-bound, but optionally operates across MRUs:

(110)	dog bite kitten	"the dog bit the kitten"
a.	[*gou* [*yao* [*xiao-mao*]]]	
	(s 3) (3 1)	TS (within MRU), obligatory
	n/a	TS (across MRUs), optional
	mind simple	"simple-minded"
b.	[*tou-nao*] [*jian-dan*]	
	(s 3) (3 1)	TS (within MRU), obligatory
	(s s 3 1)	TS (across MRUs), optional

8 Meaning-based prosodic structure

The analyses examined so far are structure-based in the sense that TS operates directly either on syntactic representations or on structurally defined prosodic units. There is another alternative we have not considered, namely meaning-based prosodic structures delimiting the scope of TS.

8.1 Sense Unit-based foot formation

Recall the original version of Foot Formation Rules proposed by Shih (1986:110), repeated below:

(111) Foot Formation Rules (FFR)

a. Immediate Constituency: Link immediate constituents into disyllabic feet.
b. Duple Meter: Scanning from left to right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.
c. Super-foot Construction: Join any left-over monosyllable to a neighboring binary foot according to the direction of syntactic branching.

Subsequently, Shih (1997) opted for a maximally simple, but non-deterministic, foot formation algorithm, in the process shifting the analytical burden to a meaning-based filter called Sense Unit Condition (SUC), to be made more precise below. Thus, she downplayed the role of *Lexical Integrity Hypothesis* she appealed to in her earlier analysis (Shih 1986), since SUC would automatically filter out a foot built over part of a word plus something else like (. . . *-jiu hao*) ". . . wine good" in (112a-i) as semantically incoherent, and therefore rhythmically ill-formed.

(112) cocktail wine good "the cocktail is good"

a. [[*ji-wei*]-*jiu*] *hao*
(1 3) (s 3) i*.
(1 s s 3) j.

wolfhound bite Xiaoming "the wolfhound bit Xiaoming"

b. *lang-gou* [*yao xiao-ming*]
(2 3) (s 3 2) i.
(2 s 3) (3 2) j.

Likewise, directionality of branching condition on clause (c) of FFR is no longer necessary, since a stray monosyllable can go left- or right-ward to form a "superfoot", the directionality of adjunction being constrained only by SUC: thus both readings (i) and (j) of (112b) are acceptable. Shih (1997) concludes: "the best strategy is to simplify the foot formation rule given in Shih (1986), and use the sense unit condition as a constraint."

On this score Shih (1997) is in agreement with Hung (1987a), who had argued earlier that FFR should be meaning-based rather than syntax-based. Some of the key examples for Hung's argument are presented in (113).

(113) cat than dog small

a. *mao* [[*bi gou*] *xiao*] "cats are smaller than dogs"
(1 (s s) 3) i.
(1 3) (s 3) j.

buy two bowl rice

b. *mai* [[*liang wan*] *mi*] "buy two bowls of rice"
(3 (s s) 3) i.
(s 3) (s 3) j*.

(114) small umbrella

a. *xiao* [*yu-san*] "a small umbrella"[44]
(3 (s 3)) i.
((s s) 3) j?

dog bite me

b. *gou* [*yao wo*] "the dog bit me"
(3 (s 3)) i.
((s s) 3) j.

(. . . (. . .) . . .) = nested prosodic structures

[44] *San* may be glossed as an "umbrella-shaped object." It occurs in other compounds like *yang-san* "sun + umbrella = parasol" and [*jiang-luo*]-*san* "fall + umbrella = parachute."

In these examples (. . . (. . .) . . .) represents nested prosodic structures, and TS operates cyclically, from the innermost bracketed units to the outermost. This accounts, for instance, for the alternative readings of (114b):

(114)

b′. *gou* [*yao wo*]

(3 (3 3))	Foot Formation	
(s 3)	TS	
(3 s 3)	TS not applicable = reading (i)	

((3 3) 3)	Alternative Foot Formation	
(s 3)	TS	
(s s 3)	TS	= reading (j)

Notice that the (a) and (b) examples of (113) and (114) are pairwise isomorphic in tree-geometry, and yet exhibit quite different sandhi behaviors. Specifically, Hung argued that *mao* "cat" and *bi* "than" in (113a-j) form an acceptable prosodic unit because they stand in a bona fide meaning relation – despite the fact that *bi* is conventionally analyzed as a preposition and a sister of *gou* "dog." This is so because *bi*, in addition to its use as a preposition, functions as a verb with the meaning "to compare."[45] Construed as a verb, it enters into a subject–predicate relation with *mao* "cat," and therefore constitutes a sense unit with *mao*. On the other hand, *mai* "buy" and *liang* "two" in (113b) do not enter into any specifiable semantic relation. Hence (*mai liang*) is not a well-formed prosodic unit either, correctly predicting reading (j) of (113b) to be unacceptable. As for (114a) and (114b), the (j) reading is more highly marked in the former than in the latter. Again, the contrast is rooted in meaning, so argues Hung, noting that while *gou yao* "dog bites" in (114b) constitutes a meaning unit, *xiao yu* "small rain" in (114a) does not make sense at all (i.e. in this context *small* modifies not the sublexical fragment *yu* (lit. rain), but *yu-san* "umbrella").

One of the most interesting cases discussed in Hung is example (115). It has two readings (i, j). According to conventional analysis, *ye* "also" is categorized as an adverb, i.e. an adjunct of VP, and therefore bracketed as indicated. However, *ye* "also" has two interpretations: it may modify the verb *you* "have" with a meaning corresponding to (116a), or, alternatively, it may have scope over the subject NP *ta* "he," with a reading

[45] Cf. *ni gen ta bi*
you with him compare
"compare yourself with him"

paraphrasable as (116b).[46] Although felicity judgments are subtle, the correspondences (115-i) = (116a) and (115-j) = (116b) seem to accord nicely with native speakers' intuition.

(115) he also have pen
ta $_{VP}$[*ye* [*you bi*]] "he also has a pen"
(1 (s s) 3) i. = (116a)
(1 3) (s 3) j. = (116b)

(116) he also have pen also have paper
a. *ta* [*ye* [*you bi*]], [*ye* [*you zhi*]] "he has a pen and paper too"
(1 ((s s) 3) ((s s) 3)

you have pen he too have pen
b. [*ni* [*you bi*]], [*ta* [*ye* [*you bi*]] "you have a pen, he too has a pen"
(3 (s 3)) (1 3) (s 3)

Each of the examples cited above instantiates a case of structural isomorphism camouflaging a multiplicity of ways of construing sense units and, according to Hung, demonstrates that it is the sense units that "appear to be more successful than ICs [immediate constituents] in predicting the full range of acceptable readings while ruling out unacceptable ones" (Hung 1987a:29).

There is therefore considerable prima facie plausibility to the meaning-based approach. Let us explore this approach in some detail. Hung makes the strong claim that the only necessary and sufficient condition on foot formation is semantic cohesion. Specifically, Hung (1987a:37) replaces Shih's FFR with the following:

(117) SUC-based Foot Formation Rules
a. Form binary-branching feet by linking syllables into disyllabic sense units wherever possible.
b. For all remaining syllables, form trisyllabic or quadrisyllabic feet by linking them to the above to form larger sense units.

Hung adopts Selkirk's (1984:291) definition of a sense unit as follows:

[46] My interpretation here differs somewhat from Hung (1987a:36). There he derives the two readings as follows:
i. *ta* [*ye* [*you bi*]]
(1 (3 (s 3)))
ii. *ta* [[*ye you*] *bi*]
(1 (s s) 3)
It is not clear whether Hung construes (i) and (ii) as bracketed above to have the two scopal interpretations corresponding to (115a, b).

(118) Sense Unit Condition (SUC)
Two constituents Ci, Cj form a sense unit if (a) or (b) is true of the semantic interpretation of the sentence:
(a) Ci modifies Cj (a head)
(b) Ci is an argument of Cj (a head)

It can be seen that the well-formed prosodic structures (113a-j), (114b-j), and (115-j) that are at odds with the tree-based predictions nevertheless satisfy SUC as defined above.

8.2 *Against meaning-based foot formation*

There is no doubt that meaning plays a part in prosodic organization. In particular, SUC is undoubtedly crucial for higher-level prosodic organizations such as intonational phrasing. The question at hand is whether meaning determines foot formation directly or through the mediation of morphosyntactic structure. For one thing, morphosyntactic constituency is based in large part on meaning relations. Meaning and structure, therefore, overlap to a large extent in the way they determine prosodic structure. However, there is evidence that the well-formedness of prosodic organization at the more local level of the foot or MRU is not reducible to SUC.

One of the principal appeals of a meaning-based approach is that it liberates foot formation from the straightjacket of syntactic constituency. Thus, an SVO sentence like *gou yao wo* "the dog bit me" (= 114b) freely lends itself to two possible prosodic organizations (S (V O)) or ((S V) O), since the pivotal verb is semantically linked to both the subject and the object, even though syntactically speaking VO, but not SV is deemed to be a constituent for standard reasons like movement and pro-verb substitution by *do so.*[47] However, the flip side of this meaning-based flexibility is that it also introduces a degree of indeterminacy that was absent in the structure-based rhythmic organization. Take the following example.

(119) I keep dog guard gate "I keep a dog to guard the gate"
wo [*yang gou* [*shou men*]]
(3 (s 3)) (3 2) i
(s 3) (s (3 2)) j*

[47] The corresponding Chinese pro-verb is *ye shi* "also be, likewise," as illustrated in a sentence like:
Mary xi-huan da wang-qiu. Wo ye shi.
Mary like play tennis. I likewise
Mary likes to play tennis, so do I.

Here we have a V_1 NP V_2 construction, where the pivotal NP serves double duty as the object of V_1 and, simultaneously the subject of V_2. By the same logic, we would expect two alternative prosodic organizations ((V_1 NP) V_2) and (V_1 (NP V_2)) as in (119-i) and (119-j) respectively. The predicted reading (119-j) is totally unacceptable. The infelicity of (119-j) is a mystery in a semantic account. On the other hand, the structure-based MRU automatically rules out (119-j) by virtue of the fact that *yang gou* "raise dog," but not *gou shou* "dog guard" are immediate constituents in a phrase structure represented as (120). (120) brings out more clearly the fact that whereas *gou* "dog" is a complement of the matrix verb *yang* "raise," it is not per se the subject of the subordinate verb *shou* "guard", even though it is the antecedent that controls the PRO which is the subject of the complement sentence.

(120) I raise dog guard gate
wo $_{VP}$[$_{V'}$[*yang gou*$_i$] $_{S'}$[*PRO*$_i$ *shou men*]]
(3 (s 3)) (3 2)

To be fair, a more finely articulated semantic representation of (119) can no doubt encode the same meaning relations represented in (120).

Regardless of how (119) is to be represented semantically, if we take SUC or some other meaning-based constraints seriously, we should expect to rule out on semantic grounds certain prosodic structures blindly generated by the tree-based algorithm of foot formation. Consider the following pair of sentences.

(121) only buy stocks, not sell stocks
a. *zhi mai gu-piao, bu mai gu-piao*
(s 3) (3 4) (2 4) (3 4)

only buy stocks, not buy bonds
b. *zhi mai gu-piao, bu mai zhai-quan*
(s 3) (3 4) (4 3) (4 4)

In (121a), The adverb *zhi* "only" has narrow scope over *mai* (T3) "buy," which is the focus of contrast with *mai* (T4) "sell." The MRU-mates *zhi mai* "only buy" therefore forms a coherent semantic unit. The prosodic organization of (121a), therefore, closely mirrors its semantic interpretation. Now turn to (121b). In this case, the adverb *zhi* "only" likewise has a narrow scope, but over a non-contiguous constituent *gu-piao* "stocks," that contrasts with *zhai-quan* "bonds." According to this reading, *mai* "buy" is outside of the scope of *zhi* "only." Under this interpretation *zhi mai* does not constitute a sense unit; nonetheless it is a perfectly

well-formed MRU. Therefore, a meaning-based prosodic organization should rule out – incorrectly – a rhythmic unit consisting of *zhi mai* "only buy" in (121b). In a syntax-based prosody, the MRU-level organization of (121b) is totally unremarkable given the conventional syntactic representation (122b), where *gu-piao* falls within the c-command domain of the scopal operator *zhi*.

(122) only buy stocks
a. [*zhi mai*] *gu-piao* . . . = (121a)
(s 3) (3 4)

only buy stocks
b. *zhi* [*mai gu-piao*] . . . = (121b)
(s 3) (3 4)

The scope of modification can be ascertained even in the absence of contrastive focus. The default reading of (123) construes the non-adjacent constituent *wu sui* "five years" as the scope of *zhi* "only." Again, the prosodic foot (*zhi (xiao ta*)) "only younger than he" cannot pass as a sense unit without voiding the term of any empirical content.

(123) only younger he 5 year "younger than he by only 5 years"
zhi [*xiao ta* [*wu sui*]]
(s (3 1)) (3 4)

The possibility of semantic links arching over discontinuous elements highlights one of the fundamental problems of a purely meaning-based approach to foot-formation, given that prosodic organization must obey other imperatives of speech production, including linearity or phonological adjacency.

Independently of the phenomenon of semantic relation in distans, SUC requires fine tuning in other respects as well. Take (124).

(124) plan write novel "plan to write a novel"
xiang [*xie xiao-shuo*]
(s 3) (3 1)

Strictly speaking, *xiang* "plan" and *xie* "write" do not form a sense unit: the argument of *xiang* is not *xie*, but the VP *xie xiao-shuo* "write a novel" as a whole. What we see here is not different in kind from the fundamental asymmetry between phonological phrasing and meaning-based syntactic constituency already noted in Chomsky and Halle (1968:372), who pointed out that neither of the first two IPs bounded by parentheses in (125) coincides with a syntactic constituent.

(125) *This is* [*the cat that caught* [*the rat that stole* [*the cheese*]]]
( ) ( ) ( ) IP

In view of the semantic saliency of heads and the fact that phonological phrasing is known to be sensitive to the special status of the head of a phrase,[48] one might plausibly revise SUC as follows:

(126) SUC by Proxy
SUC may be satisfied either by X and Y, or by their heads

SUC by Proxy, but not SUC tout court, guarantees the well-formedness of the feet in (124) and the intonation phrases in (125). Unfortunately, SUC by Proxy would wrongly license readings (i) in (127a, b), and at the same time incorrectly rule out the perfectly good readings of (128).

(127) a. democracy party elect representative — "the democratic party elects representatives"
[*min-zhu dang*] [*xuan dai-biao*]
(2 3) (s 3) (4 3) — i*
((2 s) 3) (3 (4 3)) — j

b. small bowl inside place fruit — "there are some fruits placed inside the small bowl"
[[*xiao wan*] *li*] [*bai shui-guo*]
(s 3) (s 3) (s 3) — i*
((s s) 3) (3 (s 3)) — j

(128) a. must early buy ticket — "one must buy tickets early"
dei [*zao mai piao*]
(s 3) (3 4)

b. buy pink overcoat — "buy a pink overcoat"
mai [*fen-hong da-yi*]
(s (3 2)) (4 1)

c. please punctually arrive — "please arrive punctually"
qing [*zhun-shi chu-xi*]
(s (3 1))(1 2)

d. "after I carefully pondered this matter, I decide to put it off"
I carefully ponder after decide postpone deal:with
[*wo* [*fan-fu kao-lü*]] *yi-hou*, *jue-ding zhan-huan chu-li*
(s (3 4))(3 4) (3 4) . . .

[48] As illustrated by French liaison:
un petit = endroit # obscur
a hidden little place
where liaison (symbolized as "=") occurs between the head and a pre-head modifier, but not between a head and a posthead adjunct (# marks blockage of liaison). Cf. Selkirk (1986), Nespor and Vogel (1986).

e. than I early five day arrive
bi wo [*zao* [*wu tian*]] *dao* "arrived five days earlier than I"
((s s) 3) ((3 1) 4)

Dang "party" and *xuan* "elect" of (127a) are the heads of the subject NP and the predicate VP respectively. They should therefore satisfy SUC by Proxy, and form a disyllabic foot. Likewise, *li* "inside" is the head of the locative adverbial *xiao wan li* "inside the small bowl," and should be able to join *bai* "place, put" in a sense unit. But footing in neither case yields an acceptable sandhi pattern. Conversely, each example of (128a–d) begins with a foot consisting of two semantically unrelated constituents. Thus in (128d) *fan-fu* "carefully" is obviously not the head of the VP, but an adjunct that modifies not the subject *wo* "I," but the verb *kao-lü* "ponder"; and yet it forms a foot with the former, as evidenced by the obligatory sandhi on *wo*. As for (128e), one could argue that (*bi wo zao*) "earlier than me" constitutes a sense unit to the extent that *zao* is the head of the adverbial phrase *zao wu tian* "earlier by five days." But certainly the second ternary foot (*wu tian dao*) "five days arrive" is semantically incoherent.

Finally, problems associated with SUC are magnified when we take into account the many examples that involve restructuring via cliticization. A small sample is reproduced here to illustrate the point.

(129) you it like DE
a. *ni* [*suo xi-huang*] *de* "what you like"
(s = 3)((3 1) o)

cat than dog small
b. *mao* [*bi gou*] *xiao* "the cat is smaller than the dog"
(1 = 3) (s 3)
"x = y" stands for "host + clitic"

There is no conceivable semantic link between *ni* "you" and *suo*, the pronominal enclitic "it," in (129a). (129b) is the classic example that set Hung (1987) on his quest for a meaning-based key to unlock the mystery of Mandarin tone sandhi. Curiously, while attempting to explain the footing of (*mao bi*) "cat than," Hung totally overlooked the equally odd foot-mates (*gou xiao*) "dog small." By no stretch of the imagination can we construe *gou xiao* as a semantically coherent unit, *xiao* "small" being predicated not on *gou* "dog" but on *mao* "cat" three syllables away.

How do we, then, explain the observed prosodic structures once we abandon SUC as the guiding principle for foot formation? Reading (129b)

is accounted for by means of cliticization. The prosodic structure of (114b-j) *gou yao wo* "the dog bit me" is probably due to tempo. The sequence (s s 3) does not necessarily force a left-branching foot analysis ((s s) 3), but can be seen as the result of TS applying simultaneously to a ternary foot regardless of its internal structure. We have seen this "flattening" effect in allegro speech in section 1.5. As for (115-i, j), there is no a priori reason why (115-j) – but not (115-i) – could not be represented syntactically as (130), where *ye* functions exactly like the English gloss *too*. The bracketing of (130) would yield the correct prosodic structure via Foot Formation or a structurally "congruent" MRU, consistent with No Straddling.

(130) [*ta ye*] [*you bi*]
he too has pens

It should be obvious that meaning relations and morphosyntactic structures overlap to a large extent. Where they do make different predictions, it appears that, on balance, the available evidence weighs on the side of the approach taken here, namely that prosodic organization is based directly on the morphosyntactic representation, even though the latter is ultimately rooted for the most part in semantic considerations in the first place.

Appendix Prosodic and syntactic word

Central to our analysis of Beijing Mandarin TS is the two-pass MRU formation, first lexical then phrasal. The default case is for the syntactic word (s-word) and the domain of lexical MRU to be coextensive. This has been amply demonstrated in sections 2–3. However, mismatches between s-word and the domain of lexical MRU are known to occur.

First, a syntactic word is treated like a phrase phonologically speaking. This is illustrated by the examples that follow.

(131) very have farsightedness
a. *hen # you*-[*yuan-jian*] "very farsighted"
(s 3) (3 4) i.
(3 s 3 4) j*.

very have two trick
b. *hen # you*-[*liang*-[*xia-zi*]] "very skillful"
(s 3) (3 4 o) i.
(3 s 3 4 o) j*.

The [have + NP] constructions above clearly constitute single syntactic words. The semantic non-compositionality of expressions like "have + two

+ tricks" = "skillful" (131b) suggests that they correspond to word-like idioms. In addition, their collocational restrictions attest to the syntactic opacity of the internal structure of these morphologically complex constructions. The degree adverb *hen* "very" subcategorizes for degree adjectives or stative verbs like *gao-xing* "happy" or *xi-huan* "like to"; it cannot modify verbs like *da* "play" or *you* "have, possess," as shown in (132).[49]

(132)

a. **hen da wang-qiu*
very play tennis
"play a lot of tennis"

b. **hen you xi-jun*
very have bacteria
"have a lot of bacteria (full of bacteria)"

The fact that *hen* can modify *you-yuan-jian* "have farsightedness" and *you-liang-xia-zi* "(lit.) have two tricks" means that these [have + NP] expressions are interpreted as single complex predicates equivalent to indivisible stative verbs translatable as "be farsighted" and "be skillful," and so forth. What is striking is that these syntactic words behave phonologically like phrases in that the word-internal *you* "have" can break away to form a foot with the adverb *hen.*

Conversely, syntactic phrases are treated like words for the purpose of lexical MRU. We have already seen in section 4 many cases where lexical MRU treats larger-than-word constructions such as clitic groups as if they were words, in that MRU formation on clitic groups must precede and preempt phrasal MRUs. The larger-than-word structure that is rhythmically parsed as a single lexical item is not restricted in kind to clitic groups built around one single content word. Often more fully developed syntactic phrases involving several content words may be treated by phonological rules as though they constituted word-level entities.

Such a syntax–phonology mismatch is more straightforwardly demonstrated by data from other dialects, for instance **Shanghai**. Recall that this dialect calls for left-headed stress-feet to be built over the word (chapter 7, section 3.1). Within the foot, only the leftmost prominent syllable retains its lexical tone and spreads it to its foot-mates on the right.[50] Consider a minimal pair like the one in (133):

[49] One cannot attribute the distributional pattern of (132) to semantics. Just as one can have more or less of farsightedness or skills (131a), one can also play more or less of tennis, or be more or less seriously infected by bacteria, etc.

[50] For cross-dialect comparison, I will use Pinyin transcription in the following examples.

(133) drop anchor

a.	*pao-mao*	"breakdown"
	HL LH	base tone
	(HL o)	Deletion
	(H L)	Spread

drop anchor

b.	*pao # mao*	"to drop anchor"
	HL LH	base tone
	(HL) (HL)	Deletion/Spread not applicable

(133a) instantiates a semantically opaque lexical compound; (133b), on the other hand, made up of exactly the same morphemes, is interpreted phrasally and literally. By the same token, a single stress-foot is built over (133a), but two over (133b), given the word-bound nature of stress-foot formation in Shanghai. Now consider the following:

(134) make fun

a.	*xun* [*kai-xin*]	"make fun of someone"
	LH HL HL	base tone
	(LH o o)	Deletion
	(L H o)	Spread

get-in trick

b.	*shang dang*	"be duped, fooled"
	LM MH	base tone
	(LM o)	Deletion
	(L M)	Spread

Both (134a) and (134b) contain only one stress-foot. Phonologically, therefore, constructions with an internal verb + object structure like these behave like word-size units. Syntactically, however, they display all the characteristics of syntactic phrases. Contrast a verbal complex like *xun kai-xin* "to tease" (lit. make + fun) with *dan-xin* "be concerned about" (lit. carry + heart).

(135)

a. **xun kai-xin* [*ta de tong-xue*]
make fun his schoolmate
"make fun of his schoolmate"

b. *dan-xin* [*ta de di-di*]
concerned he DE younger-brother
"be concerned about his younger brother"

Functionally both are transitive verbs subcategorizing for an object NP. Structurally, however, while *dan-xin* behaves like a simple verb and therefore takes *ta de di-di* "his younger brother" as expected, *xun kai-xin*

patterns as a verb + object phrase, with the object position already taken up by *kai-xin* "fun," and consequently can no longer take an additional, so-called "outer," object like *ta de tong-xue*. This phrase-structure constraint forces the "logical" object *ta de tong-xue* to be interposed in the middle of the [V ____ Object] construction:

(136) *xun* [*ta tong-xue de*] *kai-xin*
make he schoolmate DE fun
"make fun of his schoolmate"

Furthermore, the "internal" object NP (underlined) of expressions like *xun* <u>*kai-xin*</u> "make <u>fun</u>" and *shang* <u>*dang*</u> "be duped, fooled" (lit. get-in + trick) can show up as an empty category, as in (137a, b).

(137)

a. *peq*[51] *dang*$_i$ [*ni shang* ___$_i$]
give trick you get-in
"set a trick for you to get (trapped) in"

b. *zhe ge dang*$_i$ [*ni zhen shang* ___$_i$ *le*]
this CL trick you really get-in Asp
"this trick really fooled you"

Such blatant violations of "Lexical Integrity" have prompted most grammarians of Chinese to regard verb + object constructions as true phrases (cf. J. Huang 1984, Dai 1992). Herein lies the paradox: on the one hand, these verb + object expressions are syntactically phrases; on the other hand, they are parsed phonologically as words, as diagnosed by the fact that only one foot is built over each of them (see (134a, b) above).

This non-isomorphism between the syntactic word and the phonological word is hardly surprising or unique to Chinese. Ackerman and Webelhuth (1992, forthcoming) argue that the notion of the "word" is a multifaceted one. In essence, a word is *operationally* rather than *structurally* defined: an expression is an X-word to the extent that it constitutes the object of X-operations, where X-operations may be lexical–functional, syntactic, or phonological in nature. Thus, German [PV V] complexes (where PV = pre-verb) constitute a word-like functional unit in that they are endowed with idiosyncratic meaning (e.g. *nehmen* "take" vs. *an=nehmen* "accept"), involve valency change (*gehen* "go, intrans." vs. *ein=gehen* "enter, trans."), and exhibit other lexical effects like case-assignment (*bei=stehen* "support" takes the dative, while *ein=stellen* "hire" takes the accusative) and order of affixation (*bring-en* "bring" → **Bring-ung*

[51] *peq* "give" has no cognate in Beijing Mandarin.

"bring-ing"; *bring-en* "bring" → *an=bring-en* "to install" → *[An=bring]-ung* "installation"). In this sense, [PV V]s satisfy the functional and morphological criteria of wordhood. Remarkably, however, such [PV V]s behave not like syntactic words, but like bona fide phrases. This is so because the so-called separable prefixes are stranded when the verbs move to the second (V2) position (e.g. *Weil er das Fenster auf=macht* "because he opens the window" vs. *Er* macht *das Fenster* auf "He opens the window"), and can be topicalized (Auf *hat er das Fenster* gemacht), in clear violation of Lexical Integrity. In this connection it is worth noting that the characteristic "compound stress" pattern (*AUF=machen* "open" vs. *er-ZIEHEN* "rear, bring up") suggests that [PV V] expressions like *auf=machen* are treated like unitary phonological words (cf. stress on the first stem in the nominal compounds like *RAT-haus* "counsel + house = city hall"). In short, German [PV V] constructions behave like unitary words functionally, morphologically, and phonologically; syntactically, however, their internal structure is transparent and accessible to phrase-level operations (e.g. V2 position, topicalization).

The German case is analogous to the Shanghai verb + object expressions like *xun kai-xin* "make fun of" and *shang dan* "be duped": by virtue of being metrified as one single foot each and undergoing *lexical*, i.e. word-internal, tone sandhi, they pass the phonological test of wordhood; syntactically speaking, however, they retain the status of ordinary phrases.

Ackerman and Lesourd (1992) made a similar case with Hungarian [PV V] complex predicates. These [PV V] constructions display all the standard lexical effects, but surprisingly undergo such phrasal operations as coordinate reduction and ellipsis. Interestingly enough, PV and V also constitute separate phonological words, for instance, for the purpose of applying vowel harmony (cf. Nespor and Vogel 1986:122ff.). Thus the PV and V do not harmonize in examples such as *fel-ugrani* "up + jump" or *oda-menni* "there + go." Unlike German then, the [PV V] complex predicate in Hungarian constitutes a single lexical unit morphologically and semantically/functionally, but separate words phonologically and syntactically.

The mismatch among the functional, structural, and phonological dimensions of wordhood opens up intriguing questions concerning a linguistic construct that seems paradoxically tangible and elusive at the same time. Fresh insights may well come from a surprising source, a morphology-poor but sandhi-rich language like Chinese.

10 *Phonological phrase as a sandhi domain*

Tone sandhi in southern Min, notably **Xiamen** and the closely related dialects spoken in Taiwan, operates within a domain that is defined neither by reference to stress-feet as in Shanghai and other Wu dialects, nor in terms of "minimal rhythmic units" (MRUs) as in Beijing and other Mandarin varieties. The sandhi domain observed across a wide variety of southern Min dialects, including Xiamen and Taiwanese, approximates that of the phonological phrase, or **p-phrase**, in the prosodic hierarchy. The focal point in this chapter is to find the proper and precise characterization of this prosodic category.

I first summarize the analysis given of Xiamen in Chen (1987a) (section 1), with independent support from Shanghai (section 2). Further refinements and residual problems are discussed, in particular with reference to "domain c-command" (section 3) and lexical government (section 4). The chapter closes with some remarks on the effects of rhythm on tone sandhi.

1 End-based p-phrase

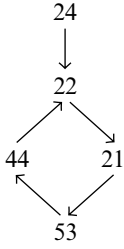
The Xiamen tone system is fairly complex, though not at all uncommon among Chinese dialects (1).

(1)	smooth syllables	44	H
		24	MH
		22	M
		21	ML
		53	HM
	checked syllables	4q	H
		32q	ML

-q marks a "checked" tone

Like other Chinese dialects, Xiamen divides syllables into smooth (or legato) and checked (or staccato) types, each with its subsystem of sandhi substitutions. Specifically, Xiamen distinguishes two positions: domain-final and everywhere else, which I will call **sandhi positions** (including domain-initial and domain-internal) – leaving for the moment undefined what this domain is. In sandhi positions, every tone is replaced by its corresponding sandhi tone. The pattern of substitution of one "smooth" tone by another has been likened to a musical chair movement, and has been referred to as the "southern Min tone circle," "tone clock," etc. The checked syllable tones toggle, so to speak, between the high and the low register: the high tone [4q] becomes low [21q], while the low switches to a high, which has two allotones, [53q] for syllables ending in a glottal stop, and [4q] elsewhere. We repeat the TS rules from Chen (1987a:111f.):

(2) smooth syllable tone sandhi (tone circle)



(3) checked syllable tone sandhi[1]

4q	→	21a
32q	→	4q for syllables ending in p,t,k
		53q for syllables ending in q

The effect of tone sandhi is illustrated below. In the examples that follow (taken from Chen 1987a:112), the key word in question (underlined for clarity) occupies a domain-final position in the left column, therefore not subject to the sandhi rules, and surfaces with its base tone unchanged; in the right column, however, the same key words occur domain-initially, therefore undergo tonal substitution, and take on a corresponding sandhi tone, consistent with the rules stated above.

[1] The glottal stop "-q" (but not a full-fledged stop p,t,k) deletes in sandhi contexts. I ignore this point in the ensuing discussion.

(4)

		base form		sandhi form	
a.		*tsin p'ang* 44	"very fragrant"	*p'ang tsui* 22	"perfume" (fragrant + water)
b.		*p'e we* 24	"leather shoes"	*we tua* 22	"shoe laces"
c.		*wi pĩ* 22	"stomach ailment"	*pĩ lang* 21	"sick person"
d.		*k'i ts'u* 21	"build a house"	*ts'u ting* 53	"roof" (house + top)
e.		*tua hai* 53	"big ocean"	*hai kĩ* 44	"ocean front"
f.	i.	*ts'ut lip* 4q	"exit and entrance"	*lip k'ao* 21q	"entry point"
	ii.	*tua beq* 4q	"barley" (lit. large wheat)	*beq hun* 21q	"flour" (lit. wheat powder)
g.	i.	*m bat* 32q	"not know"	*bat li* 4q	"literate" (lit. know characters)
	ii.	*pak aq* 32q	"Peking duck"	*aq nng* 53	"duck egg"

The relationship between the input and the output tones as stated in the sandhi rules formulated above has eluded students of tonology ever since Chiu (1931:25; cf. Bodman 1955) first discovered the Xiamen "law of tone-shifts." Theoretical issues involved in the complex tone sandhi of Xiamen (and the closely related dialects of Taiwan) have been debated in Wang (1967), R. Cheng (1968), Yip (1980), Wright (1983), Du (1988), King (1988), and Tsay (1994), among others. Its phonetic motivation is presumably shrouded in the as yet poorly understood prehistory of the Southern Min dialects. In what follows I will set aside the representational issues associated with the morphotonemic alternations exhibited by Xiamen. Instead, I will focus exclusively on the proper circumscription of the sandhi domain.

The crux of the problem can be illustrated with a few examples.

(5) wind + blow

a. *hong-ts'e* "kite"
44 44 base tone
(22 44) sandhi form

wind blow rain pour

b. *hong ts'e hɔ ak* "wind blows, and rain pours"
44 44 22 32q
(44) (44)(22)(32q)

release kite

c. *pang hong-ts'e* "to fly a kite"
21 44 44
(53 22 44)
(. . .) = domain of tone sandhi

In (5a) *hong* "wind" occurs in a non-final position, and undergoes tonal substitution. The same *hong* in (5b) also occurs in a non-final position, but retains its base tone. One cannot simply say that tone sandhi is a lexical rule, operating only on wordforms, because the word *pang* "to release" also undergoes tonal substitution, although it uncontroversially constitutes a separate word from *hong-ts'e* "kite." The final vs. elsewhere positions are defined with reference to the string of elements bounded by (. . .). How we cut up a speech continuum into strings bounded by (. . .) is the puzzle we set out to unravel.

From this point on I will dispense with the base tones and their sandhi forms, since for our present purposes the patterns of tonal substitutions are only of marginal interest, the only relevant question being where precisely the domain boundaries (as diagnosed by tone sandhi) are located. I will therefore follow the notational conventions used in Chen (1987a): # marks the boundary between p-phrases (i.e. tone groups). AB # C is equivalent to (AB) (C); that is to say, A and B join together into one single p-phrase, but B and C belong to two separate p-phrases. Hence # blocks tone sandhi which operates strictly within the p-phrase. For expository clarity I occasionally use the symbol "=" to signal the fact that the elements on either side are linked together as members of a single prosodic domain, and therefore constitute an obligatory sandhi site. Thus A = B # C is non-distinct from A B # C. The examples given above as (5a, b, c) are more parsimoniously annotated as:

(6) a. *hong-ts'e* "kite"
b. *hong # ts'e # hɔ # ak* "wind blows, and rain pours"
c. *pang hong-ts'e*, **or** "to fly a kite"
pang = hong-ts'e

The task at hand is to discover a general principle that predicts the distribution of "#" that signals the break between two sandhi domains.

1.1 Marginal relevance of stress-feet

It is easy to demonstrate that minimality is totally irrelevant as a defining feature of the sandhi domain in Xiamen: each monosyllable in example (6b = 5b) constitutes a separate sandhi domain, in effect insulated from

interacting with any surrounding tone. The domain of Xiamen tone sandhi is clearly a different entity from the MRU (minimal rhythmic unit) of Mandarin Chinese.

The same examples also demonstrate that tone sandhi in Xiamen is not confined to a phonological word (p-word), if by p-word we understand a prosodic unit organized around a lexical core. Rather, Xiamen TS spans a larger structure, as evidenced by (6c), which is arguably a full-fledged VP, consisting of two lexical heads, *pang* "release, let go" and *hong-ts'e* "kite."

It is equally simple to show that the notion of stress-foot that plays a key role in the description of Wu dialects enjoys only a marginal status in Xiamen. Recall the following contrast in **Shanghai** (from chapter 7, section 3.3).

(7)		(x . .)	compound stress
		(x .) (x)	word stress
	a.	[*h̲õ. moq*] *zã*	"redwood bed"
		LH LH LH	
		(L H) (LH)	tone deletion/association
		(x . .)	Stress reduction
→	b.	[*h̲õ. moq*] *zã*	
		LH LH LH	
		(L H o)	

Cyclic stress assignment yields a perfect grid in (7a), with two stress-feet. Foot-internal tone deletion and association yield the desired outputs. Optional stress reduction compresses the two feet into one, with predictable consequences in (7b). In sharp contrast to (7), a right-branching nominal expression like (8) gives rise to the familiar stress clash in (8a), which is resolved by destressing the offending *moq* "wooden," resulting in (8b). This explains nicely the fact that while (7) permits two alternative readings, (8) allows only one.

(8)		(x . .)	compound stress
		(x) (x .)	word stress
	a.	*h̲õ* [*moq zã*]	"red wooden bed"
		LH. LH. LH	
		(LH)(L. H)	= *
		(x . .)	Clash resolution
→	b.	*h̲õ* [*moq zã*]	
		LH LH LH	
		(L H o)	tone deletion/association

Prompted by the success of a prominence-based account of Shanghai, Duanmu (1995) extends the same principle to his description of Xiamen tone sandhi. On Duanmu's account, Xiamen represents the exact mirror image of Shanghai, that is the stress-foot in Xiamen is right-headed. Interesting consequences follow. Consider the following pair from Xiamen:

(9)
```
                    x
          (.   x)   (x)
   a.  *[ts'ian li] # be          "winged steed"
       thousand mile horse

          (.    .     x)
→  b.  [ts'ian li] = be           Clash resolution
```

(10)
```
                    x
        (x)   (.     x)
   a.  ts'u # [ting-bin]          "roof"
       housetop

        (.     .     x)
→  b.  ts'u = [ting-bin]          Stress reduction (optional)
```

Left-branching compounds like (9a) create a stress clash, triggering the obligatory de-stressing of *li* "mile." Hence, only a single stress-foot is built over (9b). By contrast, a right-branching expression like (10a) is metrically parsed into a perfect grid, with two stress-feet, collapsible into one through optional stress reduction. By equating the sandhi domain with the stress-foot, we account for the fact that while (9) constitutes one single sandhi domain, so that all syllables except *be* "horse" must replace their base tone with a sandhi tone, (10) has two alternative readings. It stands to reason to use the same analytical tool devised for Shanghai to deal with Xiamen, as proposed by Duanmu (1995). The only assumption we need to make is that Xiamen has a right-prominent prosodic structure. Everything else follows.

Or does it? There are two types of problems with extending the Shanghai-style metrical analysis to Xiamen. Firstly, a priori, clash-resolution via stress deletion is by no means the only logical possibility. Rightward clash is either freely tolerated (as in Shanghai, Danyang; see chapter 8, sections 2.3–2.4) or else is alleviated by Iambic Reversal, as in Mandarin (chapter 7, section 2.3). Therefore, even assuming a right-prominent system for Xiamen, it does not follow that stress-clash automatically leads to destressing and rebracketing. Secondly, in order to account for the alternative readings of (7) and (10), we must posit an *optional* stress reduction. Now take example (11). It has a perfect grid, hence Clash Resolution is

irrelevant. Optional Stress Reduction predicts that both (11a) and (11b) are possible. The prediction is false: only reading (11b) with obligatory tone sandhi is acceptable. Similar examples can be found ad infinitum.

(11)
```
                x
        (x)     (.      x)
```
a. **ang* # [*tsng-t'au*] "red brick"

b. *ang* = [*tsng-t'au*]

cf. *gong* [*gin-na*] "silly child"
kuã [*siau-suat*] "read a novel"
gau [*kong we*] "eloquent" (lit. skillful at talking)

It doesn't help to appeal to Weak Prohibition on degenerate feet, since a monosyllabic sandhi domain is perfectly acceptable, as demonstrated by (7a) and (10a).

The picture is further muddied by the contrast between (12) and (13). While both (12a) and (12b) are acceptable readings, (13a) is felicitous only under a contrastive reading – as in "NEWS journalist" vs. "PHOTO-journalist."

(12) gate behind
[*tua-mng*] [*au-piaq*] "behind the gate"
a. *tua-mng* # *au-piaq*
b. *tua-mng* = *au-piaq*

(13) news journalist
[*sin-bun*] [*ki-tsia*] "(print) journalist"
a. *sin-bun* # *ki-tsia* ? (contrastive reading only)
b. *sin-bun* = *ki-tsia*

The only conclusion we can draw is that Xiamen tone sandhi does not yield gracefully to a metrical analysis.

This does not mean that rhythm is irrelevant. Xiamen does exhibit fairly robust rhythmic effect (see section 5), but its core sandhi system is keyed to a totally different principle. Basically, the sandhi process in Xiamen is circumscribed by a domain that most closely approximates the p-phrase in prosodic hierarchy. As originally proposed in Chen (1987a), the characterization of this prosodic domain rests on two key concepts: maximal projection and the argument vs. adjunct distinction.

1.2 Maximal projection

Abstracting away from the phonological issues, we can restate the tonal alternations schematically as (14), which simply says that all non-final

tones (T) change into corresponding sandhi tones (T′), within the domain of a phonological phrase, or p-phrase. The p-phrase in question corresponds to the non-committal term "tone group" as used in Chen (1987a).

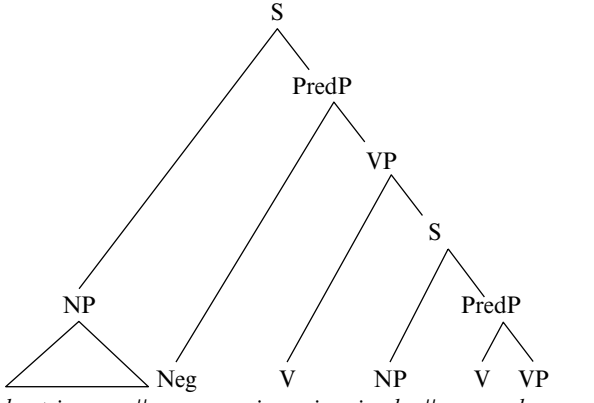
(14) Xiamen tone sandhi (TS)

$$(\ldots T \ldots T)\ \text{p-phrase} \downarrow T'$$

As a first approximation, p-phrase can be derived by demarcating the right edge of a maximal projection, X^{max}, in terms of the end-setting typology (Selkirk 1986):

(15) Xiamen p-phrase: {Right, X^{max}} (preliminary version)

To borrow a familiar example from Chen (1987a), (16) shows a sentence cut up into three p-phrases in accordance with (15).



(16) *lao tsim-a-po # m siong-sin ying-ko # e kong-we*
old lady not believe parrot can talk
"the old lady doesn't believe that parrots can talk"

Thus, the p-phrase boundary comes at the end of the matrix subject NP (*lao tsim-a-po* "old lady"), the embedded subject NP (*ying-ko* "parrot") and, of course, the end of the sentence, which is also the right edge of S, PredP, and VP. Note in particular that the middle p-phrase is a non-constituent: *m siong-sin ying-ko* "not believe that parrots."

The fact that TS applies across *siong-sin* "believe" and *ying-ko* "parrot" demonstrates that Xiamen p-phrasing is blind to the left-edge of X^{max} since *ying-ko* marks the beginning of an NP, in fact a subordinate clause. In this respect Xiamen p-phrasing functions like some sort of phonological

punctuation: the appearance of the unchanged base tone serves to signal the end of a major syntactic constituent, the X^{max} or XP (however, see further discussion in chapter 2, section 4.2). The diagnostic value of such a phonetic cue is exemplified by the contrast below, where the unchanged base tone is underlined for emphasis:

(17) a. [*yi* [*sia* *k'a* *kin*]$_{VP}$]$_S$ "he writes faster"
he write more fast
44 53 53 53
(22 44 44 <u>53</u>)

b. [*yi* [*sia*]$_{VP}$ #]$_S$ *k'a* *kin* "it would be faster for him to write"
he write more fast
44 53 53 53
(22 <u>53</u>) (44 <u>53</u>)

In (17a), *k'a kin* "faster" is an adjunct of VP. There is no internal XP bracket,[2] hence no p-phrase break. In contrast, *k'a kin* is interpreted as the predicate that takes the sentential subject *yi sia* "he writes." *Sia*, therefore, is bounded on the right by VP (and S), and hence by a phonological break.

The cross-categorial nature of Xiamen p-phrasing is shown by the fact that not only NPs, as in (16) and VPs, as in (17b), but other XPs such as QP, AP as well as AvP also induce a tone-sandhi break, as illustrated by the examples of (18).

(18) a. *ti-baq* # [*tsit kun*]$_{QP}$ # *sã k'ɔ*
pork one catty three dollar
"pork is three dollars a catty"

b. *ka* *hak-sing* # *hun-lian ka* [*ko* *ang*]$_{AP}$ # *ko tsuan*
OM student train Comp both red and expert
"train students so that they become both red[3] and expert"

c. *peq* [*lu* *kuãi*]$_{AvP}$ # *puaq lu* *t'iam*
climb more high fall more hard
"the higher you climb, the harder you fall"

OM = object marker
Comp = complementizer

[2] The personal pronoun *yi* "he" is treated as simple X^o, not a full-blown lexical projection. For a fuller treatment of pronouns in Xiamen tone sandhi, see J. Lin (1994). Analogously, wh-words in Papago phonology behave as particles, not as maximal projections, even where they are necessarily answered by XPs, cf. Hale and Selkirk (1987:168f.).

[3] I.e. ideologically orthodox and motivated.

1.3 Argument vs. adjunct

Crucial to Chen's (1987a) account of Xiamen tone sandhi is the distinction between arguments and adjuncts. In its original formulation, Xiamen p-phrasing or what I called Tone Group Formation, reproduced here as (19), says in effect that tone sandhi is blocked by an argument phrase, but not an adjunct phrase.

(19) Tone Group Formation
Mark the right edge of every XP with #, except where
XP is an adjunct c-commanding its head. (Chen 1987a:130)

Evidence for this functional distinction rests on such contrasts as (20, 21, 22) given below. Let us look at the sandhi behavior of the various pre-head phrasal units (XP or X^{max}) in (20)–(22); whether or not an XP blocks TS depends on its functional relation to the head. The classic pair of sentences (20a, b) is taken from R. Cheng (1968).

(20) $[\text{NP \# AP}]_S$ vs. $[\text{NP = A}]_{AP}$

a. $[mua\text{-}a]_{NP}$ # $[tua]_{AP}$ # *e* *sio-piã*
sesame-seed big Comp bun
"buns (on which) the sesame seeds are small"

b. $[[mua\text{-}a]_{NP} = tua]_{AP}$ # *e sio-piã*
"buns as big as sesame-seed" (tiny buns)

marks p-phrase boundaries
= signals obligatory sandhi contexts

In (20a) the NP *mua-a* "sesame seed" functions as the argument of *tua* "big," and is therefore demarcated on the right by the p-phrase boundary "#." Tone sandhi blocks at this juncture. In (20b), on the other hand, the same *mua-a* functions as the degree adjunct of the head verb *tua*, so that *mua-a tua* "big as sesame seed" exhibits the same kind of semantic relation as that which holds in such expressions as *razor-sharp*, *skin-deep*, and so forth.

The preverbal NP *p'ue* "letter" in (21a) functions as an argument, and blocks TS, while the AvP *kuã-kin* "hurriedly" in the same position allows TS to apply within the VP in (21b).

(21) $[\text{NP \# V}]_{VP}$ vs. $[\text{AP = V}]_{VP}$

a. *Ting sio-tsia* # $[[p'ue]_{NP}$ # *sia liao lo*$]_{VP}$
Ting Miss letter write Asp Prt
"Miss Ting has written the letter"

b. *Ting sio-tsia* # [[*kuã-kin*]$_{AvP}$ = *to-k'i*]$_{VP}$
hurriedly return
"Miss Ting hurriedly went home"

That the categorial distinction between NP and AvP alone does not suffice to account for their TS behavior in (21) is apparent when we compare (22a) and (b). Here we have two NPs, one functioning as the argument of the preposition (or rather, postposition) *ao-piaq* "behind," the other as the adnominal modifier of *ki-tsia* "journalist." TS blocks in the former but applies in the latter, consistent with the argument/adjunct dichotomy.

(22) [NP # P]$_{PP}$ vs. [NP = N]$_{NP}$

a. [[*tua-mng*]$_{NP}$ # *ao-piaq*]$_{PP}$
gate behind
"behind the gate"

b. [[*bi-kok*]$_{NP}$ = *ki-tsia*]$_{NP}$
U.S. journalist
"American journalist"

In view of these facts, we need to revise the Xiamen p-phrase as follows:

(23) Xiamen p-phrase: {Right, X^{max}}, X^{max} not an adjunct.
(revised version)

This concludes our summary of the analysis proposed in Chen (1987a). Although it accounts for a broad range of data, it is not entirely satisfactory. In particular, the phonological phrasing calls for a curious mix of structural (maximal projection) and functional (argument vs. adjunct) conditions. Before we look into various refinements suggested in particular by Chung (1989) and J. Lin (1994), let us consider some independent support for the prosodic category p-phrase that Shanghai has to offer.

2 Supporting evidence for p-phrase

The Wu dialect of **Shanghai** has been described in some detail in chapter 7, section 3. The core processes of tone deletion and tone association are shown to be foot-level rules, the foot being constructed in accordance with the general principles of metrical phonology. However, this is only part of the picture. In addition to metrical feet, we need a higher-order prosodic category, the p-phrase. There are two types of arguments that motivate the p-phrase: first, foot construction respects p-phrase boundaries; second, post-focus de-stressing and de-toning is a p-phrase-level rule.

2.1 Foot construction respects p-phrase boundaries

There appears to be a prima facie puzzling contrast between the (a) and the (b) sentences in Shanghai given in (24), as pointed out by Selkirk and Shen (1990: 324f.).

(24)	(x .) (x .)	
a.	*zï laq zã-he*	"live in Shanghai"
	reside in Shanghai	
	LH LHq LH MH	base tone
	(L H) (L H)	sandhi form
	(x .) (x . .) (x .)	
b.	*zï laq zã-he geq ni-tsï*	"(it) is the son who (lives) in Shanghai"
	be in Shanghai Prt[4] son	
	LH.LHq LH.MH LHq.LH.MH	
	(L H) (L H o) (L H)	= *
	(x) (x) (x . .) (x .)	
b′.	*zï laq zã-he geq ni-tsï*	
	LH.LHq LH.MH LHq.LH.MH	
	(LH) (LH) (L H o) (L H)	= ok
(25)	(x .) (x .)	
a.	*tsɤ tɔ nö-tsin*	"walk to Nanjing"
	walk to Nanjing	
	MH. MH LH.HL	
	(M H) (L H)	
	(x .) (x . .) (x)	
b.	*tsɤ tɔ nö-tsin geq lu*	"take the way (which leads) to Shanghai"
	walk to Nanjing Prt way	
	MH MH LH HL LHq LH	
	(M H) (L H o) (LH)	= *
	(x) (x) (x . .) (x)	
b′.	*tsɤ tɔ nö-tsin geq lu*	"take the way (which leads) to Shanghai"
	walk to Nanjing Prt way	
	MH MH LH HL LHq LH	
	(MH) (MH)(L H o) (LH)	= ok

(24a) and (25a) are exactly as one would expect. The function words *laq* "in" and *tɔ* "to" are prosodically weak, and cliticize left to form a p-word with the lexical host on the left, exactly as in Beijing Mandarin (see chapter 9, section 4). Stress-feet are constructed in the usual manner, and tone deletion/association apply within the metrical feet as expected. What

[4] I gloss *geq* simply as Prt. It corresponds to Mandarin *de*, and Xiamen *e*, and functions as a subordinator, relative clause marker, or complementizer.

is surprising is the fact that the same function words *laq* and *tɔ* resist the same phonological parsing in (24b) and (25b); instead they each form a separate foot, and therefore block both tone deletion and tone association, as attested by the retention of the underlying tones /LHq/ and /MH/ on *laq* and *tɔ* respectively in reading (24b′) and (25b′). To this puzzle Selkirk and Shen advance the following solution. They note that the prepositions *laq* and *tɔ* of the (a) sentences are "weakly embedded," within the functional projection of PP (= P^{max}); on the other hand, the same prepositions of the (b) sentences are "strongly embedded," that is enclosed within the PP, which is in turn contained in the NPs headed by *ni-tsï* "son" and *lu* "way, road." The phrase structures of (24b) and (25b) are spelled out more fully as (26) and (27).

```
                reside in   Shanghai
(26)  a.   VP[zï PP[laq NP[zã-he
                (     )   (     )                    p-phrase
                (x    .)  (x   .)                    foot

                be      in      Shanghai Prt   son
           VP[zï NP[ PP[laq NP[zã-he      geq N[ni-tsï
                ( )     ( )   (                  )   p-phrase
      b.*       (x        .)  (x   .       .)  (x .) foot
      b′.       (x)     (x)   (x   .       .)  (x .) foot

                walk  to    Nanjing
(27)  a.   VP[tsɤ PP[tɔ NP[nö-tsin
                (     )   (     )                    p-phrase
                (x    .)  (x   .)                    foot

                walk    to      Nanjing Prt  way
           VP[tsɤ NP, PP[tɔ NP[nö-tsin   geq N[lu
                ( )     ( )   (          )          p-phrase
      b.*       (x        .)  (x   .       .)  (x)   foot
      b′.       (x)     (x)   (x   .       .)  (x)   foot
```

Since, as we will see presently, the right edges of the syntactic constituents are irrelevant, we annotate only the left edges. In this light, the structural difference between the two occurrences of the preposition *laq* becomes apparent: in (26a) *laq* abuts the left edge of PP, a functional category, while in (26b) it lies at the left extremity of both PP and NP, the latter being a lexical projection headed by *ni-tsï* "son." It is, according to Selkirk and Shen (1990), on these lexical vs. functional projections that turns the difference in the phonological parsings of (26a) vs. (26b). Specifically, Selkirk and Shen (1990:332) propose that in addition to the p-word, over which stress-feet are built, we need to posit a higher prosodic

constituent, the p-phrase,[5] which can be derived by the same end-setting parameters as we have seen in Xiamen, except that this time, X = Lex^{max}, that is the maximal projection of a lexical category, NP, VP, AP:

(28) Shanghai p-phrase: {Left, Lex^{max}}

Given this parametric choice, *zï* "reside" and *laq* "in" of (26a) belong to the same p-phrase, while *zï* "be" and *laq* "in" of (26b) are segregated into two separate p-phrases, because of the left edge of the NP, that represents the maximal projection of *ni-tsï* "son." The same procedure yields the p-phrasing of (27a, b).

Having established the superodinate prosodic constituents, the top-down effect of p-phrases on the parsing of elements into p-words and stress-feet follows automatically from the general well-formedness constraint on prosodic structure, which states, to quote Selkirk and Shen (1990:320): "The prosodic structure of a sentence must conform to the rule schema $C^n \rightarrow C^{(n-1)}$," that is to say, a higher prosodic structure is exhaustively parsed into prosodic constituents of the next level down. In particular, the Strict Layer Hypothesis (Selkirk 1981a, b, 1984, Nespor and Vogel 1986, Hayes 1989) prohibits the joining of two prosodic constituents of a higher order into one prosodic category of a lower order. To put it in reverse, it bars a lower prosodic category from straddling two higher constituents. In the case under discussion, since *zï* and *laq* are parsed as two p-phrases (C^n), each must exhaustively dominate the prosodic constituent on the immediate level down, namely p-word (C^{n-1}) and, a fortiori, stress-foot. The readings (26b) and (27b) are rejected, because they would presuppose a phonological parsing that is antithetical to the Strict Layering principle, namely with a p-word (*zï laq)* that straddles two p-phrases {*zï*} {*laq*}. In short, independent and general well-formedness conditions on prosodic structure enforce the creation of a separate p-word / stress-foot on *laq* in (26b′), and correspondingly *dɤ* in (27b′).

2.2 *De-toning within the p-phrase*

The simple and principled account proposed above has considerable independent support. P-phrase is necessary as a domain that circumscribes the process of (optional) Post-Focus Tone Deletion, stated informally below:[6]

[5] Selkirk and Shen (1990:332) actually label this prosodic constituent as the *major phrase.*

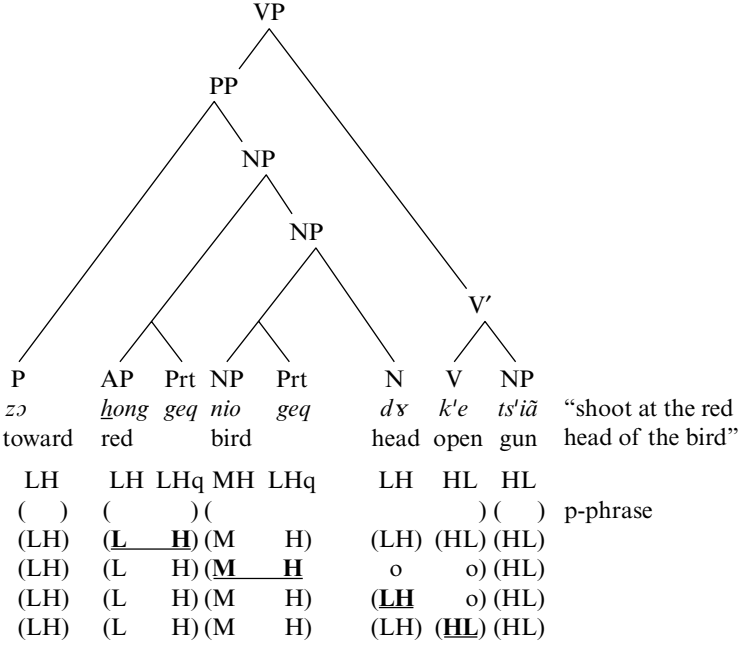
[6] (29) is a slightly modified version of Selkirk and Shen (1990:327) rule (24). T stands for the root tone; p-phrase is substituted for their "*major phrase.*"

(29) Shanghai Post-Focus Tone Deletion

$$(\ldots \text{focus} \ldots \sigma \ldots)_{\text{p-phrase}}$$
$$\quad\quad\quad\quad\quad\quad\quad\quad |$$
$$\quad\quad\quad\quad\quad\quad\quad\quad T \rightarrow \varnothing$$

(29) has the effect of deleting all tones occurring after the focus element – within a p-phrase. Without the notion of p-phrase as defined in the preceding section, one would be hard put to explain the various readings of sentence (30), taken from Selkirk and Shen (1990:330, ex. (29)):

(30)



Underlined boldface signals the element under focus. Thus reading (b) means something like "Shoot at the red head of the *bird*" (not the *cat*), and so forth. What is noteworthy is that in reading (b), the focus on *nio* "bird" triggers the metrical demotion (marked by stresslessness and atonicity, symbolized by "o") of *dɤ* "head" and *k'e* "open." Likewise, in reading (c) the focus on *dɤ* "head" causes *k'e* "open" to be de-emphasized by losing its distinctive tone. However, no post-focus tone deletion is observed in either reading (a) or reading (d): here the highlighting of *k'e* in reading (d), for instance, leaves the tone on *ts'iã* "gun" intact. By limiting post-focus to the p-phrase, we derive such focus-sensitive, phonetically contrastive readings in a straightforward manner. Assuming the

phrase structure of (30) as specified, {Left, Lex^{max}} picks out *h̲ong*, *nio*, and *ts'iã* as the beginnings of separate p-phrases. Accordingly, the sentence is parsed into the four p-phrases. On this account, both *dɤ* "head" and *k'e* "open" are within the same p-phrase as the focus *nio* "bird" in reading (b). Consequently, both *dɤ* and *k'e* are subject to Post-Focus Tone Deletion. In reading (d), on the other hand, *ts'iã* lies outside of the p-phrase that contains the focused element *k'e*, and therefore retains its tone. I take post-focus de-toning in Shanghai to be one of the strongest pieces of evidence in support of the prosodic entity we call the p-phrase.

3 M-command or domain c-command

Having confirmed the need for the prosodic category p-phrase with supporting evidence from Shanghai, let us return to the Xiamen case. Recall that p-phrase in Xiamen is defined as follows:

(31) **Xiamen p-phrase**: {Right, X^{max}}, X^{max} not an adjunct.

As noted at the end of section 1, the negative condition "X^{max} not an adjunct" complicates the p-phrasing rule, and renders it less appealing as a general explanatory principle. Furthermore, the evidence produced above in support of the argument vs. adjunct distinction is less than compelling. Take the sentences (21a, b) repeated below:

(32) a. *Ting sio-tsia* # [[*p'ue*]$_{NP}$ # *sia liao lo*]$_{VP}$
Ting Miss letter write Asp Prt
"Miss Ting has written the letter"

b. *Ting sio-tsia* # [[*kuã-kin*]$_{AvP}$ = *to-k'i*]$_{VP}$
hurriedly return
"Miss Ting hurriedly went home"

They were cited as evidence in support of the distinction between argument (the object *p'ue* "letter") and adjunct (the adverbial modifier *kuã-kin* "hurriedly") as the key to the different sandhi behaviors of the two maximal projections NP and AvP. But there is an alternative interpretation. One could construe the preverbal object NP as having been topicalized outside of the VP. In other words:

(32) a′. *Ting sio-tsia* # [*p'ue*]$_{NP}$ # [*sia liao lo*]$_{VP}$
Ting Miss letter write Asp Prt

b′. *Ting sio-tsia* # [[*kuah-kin*]$_{AvP}$ = *to-k'i*]$_{VP}$
Ting Miss hurriedly return

On this construal, the crucial difference between (32a) and (32b) is not between an argument and an adjunct, but between an XP within or without the maximal projection of the verb. Since the object NP *p'ue* "letter" is topicalized out of the VP, it is bounded on the right by a p-phrase break; on the other hand, the AvP *kuã-kin* "hurriedly" lies within the domain of the VP, hence forms a single p-phrase with the head verb. On this view, what determines p-phrasing is not functional relation, but structural configuration. If this is true, then the end-setting for p-phrasing could be reduced to simple structure, and we pare away the ungainly appendage "X^{max} not an adjunct." This is exactly the position taken by R. Chung (1989). Chung proposes to replace Chen's (1987a) Tone Group Formation rule (= (19)) with what he calls the K-condition, stated as follows:

(33) K-condition
For a rule to apply to a sequence of two words *a* and *b*
(i) *a* must domain-c-command *b* or
(ii) *b* must domain-c-command *a* (Chung 1989:194)

He adopts Kaisse's definition of domain c-command, reproduced here as (34), which is equivalent to Chomsky's (1986) definition of m-command, paraphrased here as (35).

(34) Domain c-command
In the structure [X^{max} . . . x . . .], X^{max} is defined as the domain of x. Then x c-commands any Y in its domain. (Kaisse 1985:159)

(35) M-command
A m-commands B iff A does not dominate B and every maximal projection C that dominates A dominates B. (Chomsky 1986:8)

K-condition says in effect, TS applies between two terms, *a* and *b*, so long as they stand in a head–XP relation, where XP is neutral between argument and adjunct.

M-command (or domain c-command) is a fundamental concept in GB theory; moreover, m-command has been claimed to constitute the crucial grammatical property that defines the prosodic domain of phonological rules in a variety of languages (cf. Kaisse 1985). **French** liaison is a classic example: it is restricted to between a head and its complement, namely an XP within the former's domain of m-command. Contrast (36a) and (36b), taken from Selkirk (1974):

(36) a. *Nous* [*nous heurtons = à des difficultés imprévues*]$_{VP}$
"We're running into unforeseen difficulties."

b. *Elle* [*enrageait*]$_{VP}$ # *à cause de leur insolence*
"She got angry on account of their insolence."

Liaison applies in (36a) at the site signaled by "=," as the subcategorized complement PP *à des difficultés imprévues* "into unforeseen difficulties" lies within the VP, but blocks in (36b) at the juncture marked by "#," since the adjunct PP *à cause de leur insolence* "on account of their insolence" lies outside of the maximal projection of the intransitive verb *enrageait* "got angry." To cite just one more example, there is a rule in **Kimatuumbi** that shortens the stem vowels of a lexical head when it is followed indifferently by either an argument or an adjunct – within the maximal projection of the head (Odden 1987). This accounts for the short vowels in the (a) sentences of (37) and (38) and the long vowels in the (b) examples:

(37) a. *naampei* [*kikolombe kikulu*]$_{NP}$
I-him-gave shell large
"I gave him a large shell"

b. *naampei* [*kikoloombe*]$_{NP}$ *kikulu*
I-him-gave shell large-one
"I gave the large (thing) a shell"

(38) a. [*naan-kalangile yoopata eela*]$_{VP}$
I-for-him-fried get money
"I fried for him to get money"

b. *naayuwine* [*aakalaanga*]$_{VP/S'}$ *liiso*
I-heard he'll-fry yesterday
"Yesterday I heard that he will fry"

Note that the AP *kikulu* "large" in (37a) is a nominal adjunct, within the m-command domain of *kikolombe* "shell'; together they constitute a sandhi domain. On the other hand, the same AP functions like a headless adnominal, outside of the maximal projection of the object NP *kikoloombe* in (37b); the two constituents, therefore, belong to separate sandhi domains. Consequently, vowel shortening fails in the latter case. Likewise, the purposive clause *yoopata eela* "to get money" is bracketed within the VP headed by the verb *kalangile* "fried" in (38a), while the adverb *liiso* "yesterday" modifies the matrix verb *naayuwine* "I heard" and is therefore outside the scope of the embedded clause *aakalaanga* "he'll fry" in (38b). This structural difference is reflected in their sandhi behavior. There is, therefore, a priori plausibility that it is m-command and not the argument/adjunct distinction that determines the scope of sandhi processes in

Xiamen. For this reason it is worthwhile to critically re-evaluate these two alternative analyses of Xiamen tone sandhi.

In particular, let us focus on the crucial VP-internal XPs. Specifically, we will examine two types of VP-internal XPs: preverbal arguments (section 3.1) and postverbal adjuncts (section 3.2). I will show that the relevant facts are incompatible with an analysis based on the notion of m-command.

3.1 Preverbal arguments

To put the m-command hypothesis to the test, we need a configuration $[\ldots \mathrm{XP} \ldots]_{\mathrm{YP}}$, where XP demonstrably lies within YP. The sandhi behavior of adjunct XPs is neutral between the functional and the m-command hypotheses, since both theories predict that an adjunct XP constitutes a sandhi site with a non-null element to its right within YP. What is crucial, therefore, is whether or not demonstrably YP-internal argument XPs block tone sandhi: they should, consistent with the functional hypothesis; they should not, according to the m-command account. Let us use VP-internal arguments as the test case. Since the status of preposed object NPs like *p'ue* "letter" in (32a) of the preceding section is unclear, we need to establish more clearly the VP-internal status of other arguments. In addition to bare preverbal NPs that function as objects (under focus), we have PPs or, if you will, NPs overtly carrying a direct or indirect object marker (OM) like *tsiong* and *ka* (corresponding to Mandarin *ba* and *gei* respectively). Unlike bare object NPs, these PPs or case marked NPs cannot occupy the canonical topic, namely sentence-initial, position, witness (39) and (40).[7] (*T* in the examples that follow stands for *Ting sio-tsia* "Miss Ting.")

(39) a. *T* [*tsiong p'ue-k'aq # t'iaq-k'i*]
OM envelope tear
"Miss Ting tore up the envelope"

b. *T* # [*ka yin sio-ti* # *kong*]
OM her younger-brother tell
"Miss Ting told her younger brother"

(40) a. **tsiong p'ue-k'aq # T # [t'iaq-k'i*]
(cf. 39a)

[7] In the examples to follow, [. . .] marks off the constituent VP, and *T* stands for *Ting sio-tsia* "Miss Ting."

b. **ka yin sio-ti # T # [kong]*
(cf. 39b)

[. . .] = VP
OM = object marker

Furthermore, whereas topics or preposed NPs under focus can precede sentential adverbs (ad-S), case-marked NPs cannot. This can be seen in the following examples:

(41) a. *T # p'ue # tsa-lit # [sia liao lɔ]*
letter yesterday write Asp Prt
"As for the letter, Miss Ting wrote it yesterday"

b. *T # ts'iu-sik # kai-tsai # [bo tua]*
jewelry fortunately not wear
"Fortunately, Miss Ting did not wear her jewelry"

(42) a. **T # tsiong p'ue # tsa-lit # [sia liao lo]* (cf. 41a)
OM

b. **T # tsiong ts'iu-sik # kai-tsai # [bo tua]* (cf. 41b)
OM

In addition, whereas focused NPs can precede a modal and the negative particle, PPs or overtly case-marked NPs cannot, as shown in these examples:

(43) *T # tsit king ts'u # [bue tso-ts'ut-k'i]*
this Cl house not-will rent-out
"Miss Ting will/will not rent out this house"

(44) * *T # tsiong tsit king ts'u # [bue tso-ts'ut-k'i]*
OM

The ungrammaticality of (40), (42), and (44) furnishes us with negative evidence showing that PPs or overtly case-marked NPs do not behave like exo-VP topics or focused (bare) NPs. Crucially, there is positive evidence that these PPs or case-marked NPs are VP-internal. One diagnostic of their endo-VP constituency is the position of these PPs relative to VP-adverbs (ad-V), by definition an element of VP. In each of the examples given in (45) the PP or case-marked object NP occurs between an ad-V and the head V and, therefore, within the bracketed constituent VP.

(45) a. *T # [t'iao-kang = tsiong p'ue-k'aq #t'iaq-k'i]*
purposefully OM envelope tear
"Miss Ting purposefully tore up the envelope"

b. *T # [luan-tsu = ka yin sio-ti # kong]*
mindlessly OM her brother tell
"Miss Ting mindlessly told her younger brother"

c. *T # [kuah-kin = t'ao # a-bo so # yan-tsi # a-bo buaq # to tsau-ts'ut-k'i]*
hurriedly head neither comb rouge nor put-on Prt run-out
"Miss Ting ran out in a hurry, without combing her hair or putting on any make-up"

(45c) is particularly noteworthy in that not only PPs or overtly case-marked NPs can occur in preverbal VP-internal position, but so can bare object NPs like *tao* "head" and *yan-tsi* "rouge."

Given the endo-VP status of both bare and case-marked NPs, K-condition and p-phrasing based upon the notion of m-command make falsifiable predictions. K-condition predicts that TS applies between all such preverbal NPs and the verbal head, since the former lie within the VP and therefore come under the m-command of the latter. The prediction is counterfactual. Conversely, K-condition also predicts that TS should *not* apply between the ad-Vs like *t'iao-kang* "purposefully" and the object NPs like *tsiong p'ue-k'aq* "envelope," since both of them are XPs and outside of each other's domain. The prediction is also falsified by the attested facts.

An account in terms of the argument/adjunct distinction accommodates these facts without difficulty. The object NPs, whether case-marked or not, block TS, since they are arguments. On the other hand, TS applies between the ad-Vs and whatever follows, because the ad-Vs function as adjuncts. In other words, (45c) would have a structure like (46). In as much as the AP *kuah-kin* "hurriedly" c-commands the head V *so* "to comb," it forms a single tone group with *t'ao* "head."

(46) . . . [*kuã-kin* = [*t'ao a-bo so* . . .]$_{V'}$ . . .]$_{VP}$
hurriedly head neither comb . . .

In Chung's account, TS is limited essentially to two words standing in a head/complement relation (word order irrelevant; complement = either argument or adjunct), satisfying the Sense Unit Condition (Selkirk 1984). This account cannot be right in view of the fact that TS applies between an adverb ("hurriedly") and an NP ("head") which bear no direct structural or semantic relation to each other.

Pre-verbal object NPs are not the only kind of demonstrably VP-internal arguments that block TS; other argument structures exhibit the same sandhi behavior, in particular agent phrases in passive constructions,

directional complements, and focused NPs, as illustrated in (47a, b, c) respectively.

(47) a. *T # [hɔ kao # kiã-tioq]*
Pass dog scare/startle
"Miss Ting was startled by the dog"

b. *T # [tui oq-teng # to-k'i]*
from school return
"Miss Ting returned from school"

c. *T # [liam ts'iu-sik # ke t'eng hɔ ts'at-a]*
even jewelry also remove give burglar
"Miss Ting removed even her jewelry and gave it to the burglar"

We can apply the same diagnostics to test the endo-VP status of these phrases. For instance, they do not occur in sentence-initial positions (48), or before a sentential adverb (49), but can be flanked by an ad-V and the head V (50).

(48) a. * *hɔ kao* # T # [*kiã-tioq*]
(cf. 47a)

b. * *tui oq-teng # T # [to-k'i]*
(cf. 47b)

c. * *liam ts'iu-sik # T # [ke t'eng hɔ ts'at-a]*
(cf. 47c)

(49) a. * *T # hɔ kao # kai-tsai # [bo kiã-tioq]*
Pass dog fortunately Neg startle
"Fortunately, Miss Ting wasn't startled by the dog"

b. * *T # tui oq-teng # tsa-lit # [to-k'i]*
from school yesterday return
"Miss Ting returned from school yesterday"

c. * *T # liam ts'iu-sik # kai-tsai # [ke t'eng hɔ ts'at-a]*
even jewelry fortunately also remove give burglar
"Fortunately, Miss Ting removed even her jewelry and gave it to the burglar"

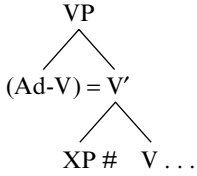
(50) a. *T # [ts'a-put-to = hɔ kao # kiã-si]*
almost Pass dog scare-to-death
"Miss Ting was almost scared to death by the dog"

b. *T # [kuã-kin = tui oq-teng # to-k'i]*
hurriedly from school return
"Miss Ting hurriedly returned from school"

c. *T # [kuã-kin = liam ts'iu-sik # ke t'eng hɔ ts'at-a*
hurriedly even jewelry also remove give burglar
"Miss Ting hurriedly removed even her jewelry and gave it to the burglar"

In summary, the examples cited in this section share the following structure:

(51)



Using the ad-Vs as a diagnostic marking the beginning of the VP, we show a number of different phrases occupying the VP-internal position of XP. This XP can be an object NP (bare or case-marked or, equivalently, PP), an agent of passive constructions, a focused NP, or a PP indicating a directional complement, etc. M-command predicts that (i) in all such cases TS would apply across XP and the head V, and (ii) TS would block between ad-V and XP, since the XP does not m-command the ad-V. The prediction is wrong on both counts. On the other hand, a function-sensitive p-phrasing as reflected in Tone Group Formation, predicts that (i) TS would apply between ad-V and XP, since ad-V is an adjunct c-commanding V, and (ii) TS would fail between XP and V, since XP is an argument of V – exactly as borne out by the facts.

3.2 Postverbal adjuncts

Needless to say, both arguments and adjuncts can also occur in postverbal positions. However, the phrase-structure constraints of Chinese severely limit what can follow the verb. Basically there are five possibilities:

(52) i. [. . . V = XP]
ii. [. . . V = NP # NP]
iii. [. . . V = NP # PP/S′]
iv. [. . . V = AvP = NP]
v. [. . . V = AvP = S′]

[. . .] = VP

(i) is neutral between K-condition and a functional account based on the argument–adjunct dichotomy: both predict the application of TS across V and XP. (ii) represents a double object construction. Both K-condition and the functional account predict the failure of TS between the two NPs, though for different reasons: according to the functional hypothesis,

because both NPs are arguments; in terms of K-condition, because neither NP m-commands the other, since they constitute a barrier to m-command. Relevant examples are given below.

(53) a. [*sang* = *tang-oq* # *tsit* *pun ts'eq*]
give schoolmate one Cl book
"give schoolmate a book"

b. [*sang* = *tsit* *pun ts'eq* # *hɔ tang-oq*]
give one Cl book OM schoolmate
(same)

Case (iii) differs from those of (53) only in that the second complement is a PP or a subordinate clause (S′), as illustrated in (54).

(54) a. [*sang* *lang-k'e* # *kau mng-k'ao*]
accompany guest to/arrive door
"see guests to the door"

b. [*kia tsĩ* # *k'i bi-kok*]
send money to U.S.
"send money to the U.S."

Again, m-command and the argument/adjunct distinction make the same prediction: TS should block between NP and PP/S′, for the same reasons as those given for (b).

There remain cases (iv) and (v) that discriminate between the two hypotheses under scrutiny. Case (iv) is instantiated by (55).

(55) a. [*k'uã* [*tsit pai*]$_{\text{AvP}}$ = *dian-yã*]
watch one time movie
"watch the movie once"

b. [*t'ak* [*puã tiam-tsing ku*]$_{\text{AvP}}$ = *ts'eq*]
read half hour long book
"read for half an hour"

The duration and frequency phrases (AvP) between the verb and the object NP function as adjuncts, and therefore not surprisingly trigger TS. However, on the alternative m-command account, (55a) and (55b) are problematic. Here the two terms (*a* and *b*) of the K-condition are the AvP and the object NP. Neither one m-commands the other, though both are m-commanded by the verb. K-condition fails; therefore TS should block, contrary to the attested facts.

Case (v), exemplified in (56), differs from (iv) only in that the second term *b* is a complement clause (S′) rather than an NP.

(56) a. [*ts'iũ* [*neng pai*]$_{AvP}$ = [*hɔ sian-sĩ # t'iã*]$_{S/VP}$]
sing two times for teacher hear
"sing twice for the teacher to hear"

b. [*tsau* [*k'aq kin*]$_{AvP}$ = [*hɔ gua k'uã*]$_{S'/VP}$]
run more fast for me see
"run faster for me to see"

c. [*liong* [*k'aq tua-siã*]$_{AvP}$ = [*ka yi kã-si*]$_{S'/VP}$]
scream more loud OM him startle
"scream louder to startle him"

It seems that the notion of m-command or domain c-command is not viable as a substitute for the argument/adjunct distinction.

4 Lexical government

We now turn to a second alternative that holds far greater promise for a unified and explanatory account based on the notion of lexical government advanced by Hale and Selkirk (1987) and J. Lin (1994). Like the m-command account, the lexical government analysis seeks to base the algorithm of Xiamen p-phrasing on a purely configurational property, rather than on an awkward mix of structural (X^{max}) and functional (argument vs. adjunct) conditions embodied in Chen's (1987a) Tone Group Formation. In a sense, lexical government exploits the same configurational properties of m-command, but superimposes on it the further stipulation that the m-commander be a lexical, rather than a functional, head.

4.1 Sentential vs. VP adverbs

The strongest argument for the lexical government approach to Xiamen tone sandhi comes from the differential behavior exhibited by sentential vs. VP adverbs, a problem that remained unresolved in Chen's (1987a) original account. Let us, therefore, begin our discussion with a word on this distinction.

Recall that in section 3.1 we used the VP-adverbs (ad-V) to locate the left boundary of VP, so that an XP that is wedged between an ad-V and the verb can be ascertained to lie within the VP. We did not provide a characterization of what an ad-V may be or, in particular, how to distinguish it from a sentential adverb (ad-S). Here we take up this question. The syntactic diagnostic of ad-V-hood vs. ad-S-hood rests principally on word order. First, only ad-S can occur in pre-subject, sentence-initial position, as in (57); ad-V cannot, witness (58).

(57) *tsa-lit/kai-tsai/k'o-ling # T # tse tsit pan ki*
yesterday/fortunately/maybe take this Cl flight
"Yesterday/fortunately/maybe Miss Ting took this flight"

(58) a. **t'iao-kang (#) T # tsiong p'ue-k'aq # t'iaq-ki*
purposefully DO envelope tear
"Miss Ting purposefully tore up the envelope"

b. **luan-tsu (#) T # ka yin sio-ti # kong*
mindlessly[8] IO her brother tell
"Miss Ting mindlessly told her brother"

c. **kuã-kin (#) T # ts'ut-k'i*
hurriedly exit
"Miss Ting left in a hurry"

d. **ts'a-put-to (#) T # hɔ kau # kiã-si*
almost by dog scare-die

"Miss Ting was almost scared to death by the dog"
DO, IO = direct/indirect object marker

Secondly, PPs or overtly case-marked object NPs can precede ad-V but not ad-S. Compare (59) and (60). Note that while *ka yin sio-ti* "to her brother" can precede or follow an ad-V *luan-tsu* "mindlessly" (with different scopal interpretations), it cannot occur in front of sentential adverbs such as *kai-tsai/tsa-lit/k'o-ling* "fortunately/yesterday/maybe" etc.

(59) a. *T # luan-tsu = ka yin sio-ti # kong*
mindlessly IO her brother tell
"T mindlessly told her brother"
(careless in choosing who to confide in)

b. *T # ka yin sio-ti # luan-tsu = kong*
(same)
(careless in telling her brother what she ought to be discrete about)

(60) a. *T # kai-tsai/tsa-lit/k'o-ling # ka yin sio-ti # kong*
fortunately/yesterday/maybe IO her brother tell
"Fortunately/yesterday/maybe Miss Ting told her brother"

b. **T # ka yin sio-ti # kai-tsai/luan-tsu/k'o-ling # kong*
(same)

Likewise, the negative particle *bo* can occur before or after the ad-V, but cannot precede the ad-S, as shown in (61) and (62).

[8] Or "unthinkingly."

(61) a. *yi kai-tsai / tsa-lit # bo tse tsit-pan ki*
he fortunately/yesterday Neg take this Cl flight
"fortunately/yesterday he didn't take this flight"

b. **yi bo kai-tsai / tsa-lit # tse tsit pan ki*
(same)

(62) a. *yi kuã-kin = bo ts'ut-siã*
he hurriedly Neg speak:up
"he quickly shut up"

b. *yi bo kuã-kin = ts'ut-siã*
he Neg hurriedly speak:up
"he did not speak up right away"

Remarkably, to this syntactic dichotomy between sentential and VP adverbs corresponds a phonological distinction in terms of their sandhi behavior: in all the examples cited above, all the adverbials diagnosed by the word-order tests as sentential in scope block TS, while those which are identified as ad-Vs by the same tests, allow TS to apply between them and whatever follows. This contrast raises an interesting issue for the adjunct-based account of TS in Xiamen. Since ad-Vs and ad-Ss are both adjuncts, one would expect them to exhibit identical TS behavior; but they don't.[9] This phonological contrast is quite robust, as illustrated by the following:[10]

(63) a. *Ting sio-tsia # kai-tsai # [tse tsit pan ki]*
Ting Miss fortunately take this Cl flight
"Fortunately, Miss Ting took this flight"

b. *yi tong-lian # [tiã yin bo # e we]*
he of-course listen his wife E talk
"Of course, he listens to his wife"

c. *li te-it-tsio # [ma tioq pue-siong]*
you at-least also must pay-for-damage
"At least you must pay for the damage"

[9] This was duly pointed out in footnote 17 of Chen (1987a:146), which reads in part as follows:

> There remains an unresolved problem in our treatment of ad-S. Given our definition of adjuncts as XP contained within another YP, namely $[\ldots XP \ldots]_{YP}$, and under the assumption that S is the maximal projection of Aux or INFL, ad-S fits our definition of adjuncthood and should not therefore block TS, since it lies within S (= $INFL^{max}$). In an earlier version (1985) of this paper I speculated about alternative solutions, which are best left out as premature for the moment.

[10] Examples of (63) and (64) are repeated from Chen (1987a:124f.) with slight modifications. For clarity I added [. . .] to mark the beginning and end of a VP, or PrP.

(64) a. *tsit-e gin-a* # [*k'un-lat* = *tak-ts'eq*]
this boy diligent study
"this boy studies hard"

b. *yi* [*luan-tsu* = *kong*]
he mindlessly talk
"he talks mindlessly"

c. *tsĩ* # *tioq* [*k'iam-ki'am-a* = *ying*]
money must sparingly use
"one must use money wisely"

d. [*tsa* = *k'i-lai*]
early rise-up
"get up early"

The underscored adverbials of (63) are sentential in scope; for instance, the adverb in (63a) can be paraphrased by a predicate taking a sentential argument, so that (63a) reads as something like "It is fortunate that Miss Ting took this flight" (rather than another flight that crashed, was canceled etc.). Syntactically speaking, these adverbials can be preposed to the sentence-initial position. Thus, (63a) exhibits the perfectly acceptable alternative word order: *kai-tsai* # *Ting sio-tsia* # *tse tsit pan ki*. In each case of (63) tone sandhi blocks. In contrast, the examples of (64) contain only ad-Vs, witness the ungrammaticality of (64a) resulting from moving the ad-V to the beginning of the sentence: **k'un-lat tsit-e gin-a tak-ts'eq*. In each of these examples tone sandhi applies freely between the ad-V and the V.

Noting that whereas ad-Vs are licensed (i.e. headed) by lexical categories (N, V, and A), ad-Ss are licensed by functional categories I(nfl) and C(omp), Chen (1992c:19) speculates that "one could perhaps amend TG (Tone Group) Formation by requiring that the XP be an adjunct phrase of a *lexical* head, thus excluding Infl and Comp as heads."[11] This alternative approach to Xiamen tone sandhi based on the notion of *lexical government* had been foreshadowed in Hale and Selkirk (1987) and Selkirk and Shen (1990), but it is J. Lin (1994) who argues most persuasively

[11] In the same vein, H-M. Zhang (1992:67) proposed the following revision of Xiamen TG Formation: "Mark the right edge of every XP with #, except where XP is an adjunct m-commanding either its head or the head of XP on the right *except Infl*" (emphasis added).

for the position. Lin proposes to revise the Xiamen p-phrasing (31) as follows:

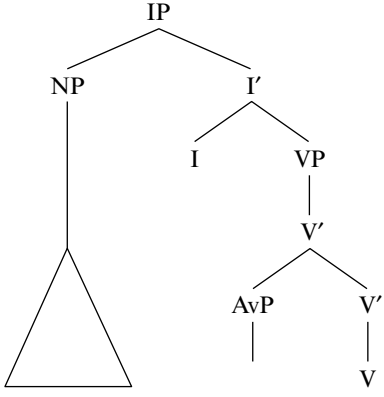
(65) Xiamen p-phrase:
{Right, X^{max}}, X^{max} not lexically governed

It specifies that the designated category X^{max} not be lexically governed. In accordance with Chomsky (1986; cf. Hale and Selkirk 1987) Government is defined as follows:[12]

(66) Government
A governs B iff A m-commands B and every barrier for B dominates A.

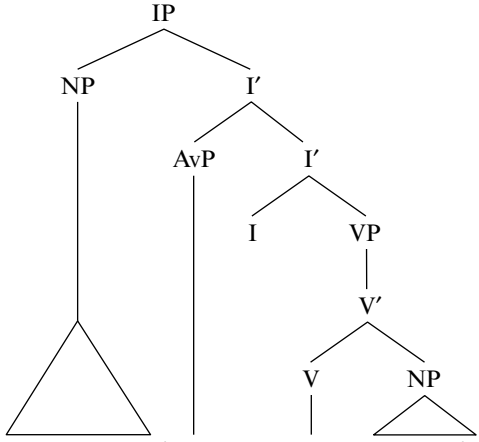
It follows immediately therefrom that whereas ad-Ss, which are governed by the *functional* head Infl, are bounded on the right by a p-phrase break, ad-Vs, which are governed by the *lexical* head V, are not. Under the lexical government hypothesis, these sandhi facts follow straightforwardly from the standard assumption that ad-Ss are adjuncts of the functional category IP, whereas ad-Vs are adjuncts within the lexical projection VP/PrP. This structural difference is brought out more clearly in (67). Note that the AvP is adjoined to V′ in (67a), but I′ in (67b).

(67)



a. *tsit-e gin-a* # *k'un-lat* = *t'ak-ts'eq*
this child diligent study
"This child studies hard."

[12] We have already encountered m-command in section 3; barriers are defined as in Chomsky (ibid.).



b. *Ting sio-tsia # kai-tsai # tse tsit pan ki*
Ting Miss fortunately take this Cl flight
"fortunately, Miss Ting took this flight"

4.2 NP as DP

The lexical government account of Xiamen tone sandhi rests on another crucial assumption, namely that NPs are actually determiner phrases, or DPs. Recall that in sections 3.1–3.2 we took pains to show that certain demonstrably VP-internal XPs (underlined) do block tone sandhi. Some of the crucial examples are repeated below for convenience.

(68) a. *T # [liam ts'iu-sik # ke t'eng hɔ ts'at-a]*$_{VP}$
even jewelry also remove give burglar
"Miss Ting removed even her jewelry and gave it to the burglar"

b. [*sang tang-oq # tsit pun ts'eq*]$_{VP}$
give schoolmate one Cl book
"give schoolmate a book"

c. [*kia tsĩ # k'i bi-kok*]$_{VP}$
send money to U.S.
"send money to the U.S."

Example (68a), with a preverbal object NP, instantiates a structure like (69a); both double object constructions like (68b) and postverbal arguments followed by a non-null element (Y) like (68c) share the configuration of (69b). In each case, the underlined VP-internal XP blocks tone sandhi.

(69) a. [. . . XP V . . .]$_{VP}$
b. [. . . V XP Y]$_{VP}$

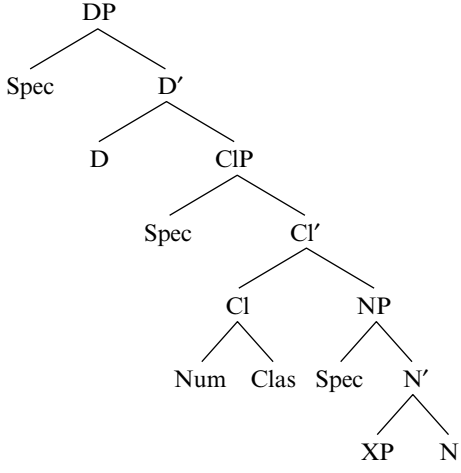
Contrast these with those of (70a) and (70b), which have *exactly* the identical structures of (69a) and (69b) respectively. But, crucially, tone sandhi *must* apply between the underlined XP of (70a, b) and the non-null string (V, Y) that follows.

(70) a. *T* # [*kuã-kin* = *to-k'i*]$_{VP}$
hurriedly return
"Miss Ting hurriedly went home"

b. [*k'uã tsit pai* = *tian-yã*]$_{VP}$
watch one time movie
"watch a movie once"

More than anything else, it was this contrast in the sandhi behavior of VP-internal XPs that forced Chen (1987a, 1992c) to adopt a functional analysis keyed to the argument–adjunct dichotomy.[13] This conclusion is no longer inescapable in view of recent developments in extended X′-theory. Following Abney (1987) and Fukui (1986), Tang (1990) proposes that the internal structure of NPs in Chinese should be fleshed out more fully in extended X′-theory as (71).

(71)



Cl and Q, and correspondingly Cl′ and Q′, ClP, and QP are used interchangeably in Lin (1994). Setting aside the optional intermediate category ClP (= QP), NPs are governed by a functional head, the determiner, if nothing else. On this account, the XPs of (68a, b, c) are NPs enclosed

[13] Unless, of course, one resorts to the NP vs. AvP distinction, that would thwart the attempt at a cross-categorial characterization of prosodic constituents.

within DPs, albeit without overt determiners. Thus, the indirect object NP *tang-oq* "schoolmate" of (68b) terminates a p-phrase by virtue of its being governed not by the verb *sang* "give," but by the empty (non-lexicalized) functional head, the determiner. In other words, (68b) is bracketed as (72).

(72) [*sang* [[*tang-oq*]$_{NP}$]$_{DP}$ # *tsit pun ts'eq*]$_{VP}$
give schoolmate one Cl book

There tone sandhi is blocked by the right edge not of DP (which is lexically governed by the verb *sang*), but of NP, which is functionally governed by the phonetically null determiner head. This is what Lin (1994:30) refers to as "string-vacuous effect." By contrast, the XPs of (70a, b) are lexically governed, by the verbs, and therefore fail to form a separate p-phrase, consistent with the exclusionary clause on Xiamen p-phrasing: "X^{max} not lexically governed." Hence, TS applies. This NP-as-DP analysis in extended X'-theory makes it possible to reinterpret the functional distinction between arguments and adjuncts in more parsimonious, purely structural or configurational terms.

The lexical government account proposed by Lin (1994) finds a dramatic confirmation in the following pair of sentences.[14]

(73) a. *m-t'ang tsiaq* [*hit liap* [*ts'ĩ-sik*]$_{AP}$ = *p'iang-ko*]
don't eat that Cl green apple
"Don't eat that green apple"

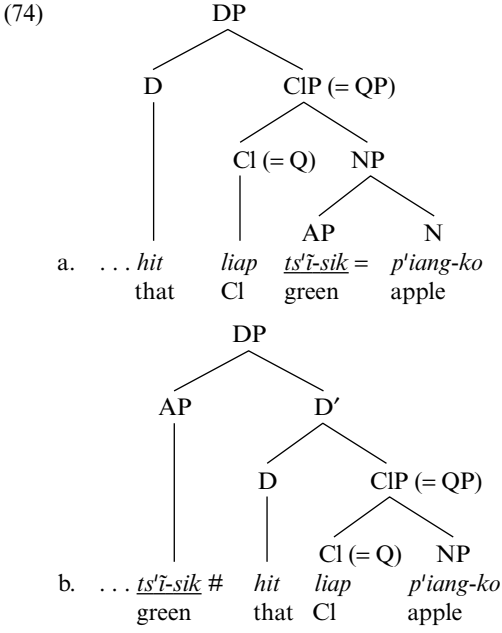
b. *m-t'ang tsiaq* [[*ts'ĩ-sik*]$_{AP}$ # *hit liap p'iang-ko*]
(same)

Note that whether or not the same AP *ts'ĩ-sik* "green" joins the constituent on the right in a single p-phrase depends on its word order relative to *tsit liap* "this + Cl." Since this AP functions as an adjunct within the maximal projection of *p'iang-ko* "apple" regardless of word order, a functional account offers no plausible explanation for this contrast – which had gone unnoticed in the literature until pointed out by Lin. By assuming a structure like (71), the contrast of (73a, b) falls out straightforwardly from lexical government. The more fully articulated tree configurations of (74a, b), corresponding to (73a, b), make it clear that the AP in (74a) is

[14] Examples (73a, b) correspond to Lin (1994, exs. (37a, b)). Transcription has been modified slightly to neutralize minor inconsistencies and subdialectal differences.

an ad-N, governed by the lexical head *p'iang-ko* "apple," whereas the same AP in (74b) is a DP-level adjunct, and hence governed by the functional category *hit* "that," the determiner head. In this respect, as in the case of sentence vs. VP adverbials, a government-theoretic approach succeeds admirably where a functional account has failed.

(74)



4.3 *Functional relations reinterpreted*

The notion of lexical government can be extended to other facts once thought to hinge on the argument/adjunct distinction. Recall the classic minimal pair discussed earlier, repeated and relabeled below:

(75) a. $[[[\underline{mua\text{-}a}]_{DP}$ # $tua]_{IP}$ # $e]_{CP}$ *sio-p'iã*
sesame-seed big Comp bun
"buns (on which) the sesame-seeds are small"

b. $[[[\underline{mua\text{-}a}]_{NP} = tua]_{AP}$ # $e]_{CP}$ *sio-p'iã*
sesame-seed big Comp bun
"buns as big as sesame-seed" (tiny buns)

The subject *mua-a in* (75a) is bounded on the right by # both because it is an NP governed by a functional category D, and because the DP is

governed by Infl.[15] On the other hand, the degree modifier *mua-a* in (75b) is non-referential in nature, and is therefore labeled as either a zero-level N or possibly an NP, but not a DP. As bare N or NP, *mua-a* is governed by the adjectival head *tua*, and therefore fails the "X^{max} not lexically governed" clause on p-phrasing.

As for objects of PPs, for instance in (76a), the nominal expression *tua-mng* "gate" is a referential DP. Not surprisingly it forms a separate p-phrase, since *tua-mng* is governed by a phonetically null determiner.[16] XPs other than NP/DP that function as either topic, subject, or predicate phrases automatically induce a tone-sandhi break, since they are constituents of IP, governed by Infl. Thus the topic NP/DP *ti-baq* "pork" and the subject QP *tsit kun* "one catty" of (76b) are marked off by #, since they are constituents of CP and IP respectively, therefore not lexically governed. Similarly, the VP *peq lu kuãi* "climb higher" is a predicate phrase in (76c), and is governed by the functional category Infl.

(76) a. [*tua-mng*]$_{DP}$ # *ao-piaq*
gate behind
"behind the gate"

b. [*ti-baq*]$_{NP/DP}$ # [*tsit kun*]$_{CIP}$ # *sã k'ɔ*
pork one catty three dollars
"pork is three dollars a catty"

c. [*peq lu kuãi*]$_{VP}$ # *puaq lu t'iam*
climb more high fall more hard
"the higher you climb, the harder you fall"

As a bonus, a lexical government account also rids TG Formation (section 1.3) of another stipulation. As originally stated, TG Formation says to the effect that the right edge of an XP creates a tone-sandhi break – "except where XP is an adjunct *c-commanding its head* [emphasis added]." The c-command condition was motivated by facts such as the following:[17]

[15] Although in this case neither one of the functional heads, D and Infl, is phonetically realized.

[16] In any case, the DP/NP is in turn governed by *ao-piaq,* a "postposition" meaning "behind," or perhaps a locative expression translatable as "the back(side) of," presumably a non-lexical head.

[17] (77a, b) and (78a, b) = examples (49a, b) and (52a, b) in Chen (1987a) respectively.

(77) a. [*pin-tuã*]$_{AP}$ = *gin-a*
lazy boy
"lazy boy"

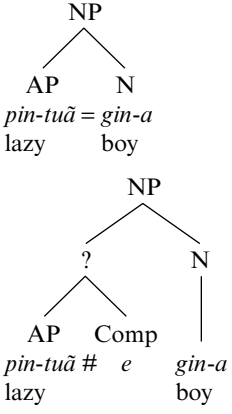
b. [[*pin-tuã*]$_{AP}$ # *e*] *gin-a*
lazy Comp boy
"lazy boy"

(78) a. [*ts'ong-bing*]$_{AP}$ # [*k'un-lat*]$_{AP}$ # *e gin-a*
smart hard-working Comp boy
"a smart and hard-working boy"

b. . . . *hun-lian kaq* [*ko ang*]$_{AP}$ # [*ko tsuan*]$_{AP}$
train Comp both red and expert
"train . . . to be both 'red'[18] and expert"

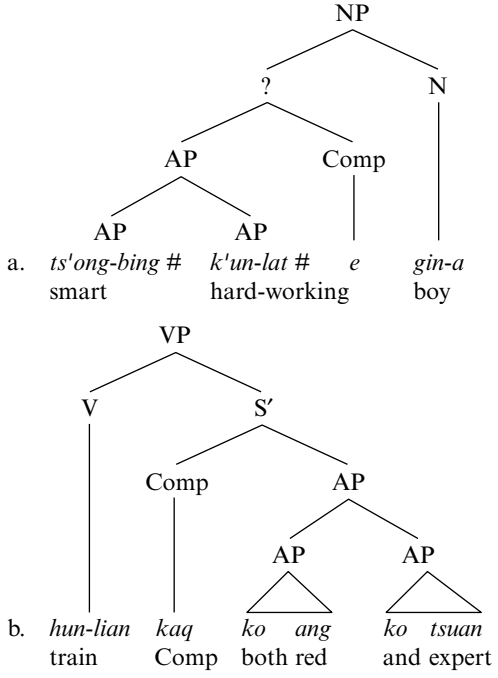
A bare AP modifier joins the nominal head in a single sandhi domain, as expected (77a); surprisingly, however, an AP linked to a nominal head via the grammatical particle *e*, variously characterized as a subordinator, complementizer, or relative-clause marker etc., is bounded by the p-phrase boundary # (77b). Equally unexpected is the fact that coordinate APs of (78a, b) also induce a sandhi break, even though in each case the AP functions as a nominal modifier, or a clausal adjunct indicating the degree or extent of the verb. Chen's (1987a:132f.) solution was to add on a condition based on c-command. Assuming (79a, b) and (80a, b) to be the corresponding tree representations of (77a, b) and (78a, b) respectively, c-command condition is met only in (79a), and fails elsewhere.

(79)



[18] Recall that "red" means "ideologically committed and motivated."

(80)



Viewed from the government-theoretic perspective, the c-command condition is rendered otiose by construing [AP (AP) + *e*] in (79b) and (80a) as a (reduced) relative clause, namely a functional projection CP of the functional head *e*, the Comp or relative clause marker. As for (80b), the APs are clearly predicate phrases contained within a resultative clause introduced by the complementizer *kaq*. In short, in all these cases the APs are either ungoverned (in coordinate constructions) or governed by the functional category Infl or Comp. Hence, they fall under the exclusionary clause of "X^{max} not lexically governed," and therefore are demarcated on the right edge by a p-phrase boundary.

4.4 Residual problems

The lexical government account rests crucially on the assumption that NPs are analyzable as DPs in extended X′-theory. Syntactic motivations aside, this NP-as-DP hypothesis is supported by independent phonological evidence, as pointed out in the preceding sections. However, there remain a number of unresolved issues. Consider the following set of examples.[19]

[19] Examples (81a, b, c) = examples (66a, b, c) of Chen (1987a:138).

(81) a. *tso* [*tsit ts'ut* = [*liok-yã-p'ĩ*]$_{NP}$]$_{ClP,DP}$ # *lai k'uã*
rent a/one Cl video-movie to watch
"rent a/one video-movie to watch"

b. *liok-yã-p'ĩ* # [*tsit ts'ut*]$_X$ # *tso lua-tsue tsĩ*
a/one Cl rent how-much money
"how much does a/one video-movie rent for?"

c. *liok-yã-p'ĩ* # [*tso* [*tsit ts'ut*]$_X$ = *lai k'uã*]$_{VP}$
rent a/one Cl to watch
"rent *a* video-movie to watch"

d. *liok-yã-p'ĩ* # [*tso* [*hit* [*ts'ut*]$_{ClP}$]$_{DP}$ (#) *lai k'uã*]$_{VP}$
that Cl
"rent *that* video-movie to watch"

e. *liok-yã-p'ĩ* # [*tso* [*tsit ts'ut*]$_X$ # *lai k'uã*]$_{VP}$
a/one Cl
"rent *one* video movie to watch"

Example (81a) is unproblematic. The NP is topicalized in (81b). The leftover X functions as the subject phrase, hence is functionally governed by Infl, regardless of how we label this X. The matter becomes a bit more complicated in (81c). In the traditional X′-theory, the constituent X in (81c) is presumably a headless argument NP. Therefore, one expects it to be bounded by # on the right, in the functional account. Chen (1987a:138) had a long and embarrassingly unconvincing story to tell about it, which does not bear repeating here. Lin's (1994) lexical government account is much more straightforward: X is a pure QP (= ClP), in the sense that the QP (i) has no internal functionally governed NP; and (ii) is not projected under DP, so that QP itself is not functionally governed by D, but directly governed by the lexical head *tso* "rent." Ergo, no tone-sandhi break. A potential problem arises when we look at (81d). Here we have an overt determiner *hit* "that" which governs the ClP (= QP) *ts'ut*. The functional government that holds between *hit* and *ts'ut* predicts a tone-sandhi break between *ts'ut* and the purposive clause *lai k'uã* "to watch." Contrary to this prediction, tone sandhi is possible, if not obligatory at this juncture.[20] We encounter a more serious problem with the contrasting pair (81c) and (81e). The numeral expressions *tsit* and *neng* can be taken literally to mean exactly "one" and "two," or interpreted loosely as meaning "a, any" and "a few, a couple of" respectively.[21] The two readings are disambiguated

[20] In my own speech, *ts'ut* blocks tone sandhi only in contrastive reading (rent *that* one rather than *this* one); otherwise, tone sandhi preferably applies in the neutral reading.

[21] This holds true not only in Xiamen, but across Chinese dialects generally, including Shanghai.

by means of tone sandhi, as indicated by the English translations of (81c, e). Lin's (1994) account of this contrast is as follows: real numbers (*one, two*) are quantifiers; existential quantifiers (*a, some, a couple of*) are determiners. Sentence (81e), therefore, annotated as (82), has an internal QP (= ClP) governed by the functional category D (though not overtly lexicalized). This accounts for the tone break in (82).

(82) real number
tso [[*tsit ts'ut*]$_{QP/ClP}$]$_{DP}$ # *lai k'uã*
rent one Cl to watch
"rent ***one*** (movie) to watch"

By the same logic, the existential quantifier reading (*a, some, a couple of*) of (81c) should be bracketed and labeled as (83). Here the number expression *tsit* occupies the D position. The QP or ClP consists of the classifier alone. This means that the QP is in fact governed by the overt determiner head *tsit* "a, some." Hence, we predict a break after the QP in (83) as well, contrary to the attested reading.

(83) existential quantifier
tso [*tsit* [*ts'ut*]$_{QP/ClP}$]$_{DP}$ = *lai k'uã*
rent Det Cl to watch
"rent ***a*** (movie) to watch"

A related problem stems from *referentiality* as a defining characteristic of DP. Recall the classic example of *[[mua-a]$_X$ = tua] e sio-piã* "buns as big as a sesame seed," where *mua-a* "sesame seed" functions as a degree modifier of the adjective *tua* "big." The constituent labeled here as X cannot be a DP – otherwise, the internal NP would be *functionally* governed by the phonetically null determiner, and therefore induce a tone-sandhi break. Instead, this X must be construed as a simple NP, directly governed by the lexical adjectival head *tua* "big." Lin (1994) justifies the labeling of X as NP by appealing to its non-referential interpretation. This seems perfectly reasonable. The trouble is that tone sandhi applies even under a patently referential reading, as in example (84).

(84) [[*hit* [*liap* [*mua-a*]$_{NP}$]$_{ClP}$]$_{DP}$ = *tua*]$_{AP}$ # *e sio-piã*
Det Cl sesame-seed big Comp bun
"buns as big as that sesame seed"

Here the degree modifier is the referential DP *hit liap mua-a* "that sesame seed." Therefore, the internal NP *mua-a* is functionally governed by Cl (the ClP/QP is in turn governed by D). We should expect a tone group

break here (double "string-vacuous effect"). Crucially tone sandhi applies here in the normal reading.[22]

Finally, let us compare Xiamen TG Formation with **Papago** tonal phrasing (Hale and Selkirk 1987), a case study that inspired Lin's (1994) very appealing analysis of Xiamen tone sandhi to begin with. The essential fact about Papago is this: XPs occurring between the Aux and the V link together into a single tonal phrase; preposed subjects and other extraposed NPs form independent tonal phrases. Each tonal phrase is characterized by the intonational pattern (L)HL. The distribution of (L)HL is illustrated by the following examples:[23]

(85) "Did the cowboy brand the calf?"
Inter-Aux art cowboy art calf brand
a. *Na-t* [[*g* *wakial*]$_{NP}$ [*g* *wisilo*]$_{NP}$ *cepos*]$_{VP}$
(L HHH HHH H L)

"Did the cowboy brand the calf?"
Inter-Aux art cowboy brand art calf
b. *Na-t* [[*g* *wakial*]$_{NP}$ *cepos*]$_{VP}$ [*g* *wisilo*]$_{NP}$
(L HHH H L) (HLL)

"The cowboy branded the calf"
cowboy Aux art calf brand
c. [*Wakial*]$_{NP}$ *'at* [[*g* *wisilo*]$_{NP}$ *cepos*]$_{VP}$
(H LL) (L HHH H L)

Details aside, Hale and Selkirk's (1987) account goes like this. Papago word order is basically head-final. Subject as well as object NPs originate as VP-internal arguments, and both are bracketed as VP-internal. In (85a), therefore, both the subject and the object NP are lexically governed by the head V *cepos* "brand." Conversely the object NP *g wisilo* "the calf" ends up in the marked, post-head position in (85b) via extraposition. Under this analysis, the object NP of (85b) lies outside of the VP projection, and

[22] An alternative pronunciation with tone sandhi break is permissible, but only under contrastive reading: roughly "buns the size of *that*, not this, sesame seed."
Parenthetically, one might construe (84) as (84′)
(84′) [*hit* [*liap* [*mua-a*$_{NP}$ = *tua*]$_{AP}$ # *e sio-piah*]$_{NP}$]]$_{CIP}$]$_{DP}$
"That bun (which is) as big as sesame seed"
where *hit* and *liap* are the determiner and the classifier for *sio-piah* instead of *mua-a*. Tone sandhi applies because aside from the lexically governed NP represented by *mua-a*, there is no other]$_{XP}$ bracket between *mua-a* and *tua*. This construal is ruled out by the selectional restrictions holding between the classifier and the head noun. *Liap* can only classify "small, rounded" objects like sand, grains, beads, etc.; a noun like *sio-piah* calls for a classifier like *te* "piece." Thus *tsit te sio-piah*, but **tsit liap sio-piah* "a bun."

[23] Examples (85a, b, c) = examples (7), (8), and (26) of Hale and Selkirk (1987).

is therefore ungoverned by the lexical verb. Likewise, the subject NP of (85c) occupies pre-Aux, sentence-initial position via movement (into Spec of IP), out from within the VP. In this case, the subject NP is governed by Infl, a non-lexical category. Exploiting this L-governed and non-L-governed distinction, Hale and Selkirk derive the tonal phrasing of Papago by setting the end parameter:

(86) Papago Phrasing Parameter (Hale-Selkirk 1987:164)
$]_X^{max}$, X^{max} not lexically governed.

which is exactly what Lin (1994) proposed for Xiamen.

Now, the NP-as-DP hypothesis introduces a wrinkle in the government-theoretic distinction on which turns Papago tonal phrasing. The NPs in (85) are projected under overt determiners or articles ("*g*" in Papago) and are presumably referential in meaning. They are, therefore, arguably DPs containing NPs. By virtue of (86), each of the DP-internal NPs in (85a) qualifies as a non-L-governed X^{max}, and therefore should be demarcated by a phonological phrase boundary on the right margin. This in effect obliterates the distinction between extraposed and VP-internal NPs – since neither is governed by the lexical verb, but instead by either the determiner or Infl, or else is ungoverned. Hale and Selkirk (1987) were acutely aware of this corollary of analyzing NP as DP. In their words,

> [T]he NP constituent within a DP . . . is not treated as a non-L-governed maximal projection by the Papago Phrasing Parameter (29) [= (86)], even though it is governed by a non-lexical category. This contrasts with the situation of a maximal projection in the specifier position of IP, i.e. in pre-AUX position. We believe that this difference is due to the fact that the Determiner–NP relationship has a special character, since in that construction the Determiner and the NP form a lexical-functional unit – i.e. an argument. (Hale and Selkirk 1987:170)

It is not clear why this "special character" or lexical–functional unity that exempts Papago D–NP constructions from (86) does not hold true for Xiamen as well. Whatever syntactic properties distinguish the Chinese D–NP construction from its Papago equivalent remain for now a mystery.

Despite these problems, the lexical government analysis retains considerable esthetic appeal. As originally formulated, Chen's (1987a) TG Formation (restated as Xiamen P-Phrase in section 1.2) combines a number of disparate and incongruous properties, some of which are structural, others functional in nature: categorial rank (X^{max}), right edge, adjuncthood,

and c-command. By comparison, the notion of lexical government relies solely on configurational properties, and rids TG Formation of the last two awkward stipulations (adjuncthood and c-command), thereby offering a simpler and far more elegant algorithm for deriving the prosodic domain of tone sandhi in Xiamen – assuming that further investigation will eventually furnish us with a principled way (i) to differentiate NPs from DPs, and (ii) to characterize the government-theoretic relation between the determiner and its nominal complement across language types.[24]

5 Rhythmic effect in Xiamen

It was stated at the outset of section 1.1 that minimality was totally irrelevant as a defining feature of the sandhi domain in Xiamen. Here I wish to qualify that statement. True, the p-phrase in Xiamen is irreducible to the MRU (minimal rhythmic unit) in Mandarin Chinese. However, that does not mean that rhythm plays no part in Xiamen. In fact, rhythmic effect is pervasive across Chinese dialects, even in those dialects, like Xiamen, which delimit their tone-sandhi domain by quite different principles. In ordinary speech, the relevant prosodic domain of tone sandhi is a p-phrase, as defined in the preceding sections. Accordingly, a sentence like (87) is cut up into chunks of irregular sizes, from one, to three, to seven syllables in length.[25]

(87) *niao*]$_{NP}$ # *ts'e-tioq tsit e wu-mo-kiu*]$_{NP}$ # *lai t'it-t'o*
cat find Perf one Cl shuttlecock to play
"the cat found a shuttlecock to play with"

On the face of it, therefore, there is no detectable effect of the rhythmic flow of speech characterized by the regular spacing of stressed and unstressed syllables. But even a language such as Xiamen, organized along quite different principles, exhibits distinctly rhythmic effects under certain special conditions: specifically, in verse, literary readings, and idiomatic expressions.

[24] Independently of Lin (1994), Bao (1995) made the following observation: In the configuration X^{max} Y,

a. if X^{max} is an argument, then Y is Y^{max}
b. if X^{max} is an adjunct, then Y is Y^{o}

This configurational restatement of the argument/adjunct dichotomy is similar to m-command and to Lexical Goverment in that if Y is Y^{max}, then X^{max} lies outside of Y and is not governed by Y.

[25] In fact, a tone group can be indefinitely long. In one example cited in Chen (1987a:113), a 12-syllable-long sentence forms one single tone group.

5.1 Verse recitation

Boyce (1980) appears to be the first to point out the rhythmic effect on Xiamen tone sandhi. As noted before, a heptasyllabic poetic line is parsed into three feet, namely [(σσ)(σσ)(σσσ)]. In verse recitation, it is often this *poetic* foot, and not the regular p-phrase, that constitutes the domain of tone sandhi. This accounts for the phenomena of overapplication (over-extension) as well as underapplication (suspension) of the tone-sandhi rule, for example, in the reading of ballads cited in Chen (1980:22f.).[26]

(88) "tears flow, the heart aches"
tear drop down heart ache
a. *bak-sai*]$_{NP}$ *lao loq*]$_{VP}$ *sim-tao*]$_{NP}$ *sng*[27]
( ) ( ) ( ) ( ) i.
( ) ( ) ( ) j.

"people of yore do not see today's moon"
old people not see today moon
b. *kɔ lin*]$_{NP}$ *put kian kim-si guat*
( ) ( ) i.
( ) ( ) ( ) j.

"woman dressed up as man; nobody suspects"
woman make-up male attire nobody suspect
c. *lu*]$_{NP}$ *ban* *lam tsong*]$_{VP}$ *bo lang gi*
( ) ( ) ( ) i.
( ) ( ) ( ) j.

(. . .) = p-phrase, tone sandhi domain
i. = expected reading in ordinary speech
j. = recorded verse recitation

In (88a), the subject NP *sim-tao* "heart" ought to form a separate p-phrase, and therefore block tone sandhi between *sim-tao* and the verb *sng* "ache." Instead, the attested reading treats the last three syllables as a single sandhi domain, and overapplies the tone-sandhi rule to *-tao*. Conversely, (88b) exemplifies the suspension of tone sandhi. The verb *kian* "see" should join its complement *kim-si guat* "today's moon" in one tone group, and therefore undergo tone sandhi. In actual oral rendition, *kian* is set off as part of a separate p-phrase, blocking tone sandhi. Finally, (88c) instantiates a simultaneous under- and overapplication of tone sandhi: it overapplies between the subject and the predicate (*lu ban* "woman

[26] (88b) is cited from Hsiao (1991:153). See Hsu (1994) for additional examples.
[27] The rhyme *ng* is a syllabic velar nasal.

dressed up"), but blocks between the verb and its complement (*ban lam tsong* "dressed up as man").[28] The patterns exemplified here clearly represent rhythm-induced restructuring of the sandhi domain.

5.2 Idioms

Rhythmic effect is not restricted to verse recitation. It is equally apparent in literary readings, notably the four-syllable idioms that abound in everyday speech. Thus, compound nouns like those of (89a, b) contain no internal $]_{XP}$ boundaries, and should therefore constitute one single sandhi domain; in actual fact, the normal reading treats them as comprising two or three separate foot-sized units. In this respect, the domain of tone sandhi is coextensive with the MRU in Beijing Mandarin.

(89) China people republic
a. *tiong-hua lin-bin kiong-ho-kok* "People's Republic of China"
( ) ( ) ( )

allergic rhinitis
b. *ke-bin-sing p'ĩ-giam* "allergic rhinitis"
( ) ( )

Idiomatic expressions exhibit certain idiosyncrasies. For instance, (90a) and (90b) share exactly the same grammatical structure ([subject [verb + complement]]), but differ in their sandhi behavior: the former obeys the indigenous principle of X^{max}-sensitive grouping, while the latter favors the Mandarin-style parsing of phrases into two- and three-syllable units. Finally, (90c), grammatically isomorphic to (90a, b), is something of a hybrid, combining both characteristics. On the one hand, it marks off the subject NP, consistent with p-phrasing; on the other hand, it introduces a rhythmic boundary between the second and the third syllable, which is otherwise uncalled-for. In any case, where connected speech is broken up into MRUs, the procedure is entirely consistent with the foot-formation rules formulated for Mandarin Chinese.

(90) momentum like split bamboo "irresistible force"[29]
a. *se*$]_{NP}$ *lu p'o tik*
( ) ( )

[28] *Bo lang* "nobody" is treated as a non-lexical item, hence not a full-fledged maximal projection, even though it functions as the subject NP. This accounts for the fact that *bo lang* is not marked off by a tone group boundary.

[29] A metaphor based on the common observation of how a bamboo stick is cleaved: the split inexorably runs along the entire length of the stick.

	age	rival	South Mountain	"longevity"
b.	*siu*]$_{NP}$	*pi*	*lam suã*	
	( )		( )	
	talent	exceed	eight measure	"extremely talented"[30]
c.	*tsai*]$_{NP}$	*ko*	*pat tao*	
	( )	( )	( )	

Expressions like (89)–(90) all have one thing in common, namely a literary flavor that sets them apart from the vernacular forms. In this connection it is worth noting that (Mandarin-based) literary pronunciation and colloquial (vernacular) variants coexist in most Chinese dialects.[31] In view of the widespread diglossia and bilingualism, one should not be surprised to find the alternative patterns of phonological parsing observed in Xiamen.

[30] *tao* is a dry unit of measure for grain (= 1 decalitre).

[31] This phenomenon is referred to as *wen bai yi du*, and is the subject of numerous philological and sociolinguistic studies.

11 *From tone to intonation*

In this final chapter I undertake to investigate the tonal behavior of **Wenzhou** starting from base tones in citation forms through morphotonemics all the way to their concatenation in intonational phrases. Setting aside the details for the moment, let us look at the overall picture. At the word-level, the relevant generalizations can be captured by means of a diagram:

(1) $(\text{proclitics}\ [\ \ldots\ \sigma\ \sigma\ \sigma]_{\text{word}}\ \text{enclitics})_{\text{p-word}}$

pretonic tonic nucleus posttonic

The (morphosyntactic) word is right-prominent: up to three of the rightmost syllables (in boldface) may carry tone and constitute the **tonic nucleus**. All syllables farther away from the end of the word become toneless. All pretonic syllables, including proclitics if any, uniformly take a default L. Posttonic enclitics, on the other hand, are linked to the last tone segment of the tonic nucleus via tone spread.

Above the p-word, which I use interchangeably with the term clitic group, is the intonational phrase (IP). What is striking about Wenzhou is that within each IP there can be one and only one tonic nucleus, with all other p-words being accentually demoted and tonally reduced (to a default L). In this respect, Wenzhou represents an analog of Shanghai post-focus de-toning, whereby all postfocus elements within a phonological phrase undergo tonal reduction (chapter 10, section 2). How a sentence is cut up into IPs (Intonational Phrasing), and which element within an IP is accorded prosodic prominence, and hence retains its tonic nucleus (Tonic Prominence), are the two issues taken up in sections 4–5. There is nothing in my analysis that calls for a prosodic category intermediate between the p-word and the IP.

Accordingly, this chapter is organized as follows: after a brief introductory section, the main body of the paper is divided into two parts.

The first part deals with two types of sandhi processes whose scopes are circumscribed, respectively, by the p-word (sections 2–3) and the intonational phrase (section 4). The second part is devoted to the issues of intonational phrasing (section 5) and the principles determining the location of tonic prominence (section 6).

1 Wenzhou tone system

Wenzhou city is located on the southeastern coast of China, about 220 miles south of Shanghai. Of the southern Wu dialects, the variety spoken in this region is both representative and the most extensively reported, mainly in the works of Zheng-Zhang (1964a, 1964b, 1980–1981). Zheng's seminal work first brought to the linguists' attention a highly intricate pattern of morphotonemics. For our primary purpose, it was necessary to create a corpus of our own consisting of full sentences in connected speech. The primary data on which the present study is based were gathered over a period of four years (1986–1990).[1]

As reported in Zheng-Zhang (1980–1981), the sound system of Wenzhou displays a full and symmetrical system of eight tones corresponding to the four Middle Chinese tonal categories evenly divided into two registers, as shown below.

(2)

	even	rising	falling	dipping
high register	33	35	42	313
low register	11	24	31	212

Given its overspecificity, it is possible to reduce this five-point scale to three pitch levels, namely H, M, L. Furthermore, in my informant's speech the difference between the two rising and the two dipping tones is barely perceptible, and can be regarded as intrinsic pitch variations keyed to the phonemic contrast between the voiced (low register) and voiceless (high register) onsets. The entering tones, once associated with checked syllables ending in an obstruent coda [p, t, k], surface synchronically as [313, 212],

[1] Both on location in China (Fall 1986 and Summer 1990) and in San Diego during my informant Pan Wuyun's visits to the U.S. (Winter and Summer of 1986, and Spring of 1989). Where there is a subdialectal difference between the pronunciations of Zheng and Pan, I have consistently followed the latter.

with a slight dipping pitch inflection.[2] I shall indicate both dipping tones as Lq (with "q" as a diacritic marking the "entering" tone [= tone IV] as a historical category). The resulting simplified tone system is then as follows:

(3)

	even	rising	falling	dipping
high register	M	MH	HM	Lq
low register	L		ML	

These eight words illustrate the tonal values of Wenzhou:[3]

(4)

even:	[*ta*]	M	"simple"
	[*da*]	L	"big"
rising:	[*ta*]	MH	"gall"
	[*da*]	MH	"bland"
falling:	[*ta*]	HM	"dawn"
	[*da*]	ML	"talk"
dipping:	[*ta*]	Lq	"ride, take"
	[*da*]	Lq	"reach"

2 Word-level tone sandhi

2.1 Disyllabic compounds

In connected speech, the eight base tones of (1) often surface in a variety of transforms. In disyllabic words, to take the simplest case, both underlying tones may undergo change, the specific resulting sandhi tone shapes being determined by the nature of both input tones. Table 11.1 summarizes the mapping relations between the base tones and their sandhi forms. The leftmost column indicates the base tone of the first syllable, while the citation tones of the second syllable are arranged on the top row. Within

[2] The corresponding syllables still end in a glottal stop in many closely related Wu dialects. Wenzhou, on the other hand, has lost not only this glottal stop, but also any vestigial contrast in vowel length. In fact, the synchronic reflexes of the Middle Chinese checked tones are nearly as long as the even tones, and significantly longer than the rising and falling tones, both derived from the so-called smooth syllables ending in a vowel or a nasal consonant.

[3] The phonetic transcriptions given in [. . .] are in IPA symbols. Elsewhere I use the "Pinyin" system for transcribing Wenzhou examples. Where a morpheme has no Pinyin equivalent in Beijing, it is transcribed in IPA symbols (underlined).

Table 11.1. *Wenzhou disyllabic lexical tone sandhi*

<table>
<tr><th colspan="2" rowspan="2">first σ</th><th colspan="6">second σ</th></tr>
<tr><th>Ia
M</th><th>Ib
ML</th><th>IIIa
HM</th><th>IIIb
L</th><th>II
MH</th><th>IV
Lq</th></tr>
<tr><td>Ia</td><td>M</td><td>M.M</td><td rowspan="2">L.L</td><td colspan="2" rowspan="2">MLM.HM</td><td rowspan="5">HM.MH</td><td rowspan="5">HM.Lq</td></tr>
<tr><td>Ib</td><td>ML</td><td>L.M</td></tr>
<tr><td>IIIa</td><td>HM</td><td rowspan="3">HM.M</td><td colspan="2" rowspan="2">HM.ML</td><td rowspan="3">HM.L</td></tr>
<tr><td>IIIb</td><td>L</td></tr>
<tr><td>II</td><td>MH</td></tr>
<tr><td>IV</td><td>Lq</td><td>Lq.M</td><td>L.L</td><td>Lq.HM</td><td>Lq.L</td><td>Lq.MH</td><td>Lq.Lq</td></tr>
</table>

I, II, III, IV = Middle Chinese tonal categories
a, b = high and low registers

the box, the sandhi forms are found where the rows and columns of the base tones intersect.

It can be seen readily from table 11.1 that out of the thirty-six logically possible two-tone combinations (6 × 6), only fourteen ditonic patterns actually occur in lexical compounds. This pattern of convergence and partial neutralization is characteristic of all Chinese dialects exhibiting tone sandhi. A few general observations can be inferred. As a rule the last syllable retains its base tone. In addition, the checked tone Lq in the initial position also remains unchanged. However the exact nature of the phonetic processes that created the patterns summarized in table 11.1 has been obscured by intervening historical changes. Thus, synchronically speaking, a disparate class of {HM, L, MH} in the penult position merge into a single sandhi shape HM, regardless of the final tone.[4] Restated in terms of Middle Chinese (MC) categories, this grouping turns out to be less arbitrary, since the set of HM, L, and MH tones corresponds to the subclass of so-called "*ze*" (lit. inflected or oblique) tones on smooth syllables, in opposition to the *ping* (lit. even) and *ru* (lit. entering) checked tones. I will not attempt to translate table 11.1 into the customary rule formalism. Henceforth I shall use the label **Disyllabic TS** (for disyllabic

[4] With the minor exception of HM + ML = [L-L].

lexical tone sandhi) as a shorthand to refer to the sum of individual processes that combine to produce the sandhi forms indicated in table 11.1. Note in passing that Disyllabic TS is a lexical rule, and applies only to lexical forms (in a sense to be made more precise, see section 2.3), not phrases, the latter being governed by the rule of Tonic Prominence, to be developed in a subsequent section.

The effects of Disyllabic TS are illustrated below. The first and second lines below the transliteration indicate the base tone and sandhi tone respectively. The capital letters A, B, C (and A′, B′, C′ . . .) refer to the fourteen actually occurring two-tone patterns generated by the application of Disyllabic TS, and correspond to the labeling in the published reports by Zheng cited in the introductory section.

(5) Aa	*yi-sheng* M M M M	"physician"			
Ab	*tian-ji* ML. M L M	"frog"			
B	*guang-zhou* MH M HM M	(place name)	B′	*bei-fang* Lq M Lq M	"North"
C	*tian-tang* M ML L L	"paradise"			
D	*xiao-mi* MH MH HM MH	"rice"	D′	*tie-bang* Lq MH Lq MH	"iron bar"
E	*kai-hui* M L MLM HM	"call a meeting"			
F	*wen-ti* L ML HM.ML	"question"	F′	*bai-huo* Lq HM Lq HM	"merchandise"
G	*zi-dan* MH L HM L	"bullet"	G′	*re-du* Lq L Lq L	"temperature'
H	*xiao-xue* MH Lq HM Lq	"elementary school"	H′	*fa-ze* Lq Lq Lq Lq	"law"

2.2 *Long compounds*

Recall that up to three of the rightmost syllables in a (morphosyntactic) word may bear tone. We have seen that the tone shapes of the last two syllables are determined by a complex set of morphotonemic alternations collectively referred to as Disyllabic TS. Disyllabic TS yields one of fourteen possible ditonic melodies. The antepenultimate tone is neutralized to an even (level) tone, which assumes a pitch height opposite to the first element of the ditonic melody by a Polarity rule, which is stated as follows:

(6) Polarity:

$$\begin{array}{ccc} \sigma & & \sigma\ \sigma]_{\text{word}} \\ | & & | \\ T & \rightarrow & L\ /\_\_\ H \\ & & H\ /\ \text{elsewhere} \end{array}$$

In longer polysyllabic compounds, all syllables farther away from the end than the antepenult lose their underlying tone, and carry a default L tone by Default, which I will annotate simply as [o]. This process is stated as (7) below.

(7) Deletion

$$\begin{array}{l} \ldots\sigma\ldots\sigma\ \sigma\ \sigma]_{\text{word}} \\ \quad | \\ \quad T \rightarrow \varnothing \end{array}$$

A few examples will illustrate the effect of Polarity and Deletion:

(8)

a.	*dai*-[*ming*-*ci*]	"pronoun"
	L ML ML	base tone
	L L	Disyllabic TS
	H L L	Polarity
b.	[*zi*-*ben*]-*jia*	"capitalist"
	M MH M	base tone
	HM M	Disyllabic TS
	L HM M	Polarity
c.	[*huo*-*che*]-*piao*	"train ticket"
	MH M HM	base tone
	MLM.HM	Disyllabic TS
	H MLM.HM	Polarity

It is worth noting that the ditonic melody is derived from the last two syllables of a compound, regardless of the internal morphological structure. Take *dai*-[*ming*-*ci*] "pronoun" and [*zi*-*ben*]-*jia* "capitalist" for instance.

As the bracketings suggest, whereas the former is created by prefixing *dai* "proxy, pro-form" to a free form *ming-ci* "noun," the latter is formed by suffixing *jia* "-ist" to the base *zi-ben* "capital, fund." However, phonology is quite indifferent to this contrast between right- and left-branching morphological tree structures. Disyllabic TS simply applies to a non-constituent like *ben-jia* in (8b) as well as to a constituent like *ming-ci* in (8a). It is clear from (8b) that Polarity must follow Disyllabic TS, that is to say the tonal value of the antepenult is determined by dissimilation not to the base tone but to the sandhi form produced by Disyllabic TS. If the antepenult were to dissimilate with the base tone, it would emerge with a high tone.

Lexical compounds more than three syllables long are somewhat less common, but do occur. Here are some examples.

(9)	*gong-zuo-dan-wei*	"work unit"
	M Lq M L	base tone
	MLM.HM	Disyllabic TS
	H.MLM.HM	Polarity
	o H.MLM HM	Deletion
(10)	*an-de-luo-po-fu*	"Andropov"
	M.Lq.ML.M.M	base tone
	M.M	Disyllabic TS (vacuous)
	H M M	Polarity
	o o H M M	Deletion
(11)	*chai-mi-you-yan-jiang-cu-cha*	"daily necessities"
	ML.MH.ML.ML.HM.HM.ML	base tone
	L L	Disyllabic TS
	H L L	Polarity
	o o o o H L L	Deletion

Even though both *gong-zuo* "work" and *dan-wei* "unit" occur as independent words, the compound behaves as one single lexical unit in (9). This accounts for the appearance of a high tone on the antepenultimate syllable by virtue of Polarity, which is a lexical level TS rule, not applicable to atonic syllables in phrasal phonology.[5] Since no pitch has been assigned to the initial atonic syllable *gong*, it takes on the value of L by Default. Examples (10) and (11) exemplify pentasyllabic and heptasyllabic compounds, where the rules developed so far combine to produce the attested sandhi forms in a straightforward manner. *An-de-luo-po-fu* is

[5] For the contrast between phrasal and lexical TS rules, see below.

monomorphemic, obviously a phonetic transliteration. The word-for-word gloss of *chai-mi-you-yan-jiang-cu-cha* (11) is "firewood-rice-oil-salt-sauce-vinegar-tea," the essential ingredients of daily life.

2.3 Lexicalized phrases

Short syntactic phrases two to three syllables in length are often treated like a single lexical unit. We may refer to such quasi-lexical expressions as **lexicalized phrases**, in recognition of their hybrid nature exhibiting syntactic properties of phrases, yet behaving phonologically like lexical compounds. We have seen similar hybrids such as *xun kai-xin* "to make fun" and *shang dang* "to be duped" etc. in Shanghai, among other dialects (see Appendix to chapter 9). Lexicalization of syntactic phrases in Wenzhou is subject to syntactic/semantic as well as phonological constraints. We may state these constraints as:

(12) Lexicalized Phrases
Trisyllabic or shorter modifier + head constructions may be treated as a single lexical compound. Longer modifier + head constructions optionally turn into p-words depending on frequency of occurrence, degree of idiomatization etc.

Some examples follow:

(13)	a.	*da shu*	"big tree"
		L L	base tone
		HM.L	Disyllabic TS
	b.	*zhen pi-chou*	"very miserly"
		M ML HM	base tone
		MLM.HM	Disyllabic TS
		H MLM HM	Polarity
		utmost run	
	c.	*pin-ming* <u>*zei*</u>[6]	"to run for all one is worth"
		HM L Lq	base tone
		HM. Lq	Disyllabic TS
		L HM. Lq	Polarity

Wenzhou treats direct modifier + noun constructions like (13a) as words both syntactically and phonologically, exactly as in Beijing Mandarin (see chapter 9, section 3.1). (13b) and (13c) are presumably phrasal construc-

[6] *Pin-ming* literally means "death-defying." The underlined <u>*zei*</u> "to run" does not have a Beijing cognate.

tions. However, they behave unequivocally as though they were lexical compounds. For instance, Disyllabic TS is blind to the internal structure of (13c), and simply treats the last two morphosyntactically unrelated morphemes as constituting the disyllabic base. If (13c) were interpreted phonologically as a phrase, the expected derivation would be as follows:

(14)	*pin-ming* <u>*zei*</u>	"run for all one's worth"
	HM L Lq	base tone
	(HM L) (Lq)	Disyllabic TS (vacuous)
	(HM L) (o)	Tonic Prominence (to be given below)

(. . .) = word or p-word

At the lexical stratum, Disyllabic TS is vacuous on *pin-ming* "utmost, for all one is worth." At the phrase level, in modifier + head constructions, the modifier occupies a metrically strong position by virtue of the Tonic Prominence rule to be given below, and therefore constitutes the tonic nucleus. By the same token, the verb *zei* "run" is demoted to a metrically weak position, and is tonally reduced to zero "o," phonetically interpreted as L. Interestingly, both readings (13c) and (14) are attested alternatives. This means that lexicalization is a gradient process, and makes allowance for free variants.

Phrasal constructions other than the modifier + head type do not generally behave like lexical compounds, regardless of syllable count. Thus verb + object expressions like the following do not typically undergo the lexical Disyllabic TS rules summarized in table 11.1:

(15)		*da shui*	"fetch water"
		MH.MH	
	a.	(HM.MH)	Disyllabic TS,*
	b.	(MH)(MH)	Disyllabic TS, not applicable
		(o) (MH)	Tonic Prominence (to be given below)

(. . .) = word or p-word

If *da shui* "to fetch water" were treated like a lexical compound, Disyllabic TS would apply to produce the unacceptable [HM.MH] as the sandhi output. Analyzed as a phrasal construction, each syllable constitutes a separate lexical domain; therefore Disyllabic TS, which is a lexical rule, must fail. Subsequently, at the phrase level, *da* "fetch" is de-accented, and tonally reduced to [o], a default L.

The foregoing discussion should not lead one to conclude that all lexical compounds in Wenzhou (or Chinese in general) must have the internal structure of modifier–head. Like other Chinese dialects, Wenzhou makes extensive use of established lexical compounds exhibiting internal morphosyntactic structures of various sorts, including subject–predicate, verb–object, coordination, even subject–verb–object: *tou-teng* "headache, trouble" (from head + to ache), *bing-bian* "coup-d'état" (from soldier + to rebel), *shi-wei* "demonstration" (from to show + force), *da-xiao* "size" (from big + small), *fei-jie-he* "tuberculosis" (from lung + to grow + cell) etc. Finally, there is a sizable class of expressions that fluctuate between lexical and phrasal status. Their ambiguous morphosyntactic status is reflected in their alternative readings, one derivable from lexical Disyllabic TS, the other from phrasal Tonic Prominence. In addition to *pin-ming* <u>*zei*</u> "run for all one is worth" noted above, here are some more alternative readings (i, ii):

(16)

	open meeting	
a.	*kai* *hui*	"to attend a meeting"
	M L	base tone
	(MLM.HM)	i. Lexical, by Disyllabic TS
	(o) (L)	ii. Phrasal, by Tonic Prominence
	invite guest	
b.	*qing* *ke*	"to throw a party"
	MH Lq	base tone
	(HM Lq)	i. Lexical, by Disyllabic TS
	(o) (Lq)	ii. Phrasal, by Tonic Prominence
	listen lecture	
c.	*ting* *ke*	"to attend a lecture"
	M HM	base tone
	(MLM.HM)	i. Lexical, by Disyllabic TS
	(o) (HM)	ii. Phrasal, by Tonic Prominence

2.4 *Special cases of lexical tone sandhi*

Before we leave the discussion of lexical tone sandhi, a brief mention must be made of two special cases of polysyllabic compounds. As mentioned earlier, checked tone Lq in penult position remains unchanged under Disyllabic TS.[7] In trisyllabic and longer compounds, however, the

[7] Recall that Lq is used diacritically to refer to a historical tonal category, corresponding to a checked syllable in Middle Chinese. Synchronically speaking, Wenzhou no longer has CVq type in its syllabic inventory.

penult Lq changes to HM, and then undergoes Disyllabic TS in the usual fashion. We therefore need a special rule which I will call Checked TS (for checked tone sandhi), ordered before Disyllabic TS.

(17) Checked TS

$$[\ldots \sigma\ \sigma\ \sigma]_{\text{word}}$$
$$\quad\ \ |$$
$$\quad \text{Lq} \rightarrow \text{HM}$$

The effect of Checked TS is illustrated by the examples of (18).

(18) a.	*da-xue-sheng*		"college student"
	L Lq	M	base tone
	HM		Checked TS
	HM	M	Disyllabic TS (vacuous)
	L HM	M	Polarity (vacuous)
b.	*da-shi-tou*		"big boulder"
	L Lq	ML	base tone
	HM		Checked TS
	L	L	Disyllabic TS
	H L	L	Polarity

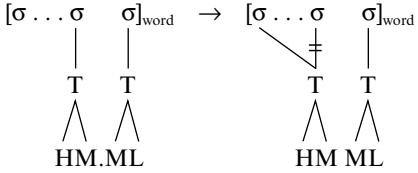
A second special case involves the ditonic pattern [HM ML] generated by Disyllabic TS.[8] Normally one expects all syllables preceding the ditonic pattern [HM ML] to be L, either by Polarity (in the case of the antepenult), or via Tone Deletion and the default L (in the case of the syllables farther to the left). In other words, the canonical form of a long compound ending in [HM ML] should be [. . . o.L.HM ML]. In fact, the attested forms suggests a process of HM-shift, whereby the high-falling [HM-] is moved to the first syllable of the compound, skipping over any number of syllables in between. In the process, this shift leaves toneless the medial syllables, which are linked with the mid-level pitch, presumably by the Well-formedness Condition (WFC).[9] The HM-shift is stated and illustrated below.[10]

[8] This is ditonic pattern F in (5).

[9] The delinking of the medial tones is presumably a by-product of the line-crossing prohibition. A similar situation is observed in Shanghai; see chapter 7, section 3.4.

[10] Alternatively, HM-Shift can be seen as a case of melody spread. That is to say, the ditonic pattern HM.ML behaves like a word melody, which spreads across the entire lexical domain in an edge-in fashion, as seen in Tangxi, Shanghai, Wuxi, and Danyang, described in chapters 7–8.

(19) HM-Shift



(20) a.	[*sun-nü*]-*xu*		"grand-daughter's husband"
	M.MH.HM		base tone
	HM.ML		Disyllabic TS
	HM. o ML		HM-Shift
	HM m.ML		WFC
b.	[*wai*-[*sun-nu:*]]-*xu*		"maternal grand-daughter's husband"
	L M MH HM		base tone
	HM ML		Disyllabic TS
	HM o o ML		HM-Shift
	HM m m ML		WFC
	wireless telephone tube		
c.	[[*wu-xian*]-[*dian-hua*]]-*tong*		"radio receiver"
	ML HM L L ML		
	HM ML		Disyllabic TS
	HM o o o ML		HM-Shift
	HM m m m ML		WFC

m = default pitch supplied by WFC

Finally, the combined effect of Checked TS and HM-Shift is illustrated below. Checked TS creates the appropriate tonal sequence which gives rise to the ditonic [HM.ML] pattern, which in turn undergoes HM-Shift, triggering WFC to produce the final output [HM.m.ML].

(21)	*gao*-[*tie-jia*]	"tall steel case"
	M Lq. HM	base tone
	HM	Checked TS
	HM.ML	Disyllabic TS
	HM o ML	HM-Shift
	HM m ML	WFC

3 Clitic groups

Often a lexical word is surrounded by one or more function words. Typically, function words (especially if they are monosyllabic) are prosodically dependent, may not carry an independent tone, and must cliticize to the lexical host bearing the tonic nucleus. Specifically, pretonic proclitics

uniformly take the default L; posttonic enclitics assume the pitch value of the last element of the tonic nucleus.

We have already come across the notion of the clitic group, which plays an important role in our analysis of Beijing tone sandhi (chapter 9, section 4.1–4.2). I will not attempt a precise and all-encompassing definition of the class of clitics, which corresponds roughly to the traditional notion of *xuci* or "empty words" (vs. *shici* or "content words"). For our purposes, suffice it to say that clitics comprise the closed set of function words and grammatical particles that do not belong to one of the major lexical categories, namely nouns, verbs, and adjectives. As for adverbs, they form a notoriously heterogeneous subset and elude a precise definition. For expository clarity, clitics are written in small capitals. Clitics do carry a lexically assigned tone. For instance, in response to "How do you say the word 'not'?" one could say [BU] with a high rising tone [MH]. However, in connected speech, clitics are metrically weak and tonally reduced. I will refer to this as Clitic Reduction.[11] Proclitics or pretonic clitics uniformly carry a low tone (marked here simply as [o]), the default value of all atonic syllables, e.g.

(22) a.	CONG *hangzhou*			"from Hangzhou (place name)"
	(ML	ML	M)	base tone, p-word
	o	ML	M	Clitic Reduction
	o	L	M	i. Disyllabic TS
	H	L	M	ii. Polarity,*
b.	BU *tai-ping*			"not peaceful"
	(MH.HM.ML)			base tone, p-word
	o	HM	ML	Clitic Reduction
	o	L	L	i. Disyllabic TS
	H	L	L	ii. Polarity,*

CONG, BU etc. (in small caps.) = clitics

(Where the underlying tones of clitics are immaterial, I will assume that Clitic Reduction has already applied in the examples that follow.) Needless to say, Clitic Reduction either precedes (and bleeds) Polarity or in any case renders the latter vacuous.

Enclitics, that is clitics in posttonic positions, assimilate to the tonal value of their host on the left via Spread.

[11] For simplicity, I will ignore the distinction between two types of function words: (i) -GI "possessive, subordinator, relative clause marker" and the like, which are inherently atonic, and unstressable; (ii) BU "not" and others which are toneless in normal connected speech, but are stressable and carry tone under focus or emphasis. In any event, clitics become toneless via Clitic Reduction.

(23) Spread (. . . σ <σ>. . .)$_{\text{p-word}}$

T

<σ> = toneless syllables
(. . .) = p-word

The effect of Spread is illustrated by two enclitics – *GE* and *GI* – which function as the classifier/measure word and the nominalizer or relative clause marker respectively.[12] In the examples that follow, the pitch values obtained by Spread are annotated in lowercase letters [h, m, l] (for high, mid, low).

(24) a. *liang GE* "two Cl"
MH o
MH h Spread

b. *san GE* "three Cl"
M o
M m

c. *si GE* "four Cl"
HM o
HM m

d. *liu GE* "six Cl"
Lq o
Lq l

red RC
e. *hong GI* "the red one/that which is red"
ML o
ML l

large RC
f. *da GI* "the large one/that which is large"
L o
L l

Cl = classifier
RC = relative clause marker

In principle the last tone of the lexical host can spread to any number of enclitic syllables. Thus, depending on the intonation break (marked by "/" – see section 5 below on intonational phrasing), the last high tone of the lexical host *lou* "basket" can spread to as many as four enclitics.

[12] Wenzhou *GI* corresponds to Mandarin *de*.

(25) think Asp

a. *xiang* QI — "think of, remember"
(MH o) — base tone, p-word
MH h — Spread

think Asp Asp

b. *xiang* QI LAI — "as one thinks about it"
(MH o o) — base tone, p-word
MH h h — Spread

basket Loc have some Cl apple

c. *lou* LI YOU JI GE *ping-guo* — "There are several apples in the basket"
(MH o o) (o o ML.MH) — base tone, p-word
MH o o / o o ML.MH — Intonational Phrasing
HM.MH — Disyllabic TS
MH h h / o o HM.MH — Spread = reading (i)

(MH o o o o)/(ML.MH) — p-word
MH o o o o / ML.MH — Intonational Phrasing
HM.MH — Disyllabic TS
MH h h h h / HM.MH — Spread = reading (j)

(. . .) = p-word
/ = intonation break
small caps indicate clitic-hood

Consistent with prosodic hierarchy, p-words or clitic groups are constituents of an IP, within which they must be wholly contained. Hence, given alternative intonational phrasings, JI GE "some Cl" pattern like proclitics of the second p-word headed by *ping-guo* "apple" in reading (i), but as enclitics of the first p-word anchored in *lou* "basket" in reading (j). As stated, Spread extends the last tone segment of the lexical host rightwards to any and all toneless syllables within the IP-bound clitic group. This accounts for the alternative readings of (25c).

Polysyllabic function words (such as JI-SHI "even if") enjoy a certain prosodic independence and behave for all practical purposes like content words.[13] Furthermore, a pair of closely related monosyllabic clitics optionally turns into a sequence of host + clitic. This process is stated as follows:[14]

(26) Host-Clitic
Optionally turn [clitic + clitic] into [host + clitic]

[13] Likewise in English: polysyllabic function words such as *under*, *over*, and so forth do not undergo destressing; cf. Selkirk (1984).

[14] Analogously, we have in English expressions like *on+to*, *there+fore* etc. which coalesce into prosodically independent wordforms with stress patterns resembling content words.

The requirement that the clitics be in close relation (essentially "in construction with" each other, symbolized by the morphosyntactic bracketing [. . .]) is illustrated by the following example involving a pronoun plus the scopal adverb *SANG* "only, just,"[15]

(27)	I only know Asp	
a.	[*WO* *SANG*] *xiao* *DE*	"only I know about it"
	MH M MH Lq	base tone
	(MH.M) (MH.Lq)	Host-Clitic
	(MH o) (MH o)	Clitic Reduction
	(MH h)	Spread
	(MH h) (o o)	Tonic Prominence (see below)
	I only know Asp	
b.	*WO* [*SANG* *xiao* *DE*]	"I just found out about it"
	MH M MH Lq	base tone, Host-Clitic not applicable
	(o o MH o)	Clitic Reduction
	(o o MH h)	Spread

As the translations suggest, the adverb *SANG* "only" is in construction with the subject NP *WO* in (27a), but with the verb *xiao* *DE* "know" in (27b). Consequently, Host-Clitic applies in (27a), but not in (27b). In the former case, as the derived lexical host, the function word *WO* acquires a certain prosodic independence, and hence is capable of carrying a tone, and adverb *SANG* acts as an enclitic, assuming the high tone of its host *WO* "I." In the latter case, *WO* and *SANG* are unrelated to each other, and both behave as proclitics to the main verb *xiao* *DE*, and therefore take on the default L tone ([o]). The directionality of clitic attachment depends in part on phonological phrasing, which will be dealt with in section 6.

4 Phrasal tone sandhi

We now turn to phrase-level sandhi phenomena. The fundamental governing principle is this: within an intonational phrase there is one and only one tonic nucleus; all other syllables are atonic. All atonic syllables which do not otherwise acquire some tonal values by Spread[16] are uniformly assigned a L tone by default.

Putting aside for the moment the questions of how to break down a sentence into IPs (see section 5), and how to locate the tonic nucleus within each IP (see section 6), let us contrast the phonetic effects of word-level vs. phrase-level tone sandhi.

[15] *SANG* is a dialect-specific morpheme corresponding to Mandarin *cai*.
[16] And in special cases, Checked TS, HM-Shift, and WFC, discussed in section 2.4.

4.1 Pretonic positions

In pretonic positions, root morphemes that form part of a lexical category behave differently from proclitics as well as phrasally non-prominent elements. This is illustrated by (28a, b). As before, p-words are enclosed within parentheses.

(28)	a.	*xiao-hang-zhou*	"Little Hang-zhou" (place name)
		(MH ML M)	base tone
		L M	Disyllabic TS
		H L M	Polarity
	b.	*CONG hang-zhou*	"from Hang-zhou"
		(o ML ML)	base tone, Clitic Reduction
		o L M	Disyllabic TS

Both (28a) and (28b) constitute single p-words. The difference is, of course, that *xiao* is a lexical root morpheme, and is therefore potentially tone-bearing – in this case H, by virtue of Polarity – while *CONG* is a clitic, and therefore inherently atonic or toneless by virtue of Clitic Reduction. Since Spread works only unidirectionally to the right, pretonic syllables remain toneless, and take L by default.

Similarly, (29a) constitutes one single lexical item in accordance with Lexicalized Phrases; consequently, Polarity assigns the antepenultimate syllable *hong* "red" a high tone. In contrast, *mai* "buy" forms a separate p-word, and therefore does not undergo Polarity, which has the p-word as its scope of operation. At the phrase-level, however, *mai* occupies a prosodically weak position, and is therefore atonic, phonetically realized as L. I use the term Tonic Prominence to refer to this phrase-level prominence, to be determined by general principles, the most important of which is the principle of Non-Head Prominence, according to which the non-head (argument, adjunct) is prosodically more prominent, while the head is accentually demoted and becomes toneless (see section 6).

(29)	a.	*hong rong-shan*	"red sweater"
		(ML ML M)	base tone, p-word[17]
		L M	Disyllabic TS
		H L M	Polarity
	b.	*mai rong-shan*	"buy sweater"
		(MH) (ML M)	base tone
		L M	Disyllabic TS
		o L M	Tonic Prominence

[17] Recall that modifer + head (but not verb + object) constructions are treated like Lexicalized Phrases, see section 2.3.

(30a, b) illustrate the same contrast between a longer compound and a phrasal construction containing a disyllabic verb.

(30)	a.	*tian-zhu-jiao-tang*	"Catholic church"
		(M MH HM ML)	base tone
		L L	Disyllabic TS
		H L L	Polarity
		o H L L	Deletion
	b.	*xiu-gai ke-wen*	"revise lecture-notes"
		(M.MH)(HM.ML)	base tone
		L L	Disyllabic TS
		o o L L	Tonic Prominence

4.2 Posttonic positions

Since lexical compounds are right-prominent, there are no posttonic syllables.[18] It goes without saying that clitic groups may have posttonic enclitics. Enclitics contrast with phrasally weak elements in their sandhi behavior: while enclitics assimilate to the tonal properties of their host, prosodically weak constituents at the phrase level simply assume the default L tone. This is illustrated below:

(31)		think Asp	
	a.	*xiang* ɥɪ	"to think of"
		(MH. o)	base tone, Clitic Reduction
		MH h	Spread
		early leave	
	b.	*zao zou*	"leave early"
		(MH)(MH)	base tone
		MH o	Tonic Prominence

In (31a), the atonic enclitic ɥɪ forms a single p-word with its lexical host *xiang* "think," and falls within the domain of Spread. In (31b), on the other hand, Tonic Prominence promotes the prehead modifier *zao* "early," while de-stressing the head *zou* "leave," in the process rendering the latter toneless.[19] Since *zou* forms a separate p-word, it is beyond the scope of Spread; instead it ends up with the default L. The examples of (32a–c) illustrate basically the same point.

[18] Except atonic suffixes. I will ignore this point for expository simplicity.

[19] This presupposes that *zao* and *zou* have not coalesced into a single word via Lexicalized Phrases, stated in section 2.3.

(32) a. *jiang-shang* GI "praise him/her"
(MH MH o) base tone, Clitic Reduction
HM. MH Disyllabic TS
HM. MH h Spread

b. *mian-qiang jiang* "reluctantly speak"
(MH.MH) (MH) base tone
HM.MH Disyllabic TS
HM.MH o Tonic Prominence

c. *chen-zao fang jia* "to go on vacation early"
(HM.MH)(HM.HM)
HM.MH Disyllabic TS (vacuous)
HM.MH o o Tonic Prominence

The table in (33) summarizes the various rules operating at the word and phrase levels.

(33)

	pretonic	tonic	posttonic
p-word	Tone Deletion Clitic Reduction Default L	Disyllabic TS Polarity Deletion (Checked TS) (HM-Shift) (WFC)	Clitic Reduction Spread
phrase	Tonic Prominence Default L	(Tonic Prominence)	Tonic Prominence Default L

Within the lexical word, the tone shapes of the last three syllables are determined by Disyllabic TS and Polarity; all syllables farther to the left undergo Tone Deletion. For clarity I segregate the rules of Checked TS, HM-Shift, and WFC from the more general rules by enclosing them in parentheses, since they have a much more limited scope of operation, applying only to specific tonal sequences. Beyond the lexical core, both proclitics and enclitics are either inherently atonic or become toneless via Clitic Reduction. However they have different surface phonetic manifestations: while proclitics consistently assume a default L, enclitics assume a variety of tone shapes in accordance with Spread. Finally at the phrase level, Tonic Prominence picks out one single tonic nucleus within each intonational phrase, and tonally reduces all other constituents to a uniform and monotonous L.

The preceding discussion presupposes that the tonic nuclei of IPs have been identified, which in turn presupposes that we know how to cut up a sentence into IPs. We will take up these questions in the next two sections.

5 Intonation phrasing

Intonational phrasing is basically governed by the Sense Unit Condition, stated by Selkirk, who credits Halliday (1967) with the original idea:

(34) Sense Unit Condition
The immediate constituents of an intonational phrase must together form a sense unit. (Selkirk 1984:286)

A "sense unit" is defined as follows:

(35) Sense Unit
Two constituents C_i, C_j form a sense unit if (a) or (b) is true of the semantic interpretation of the sentence:
a. C_i modifies C_j (a head)
b. C_i is an argument of C_j (a head). (Selkirk 1984:291)

Furthermore, intonational phrasing must respect WFC on prosodic constituents. Specifically, an IP is made up of integral p-words. In other words, parts of p-words may not belong to separate IPs higher up on the prosodic hierarchy. In this connection it is worth emphasizing the asymmetric nature of cliticization, which is made more explicit in (36).

(36) Cliticization: a function word may cliticize
Leftwards, without restriction;
Rightwards, only if both of these two conditions are met:
i. its syntactically closest host is on the right
ii. and it is IP-initial

What is intended by condition (i) of (36) is the following:

(37) a. X [c X] → i. (X c) (X)
j. (X) (c X)

b. [X c] X → i. (X c) (X)
j. *(X) (c X)

X	=	lexical host
c	=	function word, clitic
(. . .)	=	p-word
[. . .]	=	morphosyntactic constituency

The function word in (37a) can cliticize in either direction, since leftward encliticization is allowed without restriction to morphosyntactic constituency. On the other hand, the condition stated as (36) rules out rightward procliticization (j) given a structure like (37b).

Condition (ii) of Cliticization essentially rules out IP-internal proclitics. This means that (38), which in principle permits the function word "c" to cliticize to a potential host X on either side, nonetheless does not permit the parsing in (38-iv) with one single IP.

(38) X [c X]
i. (X c) / (X)
ii. (X c) (X)
iii. (X) / (c X)
iv. (X) (c X) = *

/ = intonational break

One final note before we proceed with the examples. In addition to the semantic and structural constraints, intonational phrasing is also sensitive to constraints of a purely phonological nature, which can be stated as follows:

(39) Balance
To the extent possible, intonational phrases must be balanced in terms of syllable count.

We are now ready to see how these principles combine to parse connected speech into prosodic units.

(40) "the old scholar's briefcase"
old scholar Poss briefcase
[[*lao-xue-zhe*] <u>*GI*</u>] *pi-bao*
(σ σ σ σ) / (σ σ) i.
(σ σ σ σ σ σ) j.

The possessive marker can cliticize only leftwards in accordance with Cliticization. In deliberate speech, each p-word may form an independent IP as in (40-i). Alternatively, the entire sentence may form a single IP in more casual speech (40-j). These options entail phonetic consequences from the point of view of tonal behavior: thus, (40-i) has two tonic nuclei, (40-j) has only one. The sandhi forms corresponding to the two alternative phrasings are given below as (41-i) and (41-j) respectively:

```
(41)    old scholar   Poss briefcase
        [[lao-xue-zhe]  GI] pi-bao
        (MH.Lq.MH   o)/(ML.M)        base tone, Clitic Reduction
              HM.MH                  Checked TS
              HM.MH      L    M      Disyllabic TS
         L    HM.MH                  Polarity
         L    HM.MH. h   L    M      Spread              = i.
        ------------------------------------------------------------------------------
        (L    HM.MH. h)(o    o)      Tonic Prominence    = j.
```

Recall that disyllabic and trisyllabic modifier–head constructions are (optionally) treated as lexical compounds (see Lexicalized Phrases, section 2.3), and GI may cliticize only leftwards. P-word-level rules of Checked TS, Disyllabic TS, Polarity, and Spread apply in the usual manner to produce the output of (41-i). No other rules apply to the two p-words uttered as two separate IPs in deliberate speech. However, as they join into a single IP in reading (41-j), Tonic Prominence promotes the non-head (possessive adjunct, *lao-xue-zhe* GI "the old scholar's") at the expense of the head *pi-bao* "briefcase." As a consequence, *pi-bao* is de-stressed and de-toned.

Consider next the following sentence:

```
(42)    "please wipe grand-daughter's face"
         please you  OM grand-daughter wipe face
         qing   NI   [HA   sun-nü] [cha mian-kong]
         MH     o    o    M.MH  Lq  L MH            base tone, Clitic
   a.    (σ     σ) / (σ    σ    σ) / (σ) (σ    σ)   Reduction
   b.    (σ     σ) / (σ    σ    σ)   (σ) (σ    σ)
   c.    (σ     σ    σ) / (σ    σ) / (σ) (σ    σ)
   d.    (σ     σ    σ) / (σ    σ)   (σ) (σ    σ)
   e.    (σ     σ    σ)   (σ    σ)   (σ) (σ    σ)
   f.    (σ     σ    σ)   (σ    σ) / (σ) (σ    σ)   ??
```

The point of interest here is the directionality of cliticization of the object marker HA. The phonetic consequence of the bidirectional cliticization is shown below:

```
(43)    please you OM grand-daughter
        qing NI   [HA    sun-nü] . . .
        (MH o) / (o    M.MH)         Cliticization
                     HM.MH           Disyllabic TS
         MH h     o  HM MH           Spread              i. = (42a, b)
        ------------------------------------------------------------------------------
        (MH o     o) / (M.MH)        Cliticization
                     HM.MH           Disyllabic TS
         MH h     h  HM.MH           Spread              j. = (42c, d, e)
```

Within the limits set by WFC on prosodic constituency, there is considerable freedom of phrasing, depending on the style and tempo of delivery. Thus sentence (42) can form one single IP, or break up into two to three IPs, subject to the Sense Unit Condition stated above. Thus all the alternative phrasings of (42) are permissible except (f), which is only marginally acceptable and clearly disfavored.

Sense Unit as a well-formedness condition on IP is further supported by the following

(44) "Zhang-san gave me two books yesterday"

	Zhang-san	yesterday	give	I	two	Cl	book			
	Zhang-san	*zuo-ye*	*ha*	*wo*	*liang*	*BEN*	*shu*			
	M	M	ML	L	HM	o	MH	o	M	
a.	(σ	σ)/(σ	σ)/	(σ	σ)/(σ	σ)	(σ)			
b.	(σ	σ)/(σ	σ)	(σ	σ)/(σ	σ)	(σ)			
c.	(σ	σ)	(σ	σ)	(σ	σ)/(σ	σ)	(σ)		
d.	(σ	σ)	(σ	σ)	(σ	σ)	(σ	σ)	(σ)	
e.	(σ	σ)	(σ	σ)/	(σ	σ)	(σ	σ)	(σ)	= *

Sentence (44) can be uttered in one to four IPs (a–d), except as partitioned in (e), in which the subject NP *Zhang-san* and the time adverbial *zuo-ye* "yesterday" are grouped together into an IP, in violation of Sense Unit Condition.

In addition to the semantic constraint, IP is sensitive to the balance in syllable count. Contrast the two sentences in (45) and (46):

(45) "urge Xiao-Li to study English"

	urge	Xiao-Li	study		English			
	quan	*Xiao-Li*	[*wen-xi*		*ying-yu*]			
	HM	MH.MH	M.	Lq	M	MH		
i.	(σ)	(σ σ)	/(σ σ)		(σ σ)			= ok
j.	(σ)	(σ σ)	(σ σ)/(σ		σ)			= *

(46) "urge Xiao-Li to study standard English"

	urge	Xiao-Li	study		standard		English		
	quan	*Xiao-Li*	[*wen-xi*		[*biao-zhun*		*ying-yu*]		
	HM	MH.MH	M	Lq	M	MH	M	MH	
i.	(σ)	(σ σ)	/ (σ	σ)	(σ	σ)	(σ σ)		= ?
j.	(σ)	(σ σ)	(σ	σ)/(σ		σ)	(σ σ)		= ok

In (45-j), the IP *quan xiao-li wen-xi* "urge Xiao-Li to study" is semantically acceptable, since *Xiao-Li* functions both as the object of *quan* "urge" and as the subject of *wen-xi* "study." However, the intonational break

between *wen-xi* and *ying-yu* "English" creates two phonologically unbalanced IPs (5 + 2 syllables). In (46-j), on the other hand, the phonetic balance is restored by the addition of a prenominal modifier *biao-zhun* "standard" on *ying-yu*.

Note in this connection that neither *quan* "urge" in (45, 46), nor *shu* "book" in (44) can form a monosyllabic IP, partly as a consequence of the general tendency toward syllabically balanced IPs, and partly as a general constraint on the Binarity condition on the minimal rhythmic unit we have seen at work in Beijing Mandarin (chapter 9).

As a consequence of condition (ii) on Cliticization, no IP-internal proclitics can exist. This is illustrated by the following example:

```
(47)     "give a bottle of grape wine to a friend"
         give  Cl   grape:wine   to  friend
         song [PING  pu-tao-jiu]  [HA peng-you]
         HM   o   ML.ML.MH   o  ML.MH
                  L.  HM.MH      HM.MH          Disyllabic TS, Polarity
         ---------------------------------------------------------------------
         (σ   σ)  (σ   σ   σ) / (σ   σ   σ)
                  L  HM. MH  o  HM.MH
    i.   o    o   L  HM.MH / o  HM.MH           Tonic Prominence
         ---------------------------------------------------------------------
         (σ   σ)  (σ   σ   σ   σ) (σ   σ)
                  L  HM.MH  h                   Spread
    j.   o    o   L  HM.MH  h   o   o           Tonic Prominence
         ---------------------------------------------------------------------
         (σ   σ)  (σ   σ   σ)  (σ   σ   σ)
    k.   o    o   L  HM.MH  o   o   o           Tonic Prominence = *
```

A sentence like this one may be cut up into two phrases, or spoken in one single breath group without an intonation break. Reading (47-i) represents the first option. The preposition HA "to," being IP-initial, patterns like a pretonic clitic, and therefore takes the default L. However, when the sentence is organized as a single IP, as in (47-j), the same preposition HA must cliticize leftwards, with predictable phonetic consequence: being an enclitic, it assimilates to the H of the host *pu-tao-jiu* "grape wine" via Spread. An IP-internal proclitic HA carrying a default L (annotated as [o]) is ungrammatical, as shown in reading (k).

Recall that the Host-Clitic rule creates a prosodically "autonomous" item out of a sequence of function words, provided that they stand in close structural relations (basically as immediate constituents). Here is a sentence beginning in a long string of non-lexical items.

(48) "That one (there) is taller than this one (here) by three feet"
that Loc RC than this Loc RC tall three foot[20]
[[*XU TOU*] *GI*] [*BI* [[*KI TOU*] *GI*]] [*gao* [*san chi*]]
MH.ML.HM.MH Lq.ML.HM M M Lq
(σ σ) (σ σ) . . . Host-Clitic

---

i. (σ σ σ σ)/(σ σ σ) Cliticization, IP
MH. o o o Lq. o o Clitic Reduction
MH h h h Lq. l l Spread

---

j. (σ σ σ)/ (σ σ σ σ) Cliticization, IP
MH o o o Lq. o o Clitic Reduction
MH h h o Lq. l l Spread

The Host-Clitic rule turns [*XU tou*] "that Loc, there" and [*KI tou*] "this Loc, here" into a unit, phonologically parsed as host + clitic. Under such conditions, *XU* and *KI* become the tonic nuclei around which p-words are organized. Given the syntactic bracketing, the relative clause marker *GI* can function only as an enclitic; *BI* "than," on the other hand, can cliticize in either direction, yielding two alternative phrasings (48-i, j). Clitic Reduction and Spread produce the attested readings. What is worth noting is that no other pair of clitics in (48) may serve as the tonic nucleus of a p-word, and hence an IP. None of the following phrasings is acceptable:

(49) that Loc RC than this Loc RC
[[*XU TOU*] *GI*] [*BI* [[*KI TOU*] *GI*]] . . .
MH.ML.HM MH. Lq.ML HM
(ML. o)
(HM. o)
(MH. o)
(ML. o)

6 Tonic prominence

Within each IP there is one and only one tonic nucleus, namely a lexical unit (or a derived host + clitic construction via Host-Clitic), that carries tonal specification – all other syllables being atonic, and either becoming associated with the pitch level of the tonic host (by Spread) or else eventually emerging with a L by Default. We now turn to the business of identifying the location of the tonic nucleus.

[20] Wenzhou *GI* is equivalent to Beijing Mandarin *de* and has the same grammatical function as a nominalizer, or a relative clause marker (RC) preceding a nominal head. In the illustrative example given here, we have a headless relative construction.

Within the p-word, the lexical host is the tonic nucleus. At the phrase level, the relative prosodic prominence is determined jointly by the principles indicated below. Although there is some ranking indeterminacy (see discussion below), the principles are listed approximately in descending order of importance:

(50) Tonic Prominence
  a. Stressability Hierarchy: adjunct > argument > head
  b. WSP (Weight-to-Stress Principle): stress the longer or more complex constituent.
  c. Right prominence: stress the rightmost prosodic constituent.

Curiously enough, the observed facts about Wenzhou point to a stressability hierarchy which differs in some respects from Selkirk's (1984) hypothesis, according to which argument > head > adjunct stand in a descending order of prosodic prominence. Typical examples include *FROST bitten, GERM resistant* (argument–head) vs. *lily WHITE, wafer THIN* (adjunct–head).[21] Kindred in spirit is Cinque's (1993) "universal stress rule" which embodies the following order of preference for stress placement: Comp > head > Spec. In contrast, Duanmu's (1990a; cf. 1995) Non-Head Stress principle, does not distinguish between arguments and adjuncts. Evidence in support of the adjunct > argument > head stress hierarchy is presented in section 6.1.

6.1 Stressability hierarchy

We have argued at some length in chapter 6, section 5, that the WSP (Prince 1990; cf. Peak-Prominence constraint, Prince and Smolensky 1993) must be extended beyond mora count in order to handle accent placement in New Chongming. Right Prominence manifests itself at the word level as well: within a word, the rightmost three syllables constitute the tonic nucleus. It remains to show the Stressability Hierarchy based on grammatical relations. Observe that in (51) we have an argument–head or an adjunct–head construction, while in (52) we see a head–argument or a head–adjunct sequence. In each case it is the non-head that carries the distinctive tone, while the atonic head (the verb) emerges with the default L, regardless of word order.

(51) a.	*Xiao-ming you-yong*	"Xiao-ming swims"
	(MH ML) (ML MH)	base tone, p-word
	HM ML	Disyllabic TS
	HM ML o o	Tonic Prominence

[21] Capitalization indicates stress.

		foreign shoe	
	b.	*wai-guo pi-xie*	"imported shoes"
		(L Lq) (ML.M)	base tone, p-word
		HM.Lq	Disyllabic TS
		HM.Lq o o	Tonic Prominence
(52)	a.	*can-jia kao-shi*	"take exam"
		(M.M) (MH.HM)	base tone, p-word
		HM.ML	Disyllabic TS
		o o HM.ML	Tonic Prominence
		tall three feet	
	b.	*gao san chi*	"taller by three feet"
		(M) (M Lq)	base tone, Lexicalized Phrases, p-word[22]
		HM.Lq	Disyllabic TS
		o HM.Lq	Tonic Prominence

In the case of (53) we have a typical SVO clause. The Stressability Hierarchy, therefore, does not discriminate between the subject and the object NP. By default, it is Right Prominence that picks the last argument as the locus of tonic prominence.

(53)		Xiao-ming read novel	
		Xiao-ming kan xiao-shuo	"Xiao-ming reads novels"
		(MH ML) / (L) (MH. Lq)	base tone, p-word, IP
		HM. ML HM. Lq	Disyllabic TS
	i.	HM. ML o HM. Lq	Tonic Prominence
		---	---
		(MH ML) (L) (MH. Lq)	
		HM. Lq	Disyllabic TS
	j.	o o o HM. Lq	Tonic Prominence

The examples cited so far are neutral between arguments and adjuncts. (54) provides the critical evidence for the preference to stress adjuncts over arguments. The adjunct > argument hierarchy is robust enough to override both the WSP and the default Right Prominence, at least in this case. Where syllable count does not favor overwhelmingly one over the other prosodic constituent as in this particular case (three vs. four syllables), there is a distinct preference for placing the tonic prominence on the adverbial adjunct *leng-jing* GI "calmly," despite the fact that the VP *can-jia kao-shi* "take exam" is both (marginally) longer and more complex structurally, and occupies the rightmost position.

[22] Notice that Lexicalized Phrases turns a number + nominal expression into a phonological word in (52b), so that *san chi* "three feet" undergoes Disyllabic TS like a lexical compound rather than a clitic group or a true phrasal construction.

(54) "calmly take exam"

calm -ly[23]	take	exam		
leng-jing	*GI*	[*can-jia*	*kao-shi*]	
(MH.MH.o)	(M M)	(MH.HM)		base tone, p-word
HM.MH o				Disyllabic TS
HM.MH. h				Spread
HM.MH. h o	o	o	o	Tonic Prominence

6.2 *Syllable count*

The situation is reversed where a more drastic imbalance in length and structural complexity prevails. This situation is illustrated by the following example.

(55)

calmly	take	foreign-language			
leng-jing *GI*	[*can-jia*	[*wai-yu*	*kao-shi*]$_{NP}$]$_{V'}$		
(MH.MH.o)	(M M)	(L.MH)	(MH.HM)		p-word
		HM.MH			Disyllabic TS
o o o	o o	HM.MH	o	o	Tonic Prominence

In terms of both syllable count and structural complexity, the constituent labeled as V′ makes a legitimate claim on tonic prominence by virtue of WSP. In this case, WSP seems to override Stressability Hierarchy. Within V′ it is the NP argument that dominates. By the same token, within the NP, it is the prenominal specifier *wai-yu* "foreign language" that bests the head *kao-shi* "exam" as the locus of prominence. In short, *wai-yu* emerges as the tonic nucleus for the entire sentence. Needless to say, if the sentence were to break up into two IPs, both *leng-jing* *GI* "calmly" and *wai-yu* "foreign language" would constitute the tonic nuclei for the two IPs.

The ranking indeterminacy between WSP and Stressability Hierarchy is further illustrated by simply introducing a modifier into the adverbial adjunct:

(56)

quite		calmly		take		foreign:lang				examination
[*xiang-dang*		*leng-jing*	*GI*]	[*can-jia*		[*wai-yu*		*kao-shi*]]		
(M	M)	(MH.MH	o)	(M	M)	(L.MH)				(MH.HM)
M	M	o	o	o	o	o	o	o	o	o

The addition of *xiang-dang* "quite, considerably" reduces the imbalance in syllable count and structural complexity. As a consequence, the adverbial phrase "regains" its prosodic prominence.

[23] For our purposes, I parse *GI* simply as an adverb-forming clitic.

6.3 Emphasis, contrast

Tonic Prominence embodies the general principles pertaining to the default readings. Not surprisingly, focus, and other semantic and pragmatic factors, can override the default assignment of tonic prominence, with the result that often the same phrase may have alternative phonetic realizations typically associated with different semantic interpretations. Here are some examples.

(57)		not eat rice	
		BU chi fan	
		MH.Lq.L	base tone
		(o Lq) (L)	base tone, Clitic Reduction, p-word
	i.	o o L	Tonic Prominence
		---	---
		(MH)(Lq)(L)	base tone
	j.	MH o o	Focus on *BU*
(58)		angry die	
		qi si	
		(HM)(MH)	base tone, p-word
	i.	o MH	Tonic Prominence
		---	---
	j.	HM o	Focus on *qi*

The negative *BU* in (57), being a proclitic, is normally atonic; however, under emphatic reading, it can carry the tonic nucleus. Thus, (57-i) is appropriate, for instance, as a statement about the dietary habits of the northerners as opposed to the southerners; (57-j), on the other hand, suggests an emphatic negation, for example, when a child refuses to eat in a temper tantrum. As for (58-i), the normal case is for the resultative complement *si* "to die" to assume prosodic prominence over the head *qi* "angry"; the resulting tonal contour [o.MH] has the literal or figurative interpretation of "die of anger." In contrast, (58-j) places the emphasis on *qi* "angry," in the process "bleaching" *si* "to die" of its etymological meaning, so that the phonetic realization [HM o] only has a conventional (figurative) interpretation of being simply "angry."

Concluding remarks

Back in January 1986, at the Conference on Languages and Dialects of China, held in Oakland, California (under the sponsorship of the Wang Institute of Graduate Studies), I presented a paper entitled "An overview of tone sandhi phenomena across the Chinese dialects" (Chen 1991a). Looking back across the distance of a decade, I realize how much the field has advanced on both empirical and theoretical fronts. Our store of empirical knowledge about the range and diversity of tone sandhi phenomena has been enriched enormously by recent dialectological reports. Complementing the breadth of descriptive coverage, I have attempted an in-depth analysis of a number of selected dialects including Beijing, Tianjin, Boshan (Mandarin), Shanghai, Danyang, Nantong, New Chongming, Zhenhai, Wenzhou (Wu), Changting (Hakka), and Xiamen (Min). In many cases the nature of the questions we entertained forced us to dig for new kinds of facts not previously noticed or reported even regarding well-known dialects (e.g. Beijing Mandarin or Xiamen). One of the more startling discoveries is the fact that some Chinese dialects (notably New Chongming) have developed a full-blown syndrome of classic accentual systems, as diagnosed by such standard criteria as culminativity (one tone-bearing syllable per word/unit), leveling (all tonal categories reduced uniformly to a plain H), edgemostness (head-terminal prominence), tonic clash (resolved by either de-toning or tonic shift), and the Weight-to-Accent principle (appropriately extended beyond the traditional distinction between light and heavy syllables).

From the theoretical perspective, we have examined a number of issues against the full range of tone sandhi phenomena presently known to us. In the *Preface* I alluded to three leitmotifs that underscore tone sandhi studies. First is the nature and representation of tone. I once argued for a more parsimonious tonal configuration (basically that of Yip 1989), according to which the terminal nodes of a contour tone are dependants of the register node. I am now leaning toward a more fully articulated

tonal geometry suggested by Bao (1990a), where register and contour constitute two orthogonal elements that can be manipulated independently of each other. The crucial evidence comes from the rightward contour shift in Zhenhai, on the one hand, and the leftward register spread in Chaozhou, on the other. I further brought forth evidence for whole tone displacement, whereby a high-falling tone moves leftwards to the beginning of a word across an indefinite span in Wenzhou. Under the assumption that only constituents move together, the left dislocated contour tone must form a single phonological unit.

The second recurrent theme concerns the scope of tone sandhi, namely the prosodic domains within which sandhi processes operate. Not surprisingly, we see that different languages circumscribe the scope of sandhi rules differently, ranging from sublexical stress-foot through phonological word and phonological phrase to intonational phrase. That is not to say that it is either simple or easy to circumscribe these prosodic entities. For instance, the stress-foot is constructed by means of standard metrification rules – except that clash resolution is curiously asymmetric: we have documented a robust tendency across a number of Wu dialects to tolerate stress clash in one direction but not in the other. The end-based definition of the phonological phrase (or tone group) in Xiamen proposed by Chen (1987a) has left several problems unresolved, chief among which is the inelegant mix of structural (right or left maximal projection) and functional properties (argument vs. adjunct phrases). In chapter 10 I examine two recent proposals, one based on the notion of "domain c-command," the other on "lexical government," and find the latter to be a far more promising approach, despite some residual problems arising out of the interpretation of NP as DP (therefore functionally governed by D, the determiner).

Something of a surprise is the finding that the prosodic domain of tone sandhi in Mandarin Chinese, which I refer to as the minimal rhythmic unit (MRU), turns out to be an oddity. The MRU can be a sublexical fragment, an entire word, a whole phrase, a multiclausal construction, or a non-constituent altogether. As such, it does not fit comfortably into the conventional prosodic hierarchy of {foot, p-word, clitic group, p-phrase, intonation phrase}; instead, it stands apart as a prosodic unit sui generis. If anything, MRU finds its closest analog in the foot as a unit of poetic scansion.

Finally, the question regarding how elementary tonal processes operating on two-tone substrings interact to produce polysyllabic sandhi forms has not been addressed adequately in the past. This was the topic taken

up in chapters 3–4. There I argued that the most revealing account of sandhi patterns in such languages as Tianjin must exploit constraints of a derivational nature such as Temporal Sequence and No-Backtracking. It was found that both output (Well-formedness, *Complex, etc.) and faithfulness conditions underdetermine the choice among output candidates. The derivational account I have proposed is compatible only with a conception of Optimality Theory that has the following properties: (i) GEN operates seriatim, effecting one single modification at a time; (ii) EVAL weighs alternative derivational paths against a constraint hierarchy; (iii) the relevant constraints may be representational or derivational in nature. This "harmonic serialism" was briefly alluded to in Prince and Smolensky (1993:79f.) as an alternative to the prevailing "parallel" conception of Optimality Theory. In-depth analysis of other sandhi phenomena may yet produce further evidence of a non-trivial derivational component in modeling human language.

Bibliographical appendix Tone sandhi across Chinese dialects

Note: this selective sample includes only descriptive works bearing directly on tone sandhi in specific Chinese dialects or dialect groups, excluding for the most part generic dialect surveys and tonological literature.

In compiling this bibliography we have consulted the following:

1967 William S-Y. Wang. *Bibliography of Chinese Linguistics*. Current Trends in Linguistics 2. Mouton.

1979 *Zhongguo Yuyanxue Lunwen Suoyin* [An Index to Essays on Chinese Linguistics]. Compiled by the Institute of Linguistics, Chinese Academy of Sciences. Reissued by Joint Publishing, Hong Kong. Updated edition, Beijing: Shangwu, 1983.

1981 Paul Fu-mien Yang. *Chinese Dialectology: a Selected and Classified Bibliography*. Hong Kong: Chinese University Press.

1985 Alain Lucas. *Linguistique chinoise – Bibliographie: 1975–82*. Paris: Editions Langages Croisés, Centre de Recherches Linguistiques – Asie Orientale.

1993 Jian-min Nie and Qi Li. *Hanyu Fangyan Yanjiu Wenxian Mulu* [A Bibliographical Index to Chinese Dialectology]. Nanjing: Jiangsu Jiaoyu Chubanshe.

Abbreviations used:

BIHP	*Bulletin of the Institute of History and Philology* (Academia Sinica (Taipei))
CAAAL	*Computer-aided Analysis of Asian and African Languages* (Tokyo)
CLAO	Cahiers de Linguistique – Asie Orientale (Paris)
CLS	Chicago Linguistic Society (proceedings)
FY	*Fangyan* (Beijing)
ICCL	International Conference on Chinese Linguistics (proceedings)
ISCLL	International Symposium on Chinese Languages and Linguistics (proceedings)
JCL	*Journal of Chinese Linguistics* (Berkeley)
NACCL	North-American Conference on Chinese Linguistics (proceedings)
YJY	*Yuyan jiaoxue yu yanjiu* (Beijing)
YY	*Yuyan yanjiu* (Wuhan)
ZGYW	*Zhongguo yuwen* (Beijing)

This selected bibliography consists of the following sections:

1. Doctoral theses
2. Cross-dialectal
3. Mandarin dialects
4. Jin dialects
5. Wu dialects
6. Min dialects
7. Hakka dialects
8. Yue dialects

1 Doctoral theses

1968 C-C. Cheng. *Mandarin Phonology*. University of Illinois at Urbana-Champaign. [Published as *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton, 1973]

1969 N. Woo. *Prosody and Phonology*. Massachusetts Institute of Technology.

1980 M. Yip. *The Tonal Phonology of Chinese*. Massachusetts Institute of Technology. [Garland Press]

1983 M. Wright. *A Metrical Approach to Tone Sandhi in Chinese Dialects*. University of Massachusetts.

1985 M. Chan. *Fuzhou Phonology: a Non-linear Analysis of Tone and Stress*. University of Washington.

1986 C-L. Shih. *The Prosodic Domain of Tone Sandhi in Chinese*. University of California, San Diego.

1987 T. Hung. *Syntactic and Semantic Aspects of Chinese Tone Sandhi*. University of California, San Diego. [Indiana University Linguistics Club]

1987 J.S-F. Liu. *Tone Sandhi in the Chinese Dialects*. University of Texas, Austin.

1988 T-C. Du. *Tone and Stress in Taiwanese*. University of Illinois at Urbana-Champaign.

1988 Z-S. Zhang. *Tone and Tone Sandhi in Chinese*. Ohio State University.

1988 H-B. Lin. *Contextual Stability of Taiwanese Tones*. University of Connecticut.

1989 R-F. Chung. *Aspects of Kejia Phonology*. University of Illinois at Urbana-Champaign.

1990 Z-M. Bao. *On the Nature of Tone*. Massachusetts Institute of Technology.

1990 S. Duanmu. *A Formal Study of Syllable, Tone, Stress, and Domain in Chinese Languages*. Massachusetts Institute of Technology.

1991 Y-C. Hsiao. *Syntax, Rhythm and Tone: a Triangular Relationship*. University of California, San Diego.

1992 L.M-C. Chang. *A Prosodic Account of Tone, Stress, and Tone Sandhi in Chinese Languages*. University of Hawaii.

1992 H. Lin. *On the Nature of Mandarin Tone and Tone Sandhi*. University of Victoria.

1992 H-M. Zhang. *Topics in Chinese Phrasal Tonology*. University of California, San Diego.

1993 B.X-P. Ao. *Phonetics and Phonology of Nantong Chinese*. Ohio State University.
1994 H-C. Hsu. *Constraint-based Phonology and Morphology: a Survey of Languages in China*. University of California, San Diego.
1994 J.S-C. Tsay. *Phonological Pitch*. University of Arizona.
1994 X-N. Zhu. *Shanghai Tonetics*. Australian National University.
1998 L. Chan. *Fuzhou Tone Sandhi*. University of California, San Diego.

2 Cross-dialectal

1983 L. Sagart. Contextual tone sandhi in Chinese dialects. Ms., Centre de Recherches Linguistiques sur l'Asie Orientale. Paris.
1984 P-H. Ting. Wuyu shengdiao zhi yanjiu [An investigation on tone in Wu dialects]. *BIHP* 55: 605–638.
1984 Z-H. Xiong. Zenme qiuchu liang zi zu de liandu biandiao guilyu [How to infer disyllabic tone sandhi]. *FY* 102–108.
1985 C-F. Lien. A critical survey of the tone behavior in Chinese dialects. Ms., UC Berkeley.
1986 (Wu Tai). Guanyu liandu biandiao de zai renshi [Another look at tone sandhi]. *YY* 1:1–6.
1986 A. Yue-Hashimoto. Tonal flip-flop in Chinese dialects. *JCL* 14:161–183.
1987 A. Yue-Hashimoto. Tone sandhi across Chinese dialects. In *Wang Li Memorial Volumes*, English volume, 445–474. [Paper originally presented at the 13th Sino-Tibetan Conference, University of Virginia, 1980.]
1988 W.L. Ballard. *The History and Development of Tonal Systems and Tone Alternations in South China*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
1991 M. Chen. An overview of tone sandhi phenomena across Chinese dialects. In *Languages and Dialects of China*, ed. by W. Wang, 113–158. [Paper originally presented at the conference by the same name, held in Oakland, California, 1986.]
1991 T.L. Mei. Tone sandhi and morphological relevance. In *Languages and Dialects of China*, ed. by W. Wang, 456–473. [Paper originally presented in 1986.]
1992 H. Hirayama. Cong li shi guan dian lun wu yu bian diao he Beijing qing sheng de guan xi [A diachronic perspective on the relationship between Wu tone Sandhi and Beijing neutral tones]. *ZGYW* 243–252.
1997 H-L. Soh. The syntax of phonological phrasing in Shanghai and Hokkien. Ms., Massachusetts Institute of Technology.

3 Mandarin dialects

3.1 Beijing Mandarin

1947 X-R. Zhang. Guoyu qingzhong yin zhi bijiao [Contrasting strong and weak syllables in Mandarin]. *Guowen yuekan* 57:11–15.

1957 T. Lin. Xiandai hanyu buzuyu li de qingsheng xianxiang suo fanying chulaide yufa he yuyi wenti [Syntactic and semantic issues as reflected in the neutral tone phenomena associated with complement structures in modern Chinese]. *Beijing daxue xuebao* 2.

1962 T. Lin. Xiandai hanyu qingzhongyin he jufa jiegou de guanxi [The relationship between neutral tone and sentence structure in modern Chinese]. *ZGYW* 301–11.

1963 J-Y. Fan. Dongci he quxiang xing houzhi chengfen de jiegou fenxi [Structural analysis of verbs and directional complements]. *ZGYW* 136–160.

1963 T. Lin. Beijing hua de liandu yinbian [Phonological processes in Beijing connected speech]. *Beijing daxue xuebao* 6:19–28.

1963 W. Wang. Mandarin phonology. *Project on Linguistic Analysis* 6:1–6.

1964 Y. Shen. Some experiments on Chinese tone sandhi. *Proceedings of the 5th International Congress of Phonetic Sciences*, Münster, 525–527.

1965 W. Wang. TRIP report. Ms., Ohio State University.

1966 R.L. Cheng. Mandarin phonological structure. *Journal of Linguistics* 2:135–158.

1970 C-C. Cheng. Domain of phonological rule application. In *Studies Presented to Robert B. Lees by his Students*, ed. by Sadock and Vanek, 39–60.

1972 F. Dow. A discussion on tone sandhi problems in Chinese. *Journal of International Phonetic Association* 2:13–19.

1973 C-C. Cheng. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton (= Monograph on Linguistic Analysis).

1976 A-C. Ho. Mandarin tones in relation to sentence intonation and grammatical structure. *JCL* 4:1–13.

1977 Y-L. Lee. Correlation among attitudinal factors, speed, and tone sandhi in Chinese. *Studies in Linguistic Science* 7:129–141.

1977 T-L. Mei. Tones and tone sandhi in 16th century Mandarin. *JCL* 5:237–260.

1979 J-W. Lu. Guanyu sange shangsheng liandu biandiao wenti shangque. Fu Hu Bing-zhong, sange biandiao ji qi jiaoxue [Comments on three questions regarding the tone sandhi of "shang" tone. With an appendix by B-Z. Hu on the changed tone of the third tone and related pedagogical issues]. *YJY* 2:47–59.

1980 Y-Zh. Gao. Beijing hua de qingsheng wenti [The problem of neutral tone in the Beijing dialect]. *YJY* 2:82–98.

1980 F-H. Liu. Mandarin tone sandhi: a case of interaction between syntax and phonology. Paper presented at the summer institute of the Linguistic Society of America, Albuquerque, New Mexico (Ms., University of Iowa).

1981 N-L. Jing (transliteration). Zai biandiao fangmian de yige xiao wenti [A minor question regarding tone sandhi]. *YJY* 4:91–99. (English original in *Journal of Chinese Language Teachers Association* 6:41–57 (1971).)

1982 Y-J. Jin. Beijing hua "shang sheng + qing sheng" de biandiao guilu [Tone sandhi rules governing T3 + neutral tone expressions in Beijing]. In *Yuhai Xintan* [New explorations in linguistics], vol. I, Shandong Jiaoyu Chubanshe.

1982 Z. Meng. Yixie yu yufa youguan de Beijing hua qingsheng xianxiang [Some syntax-related neutral tone phenomena in the Beijing dialect]. *Yuyanxue luncong* 9:25–59.
1982 Z-J. Wu. Putonghua yuju zhong de shengdiao bianhua [Sentence level tone change in Putonghua]. *ZGYW* 439–450.
1983 M. Hoa. Une particularité du troisième ton sur le plan accentuel en pékinois. Ms., Paris.
1983 M. Li. Guanyu qingsheng yusu qian de shangsheng biandiao wenti [On tone sandhi before atonic morphemes]. *YJY* 2:95–101.
1983 J-W. Lu. Guanyu shangsheng he qingsheng de biandiao wenti [Concerning the tone sandhi of "shang" and "neutral" tones]. *Zhongzhou xuekan* 4.
1983 C-L. Shih. Foot as a prosodic unit in Mandarin. Ms., University of California, San Diego.
1984 W. Cai. Shangsheng liandu biandiao xintan [A re-examination of shangsheng tone sandhi]. *Yuwen yuandi* 5.
1984 C-Y. Chen. Neutral tone in Mandarin: phonotactic description and the issue of the norm. *JCL* 12:299–333.
1984 Q-Y. Gao. Hanyu putonghua de yufa zhongyin [Grammatical accent in Putonghua]. *YJY* 2.
1984 M-C. Lin, Jingzhu Yan, and Guohua Sun. Beijing hua lianzizu zhengchang zhongyin de chubu shiyan [Experimental explorations in the normal accentuation of disyllabic constructions in Beijing]. *FY* 57–73.
1984 Z. Yue. Ye tan sange shangsheng liandu biandiao de wenti [Three aspects of shangsheng tone sandhi]. *YJY* 4.
1985 D.A. Beattie. *Mandarin Chinese Tone Sandhi and the Interaction of Phonology and Syntax*. M.A. thesis, University of Toronto.
1985 T. Lin. Beijing hua qusheng liandu biandiao xin tan [A re-examination of tone sandhi in connection with the going tone in the Beijing dialect]. *ZGYW* 99–104.
1985 S. Poteet. Tone sandhi and the phonological structure of Mandarin. Ms., University of California, San Diego.
1985 Z-J. Wu. Putonghua sanzizu biandiao guilü [Patterns of trisyllabic tone sandhi in Putonghua]. *Zhongguo Yuyan Xuebao* 2:70–92.
1986 J-L. Wang. On tone sandhi in Chinese (Putonghua). In N. Hasegawa and Kitagawa, eds., *University of Massachusetts Occasional Papers in Linguistics* 11:311–330.
1987 L. Cheng. Derived domains and Mandarin third tone sandhi. *CLS*, Parasession, 16–28.
1988 S. Duanmu. Two issues in Mandarin tone. Ms., Massachusetts Institute of Technology.
1989 S. Duanmu. Cyclicity and the Mandarin third tone sandhi. Ms., Massachusetts Institute of Technology.
1989 S. Milliken. Why there is no third tone sandhi rule in Standard Mandarin. Paper presented at the Tianjin International Conference on Phonetics and Phonology.

1989 S. Speer, C.L. Shih, and M. Slowiaczek. Prosodic structure in language understanding: evidence from tone sandhi in Mandarin. Ms., Bell Laboratories, Lucent Technologies.
1992 M. Chen. Phonological diagnostics of morphological structure. In Chiuyu Tseng (ed.), the *Proceedings of the Third International Symposium on Chinese Languages and Linguistics* (Tsinghua University, Taiwan, July 1992).
1992 R-F. Chung and Y-C. Chen. Influence of sex, age, and education on Mandarin tone sandhi. Paper presented at the 1st ICCL (Singapore, June 1992).
1992 H. Lin. *On the nature of Mandarin tone and tone sandhi*. Ph.D. dissertation, University of Victoria.
1992 J. Packard. Why Mandarin morphology is stratum-ordered. Paper presented at the 4th NACCL (Ann Arbor, May 1992).
1992 G. Shyi et al. The tone sandhi rule and short-term memory. Paper presented at the 1st ICCL (Singapore, June 1992).
1992 R. Sproat and C.L. Shih. Mandarin 2nd tone sandhi is not a phonological rule. Paper presented at the 1st ICCL (Singapore, June 1992).
1995 N. Zhang. Phrasal stress and Mandarin third tone sandhi. Ms., University of Toronto.
1997 C.L. Shih. Mandarin third tone sandhi and prosodic structure. In J.L. Wang and N. Smith, eds., *Studies in Chinese Phonology*, Berlin and New York: Mouton de Gruyter, 81–124.

3.2 Tianjin

1985 X-J. Li and S-X. Liu. Tianjin fangyan de liandu biandiao [Tone sandhi in the Tianjin dialect]. *ZGYW* 76–80.
1985 M. Chen. Tianjin tone sandhi: erratic rule application? Ms., University of California, San Diego.
1986 M. Chen. The paradox of Tianjin tone sandhi. *CLS* 22:98–154.
1987 M. Chen. Introductory remarks to A symposium on Tianjin tone sandhi. *JCL* 15:203–227.
1987 D.S. Davison. Tianjin Tone Sandhi Phonetics. 20th International Conference on Sino-Tibetan Languages and Linguistics (Vancouver).
1987 T. Hung. Tianjin tone sandhi: towards a unified approach. *JCL* 15:274–305.
1987 F. Tan. Tone sandhi in the Tianjin dialect. *JCL* 15:228–246.
1987 Z-S. Zhang. The paradox of Tianjin: another look. *JCL* 15:247–273.
1988 K. Field. *Tonal strength in Tianjin tone sandhi*. M.A. thesis, Michigan State University.
1988 F. Shi. Shi lun Tianjin de shengdiao ji qi bianhua [On Tianjin tones and their sandhi behavior]. *ZGWY* 351–360.
1990 F. Shi. Zai lun Tianjin de shengdiao ji qi bianhua [Tianjin tones and their sandhi behavior – revisited]. *Yuyan Yanjiu* 2 [Reprinted in S. Feng, *Studies in Tone and Stops*. Beijing University Press, 101–117.]
1997 S. Milliken, G-P. Zhang, X-Y. Zhang, Z-Q. Li, and Y. Lü. Resolving the paradox of Tianjin tone sandhi. In J. Wang and N. Smith, eds., *Studies in Chinese Phonology*, Berlin and New York: Mouton de Gruyter, 53–80.

3.3 Other Mandarin dialects

1939 F. Giet. *Die Töne des Südshantung-Dialekts in Wortverbindungen*. Yenchowfu: Regionshaus Taikiachwangbei Tsining.
1950 F. Giet. *Zur Tonität Nordchinesischer Mundarten*. (Studia Instituti Anthropos v.2.) Wien and Mödling: Missionsdruckerei St.Gabriel.
1961 F-Q. Sun. Xian fangyan de biandiao [Tone sandhi in Xian]. *ZGYW* 28–29.
1979 W. He. Huojia fangyan de liandu biandiao [Tone sandhi in Huojia (Henan)]. *FY* 122–136.
1982 R. Iwata. Lianyungang shi fangyan de liandu biandiao [Tone sandhi in the city of Lianyungang]. *FY* 285–296.
1983 S-Y. Zhang. Taiping (Xianyuan) fangyan lianzizu de liandu bian-diao [Disyllabic tone sandhi in Taiping (Xianyuan, Anhui)]. *FY* 170–174.
1984 H-Q. Chen. Xiaogan fangyan de yunmu biandu [Final sandhi in Xiaogan]. *YY* 1.
1985 S.H-N. Cheung. Zhenjiang fangyan de liandu biandiao [Tone sandhi in Zhenjiang]. *FY* 191–204.
1985 X-Q. Su. Xuzhou fangyan de liandu biandiao [Tone sandhi in Xuzhou]. *Xuzhou Shifan Xueyuan Xuebao* 1:154–165.
1987 S-Y. Zhang. Yinchuan fangyan sanzizu de liandiao [Sandhi tone in the Yinchuan dialect]. *FY* 134–139.
1989 D.S. Davison. *Lexical prosodies of Mandarin*. Ph.D. dissertation, University of California, Berkeley.
1989 C-C. Zhang. Shangxian fangyan liangzizu de liandu biandiao [Disyllabic tone sandhi in the Shangxian dialect]. *FY* 208–212.
1990 Q-Z. Yang. *Lijin fangyan zhi*. Beijing: Yuwen Chubanshe.
1991 D.S. Davison. Tone Sandhi and the dissimilation of phonation types: reflexes of the Beijing Mandarin 3rd tone sandhi rule in Northern Chinese dialects. ISCLL- 2 (Taipei), 180–200.
1991 Y. Jiang. Shuozhou shi Shuochengqu fangyan de liandu biandiao [Tone sandhi of the urban Shuozhou dialect]. *FY* 255–258.
1993 S-K. Dong. Shandong Yanggu fangyan de biandiao [Tone change in the Yanggu dialect of Shandong]. *FY* 57–58.
1993 Q-J. Huang. Hubei Yangxin fangyan de xiaocheng yinbian [Morphophonemics of the Yangxin dialect of Hubei]. *FY* 59–64.
1993 Z-Y. Qian. *Boshan fang yan yan jiu*. Beijing: Shehuikexue wenxian chubanshe.

4 Jin dialects

1980 J-Y. Hou. Pingyao fangyan de sizige shili [Quadrisyllabic constructions in Pingyao]. *Yuwen yanjiu* 1.
1980 J-Y. Hou. Pingyao fangyan de liandu biandiao [Tone sandhi in Pingyao]. *FY* 1980, 1–14; 1982, 7–14 and 85–99.
1981 J-Y. Hou. Pingyao fangyan de dongbushi [Verb complements in Pingyao]. *Yuwen yanjiu* 1.

1982 J-Y. Hou. *Pingyao Fangyan Jianzhi* [A Sketch of the Pingyao dialect]. Institute of Social Sciences, Shanxi Province. Supplement to Yuwen yanjiu.
1983 J-Y. Hou. Changzhi fangyan jilüe [A sketch of the Changzhi dialect (Shanxi)]. *FY* 260–275.
1988 Y-M. Shen. A tentative hypothesis regarding trisyllabic tone sandhi in Pingyao. Ms., University of California, San Diego.
1991 Q-H. Meng. Yangqu [Shanxi] fangyan lianzizu liandu biandiao [Disyllabic tone sandhi in the Yangqu dialect of Shanxi]. *FY* 50–55.
1994 W-T. Tsai. On matching syntax – phonology mismatches. *MIT Working Papers in Linguistics* 21:355–371.
1997 J. Zhang. *Duration in the tonal phonology of Pingyao Chinese*. M.A. thesis, University of California, Los Angeles.

5 Wu dialects

5.1 Shanghai

1972 M. Sherard. *Shanghai Phonology*. Dissertation, Cornell University.
1979 M. Sherard. Syntactic constraints on tone sandhi in Shanghai. *CAAAL* 10:23–38.
1980 M. Sherard. A Synchronic Phonology of Modern Colloquial Shanghai. *CAAAL* 5.
1980 E. Zee and I. Maddieson. Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *Glossa* 14:45–88 (earlier version appeared in *UCLA Working Papers in Phonetics* 45:93–129).
1981 T. Shen. Shanghai hua laopai xinpai de chabie [Differences between Old and New Shanghai]. *FY* 275–283.
1981 T. Shen. Laopai Shanghai fangyan de liandu biandiao [Tone sandhi in Old Shanghai]. *FY* 1981, 131–44; 1982, 100–114.
1981 B-H. Xu, Z-Z. Tang, and N-R. Qian. Xinpai Shanghai fangyan de liandu biandiao [Tone sandhi in New Shanghai]. *FY* 1981, 145–155; 1982, 115–128; 1983, 197–201.
1985 S-D. Jin. *Shanghai morphotonemics*. Indiana University Linguistics Club.
1985 T. Shen. Xinpai Shanghai hua shengdiao de diceng xingshi [The underlying representation of New Shanghai]. *YY* 2:85–101.
1986 Z-J. Lu. Tonal change: interplay between tone and tone sandhi, a case study of the Shanghai dialects. *Studies in Linguistic Sciences* 16:97–111.
1987 Z-J. Lu. Yi zhong shengdiao yanbian de moxing: Shanghai fangyan shengdiao yanbian yanjiu [A model of tonal development: a study of the tonal evolution of the Shanghai dialect]. *Xiandai Yuyan* 10:15–34.
1987 Z-J. Lu. Shanghai tones: a non-linear analysis. *Studies in the Linguistic Sciences* 17.2:93–113.
1987 T. Shen. The formation of tone groups in Shanghai. Ms., University of Massachusetts.
1988 E. Zee. Zizhu yinduan yinyunxue lilun yu Shanghai shengdiao de biandu [Autosegmental theory and Shanghai tone sandhi]. *ZGYW* 331–350.

1989 S. Duanmu. Tone in connected speech: a case study of Shanghai. Ms., Massachusetts Institute of Technology.
1990 E. Selkirk and T. Shen. Prosodic domains in Shanghai Chinese. In S. Inkelas and D. Zec, eds. *The phonology–syntax connection*, 313–338.
1992 S. Duanmu. *End-based theory, cyclic stress, and tonal domains*. CLS 28.
1993 P. Rose. A linguistic–phonetic acoustic analysis of Shanghai tones. *Australian Journal of Linguistics* 13:185–220.
1994. X-N. Zhu. *Shanghai tonetics*. Ph.D. dissertation, Australian National University.
1995 Z-M. Bao. Local tree geometry and the syntax–phonology interface. Ms., National University of Singapore.
1995 S-D. Jin. Shilun Shanghai fangyan de shengdiao yinxi [On the tonemic system of Shanghai]. In *Wuyu Yanjiu* ed. by E. Zee, Chinese University of Hong Kong, 195–215.
1997 S-D. Jin. Toward a systematic account of Shanghai tonal phonology. In J. Wang and N. Smith, eds., *Studies in Chinese Phonology*, Berlin and New York: Mouton de Gruyter, 125–156.

5.2 Suzhou

1979 X-L. Ye. Suzhou fangyan de liandu biandiao [Tone sandhi in Suzhou]. *FY* 30–46.
1979 X-L. Ye. Zai lun Suzhou fangyan shangsheng he yinqu de liandu biandiao [Suzhou shangsheng and yinqu tone sandhi revisited]. *FY* 306–308.
1979 J-M. Zhang. Suzhou de liandu biandiao taolun [Concerning Suzhou tone sandhi]. *FY* 304–306.
1982 Z-L. Xie. Suzhou liangzizu de liandu biandiao [Disyllabic tone sandhi in Suzhou]. *FY* 245–263.
1983 R-R. Liao. Suzhou danzi diao shuangzi diao de shiyan yanjiu [Experimental studies of monosyllabic and disyllabic tone patterns in Suzhou]. *YY* 2:41–83.
1983 N-R. Qian and R-J. Shi. Guanyu Suzhou fangyan liandu biandiao de yijian [Proposals concerning tone sandhi in Suzhou]. *FY* 275–285.
1983 P. Wang. Suzhou fangyan liangzizu de liandiao geshi [Disyllabic tone sandhi patterns in Suzhou]. *FY* 286–96.
1986 M. Hashimoto. Suzhou hua biandiao xianxiang zhi benzhi [The nature of Suzhou tone sandhi]. Paper presented at the 2nd conference on Wu dialects.
1987 P. Wang. Reanalysis of Suzhou phonology. *YY* 1:41–48.
1988 X-L. Ye 1988. *Suzhou Fangyanzhi* [The Suzhou dialect]. Shanghai: Jiangsu Jiaoyu chubanshe.
1995 T. Shen. The underlying representation of Suzhou tones. In *Wuyu Yanjiu*, ed. by E. Zee, Chinese University of Hong Kong, 129–144.
1996 L. Chan. Contour tones: tonal cluster vs. single melodic unit. Ms., University of California, San Diego.

5.3 Chongming

1979 H-Y. Zhang. Chongming fangyan de liandu biandiao [Tone sandhi in Chongming]. *FY* 284–302.
1980 H-Y. Zhang. Chongming fangyan sanzizu de liandu biandiao [Trisyllabic tone sandhi in Chongming]. *FY* 15–34.
1988 H-M. Zhang. A syntactic or prosodic domain? – on tone sandhi in Chongming. Ms., University of California, San Diego.
1990 M. Chen. Competing strategies and the harmonic principle. Ms., University of California, San Diego.
1997 M. Chen and H-M. Zhang. Lexical and postlexical tone sandhi in Chongming. In J. Wang and N. Smith, eds., *Studies in Chinese Phonology*, Berlin and New York: Mouton de Gruyter, 13–52.

5.4 Wenzhou/Wenling

1958 G-Q. Cao. Wenling hua rusheng biandiao tong yufa de guanxi [Entering tone sandhi and its relation to syntax in Wenling]. *ZGYW* 340–341.
1964 S-F. Zheng-Zhang. Wenzhou yinxi [The phonology of Wenzhou]. *ZGYW* 28–60.
1964 S-F. Zheng-Zhang. Wenzhou fangyan de liandu biandiao [Tone sandhi in Wenzhou]. *ZGYW* 106–152.
1973 W.L. Ballard. Wenzhou tone sandhi revisited. Paper presented to Sino-Tibetan Conference.
1978 R. Li. Wenling fangyan de bianyin [Morphophonemics in Wenling]. *FY* 96–103.
1979 R. Li. Wenling fangyan de liandu biandiao [Tone sandhi in Wenling]. *FY* 1–29.
1980 S-F. Zheng-Zhang. Wenzhou fangyan erweici de yuyin bianhua [Morphophonemics involving er-suffixation in Wenzhou]. *FY* 1980, 245–262; 1981, 40–50.
1987 W.L. Ballard. Implications of tone sandhi in Wenzhou and Wenling. In *The Wang Li Memorial Volume*, ed. by Chinese Language Society of Hong Kong. Hong Kong: Joint Publishing, 17–24.
1989 M. Chen. Wenzhou tone sandhi: compounds and clitic groups. Ms., University of California, San Diego.
1990 M. Chen. Intonational phrasing and tone sandhi in Wenzhou. Ms., University of California, San Diego.
1992 R. Li. Wenling fangyan de qing sheng. *FY* 1–8.

5.5 Other Wu dialects

1947 S-X. Lü. Danyang hua li de lian ci bian diao [Lexical tone sandhi in Danyang]. *Zhongguo wenhua yanjiu huikan* 7, Chengdu.
1953 G. Kennedy. Two tone patterns in Tangsic. *Language* 29:367–373.

1959 M-Y. Hu. Haiyan Tongyuan fangyan zhong biandiaoqun de yufa yiyi [The structural meaning of tone sandhi groups in Haiyan (Tong-yuan)]. *ZGYW* 372–6.

1959 F-T. Wang. Shaoxinghua jiyin [A phonetic description of Shaoxing]. *Yuyanxue Luncong* 3:73–136.

1979 C-R. Chen. Pingyang fangyan jilüe [A brief sketch of the Pingyang dialect]. *FY* 47–74.

1980 W.L. Ballard. On some aspects of Wu tone sandhi. *Journal of Asian–African Studies* 19:83–164.

1980 S-S. Lü. Danyang fangyan de shengdiao xitong [The tonal system of Danyang]. *FY* 85–122.

1982 W-Y. Pan. Qingtian fangyan liandu zhong de biandiao he bianyin xianxiang. [Morphophonemics and tone sandhi in the connected speech of Qingtian]. *Wuyu Yanjiu* 1.

1983 X-L. Ye. Wujiang fangyan shengdiao de zaidiaocha [Wujiang tones revisited]. *FY* 32–35.

1983 H-M. Ren. A "copying" type of tone sandhi in Wuxi Chinese. Proceedings of the 19th International Conference on Sino-Tibetan Languages and Linguistics, ed. by M. Chan. Seattle: University of Washington.

1984 W. Ballard. Wu, Min and a little Hakka lexical tone sandhi: right and left. *CLAO* 13:3–34.

1984 G-T. Fu. Wuyi fangyan de liandu biandiao [Tone sandhi in Wuyi]. *FY* 2, 109–127.

1985 N-P. Chan. Tone sandhi phenomenon in Ningponese. *FY* 15–27.

1987 S-X. Fang. Tone sandhi of disyllabic words in the Zhoushan dialect. *FY* 116–123.

1988 N.R. Qian. Wuyu shengdiao xitong de leixing ji qi bianqian [The typology and evolution of the tonal systems of Wu dialects]. Ms., Shanghai University.

1989 M. Chan and H-M. Ren 1989. Wuxi tone sandhi: from last to first syllable dominance. *Acta Linguistica Hafniensia* 21:35–64. [An earlier version appeared in *UCLA Working Papers in Phonetics* 63:48–70.]

1989 P. Rose. Phonetics and phonology of yang tone: phonation types in Zhenhai. *CLAO* 18:229–245.

1989 P. Rose. Tonology through acoustic phonetics – an analysis of disyllabic lexical tone sandhi in Zhenhai. Ms., Australia National University. To appear in B.H. Xu, ed., *Collected papers commemorating the 60th anniversary of the publication of Y.R. Chao's Studies in modern Wu dialects*.

1989 T. Toda. Acoustic quantification of tone spreading in Daishan and Mende. *CLAO* 18:63–80.

1990 P. Rose. Acoustics and phonology of complex tone sandhi. An analysis of disyllabic lexical tone sandhi in the Zhenhai variety of Wu Chinese. *Phonetica* 47:1–35.

1992 F. Shi. Wu Jiang fangyan sheng diao geju de fenxi. *FY* 189–194.

1993 B.X-P. Ao. *Phonetics and phonology of Nantong Chinese*. Ph.D. dissertation, Ohio State University.

1993 Z-M. Chen. Ningbo fang yan sheng diao bian yi. *ZGYW* 367.
1993 Z-Z. Tang and Z-M. Chen. *Jiading fangyan yanjiu*. Shehuikexue wenxian chubanshe 4:22–35.
1994 J-Y. Lü. Lüsi fangyan liangzizu liandu biandiao [Disyllabic tone sandhi in Lüsi]. *FY* 46–55.
1995 M. Chan. An autosegmental analysis of Danyang tone sandhi. In *Wuyu Yanjiu*, ed. by E. Zee, Chinese University of Hong Kong, 145–184.

6 Min dialects

6.1 Fuzhou

1930 Y-M. Tao. Min yin yanjiu. [Studies in Min phonology]. *BIHP* 1.4:445–470.
1947 M-K. Gao. Fuzhou zhi yucong shengmu tonghua [Consonantal assimilation in Fuzhou phrases]. *Yanjing xuebao* 33:129–144.
1953 Y-X. Lan. Fuzhou yinxi [The phonology of Fuzhou]. *Wenshizhe* 5:241–331.
1965 L. Chen. Foochow reduplication. *Qinghua xuebao* 5:200–213.
1965 L. Chen and J. Norman. Introduction to Foochow dialect. San Francisco State University.
1976 I. Maddieson. The intrinsic pitch of vowels and tone in Foochow. *UCLA Working Papers in Phonetics* 33:191–202.
1980 M. Chan. Syntax and phonology interface: the case of tone sandhi in the Fuzhou dialect of Chinese. Ms., University of Washington, Seattle.
1980 M. Wright. A treatment of tone in two Min dialects. *Cahiers Linguistiques d'Ottawa* 413–421.
1981 T-Q. Chen, R-L. Li, and Y-Z. Liang. Fuzhou hua shengmu leihua yinbian de zai tantao [Fuzhou consonantal sandhi revisited]. *ZGYW* 231–237.
1982 M. Wright. Tone sandhi and vowel changes in Fuzhou. Papers of *North East Linguistic Society* 12:258–273.
1983 Y-Z. Liang. Fuzhou fangyan liandu yinbian yu yuyi fenbie [Morphophonemics and semantic contrast in Fuzhou]. *FY* 166–169.
1983 Y-D. Zheng. Fuzhou fangyan danyin dongci chongdieshi [Reduplication of monosyllabic verbs in Fuzhou]. *ZGYW* 30–39.
1983 C-M. Zhang. Fuzhou hua de biandiao zhuangui [Track shift: one of the rules in Fuzhou tone sandhi]. Paper presented at the Sino-Tibetan conference, Seattle, Washington.
1983 M. Wright. A metrical approach to tone sandhi in Chinese dialects. Dissertation, University of Massachusetts.
1984 M. Chan. Stress and vowel quality changes in the Fuzhou dialect. *University of Washington Working Papers in Linguistics* 8:19–38.
1985 M. Chan. *Fuzhou Phonology: a Non-linear Analysis of Tone and Stress*. Ph.D. dissertation, University of Washington, Seattle.
1990 M. Chan. Prelinked and floating glottal stops in Fuzhou Chinese. *Canadian Journal of Linguistics* 35:331–349.
1993 A-Z. Feng. Fuzhou fangyan de ru sheng. *FY* 101–118.

1994 P. Jiang-King. A prosodic account of tone-vowel interaction in Fuzhou. Ms., University of British Columbia.
1995 P. Jiang-King. An Optimality account of tone-vowel interaction in Fuzhou. Paper presented at Arizona Phonology Conference.
1995 J. Tsay. Physical constraints in the phonology of Fuzhou. Paper presented at ICCL 4/NACCL 7.
1998 L. Chan. *Fuzhou Tone Sandhi*. Ph.D. dissertation, University of California, San Diego.

6.2 Xiamen/Taiwanese

1930 C-P. Luo. *Xiamen yinxi* [Xiamen phonology]. *BIHP* monograph 4.
1931 B-M. Chiu. The phonetic structure and tone behavior in Hagu. *T'oung-Pao* 28:245–342.
1955 N. Bodman. *Spoken Amoy Hokkien*. Kuala Lumpur, 2 volumes.
1957 D-C. Ge. Xiamen shengdiao yu juzi jiegou ji yingyu jiaoxue de guanxi. [Xiamen tones and sentence structure and their relation to the teaching of English]. *Xiamen daxue xuebao* 167–193.
1968 R.L. Cheng. Tone sandhi in Taiwanese. *Linguistics* 41:19–42.
1968 M. Tay. *A Phonological Study of Hokkien*. Dissertation, University of Edinburgh.
1970 H-I. Hsieh. The psychological reality of tone sandhi in Taiwanese. *CLS* 489–503
1971 C-C. Liao. The Syntactic Environments of Tone Sandhi in Taiwanese. M.A. thesis, National Taiwan University.
1972 M. Nakajima. Tone sandhi in Fukien dialect. *Journal of Asian and African Studies* 5:43–60.
1973 R.L. Cheng. Some notes on tone sandhi in Taiwanese. *Linguistics* 100:5–25.
1977 R.L. Cheng and S. Xie-Cheng. *Taiwan Fujianhua de Yuyin Jiegou ji Biaoyinfa* [The phonological structure and transcription of Taiwanese Hokkien]. Taipei: Xuesheng.
1982 M. Hashimoto. The so-called "original" and "changed" tones in Fukienese. *BIHP* 53:645–659.
1983 X-T. Li. Xiamen fangyan yinping he yangping de biandiao [The tone sandhi of yinping and yangping in Xiamen]. *FY* 120–122.
1983 P-H. Ting. Derivation by tone-change in colloquial Min. Paper presented at Association of Asian Studies, San Francisco.
1984 R.L. Cheng. Beiping hua he Taiwan hua qingsheng chuxian de yitong, lishi youlai, he Taiwan xinshengdai guoyu de xingcheng [A comparison of neutral tones in Beijing and Taiwanese, their historical origins, and the emergence of a new Taiwanese Mandarin]. Ms., University of Hawaii.
1985 W-R. Hong. *Taiwan gelao huayu shengdiao yanjiu* [Studies in the Tones of Taiwan Gelao speech]. Taipei: Zili Wanbaoshe.
1985 M. Chen. The syntax of phonology: Xiamen tone sandhi. Ms., University of California, San Diego.

1987 M. Chen. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4:109–149.
1988 R. King. Theoretical problems in the description of Taiwanese tone. Ms., Harvard University.
1988 T-C. Du. *Tone and stress in Taiwanese*. Ph.D. dissertation, University of Illinois.
1988 H-B. Lin. *Contextual Stability of Taiwanese Tones*. University of Connecticut.
1990 J. Tsay. The distribution of tone in Taiwanese. Proceedings of the 1989 Western Conference on Linguistics, Fresno California State University.
1991 S-P. Wang. Taiyu biandiao guilü xinli zhenshixing [Psychological reality of Taiwanese tone sandhi]. To appear in Y.R. Zhang, S.H. Wang, and R. Cheng, eds., Xiandai Taiyu yanjiu lunwenji [Studies on Modern Taiwanese], vol. 2.
1991 R.L. Cheng. Interaction, modularization, and lexical diffusion – Taiwanese tone sandhi revisited. Paper presented at the NACCL 3 (Cornell University).
1991 S-P. Wang. Tonal neutralization and paradigmatic replacement in Taiwanese and other languages: a note on OCP, register and tone sandhi. ISCLL 2, 409–426.
1992 M. Chen. Argument v. adjunct: Xiamen tone sandhi revisited. Ms., University of California, San Diego.
1992 H-C. Hsu. Psychological reality of tone sandhi in Taiwanese: evidence from idioms, child language acquisition, and loan words. Paper presented at ICCL 1 (Singapore).
1992 H-C. Hsu. Tonal parsing in idioms: a tug of war between the foot formation rule and the tone group formation. Paper presented at ISCLL 3 (Tsinghua).
1992 H-M. Zhang. Reanalysing tone sandhi domain in Xiamen. Paper presented at ICCL 1 (Singapore).
1993 S-H. Peng. Effects of prosodic position and tonal context on Taiwanese tones. *Ohio State University Working Papers in Linguistics* 44:166–190.
1994 R.L. Cheng. Yixingyu de gongcun yu taotai – Taiwanhua de qingsheng yu shuliangyu [The coexistence and obsolescence of synonyms – neutral tone and measure words in Taiwanese]. In ISCLL 2 (August 1991, Taipei), ed. by P.J-K. Li, C-R. Huang, and J.C-C. Tang, 453–486.
1994 J-W. Lin. Lexical government and tone group formation in Xiamen Chinese. *Phonology* 11:237–276.
1995 Z-M. Bao. Local tree geometry and the syntax–phonology interface. Ms., National University of Singapore.
1996 J. Tsay and J. Myers. Taiwanese tone sandhi as allomorph selection. Paper presented at the 22nd Annual Meeting of Berkeley Linguistic Society.
1996 J. Tsay and J. Charles-Luce. An acoustic study of Taiwanese tone sandhi with implications for language acquisition. Ms., National Chung Cheng University and State University of New York, Buffalo.

6.3 Other Min dialects

1956 S. Egerod. *The Lungtu Dialect*. Copenhagen: Ejnar Munksgaard.
1958 Q-X. Dai. Minyu Xianyou hua de biandiao guilü [Tone sandhi patterns in the in dialect of Xianyou]. *ZGYW* 10:485.
1962 J-H. Huang. Putian hua de liangzizu liandu yinbian [Disyllabic morphophonemics in Putian]. *ZGYW* 510–516.
1972 Y-H. Chang. *The Hinghwa Dialects of Fukien: a Descriptive Linguistic Study*. Dissertation, Cornell University.
1978 M. Clark. On the alternation between isolation and combination tones in Chaochow. *CLS* 14:55–64.
1978 Y-Z. Gao. Fuqing fangyan de shengmu liandu yinbian [Consonantal sandhi in Fuqing]. *ZGYW* 258–259.
1979 W-L. Woon. A synchronic phonology of Hainan dialect. *JCL* 7:65–100; 268–302.
1979–80 S-Y. Zhang. Chaoyang fangyan de liandu biandiao [Tone sandhi in Chaoyang]. *FY* 1979, 93–121; 1980, 123–136.
1981 B-S. Yang and J-M. Chen. Haifeng hua yuju zhong de shengdiao wenti [Issues concerning tone sandhi in Haifeng sentences].
1982 Y-W. Chang. *The Min tone sandhi*. M.A. thesis, University of Georgia.
1983 P-H. Ting. Derivation by tone change in colloquial Min. Paper read at AAS, San Francisco.
1983 Z-X. Zhang. Zhangping (Yongfu) fangyan de liandu biandiao [Tone sandhi in Zhangping (Yongfu)]. *FY* 3:175–196.
1984 M. Chan. Fuan tone sandhi. Paper presented at the Sino-Tibetan conference, Eugene, Oregon.
1984 D-A. He. Biandu xianxiang de liang zhong guanshi yiyi – jianlun Jinjiang fangyan de gu diaozhi [The diachronic significance of sandhi phenomena, and the phonetic value of ancient Jinjiang tones]. *BIHP* 55:115–132.
1993 A-Z. Feng. *Fuqing fangyan yanjiu* [A study in the Fuqing dialect]. Shehui kexue wenxian chubanshe, 38–49.

7 Hakka dialects

1959 F-C. Li. Yuedong Taoyuan hua de teshu biandiao guilü [Special tone sandhi rules in the Taoyuan dialect of east Guangdong]. *ZGYW* 377–379.
1965 R-L. Li. Changting hua liangyinjie sanyinjie de liandu biandiao [Disyllabic and trisyllabic tone sandhi in changting]. *Xiamen daxue xuebao*.
1982 M-Z. Luo. Fujian Changting Kejia hua de liandu biandiao [Tone sandhi in the Hakka dialect of Changting, Fujian]. *YY* 188–197.
1987 C-R. Rao. Fujian Changting (Kejia) fangyan de liandu biandiao [Tone sandhi in the Kejia dialect of Changting, Fujian]. *ZGYW* 191–200.
1989 X-Z. Huang. Kejia fangyan shengdiao de tedian xulun [Further discussions on the characteristics of Hakka tones].

1989 R-F. Chung. *Aspects of Kejia Phonology*. Ph.D. dissertation, University of Illinois.
1992 L-W. Xie. Jiangxi sheng Yudu fangyan liangzizu liandu biandiao [Tone sandhi in the Jiangxi dialect of Yudu]. *FY* 222–227.

8 Yue dialects

1955 K. Whitaker. A study, on the modified tones in spoken Cantonese. *Asia Major*, ser. 2, 5:9–36, 184–207.
1966 X-W. Chen. Taishan fangyan teshu biandiao chutan [A preliminary exploration of unusual tone sandhi processes in Taishan]. *ZGYW* 34–36.
1969 Y-S. Cheung. Xianggang Yueyu yinping ji biandiao wenti ["Yang-ping" tone and tone sandhi in Hong Kong Cantonese]. *Journal of the Institute of Chinese Studies* 2:81–107.
1977 T-H. Kam. Derivation by tone change in Cantonese. *JCL* 5:186–210.
1980 T-H. Kam. Semantic–tonal processes in Cantonese, Taishanese, Bobai, and Siamese. *JCL* 8:205–240.
1981 W-R. Bai. Guangzhou fangyan liandu yinbian juli [Examples of phonetic alternations in the connected speech of Guangzhou]. *FY* 19–20.
1981 C-S. Leung. A Study of Tone Change in Spoken Cantonese. M.A. thesis, University of Hong Kong.
1982 G-Q. Ye and Z-D. Tang. Xinyi fangyan de bianyin [Phonetic alternation in Xinyi]. *FY* 47–51.
1983 B. T'sou. Guangzhou hua zhong wenbai yidu he liandu biandiao de shehui yuanyin [The sociological causes of literary vs. colloquial readings and tone sandhi in Guangzhou]. Paper presented at the Sino-Tibetan conference, Seattle, Washington.
1983 F-B. Zong. Guanyu Guangzhou hua zidiao biandu wenti [Concerning changed tones in Guangzhou]. *Wuhan daxue xuebao* 4:79–89.
1983 F-B. Zong. Guanyu Guangzhou hua zidiao biandu wenti [Concerning the sandhi forms of lexical tones in Guangzhou]. *Wuhan daxue xuebao* 4.
1987 W-R. Bai. Guangzhou hua zhong de shenglüe xing bianyin [Elliptic sandhi in Cantonese]. *FY* 114–120.
1987 M. Chan. Tone and melody in Cantonese. Proceedings of the 13th Berkeley Linguistic Society meeting.
1989 W-R. Bai. Guangzhouhua zhong de shenglüexing bianyin [Ellipsis-triggered tone change in the Guangzhou dialect]. *FY* 114–120.
1990 G-Q. Wei. *Pingxiang fanyan zhi* [The Pingxiang dialect]. Beijing: Yuwen chubanshe.

References

Abney, S. 1987. The English noun phrase in its sentential aspect. Ph.D. dissertation, Massachusetts Institute of Technology.

Ackerman, Farrell and Phil Lesourd 1992. Preverbs and complex predicates. Ms., University of California, San Diego.

Ackerman, Farrell and Gert Webelhuth 1992. Topicalization and German complex predicates. Ms., University of California, San Diego and University of North Carolina.

forthcoming. *Word and Grammar: a Theory of Natural Predicates.* Ms., University of California, San Diego and University of North Carolina.

Allen, W. Sidney 1962. *Sandhi: the Theoretical, Phonetic, and Historical Bases of Word-Junction in Sanskrit*. 'S-Gravenhage: Mouton.

Andersen, Henning 1986. Introduction: sandhi. In H. Andersen (ed.), *Sandhi Phenomena in the Languages of Europe*. Berlin-New York-Amsterdam: Mouton de Gruyter, pp.1–8.

Anderson, Stephen R. 1978. Tone features. In Victoria A. Fromkin (ed.), *Tone: a Linguistic Survey*. Academic Press, pp.133–176.

1981. Why phonology isn't "natural." *Linguistic Inquiry* 12:493–540.

Ao, Benjamin Xiaoping 1992. Metrical constituents as domains of tone sandhi in Nantong Chinese. CLS 28.

1993. Phonetics and phonology of Nantong Chinese. Ph.D. dissertation, Ohio State University.

Archangeli, Diana and Douglas Pulleyblank 1994. *Grounded Phonology*. Cambridge, Mass.: MIT Press.

Bai, Wanru 1989. Guangzhouhua zhong de shenglüexing yinbian [Tone change triggered by ellipsis in Cantonese]. *Fangyan* pp.114–120.

Baker, Mark 1988. *Incorporation: a Theory of Grammatical Function Changes*. University of Chicago Press.

Ballard, William L. 1985. Whence putative tonogenesis? Paper presented at the eighteenth International Conference on Sino-Tibetan Languages and Linguistics, Bangkok (August, 1985).

1988. *The History and Development of Tonal Systems and Tone Alternations in South China*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Bao, Zhiming 1990a. On the nature of tone. Ph.D. dissertation, Massachusetts Institute of Technology.

1990b. Fanqie languages and reduplication. *Linguistic Inquiry* 21:317–350.
1992. Toward a typology of tone sandhi. In Buszard-Welcher et al., pp.1–12.
1995. Local tree geometry and the syntax–phonology interface. Ms., National University of Singapore.
1996a. The syllable structure in Chinese. *Journal of Chinese Linguistics* 24:312–354.
1996b. Teochow tone sandhi and the representation of tone. *Proceedings of the Fourth International Symposium on Languages and Linguistics*, 1:25–46.

Beattie, David A. 1985. Mandarin Chinese tone sandhi and the interaction of phonology and syntax. M.A. thesis, University of Toronto.

Beckman, Mary E. 1986. *Stress and Non-stress Accent*. Dordrecht: Foris.

Beckman, Mary E., Maria G. Swora, Jane Rauschenberg, and Kenneth DeJong 1990. Stress shift, stress clash, and polysyllabic shortening in prosodically annotated discourse. *Proceedings of the 1990 International Conference on Spoken Language Processing* 1:5–80.

Beaudouin de Courtenay, Jan 1972. *The Beginnings of Structural Linguistics*. Translated and edited by Edward Stankewicz. Bloomington: Indiana University Press.

Beijing University 1962. *Hanyu Fangyin Zihui*. Beijing: Wenzi Gaige Chubanshe. [2nd augmented edition, 1989].

Bloomfield, Leonard 1933. *Language*. Holt, Rinehart and Winston.

Bi, Ke-zhao 1994. Guangxi Heng xian Pinghua cihui (1) [The vocabulary of Ping speech of Heng county in Guangxi, part 1]. *Fangyan* 70–77.

Bickmore, Lee S. 1995. Tone and stress in Lamba. *Phonology* 12:307–342.

Blevins, Juliette 1993. A tonal analysis of Lithuanian nominal accent. *Language* 69:237–273.

Bodman, Nicholas C. 1955. *Spoken Amoy Hokkien*. Kuala Lumpur. (2 volumes)

Bogers K., H. van der Hulst, and M. Mous (eds.) 1986. *The Phonological Representation of Suprasegmentals*. Dordrecht: Foris.

Bolinger, Dwight 1958. A theory of pitch accent in English. *Word* 14:109–149.
1982. *Intonation and its Parts: the Melody of Speech*. Stanford University Press.

Boyce, Conal 1980. Min sandhi in verse recitation. *Journal of Chinese Linguistics* 8:1–14.

Bresnan, Joan 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, pp.3–86. Cambridge, Mass.: MIT Press.

Bromberger, Sylvain and Morris Halle 1989. Why phonology is different. *Linguistic Inquiry* 20:51–70.

Buszard-Welcher, Laura A. et al. (eds.), 1992. *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society*, special session on The Typology of Tone Languages.

Carter, H. 1980. The Kongo tonal system revisited. *African Language Studies* 17:1–32.

Chafe, Wallace 1968. The ordering of phonological rules. *International Journal of American Linguistics* 24:115–136.

Chan, Marjorie 1985. Fuzhou phonology: a non-linear analysis of tone and stress. Ph.D. dissertation, University of Washington.

1989. On the status of "basic" tones. *Acta Linguistica Hafniensia* 21:35–64. [An earlier version appeared in *UCLA Working Papers in Phonetics* 63:71–94.]
1991. Contour-tone spreading and tone sandhi in Danyang Chinese. *Phonology* 8:237–259.
1995. An autosegmental analysis of Danyang tone sandhi: some historical and theoretical issues. In Zee, pp.145–184.

Chan, Marjorie and Hongmao Ren 1989. Wuxi tone sandhi: from last to first syllable dominance. *Acta Linguistica Hafniensia* 21:35–64. [An earlier version appeared in *UCLA Working Papers in Phonetics* 63:48–70.]

Chang, Kun 1975. Tonal developments among Chinese dialects. *Bulletin of the Institute of History and Philology* 46:636–710.

Chang, Laura Mei-chih 1992. A prosodic account of tone, stress, and tone sandhi in Chinese languages. Ph.D. dissertation, University of Hawaii. [Published by Crane, Taipei (1993).]

Chao, Yuenren. 1928. *Xiandai Wuyu de Yanjiu* [*Studies in the Modern Wu Dialects*]. Beijing: Qinghua University.
1930. A system of tone letters. *Le maître phonétique* 45:24–27.
1935. Zhongguo fangyan zhong baofayin zhonglei [Types of plosives in Chinese dialects]. *Bulletin of the Institute of History and Philology* 5:515–520.
1947. *Cantonese Primer*. Harvard University Press.
1956. Formal and semantic discrepancies between different levels of Chinese structures. *Bulletin of the Institute of History and Philology* 28.1:1–16.
1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.

Chen, Chungyu 1984. Neutral tone in Mandarin: phonotactic description and the issue of the norm. *Journal of Chinese Linguistics* 12:299–333.

Chen, Matthew Y. 1976. From Middle Chinese to Modern Peking. *Journal of Chinese Linguistics* 4:113–277.
1979. Metrical structure: evidence from Chinese poetry. *Linguistic Inquiry* 10:371–420.
1980. The primacy of rhythm in verse: a linguistic perspective. *Journal of Chinese Linguistics* 8:15–41.
1984. Unfolding latent principles of literary taste: poetry as a window onto language. *Tsinghua Journal of Chinese Studies* 16:203–240.
1986. The paradox of Tianjin tone sandhi. *Proceedings of the Chicago Linguistic Society Meeting* 22:98–114.
1987a. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4:109–150. [An earlier version circulated as The syntax of phonology: Xiamen tone sandhi, ms. University of California, San Diego, 1985.]
1987b. Introductory remarks to A Symposium on Tianjin Tone Sandhi. *Journal of Chinese Linguistics* 15:203–227.
1991a. An overview of tone sandhi phenomena across Chinese dialects. In W. Wang (ed.), *Languages and Dialects of China*, pp.113–158. *Journal of Chinese Linguistics*, monograph series n.3. [Original version presented at the Conference on the Languages and Dialects of China, 1985, Oakland, California.]

1991b. Recent advances in tone sandhi studies. Paper presented at LSA Summer Institute, July 13, 1991, University of California, Santa Cruz. [Revised and published as Chen 1996.]
1991c. Competing strategies and harmonic principle. Ms., University of California, San Diego.
1992a. Tone rule typology. In Buszard-Welcher et al., pp.54–66.
1992b. Phonological diagnostics of morphological structure. In Chiuyu Tseng (ed.), the *Proceedings of the Third International Symposium on Chinese Languages and Linguistics* (Taipei, July 1992).
1992c. Argument v. adjunct: Xiamen tone sandhi revisited. Ms., University of California, San Diego.
1994. Synchronic analogs of competing sound changes. In Chen and Tzeng, pp.47–70.
1996. Tonal geometry – a Chinese perspective. In Huang and Li, pp.21–48.
1997. Competing strategies and derivational economy. To appear in *Journal of Chinese Linguistics*.

Chen, Matthew Y. and Ovid J.L. Tzeng (eds.) 1994. *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*. Taipei: Pyramid.

Chen, Matthew Y. and Hongming Zhang 1997. Lexical and postlexical tone sandhi in Chongming. In Jialing Wang and Norval Smith (eds.), *Studies in Chinese Phonology*. North-Holland, pp.13–52.

Chen, Wen-bin 1953. Pinyin wenzi zhong de tongyinci wenti de chubu yanjiu [A pilot study on homophony in Pinyin]. *Zhongguo Yuwen* 2:7–10.

Cheng, Chinchuan 1968. English stresses and Chinese tones in Chinese sentences. *Phonetica* 18:77–88.
1970. Domain of phonological rule application. In *Studies Presented to Robert B. Lees by his Students*.
1973a. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.
1973b. A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1:93–110.
1991. Quantifying affinity among Chinese dialects. In Wang, pp.78–112.
1994. DOC: its birth and life. In Chen and Tzeng, pp.71–86.

Cheng, Chinchuan and Charles Kisseberth 1979–81. Ikorovere Makua tonology. *Studies in the Linguistic Sciences*, 9.1:31–63; 10.1:15–44; and 11.1:181–202.

Cheng, Lisa 1987. Derived domains and Mandarin third tone sandhi. *Proceedings of the Chicago Linguistic Society Meeting*, Parasession, 16–28.

Cheng, Robert L. 1968. Tone sandhi in Taiwanese. *Linguistics* 41:19–42.

Cheng, Tsaifa 1983. The tonal features of Proto-South Min. *East Asian Languages* 1:66–90.

Cheung, Yatsheng 1986. Xianggang guangzhouhua yingyu yinyi jieci de shengdiao guilü [The tonal pattern of English loanwords in HongKong Cantonese]. *Zhongguo Yuwen* 1:42–50.

Chiu, Bien-ming 1931. *The Tone Behaviour in Hagu: an Experimental Study*. Archives Néerlandaises de Phonétique Expérimentale.

Chomsky, Noam 1986. *Barriers*. MIT Press.
1994. Bare phrase structure. *MIT Occasional Papers in Linguistics* 5:1–48.

Chomsky, Noam and Morris Halle 1968. *The Sound Pattern of English*. New York: Harper and Row.

Chung, Raung-fu 1989. Aspects of Kejia phonology. Ph.D. dissertation, University of Illinois, Urbana-Champaign.

Cinque, Guglielmo 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239–297.

Clements, George N. 1978. Tone and syntax in Ewe. In Donna Jo Napoli (ed.), *Elements of Tone, Stress, and Intonation*, pp.21–99. Washington, D.C.: Georgetown University Press.

1979. The description of terraced-level tone languages. *Language* 55:536–558.

1985. The geometry of phonological features. *Phonology Yearbook* 2:225–252.

1990a. The status of register in intonation theory. In Kingston and Beckman, pp.58–71.

1990b. The role of the sonority cycle in core syllabification. In Kingston and Beckman, pp.283–333.

Clements, George N. and Ford 1979. Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10:179–210.

1981. On the phonological status of downstep in Kikuyu. In D.L. Goyvaerts (ed.), *Phonology in the 1980's*, pp.309–357. Ghent: Story-Scientia.

Clements, George N. and John Goldsmith (eds.) 1984. *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris.

Clifton, C., L. Frazier, and K. Rayner (eds.) 1994. *Perspectives on Sentence Processing*. Hillsdale, N.J.: Erlbaum.

Cole, Jennifer 1995. The cycle in phonology. In Goldsmith 1995a, pp.70–113.

Cole, Jennifer and Charles W. Kisseberth 1995. Restricting multi-level constraint evaluation: opaque rule interactions in Yawelmani vowel harmony. Ms., University of Illinois, Urbana-Champaign.

Connell, Bruce and Amalia Arvanti 1995. Phonology and Phonetic Evidence, Papers in Laboratory Phonology IV. Cambridge University Press.

Cooper, W.E. and S.J. Eady 1986. Metrical phonology in speech production. *Journal of Memory and Language* 25:369–384.

Cowper, Elizabeth A. and Keren D. Rice 1987. Are phonosyntactic rules necessary? *Phonology Yearbook* 4:185–194.

Crowhurst, Megan and Mark S. Hewitt 1995. Directional footing, degeneracy, and alignment. Ms., University of North Carolina and University of British Columbia. (To appear in NELS 25.)

Dai, John X.L. 1992. Chinese morphology and its interface with the syntax. Ph.D. dissertation, Ohio State University.

1994. Syntactic, phonological, and morphological words in Chinese. In Packard.

Davis, Stuart 1995. Some matters regarding the derivational residue. Paper presented at the Conference on the Derivational Residue in Phonology, Tilburg October, 1995.

Dell, François and Mohamed Elmedlaoui 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7:105–130.

Di Sciullo, Anna-Maria and Edwin Williams 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Dong, Tung-ho 1960. Sige Minnan fangyan [Four South Min Dialects]. The Bulletin of the Institute of History and Philology, *Academia Sinica* 30:729–1042.

Downer, Gordon B. 1959. Derivation by tone-change in Classical Chinese. *Bulletin of the School of Oriental and African Languages* 22:258–290.

Downing, Laura 1990. Local and metrical tone shift in Nguni. *Studies in African Linguistics* 21:261–317.

Dreher, John J. and P.C. Lee 1968. Instrumental investigation of single and paired Mandarin tonemes. *Monumenta Serica* 27:343–373.

Dressler, Wolfgang U., Hans C. Luschützky, Oskar E. Pfeiffer, and John R. Rennison (eds.) 1987. *Phonologica 1984*. Cambridge University Press.

Du, Tsai-chwun 1988. Tone and Stress in Taiwanese. Ph.D. dissertation, University of Illinois, Urbana-Champaign.

Duanmu, San 1990a. A formal study of syllable, tone, stress and domain in Chinese languages. Ph.D. dissertation, Massachusetts Institute of Technology.

1990b. Cyclicity and the Mandarin third tone sandhi. Ms., Massachusetts Institute of Technology.

1991. Stress and syntax–phonology mismatches: tonal domains in Danyang and Shanghai. In D. Bates (ed.), *Proceedings of the Tenth West Coast Conference on Formal Linguistics*, Stanford Linguistic Association, pp.127–137.

1992a. End-based theory, cyclic stress, and tonal domains. *Proceedings of the Chicago Linguistic Society Meeting*.

1992b. Re-examining contour tone units in Chinese languages. In Buszard-Welcher et al., pp.80–91.

1993a. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2:1–44.

1993b. Distinctions between word and phrase in Chinese. Ms., University of Michigan. [To appear in Jerry Packard (ed.), *Word Formation in Chinese*.]

1994a. Against contour tone units. *Linguistic Inquiry* 25:555–608.

1994b. Wordhood in Chinese. In Packard.

1995. Phonology of compounds in two Chinese dialects. *Language* 71:225–259.

1997. Recursive constraint evaluation in Optimality Theory: evidence from cyclic compounds in Shanghai. To appear in *Natural Language and Linguistic Theory*.

Duanmu, San and Bing-fu Lu 1990. Word length variations in Chinese two-word constructions. Ms., MIT and University of Connecticut.

Duncan, Erin 1994. Directionality effects through alignment. PASC 3.

Egerod, Sören C. 1967. Dialectology. *Current Trends in Linguistics*, pp.91–129.

1976. Benedict's Austro-Thai hypothesis and the traditional views on Sino-Thai relationship. In Hashimoto, pp.51–60.

Endô, Mitsuaki 1994. Yuanyin yu shengdiao [Vowel and tone]. *Proceedings of the Second International Symposium on Chinese Languages and Linguistics* (August 1991, Taipei) pp.487–516.

Everett, Daniel 1988. On metrical constituent structure in Pirahã phonology. *Natural Language and Linguistic Theory* 6:207–246.

Fang, Songxi 1987. Tone sandhi of disyllabic words in the Zhoushan dialect. *Fangyan* 116–123.

Feng, Aichen 1993. *Fuqing Fangyan Yanjiu* [*A Study of the Fuqing Dialect*]. Beijing: Shehui Kexue Wenxian Chubanshe.

Fromkin, Victoria A. (ed.) 1978. *Tone: a Linguistic Survey*. New York: Academic Press.

Fry, Dennis 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27:765–769.

1958. Experiments in the perception of stress. *Language and Speech* 1:120–150.

Fukui, Naoki. 1986. A theory of category projection and its application. Ph.D. dissertation, Massachusetts Institute of Technology.

Furere, R. and A. Rialland 1985. Tons et accents en Kinyarwanda. In D.L. Goyvaerts (ed.), *African Linguistics*, pp.99–166. Amsterdam: Benjamins.

Gandour, Jack 1974. On the representation of tone in Siamese. *UCLA Working Papers in Phonetics* 27:118–146.

1978. The perception of tone. In Fromkin, pp.41–76.

Gandour, Jack and Victoria A. Fromkin 1978. On the phonological representation of contour tones. *Linguistics of the Tibeto-Burman Area* 4.1:73–74.

Gao, Mingkai and Anshi Shi 1963. *Yuyanxue Kailun* [*Introduction to Linguistics*]. Beijing: Zhunghua shudian.

Garde, Paul 1968. *L'Accent*. Presses Universitaires de France.

Gårding, Eva 1984. Chinese and Swedish in a generative model of intonation. Nordic Prosody III, *Acta Universitatis Umensis* 59:79–91.

1987. Speech act and tonal pattern in Standard Chinese: constancy and variation. *Phonetica* 44:13–29.

Gleason, Henry 1955. *Workbook in Descriptive Linguistics*. New York: Holt.

Goldsmith, John 1976a. Autosegmental phonology. Ph.D. dissertation, MIT. [New York: Garland, 1979.]

1976b. An overview of autosegmental phonology. *Linguistic Analysis* 2:23–68.

1981. English as a tone language. In D. Goyvaerts (ed.), *Phonology in the 1980's*, pp.287–307. Ghent: Story-Scientia. [Also in *Communication and Cognition* 11:453–476.]

1982. Accent systems. In van der Hulst and Smith, pp.47–64.

1984. Tone and accent in Tonga. In G.N. Clements and J. Goldsmith (eds.), pp.19–52.

1987. The rise of rhythmic structure in Bantu. In Dressler et al.

1988. Prosodic trends in the Bantu languages. In Hulst and Smith, pp.81–94.

1990. *Autosegmental and Metrical Phonology*. Oxford, U.K. and Cambridge, MA: Basil Blackwell.

(ed.) 1993a. *The Last Phonological Rule*. University of Chicago Press.

1993b. Harmonic phonology. In Goldsmith 1993a, pp.21–60.

Goldsmith, John and Gary Larson. 1990. Local modeling and syllabification. In K. Deaton et al. (eds.), *Papers from the 26th Annual Regional Meeting of the Chicago Linguistic Society*, Part 2. Chicago Linguistic Society.

Goldsmith, John and F. Sabimana 1989. The Kirundi verb. In F. Jouannet (ed.), *Modèles en tonologie (Kirundi et Kinyarwanda)*, pp.19–62. Paris: CNRS.

Goodman, Beverley 1994. On the variability of the syllable coda. Ms., University of California, San Diego.

Grabe, Esther and Paul Warren 1995. Stress shift: do speakers do it or do listeners hear it? In Connell and Arvaniti (eds.), pp.95–110.

Gruber, J. 1964. The distinctive features of tone. Ms.

Hale, Kenneth and Elisabeth O. Selkirk 1987. Government and tonal phrasing in Papago. *Phonology Yearbook* 4:151–183.

Halle, Morris and George N. Clements 1985. *Problem Book in Phonology*. Cambridge, Mass.: MIT Press.

Halle, Morris and William Idsardi 1995. General properties of stress and metrical structure. In Goldsmith, pp.403–443.

Halle, Morris and Kenneth N. Stevens 1971. A note on laryngeal features. *Quarterly Progress Reports*, Research Lab of Electronics, MIT, 101:198–213.

Halle, Morris and Jean-Roger Vergnaud 1987. *An Essay on Stress*. Cambridge, Mass.: MIT Press.

Halliday, M.A.K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.

Hammond, Michael 1985. Main stress and parallel metrical planes. In M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Jeder (eds.), *Proceedings of Berkeley Linguistic Society* 11:417–28. (Parasession on Poetics, Metrics, and Prosody)

Hanyu Fangyin Zihui [*A Pronouncing Dictionary of Chinese Dialects*]. Compiled by Beijing University, 1962. [Second emendated edition, 1989.]

Hart, J. 't 1974. Discriminability of the size of pitch movements in speech. *IPO Annual Progress Report* 9:55–63.

Hashimoto (Yue.), Ann Oikan 1971. *Mandarin Syntactic Structures*. Princeton: Unicorn, n.8

1972. *Studies in Yue Dialects* I: *Phonology of Cantonese.* Cambridge University Press.

1986. Tonal flip-flop in Chinese dialects. *Journal of Chinese Linguistics* 14:161–183.

1987. Tone sandhi across Chinese dialects. In *The Wang Li Memorial Volumes* pp.445–474. [Original paper delivered at the thirteenth Sino-Tibetan Conference, University of Virginia, 1980.]

Hashimoto, Mantaro 1982. The so-called "original" and "changed" tones in Fukienese: a case study of Chinese tone morphophonemics. *Bulletin of the Institute of History and Philology* 53:645–659.

1984. Linguistic diffusion of Chinese tones. *Journal of Asian and African Studies* 27:1–51.

Haudricourt, A.G. 1954a. De l'origine des tons en Vietnamien. *Journal Asiatique* 242:69–82.

1954b. Comment reconstruire le chinois archaïque. *Word* 10:351–364.

1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrème-Orient. *Bulletin de la Société Linguistique de Paris*, 56:163–180.

Hayes, Bruce 1980. A metrical theory of stress rules. Ph.D. dissertation, MIT. (New York: Garland, 1985).

1982. Extrametricality and English stress. *Linguistic Inquiry* 13:227–276.

1984. The phonology of rhythm in English. *Linguistic Inquiry* 15:33–74.

1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17:467–499.

1989. The prosodic hierarchy in meter. In Paul Kiparsky and G. Youmans (eds.), *Rhythm and Meter*. Orlando: Academic Press.

1990. Precompiled phrasal phonology. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, 109–126. University of Chicago Press.

1995. *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press.

He, Wei 1989. *Huojia Fangyan Yanjiu* [*A Study of the Huojia Dialect*]. Beijing: Shangwu Yinshuguan.

Hirayama, Hisao 1974. [in Japanese] Zhongguoyu minnan minbei zufangyan de shengdiao diaozhi. Dongjing daxue wenxuebu yanjiu baogao 5, *Wenxue zhexue lunwenji* pp.193–248.

1975. Xiamenhua gudiaozhi de neibu gouni [Internal reconstruction of the phonetic values of proto-Xiamen tones]. *Journal of Chinese Linguistics* 3:3–15. [Also in *Yinyunxue Yanjiu Tongxun* 1:19–28 (1981).]

Ho, Da-an 1989. Songqi fendiao ji xiangguan wenti [Aspiration-triggered tone split and related issues]. *Bulletin of the Institute of History and Philology* 60:765–778.

1996. Stages and strata in dialectal history – case studies of Heng County, Da County and Shipo. In Huang and Li, pp.215–234.

Hoa, Monique 1983. *L'accentuation en pékinois*. Paris: Editions Langages Croisés.

Hollenbach, Barbara E. 1988. The asymmetrical distribution of tone in Copala Trique. In van der Hulst and Smith, pp.167–182.

Hombert, Jean-Marie 1975. Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in language. Ph.D. dissertation, University of California, Berkeley.

1978. Consonant types, vowel quality, and tone. In Fromkin, pp.77–112.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan 1979. Phonetic explanations for the development of tones. *Language* 55:37–58. [An earlier version entitled "Tonogenesis: theories and queries" was presented at the Eighth Sino-Tibetan Conference, Berkeley.]

Hou, Jing-yi 1989. *Jinyu Yanjiu* [Studies of the Jin Dialects]. Institute of the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Hsiao, Yuchau 1991. Syntax, rhythm and tone: a triangular relationship. Ph.D. dissertation, University of California, San Diego.

Hsu, Huichuan 1994. Constraint-based phonology and morphology: a survey of languages in China. Ph.D. dissertation, University of California, San Diego.

Hu, Tan 1979. Study on Tibetan tones. Paper presented at the Twelfth Sino-Tibetan Linguistic Conference.

Huang, Churen 1987. Mandarin Chinese NP DE: a comparative study of current grammatical theories. Ph.D. dissertation, Cornell University.

Huang, James C-T. 1982. Logical relations in Chinese and the theory of grammar. Ph.D. dissertation, Massachusetts Institute of Technology.

1984. Phrase structure, lexical integrity and Chinese compounds. *Journal of Chinese Language Teachers Association* 4:

Hulst, Harry van der and Norval Smith (eds.) 1988. *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris.

Hulst, Harry van der and Keith Snider (eds.) 1993. *The Phonology of Tone: the Representation of Tonal Register*. Berlin and New York: Mouton de Gruyter.

Hung, Tony T.N. 1987a. Syntactic and semantic aspects of Chinese tone sandhi. Ph.D. dissertation, University of California, San Diego.

1987b. Tianjin tone sandhi: towards a unified approach. *Journal of Chinese Linguistics* 15:274–305.

Hunter, Gear and Eunice V. Pike 1969. The phonology and tone sandhi of Molinos Mixtec. *Linguistics* 47:24–40.

Hyman, Larry M. 1975. Review of C. Cheng: A Synchronic Phonology of Mandarin Chinese. *Journal of Chinese Linguistics* 3:88–99.

1977. On the nature of linguistic stress. In Larry M. Hyman (ed.), *Studies in Stress and Accent*. Southern California Occasional Papers in Linguistics, n.4, University of Southern California.

1978. Tone and/or accent. In D.J. Napoli (ed.), *Elements of Tone, Stress and Intonation*, pp.1–20.

1979. A reanalysis of tonal downstep. *Journal of African Languages and Linguistics* 1:9–29.

1986. The representation of multiple tone heights. In Bogers, Hulst and Mous, pp.109–152.

1989. Accent in Bantu: an appraisal. *Studies in the Linguistic Sciences* 19:111–128.

1993. Register tone and tonal geometry. In Hulst and Snider, pp.75–108. [First circulated in 1989.]

Hyman, Larry M. and Ernest Byarushengo 1984. A model of Haya tonology. In George N. Clements and John Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, pp.53–103. Dordrecht: Foris.

Hyman, Larry M. and Francis Katamba 1992. A new approach to tone in Luganda. Ms., University of California, Berkeley.

Hyman, Larry M., Francis Katamba, and L. Walusimbi 1987. Luganda and the strict layer hypothesis. *Phonology Yearbook* 4:87–108.

Hyman, Larry M. and Russell G. Schuh 1974. Universals of tone rules. *Linguistic Inquiry* 5:81–115.

Inkelas, Sharon and William R. Leben 1990. Where phonology and phonetics intersect: the case of Hausa intonation. In Kingston and Beckman, pp.17–34.

Inkelas, Sharon and Draga Zec 1988. Serbo-Croatian pitch accent. *Language* 64:227–248.

(eds.) 1990. *The Phonology–Syntax Connection*. University of Chicago Press.

Itô, Junko 1986. Syllable theory in prosodic phonology. Ph.D. dissertation, University of Massachusetts, Amherst.

1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7:217–259.

1990. Prosodic minimality in Japanese. *Proceedings of the Chicago Linguistic Society Meetings* 26 *Parasession on the Syllable in Phonetics and Phonology*, pp.213–239.

Iverson, Gregory K. 1995. Rule ordering. In Goldsmith, pp.605–614.

Jiangsusheng he Shanghaishi Fangyan Gaikuang [*A Survey of Jiangsu Province and Metropolitan Shanghai*]. Nanjing: Jiangsu Renmin Chubanshe, 1960.

Jin, Shunde 1986. *Shanghai Morphotonemics – a Preliminary Study of Tone Sandhi Behavior across Word Boundaries*. Bloomington: Indiana University Linguistic Club. [= M.A. thesis, University of Pittsburgh, 1985]

1995. Shilun Shanghai fangyan de shengdiao yinxi [On Shanghai tonal system]. In Zee, pp.195–218.

Johnson, C. Douglas 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.

Kager, René 1994. On eliminating directional foot parsing. *GLOW* 32.

1995. The metrical theory of word stress. In Goldsmith 1995a, pp.367–402.

Kaisse, Ellen M. 1985. *Connected Speech*. New York: Academic Press.

1990. Toward a typology of postlexical rules. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, pp.127–144. University of Chicago Press.

Kam, Tak-him 1977. Derivation by tone change in Cantonese: a preliminary survey. *Journal of Chinese Linguistics* 5:186–210.

1980. Semantic-tonal processes in Cantonese, Taishanese, Bobai and Siamese. *Journal of Chinese Linguistics* 8:205–240.

Kaplan, Ronald M. and Martin Kay 1981. Phonological rules and finite-state transducers [abstract]. *Linguistic Society of America Meeting Handbook*. 56th annual meeting, December 27–30, 1981. New York: Linguistic Society of America.

1994. Regular models of phonological systems. *Computational Linguistics* 20:3.

Karttunen, Lauri 1993. Finite-state constraints. In Goldsmith 1993a, pp.173–94.

Karttunen, Lauri, Kimmo Koskenniemi, and Ronald M. Kaplan. 1985. A compiler for two-level phonological rules. In Mary Dalrymple et al. (eds.), *Tools for Morphological Analysis*. Report no.CSLI-87-108. Center for the Study of Language and Information, Stanford.

Kelly, M.H. and J.K. Bock 1988. Stress in time. *Journal of Experimental Psychology: Human Perception and Performance* 14:389–403.

Kennedy, George A. 1953. Two tone patterns in Tangsic [= Tangxi]. *Language* 29:367–373.

Kenstowicz, Michael 1994. *Phonology in Generative Grammar*. Cambridge, Mass. and Oxford, U.K.: Blackwell.

1995. Cyclic vs. non-cyclic constraint evaluation. *Phonology* 12:397–436.

1996. Base identity and uniform exponence: alternatives to cyclicity. In J. Durand and B. Laks (eds.), Current Trends in Phonology: Models and Methods. Salford: vol.I, pp.363–393. University of Salford.

Kenstowicz, Michael and Charles Kisseberth 1977. *Topics in Phonological Theory*. New York: Academic Press.

1979. *Generative Phonology: Description and Theory*. New York: Academic Press.

1990. Chizigula tonology: the word and beyond. In Inkelas and Zec, pp.163–194.

King, Ross 1988. Theoretical problems in the description of Taiwanese tone. Ms., Harvard University.

Kingston, John and Mary E. Beckman (eds.) 1990. *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge University Press.

Kiparsky, Paul 1971. Historical linguistics. In W. Dingwall (ed.), *A Survey of Linguistic Science*, pp.576–649. College Park: University of Maryland. [Reprinted in Kiparsky 1982.]

1973. Abstractness, opacity, and global rules. In Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*, pp.57–86. Tokyo: TEC.

1975. Stress, syntax and meter. *Language* 51:577–616.

1982a. From cyclic phonology to lexical phonology. In H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations* vol. I; pp.131–175. Dordrecht: Foris.

1982b. Lexical phonology and morphology. In I.S. Yang (ed.), *Linguistics in the Morning Calm*, vol. II, pp.3–91. Seoul: Hanshin.

Kirchner, Robert 1995. Going the distance: synchronic chain shifts in Optimality Theory. Ms., University of California, Los Angeles.

Kisseberth, Charles W. 1973. Is rule ordering necessary in phonology? In B. Kachru (ed.), *Issues in Linguistics: Papers in Honor of Henry and Renée Kahane*, pp.418–441. Urbana: University of Illinois Press.

1991. Metrical structure in Chizigula tonology. Ms., University of Illinois.

Klatt, Dennis H. 1974. Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *Journal of the Acoustical Society of America* 53:8–16.

Klavans, Judith L. 1982. *Some Problems in a Theory of Clitics*. Bloomington: Indiana University Press.

1985. The independence of syntax and phonology in cliticization. *Language* 61:95–120.

Klein, Thomas 1993. The directional non-iteration of German umlaut: a case against Generalized Alignment. Ms.

Koskenniemi, Kimmo 1983. *Two-level Morphology: a General Computational Model for Word-form Recognition and Production*. Publication no.11, Department of General Linguistics, University of Helsinki.

Kratochvil, Paul 1964. Syllabic stress patterns in Peking dialect. *Archiv Orientální*, 32.3:383–402.

1967. On the phonology of Peking stress. *Transactions of the Philological Society*, 154–178.

1969. Syllabic volume as acoustic correlate of perceptual prominence in Peking dialect. *Unicorn* (Princeton Chinese Linguistics Project) 5:1–28.

1974. Stress shift mechanism and its role in Peking dialect. *Modern Asian Studies* 8.4:433–458.

1987. The case of the third tone. In *The Wang Li Memorial Volumes*, pp.253–276.

Ladd, D. Robert 1980. *The Structure of Intonational Meaning*. Bloomington: Indiana University Press.

1984. Declination: a review and some hypotheses. *Phonology Yearbook* 1:53–74.

1990. Metrical representation of pitch register. In Kingston and Beckman, pp.35–57.

Lakoff, George 1993. Cognitive phonology. In Goldsmith 1993a, pp.117–145.

Language Atlas of China. 1987. Compiled jointly by Australian Academy of the Humanities and Chinese Academy of Social Sciences. Hong Kong: Longman.

Leben, William R. 1973. Suprasegmental phonology. Ph.D. dissertation, Massachusetts Institute of Technology.

Lee, Kai-fat 1976. Polysyllabicity in the Modern Chinese verb: an attempt to quantify a linguistic drift. *Journal of Chinese Linguistics* 4:24–46.

Lehiste, Ilse 1980. Phonetic manifestation of syntactic structure in English. *Annual Bulletin* 14:1–27. Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Levelt, Willem J.M. 1989. *Speaking: from Intention to Articulation*. Cambridge, Mass.: MIT Press.

Li, Rong 1979. Wenling fangyan de liandu biandiao [Tone sandhi in the Wenling dialect]. *Fangyan* 1:1–29.

1985. Guanhua fangyan de fenqu [Subgrouping of Mandarin dialects]. *Fangyan* 1:2–5.

1989. Zhongguo de yuyan he fangyan [Languages and dialects of China]. *Fangyan* 4:161–167.

Li, Weimin 1981. Shilun qingsheng he zhongyin [On neutral tone and stress]. *Zhongguo Yuwen*, 35–40.

Li, Xingjian and Liu Sixun 1985. Tianjin fangyan de liandu biandiao [Tone sandhi in the Tianjin dialect]. *Zhongguo Yuwen* 1:76–80.

Liao, Rong-rong 1994. Pitch Contour Formation in Mandarin Chinese: a Study of Tone and Intonation. Ph.D. dissertation, Ohio State University.

Liberman, Mark, B-H. Juang, and Frank Soong. Chinese text-to-speech: how to get there. Technical report, Murray Hill: AT&T Bell Laboratories.

Liberman, Mark and Janet Pierrehumbert 1984. Intonational invariance under changes in pitch range and length. In Mark Aronoff and Richard T. Oehrle (eds.), *Language, Sound, Structure*, pp.157–233. Cambridge, Mass.: MIT Press.

Liberman, Mark and Alan Prince 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8:249–336.

Lin, Jowang 1994. Lexical government and tone group formation in Xiamen Chinese. *Phonology* 11:237–276.

Lin, Mao-can 1983. The stress pattern and its acoustic correlates in Beijing Mandarin. *Proceedings of the Tenth International Congress of Phonetic Sciences*, pp.504–514.

Lin, Mao-can, Yan Jing-zhu, and Sun Guo-hua 1984. Beijinghua liangzizu zhengchang zhongyin de chubu shiyan [A preliminary experimental study on the normal stress of disyllabic words in Beijing]. *Fangyan* 57–73.

Lin, Tao 1985. Tantao beijinghua qingsheng xingzhi de chubu shi yan [A preliminary experimental study of neutral tone in Beijing]. In Lin, pp.1–26.

(ed.) 1985. *Working Papers in Experimental Phonetics*. Beijing University Press.

Lin, Yenhui 1989. Autosegmental treatment of segmental processes in Chinese phonology. Ph.D. dissertation, University of Texas, Austin.

Lisker, Leigh and Arthur Abramson 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10:1–28.

Liu, Fenghsi 1980. Mandarin tone sandhi: a case of interaction between syntax and phonology. Paper delivered at the summer meeting of the Linguistic Society of America, Albuquerque, New Mexico.

Lü, Shuxiang 1963. Xiandai hanyu danshuang yinjie wenti chutan [On monosyllabicity and disyllabicity in contemporary Chinese]. *Zhongguo Yuwen* pp.10–22.

1980. Danyang fangyan de shengdiao xitong [The tonal system of Danyang]. *Fangyan* 2:85–122.

Lü, Shuxiang et al. 1980. *Xiandai Hanyu Babai Ci* [*800 [Function] Words in Contemporary Chinese*]. Beijing: Shangwu Yinshuguan.

Luo, Chang-pei 1930. *Xiamen Yinxi. Bulletin of the Institute of History and Philology*, *Academia Sinica*. [Reissued by the Chinese Academy of Science, Shanghai, 1956.]

Luo, Meizhen 1982. Fujian Changting Kejiahua de liandu biandiao [Tone sandhi in the Changting Hakka dialect in Fujian]. *Yuyan Yanjiu* 3:188–197.

Ma, Xin-min 1990. Le Parler de Nanzhuang (Guanzhong, Shaanxi). Doctoral thesis, University of Lausanne.

Ma Newman, Roxana 1971. Downstep in Ga'anda. *Journal of African Languages* 10:15–27.

Maddieson, Ian 1972. Tone system typology and distinctive features. *Proceedings of the Seventh International Congress of Phonetic Sciences*: 958–961.

1978a. Universals of tone. In J.H. Greenberg (ed.), Universals of Human Language, vol. II, pp.335–366. Stanford University Press.

1978b. The frequency of tones. *UCLA Working Papers in Phonetics* 41:43–52.

Manomaivibool, Prapin 1976. Tonal derivation in Thai. Paper presented at the Tenth International Sino-Tibetan Conference, Georgetown University.

Matisoff, James A. 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study. *JAOS* 90.1:13–44.

1973. Tonogenesis in Southeast Asia. In Hyman, pp.71–96.

Mazaudon, Martine 1977. *Tibeto-Burman Tonogenetics*. Linguistics of the Tibeto-Burman Area 3.2:1–123.

McCarthy, John 1979. On stress and syllabification. *Linguistic Inquiry* 10:443–466.

1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17:207–264.

1988. Feature geometry and dependency: a review. *Phonetica* 43:84–108.

1989. Linear order in phonological representation. *Linguistic Inquiry* 20:71–99.

1993. The parallel advantage: containment, consistency and alignment. Paper presented at the Rutgers Optimality Workshop, Rutgers University.

1995a. Remarks on phonological opacity in Optimality Theory. In Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky (eds.), *Proceedings of the Second Colloquium on Afro-Asiatic Linguistics* (Sophia Antipolis, June 1994).

1995b. Some misconceptions about Optimality Theory. Ms., University of Massachusetts, Amherst.

McCarthy, John and Alan Prince 1986. Prosodic morphology. Ms., Waltham, Mass: Brandeis University.

1990. Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory* 8:209–84.
1993a. Prosodic morphology. Ms., Rutgers University.
1993b. Generalized alignment. Ms., University of Massachusetts, Amherst, and Rutgers University.
1994a. *Prosodic morphology*, parts 1 and 2. Prosodic morphology workshop, Utrecht.
1994b. Optimality in prosodic morphology: the emergence of the unmarked. NELS.
1995a. Prosodic morphology. In Goldsmith, pp.318–366.
1995b. Faithfulness and reduplicative identity. Ms., University of Massachusetts and Rutgers University.

McCawley, James 1970. A note on tone in Tiv. *Studies in African Linguistics* 2:257–270.
1978. What is a tone language? In Fromkin, pp.113–131.

Mei, Tsulin 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. *Harvard Journal of Asiatic Studies* 30:86–110.
1977. Tones and tone sandhi in sixteen century Mandarin. *Journal of Chinese Linguistics* 5:237–260.

Meredith, Scott 1990. Issues in the phonology of prominence. Ph.D. dissertation, Massachusetts Institute of Technology.

Mester, R. Armin and Junko Itô 1989. Feature predictability and underspecification: palatal prosody in Japanese mimetics. *Language* 65:258–293.

Mester, R. Armin and Jaye Padgett 1993. Directional syllabification in generalized alignment. Paper presented at the Rutgers Optimality Workshop, Rutgers University.

Meyer, Antje S. 1990. The time course of phonological encoding in language production: the encoding of successive syllables of a word. *Journal of Meaning and Language* 29:524–545.
1991. The time course of phonological encoding in language production: phonological encoding inside a syllable. *Journal of Meaning and Language* 30:69–89.

Mock, Carol C. 1988. Pitch accent and stress in Isthmus Zapotec. In Hulst and Smith, pp.197–224.

Mohanan, K.P. 1982. Lexical phonology. Ph.D. dissertation, Massachusetts Institute of Technology. [Revised as *The Theory of Lexical Phonology*. Dordrecht: Reidel. 1986].

Mohanan, K.P. 1995. The organization of the grammar. In Goldsmith, pp.24–69. Cambridge, Mass. and Oxford: Blackwell.

Mok, Suisang 1993. The unstructured foot in Mandarin third tone sandhi. Ms., University of California, San Diego.

Myers, Scott 1991. Persistent rules. *Linguistic Inquiry* 22:315–344.

Napoli, Donna J. and Marina Nespor 1979. The syntax of word initial consonant gemination in Italian. *Language* 55:812–841.

Nespor, Marina and Irene Vogel 1982. Prosodic domains of external sandhi rules. In Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations*, part 1, pp.225–256. Dordrecht: Foris.
1986. *Prosodic Phonology*. Dordrecht: Foris.

Newman, John and Robert G. Petterson 1990. The tones of Kairi. *Oceanic Linguistics* 29:49–76.

Norman, Jerry 1973. Tonal development in Min. *Journal of Chinese Linguistics* 1:222–238.

1988. *Chinese*. Cambridge University Press.

Noske, Roland 1993. *A Theory of Syllabification and Segmental Alternation*. Tübingen: Max Niemeyer.

Odden, David 1985. An accentual approach to tone in Kimatuumbi. In D. Goyvaerts (ed.), *African Linguistics: Studies in Memory of M.W.K. Seminenke*, pp.345–419. Amsterdam: John Benjamins.

1987. Kimatuumbi phrasal phonology. *Phonology* 4:13–36.

1988. Predictable tone systems in Bantu. In Hulst and Smith, pp.225–251.

1995. Tone: African languages. In Goldsmith 1995a, pp.444–475.

Ohala, John 1978. Production of tone. In Fromkin, pp.5–39.

Orgun, (Cemil) Orhan 1995. Correspondence and identity constraints in two-level Optimality Theory. Ms., University of California, Berkeley.

Packard, Jerome L. 1990. A lexical morphological approach to word formation in Mandarin. *Yearbook of Morphology* 3:21–37.

1994. Packard, Jerome L. (ed.) forthcoming. *Word Formation in Chinese*. Berlin: Mouton de Gruyter.

Padgett, Jaye 1995. Review of Goldsmith 1993a. *Phonology* 12:147–155.

Pankratz, Leo and Eunice V. Pike 1967. Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics* 33:287–299.

Paradis, Carole 1988. On constraints and repair strategies. *Linguistic Review* 6:71–91.

Pierrehumbert, Janet. 1980. The phonology and phonetics of English intonation. Ph.D. dissertation, Massachusetts Institute of Technology.

Pierrehumbert, Janet and Mary E. Beckman 1988. *Japanese Tone Structure*. Cambridge, Mass.: MIT Press.

Pike, Kenneth 1948. *Tone Languages*. Ann Arbor: University of Michigan Press.

Poteet, Stephen 1985. Tone sandhi and the phonological structure of Mandarin. Ms., University of California, San Diego.

Prince, Alan 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11:511–562.

1983. Relating to the grid. *Linguistic Inquiry* 14:19–100.

1985. Improving tree theory. In M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Jeder (eds.), *Proceedings of Berkeley Linguistic Society* 11:471–490.

1990. Quantitative consequences of rhythmic organization. In Michael Ziolkowski, Manuela Noske, and Karen Deaton (eds.), *Papers from the 26th Regional Meeting of the Chicago Linguistic Society* 12.2:355–398 (Parasession on the Syllable in Phonetics and Phonology).

Prince, Alan and Paul Smolensky 1993. *Optimality Theory*. Technical Report no.2, Rutgers Center for Cognitive Science, Rutgers University.

Pritchett, B.L. 1992. *Grammatical Competence and Parsing Performance*. University of Chicago Press.

Pulleyblank, Douglas 1986. *Tone in Lexical Phonology*. Dordrecht: Reibel.

Pulleyblank, Edwin G. 1978. The nature of the Middle Chinese tones and their development to Early Mandarin. *Journal of Chinese Linguistics* 6:173–203.

1991. Chinese dialect studies. In Wang, pp.431–455.

Pyles, Thomas and John Algeo 1982. *The Origins and Development of the English Languages*. Harcourt Brace Jovanovich.

Qian, Nairong 1992. *Dangdai Wuyu Yanjiu* [*Studies in the Contemporary Wu Dialects*]. Shanghai: Jiaoyu chubanshe.

Qian, Zengyi et al. 1981. *Yantai fangyan baogao* [*A Report on the Dialect of Yantai*]. Jinan: Qilu Shushe.

1993. *Boshan Fangyan Yanjiu* [*A Study of the Boshan Dialect*]. Beijing: Shehui Kexue Wenxian Chubanshe.

Rao, Changrong 1987. Fujian Changting (Kejia) fangyan de liandu biandiao [Tone sandhi in the (Kejia) dialect of Changting, Fujian]. *Zhongguo Yuwen* 3:191–200.

Rose, Phil 1990. Acoustics and phonology of complex tone sandhi. *Phonetica* 47:1–35. [A Chinese version also appears in Zee 1995, pp.239–260.]

Rose, Sharon 1994. Plane conflation, underspecification and palatalization in Chaha. *Proceedings of the West Coast Conference on Formal Linguistics* 12:101–116.

Rossi, Mario 1971. Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23:1–33.

Sagart, Laurent 1986. On the departing tone. *Journal of Chinese Linguistics* 14:90–113.

Sampson, Geoffrey 1969. A note on Wang's phonological features of tone. *International Journal of American Linguistics* 35:62–66.

Schuh, Russell G. 1978. Tone rules. In Victoria A. Fromkin (ed.), *Tone: a Linguistic Survey*, pp.221–256. Academic Press.

Selkirk, Elisabeth O. 1974. French liaison and the X′ notation. *Linguistic Inquiry* 5:574–590.

1980. Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff and M.-L. Kean (eds.), *Juncture*. Saratoga, Calif.: Anma libri.

1981a. On prosodic structure and its relation to syntactic structure. In T. Freiheim (ed.), *Nordic Prosody* II. Trondheim: TAPIR. [Also available through the Indiana University Linguistics Club, 1980.]

1981b. On the nature of phonological representation. In J. Anderson, J. Laver, and T. Myers (eds.), *The Cognitive Representation of Speech*, Amsterdam: North Holland.

1984. *Phonology and Syntax*. Cambridge, Mass.: MIT Press.

1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371–405.

Selkirk, Elisabeth O. and Shen Tong 1990. Prosodic domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, pp.313–338. University of Chicago Press.

Shattuck-Hufnagel, Stephanie 1995. The importance of phonological transcription in empirical approaches to "stress shift" versus "early accent": comments on Grabe and Warren, and Vogel, Bunnell, and Hoskins. In Connell and Arvaniti, pp.128–140.

Shen, Jiong 1985. Beijinghua shengdiao de yinyu he yudiao [The pitch range and intonation in Beijing]. In Lin, pp.73–130.

Shen, Xiao-nan 1990a. *The Prosody of Mandarin Chinese*. Berkeley: University of California Press.

1990b. Tonal coarticulation in Mandarin. *Journal of Phonetics* 18:281–295.

1992. On tone sandhi and tonal coarticulation. *Acta Linguistica Hafniensia* 24:131–152.

Shen, Zhongwei and William S-Y. Wang 1995. Wuyu zhuoseyin de yanjiu – tongji shang de fenxi he lilun shang de kaolü [An investigation into voiced stops in Wu dialects – statistical analysis and theoretical considerations]. In Zee, pp.219–238.

Sherard, Michael 1972. Shanghai phonology. Ph.D. dissertation, Cornell University.

1979. Syntactic constraints on tone sandhi in Shanghai. *Computer-aided Analysis of Asian and African Languages* 5.

1980. *A Synchronic Phonology of Modern Colloquial Shanghai*. Tokyo: Computer-aided Analysis of Asian and African Languages, monograph n.5.

Shi, Feng 1990. Hanyu he Dong-Tai yu de Shengdiao Geju [Tone Paradigms in Chinese and Kam-Tai languages]. Ph.D. dissertation, Nankai University.

1994. Songqi shengmu duiyu shengdiao de yingxiang [The influence of aspirated consonants on tones]. Ms., Nankai University.

Shi, Feng, Lin Shi, and Rongrong Liao 1987. An experimental analysis of the five level tones of the Gaoba Dong language. *Journal of Chinese Linguistics* 15:335–361.

Shih, Chilin 1986. *The Prosodic Domain of Tone Sandhi in Chinese*. Ph.D. dissertation, University of California, San Diego.

1987. The phonetics of the Chinese tonal system. Technical memorandum, AT&T Laboratories.

1988. Tone and intonation in Mandarin. Ms., Cornell University and AT&T Laboratories.

1991. Pitch variation across word boundary. Paper presented at the Third North American Conference on Chinese Linguistics, Cornell.

1997 [= 1990]. Mandarin third tone sandhi and prosodic structure. In J. Wang and N. Smith (eds.), *Studies in Chinese Phonology*, pp.81–124. Dordrecht: Foris.

Sietsema, Brian 1989. Metrical Dependencies in Tone Assignment. Ph.D. dissertation, Massachusetts Institute of Technology.

Silverman, Daniel 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 2:289–328.

Singh, Rajendra 1987. Well-formedness conditions and phonological theory. In Wolfgang Dressler et al. (eds.), *Phonologica 1984*. Cambridge University Press.

Smolensky, Paul 1986. Information processing in dynamical systems: foundations of a harmony theory. In David E. Rumelhart and James L. McClelland (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, Mass.: MIT Press.

Snider, Keith 1988. Towards the representation of tone: a three-dimensional approach. In Harry van der Hulst and Norval Smith (eds.), *Features, Segmental Structure and Harmony Processes*, part 1, pp.237–267. Dordrecht: Foris Publications.

1990. Tonal upstep in Krachi: evidence for a register tier. *Language* 66:453–474.

Sproat, Richard and Chilin Shih 1991. The cross-linguistic distribution of adjective ordering restrictions. In Carol Georgopoulos and Roberta Ishihara (eds.), *Interdisciplinary Approaches to Language: Essays in Honor of S-Y. Kuroda*, pp.565–594. Dordrecht: Kluwer Academic.

1992. Mandarin morphology still is not stratum-ordered. Ms., AT&T Bell Labs.

1993. Why Mandarin phonology is not stratum-ordered. *Yearbook of Morphology*, 185–217.

Steriade, Donca 1982. Greek prosodies and the nature of syllabification. Ph.D. dissertation, Massachusetts Institute of Technology.

1986. Locality conditions and feature geometry. *NELS* 17:595–617.

Su, Xiaoqing 1985. Xuzhou fangyan de liandu biandiao [Tone sandhi in the Xuzhou dialect]. *Xuzhou Shifan Xueyuan Xuebao* 1:154–165.

Tan, Fu 1987. Tone sandhi in the Tianjin dialect. *Journal of Chinese Linguistics* 15:228–246.

Tanenhaus, Michael K. and John C. Trueswell 1995. Sentence comprehension. To appear in Joanne Miller and Peter Eimas (eds.), *Handbook of Perception and Cognition*, vol. XI. Academic Press.

Tang, Jane C-C. 1990. Chinese phrase structure and the extended X′-theory. Ph.D. dissertation, Cornell University.

Tang, Zuo-fan 1960. Hunan Dongkouxian Huangqiaozhen fangyan [The dialect of Huangqiao town in Dongkou county, Hunan Province]. *Yuyanxue Luncong* 4:63–133.

Thompson, Sandra 1973. Resultative verb compounds in Mandarin: a case for lexical rules. *Language* 49:361–379.

Thongkum, Theraphan L. 1987. Another look at the register distinction Mon. *UCLA Working Papers in Phonetics* 67:132–165.

Thurgood, Graham 1992. From atonal to tonal in Utsat (a Chamic language of Hainan). In L. Buszard-Welcher et al. (eds.), *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society*, (Special Session on the Typology of Tone Languages, pp.145–156).

Ting, Pang-hsin 1981. Hanyu shengdiao yuanyu yunwei shuo zhi jiantao [A critical evaluation of the tonogenetic hypothesis of Chinese]. *Proceedings of the International Conference on Sinology*, pp.267–283. Taipei: Academia Sinica.

1983. Derivation by tone-change in colloquial Min. Paper presented at the American Association of Asian Studies, San Francisco.

1984. Wuyu shengdiao zhi yanjiu [A study of Wu tones]. *Bulletin of the Institute of History and Philology, Academia Sinica* 55:605–638.

1996. Tonal evolution and tonal reconstruction. In Huang and Li, pp.141–159.

Trubetzkoy, N.S. 1969. *Principles of Phonology*. (Transl. by Christiane A.M. Baltaxe.) Berkeley: University of California Press.

Tsay, Jane Suchuan 1994. Phonological pitch. Ph.D. dissertation, University of Arizona.

Tseng, Chiu-yu. 1981. An acoustic phonetic study on tones in Mandarin Chinese. Ph.D. dissertation, Brown University.

Vogel, Irene, Timothy Bunnell, and Steven Hoskins 1995. The phonology and phonetics of the Rhythm Rule. In Connell and Arvaniti, pp.111–127.

Wang, Jun et al. 1984. *Zhuangdong Yuzu Yuyan Jianzhi* [*A Brief Description of Kam-Tai Languages*]. Beijing: Minzu chubanshe.

Wang, Li 1932. *Une prononciation chinoise de Po-pei (province de Kouang-si), étudiée à l'aide de la phonétique expérimentale*. Imprimerie des Presses Universitaires de France. [Po-pei = Bobai]

Wang, William S-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33:93–105.

1969. Competing changes as a cause of residue. *Language* 45:9–25.

(ed.) 1991. *Languages and Dialects of China.* (Proceedings of the Conference on Languages and Dialects of China, Oakland, Calif.; sponsored by the Wang Institute of Graduate Studies.) *Journal of Chinese Linguistics* Monograph series no.3.

Wang, William S-Y. and Kung-pu Li 1967. Tone 3 in Pekingese. *Journal of Speech and Hearing Research* 4:130–136.

Welmers, William E. 1959. Tonemics, morphotonemics and tonal morphemes. *General Linguistics* 4:1–9.

1962. The phonology of Kpelle. *Journal of African Languages* 1:69–93.

Whitaker, Katherine P.K. 1955–56. A study on the modified tones in spoken Cantonese. *Asia Major* 5:9–36; 6:184–207.

Wong, Maurice K-S. 1982. Tone change in Cantonese. Ph.D. dissertation, University of Illinois, Urbana-Champaign.

Woo, Nancy H. 1969. Prosody and phonology. Ph.D. dissertation, Massachusetts Institute of Technology.

Wright, Martha S. 1983. A metrical approach to tone sandhi in Chinese dialects. Ph.D. dissertation, University of Massachusetts, Amherst.

Wu, Weishan 1986. Xiandai hanyu sanyinjie zuhe guilü chutan [On trisyllabic constructions in contemporary Chinese]. *Hanyu Xuexi* 5:1–2.

Wu, Zong-ji 1982. Putonghua yuju zhong de shengdiao bianhua [Pitch changes in the connected speech of Standard Mandarin]. *Zhongguo Yuwen* 439–450.

1985. Putonghua sanzizu biandiao guilü. *Zhongguo Yuyan Xuebao* 2:70–92.

Wu, Zong-ji and Cao, Jian-fen 1979. Shiyan yuyinxue zhishi jianghua [Lectures on experimental phonetics], part 4, *Zhongguo Yuwen*. 393–398.

Xiamen University 1982. *Putonghua Minnan Fangyan Cidian*. Hong Kong: Sanlian Shuju.

Xu, Bao-hua, Tang Zhen-zhu, and Nai-rong Qian 1981–83. Xinpai shanghai fangyan de liandu biandiao [Tone sandhi in the New Shanghai dialect]. *Fangyan* 1981:145–55; 1982:115–128; 1983:197–201.

Xu, Bao-hua, Tang Zhen-zhu et al. 1988. *Shanghai Shiqu Fangyan Zhi* [*Urban Shanghai Dialects*]. Shanghai Jiaoyu Chubanshe.

Yan, Jing-zhu and Mao-can Lin 1988. Beijinghua sanzizu zhongyin de shengxue biaoxian [The phonetic realization of stress in trisyllabic expressions of Beijing]. *Fangyan* 227–237.

Yan, Sen 1981. Gao'an (Laowu Zhoujia) fangyan de yuyin xitong [The phonological system of the Laowu Zhoujia variety of the Gao'an dialect]. *Fangyan* 104–121.

Ye, Xiangling 1979. Suzhou fangyan de liandu biandiao [Tone sandhi in Suzhou]. *Fangyan* 30–46.

1983. Wujiang fangyan shengdiao zai diaocha [Wujiang tones revisited]. *Fangyan* 32–35.

Yin, Zuoyan 1982. Guanyu putonghua shuangyin changyongci qingzhong yin de chubu kaocha [A preliminary study of stress in common disyllabic expressions in Putonghua]. *Zhongguo Yuwen* 168–173.

Yip, Moira 1980. The tonal phonology of Chinese. Ph.D. dissertation, Massachusetts Institute of Technology.

1989. Contour tones. *Phonology Yearbook* 6:149–174.

1990. The phonology of Cantonese loanwords: evidence for unmarked settings for prosodic parameters. In Jerry Packard (ed.), *New Directions in Chinese Linguistics*. Dordrecht: Kluwer Academic.

1991. Prosodic morphology of four Chinese dialects. *Journal of East Asian Linguistics* 1:1–35.

1992. The spreading of tonal nodes and tonal features in Chinese dialects. In Buszard and Welcher et al., pp.157–166.

1993. Tonal register in East Asian languages. In K. Snider and H. van der Hulst (eds.), *The Representation of Tonal Register*. Dordrecht: Foris, pp.245–268.

1995. Tone in East Asian languages. In Goldsmith, pp.476–494.

You, Ru-jie 1992. *Hanyu Fangyanxue Daolun* [*An Introduction to Chinese Dialectology*]. Shanghai: Jiaoyu Chubanshe.

Yuan, Jiahua 1960. *Hanyu Fangyan Gaiyao* [*An Introduction to Chinese Dialectology*]. Beijing: Wenzi Gaige Chubanshe.

Zee, Eric 1980. A spectrographic investigation of Mandarin tone sandhi. *UCLA Working Papers in Phonetics* 49:98–116.

(ed.) 1995. *Wuyu Yanjiu* [*Studies of the Wu dialects*]. New Asia Academic Bulletin, vol. XI. Hong Kong.

Zee, Eric and Ian Maddieson 1980. Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *Glossa* 14:45–88. [Also in *UCLA Working Papers in Phonetics*, March 1979, pp.93–129.]

Zhang, Chengcai and Shuikui Zhu 1987. *Xining Fangyan Zhi* [Records of the Xining dialect]. Xining: Qinghai Renmin Chubanshe.

Zhang, Hongming 1988. Xinpai Shanghai shiqu fangyan liandu yinbian zhong de zhuoyin shengmu qinghua [The devoicing of voiced onsets in the connected speech of New Shanghai]. *Wuyu Luncong* 140–153.

1992. Topics in Chinese phrasal phonology. Ph.D. dissertation, University of California, San Diego.

Zhang, Huiying 1979. Chongming fangyan de liandu biandiao [Chongming tone sandhi]. *Fangyan* 4:284–302.

1980. Chongming fangyan de sanzizu liandu biandiao [Trisyllabic tone sandhi in Chongming]. *Fangyan* 1:15–34.

Zhang, Zhengsheng 1988. Tone and tone sandhi in Chinese. Ph.D. dissertation, Ohio State University.

Zhang, Zhenxing 1982. Zhangping (Yongfu) fangyan de liandu biandiao. [Tone sandhi in Zhangping (Yongfu)]. *Fangyan* 175–196.

Zheng-Zhang, Shangfang 1964a. Wenzhou yinxi [The phonology of Wenzhou]. *Zhongguo Yuwen* 28–60.

1964b. Wenzhou fangyan de liandu biandiao [Tone sandhi in Wenzhou]. *Zhongguo Yuwen* 106–152.

1980–81. Zheng-Zhang. Wenzhou fangyan erweici de yuyin bianhua [Morpho-phonemics involving er-suffixation in Wenzhou]. *Fangyan* 1980, 4:245–262; 1981, 1:40–50.

Zhou, Youguang 1964. *Hanzi Gaige Gailun* [*On Chinese Orthographic Reform*]. Beijing: Wenzi Gaige Chubanshe. [Reissued in 1978 by Macao: Erya Chubanshe.]

Zhu, Dexi 1956. Xiandai hanyu xingrongci de yanjiu [A study on adjectives in modern Chinese]. *Yuyan yanjiu* 1:83–112.

Zoll, Cheryl 1994. Directionless syllabification in Yawelmani. ROA-28.

1995. Conflicting directionality. Ms., University of California, Berkeley.

Zwicky, Arnold 1977. On clitics. Bloomington: Indiana University Press.

1985. Clitics and particles. *Language* 61:283–305.

Zwicky, Arnold M. and Geoffrey K. Pullum 1983. Cliticization vs. inflection: English *n't*. *Language* 59:502–513.

Subject index

Author index