

Although we have concentrated in this section on language-internal motivations for change, the Structuralists were very much aware that external considerations could also influence the onset and direction of linguistic changes. For example, Martinet notes that gaps in a system may remain empty for lengthy periods, and may be filled at a specific time due to external factors such as language contact. Similarly, Meillet accepts that some similarities between related languages appear only after divergence from the common ancestor, so that resemblances among languages may reflect parallel development rather than genetic relationship alone. Finally, Jakobson (in Keiler 1972) is at pains to point out that borrowing and linguistic convergence may have profound effects on language structure. We shall return to matters of language contact in Chapters 7-11, but should be aware that the Structuralists invoked external as well as internal causation in their search for explanations of change.

There are, however, two problems for Structuralist accounts of sound change. First, it is unclear how change is ever to happen, if every element in the system is dependent on every other element; and second, if units have no meaning in isolation, but only gain their significance from their place in the system of a certain language at a certain time, how can we compare different languages or different stages of the same language? Presumably comparison presupposes complete structural equivalence, which is extremely unlikely to exist; to pursue our analogy of colours, *blue* in English cannot, in strict Structuralist terms, be compared with the word for 'blue' in Russian or Welsh, since these cover slightly different areas of the spectrum, and are opposed to different sets of terms.

2.4 The Generativists

What really changes is not sounds, but grammars (Postal 1968: 270)

2.4.1 Generative Grammar

In the last two sections, we have focussed on schools of linguistic thought which essentially constitute closed chapters in the history of the discipline. Diachronic (and synchronic) linguists have learned, and still can learn much from the approaches of the Neogrammarians and Structuralists; and indeed, the Generativists themselves might be described as Structuralist in their concentration on linguistic systems. However, there are now very few linguists working entirely within the Neogrammarian or Structuralist paradigms.

The Generative school, however, has not so far been superseded by a new, dominant paradigm, so that numerous linguists would still describe themselves as Generativists. In the last quarter-century, Generative syntax

and phonology have both fragmented into a number of related but distinct synchronic and diachronic models. Since not all of these are yet fully developed, an up-to-date survey would necessarily be incomplete and also very complex theoretically. Consequently, this section will focus on early Generative work in the so-called Standard Model, as practised primarily in the 1960s. This has the advantage of allowing a relatively finished picture to be drawn, since Generativists today only very rarely adhere to the Standard Model. The picture might be seen as a rather old-fashioned one, but only if it is taken as an attempt to portray the current state of Generative linguistics, rather than a study of the beginnings of Generative theory included to maintain the historical perspective of this chapter. More recent aspects of Generative theory will be explored briefly in Chapter 3 (for phonological change) and Chapter 5 (for historical syntax).

Generative Grammar, with Noam Chomsky as its leading figure, came to prominence from the late 1950s with works on synchronic syntax (Chomsky 1957, 1965); the Generative approach spread to phonology (Halle 1959, Chomsky and Halle 1968), and later to sound change (King 1969a). One of its primary innovations was the realisation that the class of well-formed or grammatical sentences in any language is potentially infinite; consider the English sentence *I know Sarah and Clive and Aidan and Kathy...*, which obviously could be extended *ad infinitum*. Native speakers are also constantly producing novel sentences which they have not previously uttered. If I say *Alexandra fed forty-two lletuces to the llama*, I can feel reasonably confident that I have not uttered this sentence before; indeed, this may be the first time anyone has produced it. Even so, other native speakers of the same language can readily understand such newly-minted sentences, provided that they are grammatical.

Sentences in a human language will be grammatical if they follow particular patterns; they are not simply random collections of words thrown together in random order. Furthermore, the repeated patterns found in a given language will tend to be rather few, and this insight lies behind the Generative analysis of language and language acquisition. Many earlier studies had assumed that children acquire their native language by listening, memorising and repeating; however, this would mean that children could not move beyond the repertoire of sentences they had already heard, whereas in fact they often produce novel utterances. The Generative approach to this problem crucially involved regarding language as a rule-governed phenomenon: in acquiring a language, children would have to learn a small set of rules, and a larger but still finite set of words, and could then from these finite resources produce or generate a potentially infinitely large set of grammatical sentences.

Acquisition of even these rules or patterns, however, is not strictly a

learning task, given Chomsky's further assumption that all human languages are analysable using the same sorts of rules and grammars. If the same types of structures are common to all languages, Chomsky reasons, then why should they not be already present in the mind at birth? A genetically inherited outline of language universals is consequently assumed to be innate. Children learn their own language by listening to linguistic data and filling language-specific details into this largely pre-set plan. This interaction between the innate Universal Grammar and heard, language-specific data allows the child to build a personal, internalised mental grammar. This internal grammar, which Chomsky calls competence, allows native speakers to understand and produce sentences, and represents their tacit or subconscious knowledge of their native language. The Generative linguist's task is then to build an explicit rule system which does the same job as this internal grammar. The Generativists were consequently concerned with the notion of a linguistic system, but while the Structuralists considered actual linguistic data, the Generativists concentrated on the underlying system of rules; whereas the Structuralists said 'phonemes change', the Generativists said 'rules change'. Indeed, actual language data, which Chomsky calls performance, was of strictly limited interest to the early Generativists, precisely because slips of the tongue, tiredness, lack of concentration and numerous other such factors ensure that many actual utterances do not entirely follow the rules which generate the idealised forms of competence. Chomsky hypothesises that children can abstract away from the limited and imperfect data they receive, in constructing their grammar. However, the primacy of competence is clear from the simplifying assumption that 'Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors... in applying his knowledge of the language in actual performance' (Chomsky 1965: 3). It is easy to see that issues of variation were not addressed in early Generative theory.

Two short illustrative examples might be useful at this point. A Generative syntactician would assume that a native English speaker would regard the sentences in (22) as related.

(22)

- Alexandra fed the llama.*
Did Alexandra feed the llama?
Alexandra did not feed the llama.
The llama was fed by Alexandra.
Was the llama fed by Alexandra?
The llama was not fed by Alexandra.

Early Generative syntax assumed that these sentences share essentially the same, rather abstract deep structure form, which might look something like ALEXANDRA FEED LLAMA, although each would also have markers to specify whether it would end up as a statement, negative or question on the surface. Transformational rules then turned the deep structures into the different surface structures; for instance, the rule for making questions said roughly 'add an appropriate form of the verb *do* at the beginning of the sentence'.

Similarly, in Generative phonology the goal was to assign to each morpheme, each meaningful part of a word, a single unique shape called the underlying form. The surface forms which actually appear in different contexts were then derived from this by rule. For instance, we might assume that native speakers of English know a unit meaning 'negative', with the underlying form /n/. When this prefix is added to different adjectives, the nasal assimilates partly or completely to the first consonant of the adjective stem, as shown in (23). This variation is secondary and rather superficial; what really matters is that all these ostensibly different prefixes can be related, and the variations derived by rule.

(23)

- /n/ = negative prefix
 [ɪn] *temperate*
 [ɪm] *plausible*
 [ɪk] *coherent*
 [ɪl] *relevant*
 [ɪlɪ] *literate*

Such alternations of sounds are very common; another example from English is *electric*[k] – *electricity* – *electrician*, where the stem morpheme would again have a single underlying form, say /ɛlektrɪk/. Generative phonological reference to the morpheme, a non-phonological unit, also illustrates the Generative idea that phonology interacts with other components of the grammar, such as the morphology or syntax, so that sound change may be non-phonetically conditioned.

The Generative linguist, then, aims to write a grammar mirroring the native speaker's competence. However, the number of grammars which could generate the same set of sentences is potentially infinite, and it is unclear which we should prefer. The primary early Generative criterion was simplicity: we should assume that native speakers are creatures whose minds work on a principle of least effort. Their internalised grammars must therefore operate with minimal apparatus and fuss, generating all the necessary data with the minimum complexity; that is, they must be maximally simple. This requirement is a version of Occam's Razor, which states that *entia non sunt multiplicanda praeter necessitatem*, 'entities are not to be multiplied beyond necessity' – or, less formally, don't ask for

more units than you need. If the linguist's formal grammar is to match the internal one, it must also be maximally simple, and in early Generative terms, simplicity was measured by counting the number of rules and units in the grammar, with lower numbers preferred (this, of course, is itself a slightly simplified outline).

2.4.2 *The Generative theory of sound change*

2.4.2.1 Introduction

In Generative theory, then, all sound change was seen as change in the grammar. Change between two related languages or varieties was established by constructing and comparing systems of rules and underlying forms for each stage or language. Logically, phonological change could only occur in the form, order or inventory of rules, or in the underlying representations.

The early Generative belief in evaluation procedures based on measurements of simplicity also led to the belief that all change must be simplifactory, translating more complex to simpler and more economical grammars. The Generativists therefore denied the functional motivation advocated by the Structuralists. For instance, King (1967) measured the functional load of various oppositions of sounds (that is, roughly, the number of pairs of words distinguished by the difference between these sounds). He used data from a number of Germanic languages, concentrating on sounds which are known to have subsequently undergone sound change, and his results contradict in part the predictions of the functional load hypothesis. For instance, /y/ has been lost by merger in Icelandic. There are two other Icelandic vowel phonemes close to /y/ in terms of their component features, with which /y/ might have merged; one is /u/, and the other /i/. The opposition /y/-/i/ had a functional load more than four times higher than that of /y/-/u/, but /y/ nevertheless merged with /i/.

In the next two subsections, we shall consider the mechanisms of phonological change proposed by the Generativists: first, the different types of rule change; then, restructuring of the underlying forms.

2.4.2.2 Rule change

2.4.2.2.1 Rule addition

Rule addition, which is also simply called innovation, was the most basic type of change recognised by the Generativists, and also the only sort of change which could affect the grammar of adult speakers; all the other changes to be discussed here take place between generations, as children acquiring language internalise a slightly different grammar from their

parents'. In rule addition, the speaker adds a new rule to the end of his phonological rule system; this constitutes a response to the introduction of a sound change, and the change and the resulting phonological rule will generally be identical. Rule addition only changes the synchronic rule system, not the underlying forms of morphemes; any change at the underlying level will only take place later, if at all.

An example of an added rule is given in (24).

$$(24) \left[\begin{array}{c} + \text{obstruent} \\ - \text{continuant} \\ + \text{voice} \end{array} \right] \longrightarrow [-\text{voice}]$$

This rule, although formulated using Generative phonological notation, has identical results to the second part of Grimm's Law, in (10b) above, and states that voiced stops (or voiced, obstruent non-continuant) become voiceless; their value for voicing changes. A Generative statement of the diachronic sound change would differ from the synchronic phonological rule only in the conventional use of an arrow (\longrightarrow) in the latter, and a shaftless arrowhead ($>$) in the former.

The consensus view in later Generative work was that such rules would only be added at the end of the phonological rule system. Earlier, a process called rule insertion (King 1969a: 43) had allowed new rules to be added anywhere in the grammar. This was rejected because a rule introduced early in the inventory might operate cumulatively with later phonological rules to produce an output rather remote from the previous pronunciation, potentially impairing communication between generations. However, a rule added at the end of the rule component would alter pronunciations only according to its own effect, allowing communication with speakers as yet lacking the change.

2.4.2.2.2 Rule loss

As well as appearing in a grammar, rules were said to disappear. King's (1969a: 46ff) example of rule loss involves final devoicing in Yiddish.

Yiddish ultimately derives from Middle High German, which underwent a sound change (and therefore added a rule) devoicing final stops and fricatives. This produced the alternations shown in (25b), which were not present in Old High German (25a) when the change had not yet operated. Most Modern German dialects retain this final devoicing rule, but Yiddish does not; as (25c) shows, voiced obstruents may appear finally in Yiddish.

- (25) a. Old High German
 gab 'he gave' - *gabum* 'we gave'
 tag 'day' - *tage* 'days'

- b. Middle High German
gap – *gaben*
tac – *tage*
- c. Standard Yiddish
tog 'day' – *teg* 'days'
noz 'nose' – *mezer* 'noses'

We can hypothesise that final devoicing *did* apply in Yiddish, since isolated words like *avek* 'away' and *gel* 'money' have voiceless final stops in Modern Yiddish, but had voiced ones in Old High German. Some Yiddish dialects also maintain final devoicing. In Generative terms, then, Yiddish did have a devoicing rule, but has lost it. Such loss probably should not be interpreted as the instantaneous disappearance of a rule, but rather as a gradual reduction in its scope; the process applies to fewer and fewer forms and ultimately fades out of the system.

2.4.2.2.3 Rule reordering

Two successive stages of a language, or two different dialects, sometimes seem to have the same two phonological rules applying in different orders. For instance, after German final devoicing was introduced, around AD 1000, another change lengthened vowels before voiced obstruents; so, in the fifteenth and sixteenth centuries, words like *lob* 'praise' and *veg* 'path' surfaced with short vowels in the nominative forms and long ones in the genitive (26a). This distribution is best accounted for if we assume that final devoicing preceded vowel lengthening, and that the synchronic order of rule application matched the chronological order of introduction of the sound changes. In Modern German, however, both nominatives and genitives have long vowels, suggesting that the rules have come to apply in the opposite order.

- (26) a. underlier: *lob* *lobes* *veg* *vege*
 final devoicing: *lop* — *vek* —
 V lengthening: — *lo:bes* — *ve:ge*
- b. underlier: *lob* *lobes* *veg* *vege*
 V lengthening: *lo:b* *lo:bes* *ve:g* *ve:ge*
 final devoicing: *lo:p* — *vek* —

One possible reason for such a rule reordering is suggested by Kiparsky (1978), who invokes the notion of rule simplification. Generative rules are regarded as simpler if they apply maximally and can be stated economically: an obvious case of simplification is shown in (27), where the

simplified version has fewer features and applies potentially to more segments. The symbol # indicates a word boundary.

- (27) a. $\left[\begin{array}{c} + \text{obstruent} \\ + \text{continuant} \end{array} \right] \longrightarrow [-\text{voice}] / \text{ — } \#$

- b. $[+ \text{obstruent}] \longrightarrow [-\text{voice}] / \text{ — } \#$

Kiparsky argues that rule reordering also depends on the notion of maximal application, the requirement that each rule should do as much work as possible. In Kiparsky's terms, two rules may be ordered in feeding order, or in bleeding order. In feeding order, the first rule provides contexts for the next, as shown in (28a). But in bleeding order, the first rule robs the second of segments to which it could otherwise have applied, as in (28b).

- (28) a. Feeding order:

1. $v \longrightarrow f$
 2. $f \longrightarrow \phi$

- b. Bleeding order:

1. $v \longrightarrow f$
 2. $v \longrightarrow w$

In (28a), rule 1 provides cases of /f/ for rule 2 to delete, but in (28b), rule 1 shifts all instances of /v/ to /f/, leaving no /v/s for rule 2 to apply to. Kiparsky argues that, to allow maximal application and thus increased simplicity, rules will tend to be reordered either into feeding order, or out of bleeding order. In the German final devoicing case, we see reordering out of bleeding order.

2.4.2.2.4 Rule inversion

The final type of rule change is rule inversion (Vennemann 1972), which involves the reinterpretation of original surface forms as underlying forms. One example involves the treatment of /r/ in certain accents of English, including Southern British English Received Pronunciation (RP). The operation of a number of eighteenth century sound changes in the ancestor of this variety meant that /r/ appeared after a restricted set of vowels, namely /ɔ:/, /ɑ:/, /ə/ and /ɜ:/, which we can regard as a long schwa (/ə/). Some examples are shown in (29).

- (29)
- | | |
|-------------|-----------------------------------------|
| /ɔ:/ | <i>our, floor, for, love, shore ...</i> |
| /ɑ:/ | <i>star, bazaar, far ...</i> |
| /ɪə/ | <i>beer, fear, near, here ...</i> |
| /eə/ | <i>care, there, air, square ...</i> |
| /ʊə/ - /ɜ:/ | <i>assure, pure, lure, poor ...</i> |
| /ɜ:/ | <i>stir, fur, world, err, heard ...</i> |

/aɪə/ choir, fire...
 /aʊə/ flower, tower...
 /ɔ/ letter, father, sugar, figure...

Subsequently, a process of /r/-deletion operated in a number of English accents, whereby /r/ was lost before consonants and pauses, as shown in (30). The resulting distribution of [r] is characteristic of so-called non-rhotic accents: rhotic varieties maintain [r] wherever it appears in the spelling.

- (30) *hr*-Deletion: *hr* → ∅ $\left\{ \begin{array}{l} \text{##} \\ \text{C} \end{array} \right\}$

[r] retained initially: red, robe, rule...
 [r] retained intervocally: very, hurry, soaring...
 [r] lost: beard, heart, car...

This conditioned loss of /r/ led to alternations of [r] and zero, with [r] appearing only in forms of a word with a following vowel: this pattern, shown in (31), is still very common among RP speakers, and is referred to as the use of linking [r].

- (31) Alternations of [r] - ∅:

soar[Ø] - *soar*[ɪŋ] - *soar*[r] in the sky
feared[Ø] - *feared*[ɪ] - *feared*[r] of flying
for[Ø] - *for*[ɪ] Anna
star[Ø] - *star*[ɪ] - *sugar*[Ø] - *sugar*[r]ly
letter[Ø] - *put the letter*[r] in here
Peter[Ø] - *Peter*[r] isn't my favourite person

However, instead of learning that words like *soar* have underlying /r/, which is then deleted before consonants and pauses, some speakers seem to have undergone a rule inversion. These speakers analyse *soar* as having no final /r/ at the underlying level (after all, *soar* in isolation is never pronounced with [r] in their accent) and insert [r] instead after the vowels /ɔ: a: ə/ when another vowel follows. This [r]-insertion rule is given in (32).

- (32) [r]-Insertion:
 ∅ → [r] / ɔ: a: ə / ... V

If *soar* is just /sɔ:/ underlyingly for speakers with [r]-insertion, they should be unable to distinguish it from *saw*, which is also /sɔ:/; that is, they should apply [r]-insertion to forms which had final /r/ historically and to forms with the same final vowel which did not. In fact, this prediction is borne out, since a growing number of people are producing so-called intrusive [r] - instances of [r], regularly inserted by [r]-insertion, but in words which had no historical [r]. Some cases are given in (33).

- (33) *saw*[Ø] - *saw*[ɪŋ] - *withdraw*[Ø] - *withdraw*[ɪ]al
banana[Ø] - *banana*[ɪ] - *magenta*[Ø] - *magenta*[ɪ]ish
Kafka[Ø] - *Kafka*[ɪ]esque; *Shaw*[Ø] - *Shaw*[ɪ]ism
law[Ø] - *law*[ɪ] and order
comma[Ø] - *put the comma*[r] in there
idea[Ø] - *the idea*[r] is
Anna[Ø] - *Anna*[r] isn't my favourite person

Speakers who use intrusive [r] may also insert it when they are using foreign words or acronyms, or even speaking (or singing, in the case of the Latin) a foreign language (34).

- (34) Foreign words:
 the social milieu [mi:ljɜ:ɪ] of Alexander Pope
 the junta [xunta] in Chile
 Acronyms:
 as far as BUPA[r] is concerned

Foreign languages:
 French: *j'étais déjà*[ɪ] ici
 German: *ich habe*[ɪ] einen Hund
 Latin: *hosanna*[r] in excelsis,
dona[r] eis requiem

(Data partly from Wells 1982)

Furthermore, when vowels which would not otherwise trigger [r]-insertion are reduced to schwa, [r] is regularly inserted, as shown in (35).

- (35) *tomato*[r] and *cucumber* production
 the *window*[r] isn't clean
eyeshadow[r] and *make-up*
 Cockney: *I'll tell you how* [ɪrɪ:]
 to it [ɪrɪ:]
 Norwich: *run over by a* [bɪrɪ] bus
out to [ɪrɪ] eat, *quarter to* [ɪrɪ] eight
 (data from Wells 1982, Trudgill 1974)

The data given in (33)-(35) above can only be accounted for if we assume that some speakers have undergone a rule inversion, so that, instead of analysing *soar* as /sɔ: r/ and deleting /r/ in certain contexts, they internalise it as /sɔ:/ and insert [r] in the opposite set of environments. For speakers who produce linking but not intrusive [r], we can either assume that the inversion has not taken place, or that they succeed in inserting [r] in *soaring* but not *sawing* by referring to the spelling.

2.4.2.3 Restructuring

As we have already seen, all rule changes apart from rule addition were ascribed to the child in early Generative theory: the child was seen as creative, constructing a grammar which is flexible for a time, then becomes

fixed. Children were also said to be responsible for the second major type of phonological change, restructuring of the underlying representations.

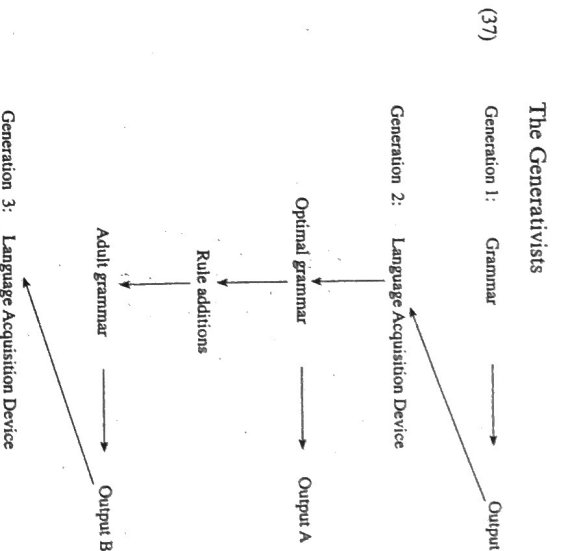
King (1969a) provides an extended example of restructuring involving English /hw/ versus /w/. At an earlier stage of English (36a), adult speakers had contrasting /hw/ in *whales*, *which* and /w/ in *Wales*, *wich*, and this situation still persists in certain Modern English dialects, such as Scottish English. However, /hw/ gradually merged with /w/ in varieties like Southern British English, leading to the addition of a rule as in (36b). Adults, who can only add rules, would retain /hw/ in the underlying forms of *whales*, *which*, but cease to produce it.

The next generation of children have the target of constructing a maximally simple grammar of their variety of English. Since they will never hear [hw] from their parents, there is no motivation for them to postulate contrastive /hw/ and /w/, plus a merger rule; instead, they will hypothesise underlying /w/ in all previous /w/ and /hw/ words, as in (36c). This restructuring produces a revised, simpler grammar, without the rule merging /hw/ with /w/, but with the effects of that rule transferred into the underlying representations.

- (36) a. Adults: /hw/ [hw] /w/ [w]
 whales *Wales*
 which *wich*
- b. New rule: /hw/ → [w]
 Underlier: /hwɪ/ *which*
 Surface: [wɪ] *which*
- c. Children: /w/ [w]
 which
 wich
 Wales
 whales

2.4.2.4 Summary

A schematic representation of the Generative theory of sound change is given in (37).



This Generativist model of change raises a number of difficulties. First, early Generative theorists tended not to consider the effects of change on the system of sounds in a language (the phoneme system in Structuralist terms). They consequently omitted from consideration Structuralist explanations of change, which centred on the structure and function of the system. Instead, the Generativists wrote formal phonological rules which reflect completed changes; but these are only restatements of the effects of the change, and are essentially non-explanatory.

Secondly, the notion of simplification, although it might be a candidate for explanation in the case of some changes, cannot be the cure-all the Generativists claimed it to be. Not all changes can be construed as simplification; it is hard, for instance, to see what a sound shift like Grimm's Law or the Great Vowel Shift might simplify, and Neogrammarian changes of the irregularity-creating variety certainly introduce complexity into the grammar. Even more basically, how can the addition of a new rule ever be considered a simplification? Furthermore, we have no entirely clear idea of what constitutes simplification, which is a relative rather than an absolute term. In other words, a particular change might be simplification under certain circumstances, but induce complexity in others. Finally, the characterisation of all change as simplification presumes a view of change as constantly creating ever simpler grammars. Languages, on this view, must have been getting gradually simpler ever since language began, a claim for which there is no evidence at all. Models promoting a single explanation in this way can all too easily be hijacked to promote a view of some languages as simpler or better than others, and of