

---

## Chomsky and Halle's Revolution in Phonology\*

B. Elan Dresher

University of Toronto

### 0. Introduction

Chomsky and Halle's approach to phonological theory, as with other components of generative grammar, represented a sharp break with the main currents of American linguistics that immediately preceded them. The differences were conceptual as well as technical. Accounts of the development of phonology emphasize technical issues, such as arguments over the existence of a 'taxonomic phonemic level', or whether it is permissible to 'mix levels' in a phonological analysis. Lying behind discussion of these issues, however, were assumptions about psychology and the practice of science. Indeed, throughout the development of phonology, major changes came about not only through technical breakthroughs, but also by reinterpreting the significance of existing technical devices. This was also the case with Chomsky and Halle's innovations.

In this chapter I discuss Chomsky and Halle's contributions to phonological theory by putting their views in the context of the theories that prevailed before them. I will also try to connect the technical issues to the larger conceptual ones concerning the nature of language acquisition and the mind.<sup>1</sup> I will be treating Chomsky and Halle's contributions together, without attempting to distinguish who contributed precisely which ideas. Their early work in generative phonology, culminating in the major work *The Sound Pattern of English* (Chomsky and Halle 1968, henceforth *SPE*), was done jointly.

---

Nevertheless, some indication of what each brought to the enterprise can be gleaned from Chomsky's 1957b review of Jakobson and Halle's *Fundamentals of Language* (Jakobson and Halle 1956). Chomsky finds that "much can be said" for Jakobson and Halle's approach to phonology. In particular, he approved of the hypothesis that the sound systems of all languages could be characterized in terms of a limited number of universal distinctive features. Second, he preferred their approach to identifying phonemes over others then current. They assigned two segments to the same phoneme if they have the same feature specifications. Most other approaches to phonemic analysis prevailing at the time assigned sounds to phonemes if they are in complementary distribution (or in free variation) and phonetically similar, appealing to a notion of similarity that is difficult to define. Finally, Chomsky seconds the authors' emphasis (advanced over the years by Jakobson) on the importance of extending phonological theory to account for language acquisition, disorders, and other aspects of linguistic behaviour.

On the other side, Chomsky observes that many of Jakobson and Halle's proposals need to be made more explicit and precise before they can be empirically tested. He further proposes an amendment to their conception of how phonemes are related to speech. He found their requirement that the distinctive features assigned to phonemes be present in their correct sequence in the phonetics too strict. He proposes that distinctive feature specifications form instead an "abstract underlying system of classification related, perhaps indirectly, to the physical facts of speech." Finally, Chomsky proposes that general criteria of simplicity play an important role in the evaluation of particular phonological analyses.

---

One can say, then, that Chomsky and Halle's theory of generative phonology was a synthesis of Jakobson and Halle's theory of distinctive features and phonemic analysis, revised in the light of Chomsky's emphasis on formal explicitness, simplicity, and abstractness and autonomy of mental representations.

### 1. Rules and derivations

When first introduced, the centrality of *rules* in Chomsky and Halle's approach to phonology appeared revolutionary. A grammar of a language must merely *list* many things – for example, the English word *tide* begins with a *t*, ends with a *d*, and has a vowel sound represented by *i*. A person who knows English but who happens never to have encountered this word cannot derive this information. It is a particular fact about English that must be learned and committed to memory.

Other facts about the pronunciation of this word are more systematic. For example, the *t* in *tide* is pronounced with a puff of air, called aspiration (represented as *t<sup>h</sup>*), in contrast to the *t* in *style*, which is not aspirated. Any speaker of English told that *tide* begins with *t* would automatically know that the *t* must be pronounced with aspiration. That is, the aspiration of *t* is not an idiosyncratic fact that must be listed in the lexical entry of *tide*, but can be encoded in a rule. Thus, the lexical, or *underlying*, form of the word *tide* need only specify that the initial sound is a /t/, where slant brackets represent *phonemic* forms; this form is then subject to the rule of aspiration, which derives the *phonetic*, or *surface*, form [t<sup>h</sup>] (where square brackets represent phonetic forms).

The vowel written *i* is a diphthong, phonetically [a:j], where : indicates that the vowel is long, and *j* represents a glide. The length of the vowel is

predictable: *tide* ends in a *d*, which is a voiced sound, and in English, stressed vowels lengthen before voiced sounds. Thus, the vowel of *bid* is longer than the vowel of *bit*, which ends in a voiceless sound, *t*. Similarly, the diphthong in *tight* is shorter than in *tide*. In some dialects, such as Canadian English, the first part of this diphthong is pronounced with a higher and more centralized tongue position before a voiceless consonant, and is phonetically transcribed as [ʌj].

Therefore, speakers of these English dialects need only learn that the diphthongs in *tide* and *tight* are both /aj/. General rules then apply to lengthen /aj/ to [a:j] before voiced sounds and to raise it to [ʌj] before voiceless sounds.

One might suppose, as it generally was in pre-generative phonology, that the distribution of the various phonetic realizations, or *allophones*, of a phoneme could be represented by an unordered set of statements. The diphthong /aj/, for example, appears as [a:j] before voiced sounds (a process we will call Lengthening) and as [ʌj] before voiceless sounds (Raising).

Chomsky and Halle proposed, however, that rules must be *ordered* if they are to give correct results and be stable in the simplest, most general way. Consider, for example the words *ride* and *write*. They undergo the rules of Lengthening and (in the dialects under consideration) Raising. The rules can be written as follows:<sup>2</sup>

- (1) Lengthening V ---> [+long] / \_\_\_\_\_ (glide)  $\left[ \begin{array}{c} \text{C} \\ +\text{voiced} \end{array} \right]$
- (2) Raising /a/ ---> ʌ / \_\_\_\_\_ glide  $\left[ \begin{array}{c} \text{C} \\ -\text{voiced} \end{array} \right]$

These rules also apply in *rider* and *writer*. In the pronunciation of North American English, the *t* and *d* in these words are pronounced with an alveolar ‘flap’, in phonetic transcription [ɾ], a quick tap of the tongue rather than a sustained occlusion.

(3) Flapping {t,d} ---> ɾ / V (glide) \_\_\_\_\_  $\left[ \begin{array}{c} \text{V} \\ \text{-stressed} \end{array} \right]$

The result of applying the three rules is shown below.

(4) A simple derivation

	<i>writer</i>	<i>rider</i>
Underlying	/rajtər/	/rajdər/
Lengthening	–	ra:jdər
Raising	rʌjtər	
Flapping	rʌjrər	ra:jrər
Phonetic	[rʌjrər]	[ra:jrər]

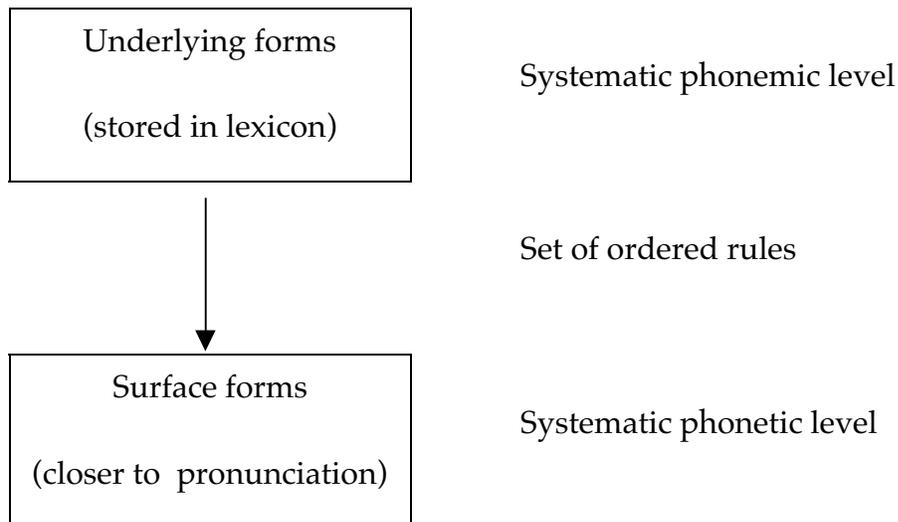
Note that the Flapping rule must follow the other two rules. If Flapping were to apply first, an incorrect form would be generated: as shown in (5), *writer* would be pronounced just like *rider*, which is not the case in this dialect.<sup>3</sup>

(5) An incorrect derivation

	<i>writer</i>	<i>rider</i>
Underlying	/rajtər/	/rajdər/
Flapping	rajrər	rajrər
Lengthening	ra:jdər	ra:jdər
Raising	–	–
Phonetic	*[ra:jrər]	[ra:jrər]

Therefore, the basic architecture of the phonological theory of *SPE* can be diagrammed as in (6).

(6) Basic architecture of phonological component (*SPE*)<sup>4</sup>



In hindsight, one might wonder why the rather simple model in (6) would have ever been considered revolutionary. None of the basic ingredients were novel: not the idea of two basic levels, nor even the idea of a derivation mediated by ordered rules. However, in the context of phonological theory in America in the 1950s, it represented a significant new departure. To see why this was so

---

requires a brief excursion to the nineteenth century origins of modern phonetics and phonology.

### 1.1. Two levels: broad and narrow transcription

That at least two levels of representation are required to represent the sounds of a language was becoming apparent already in the nineteenth century.

Phonologists and phoneticians realized that a degree of precision in the representation of sounds was required that was unattainable using conventional alphabets. They aimed to develop a system in which one sound was always represented by one symbol (unlike English, where the sound [s] is sometimes represented by <c> as in <city>, and sometimes by <s>, as in <sit>), and in which one symbol is used for only one sound (again unlike English, where the letter <c> sometimes represents [s], and sometimes [k] (<electricity>)). This movement ultimately led to the development of the International Phonetic Alphabet (IPA), a transcription system that approaches the goal of a one-to-one relation between sounds and symbols. This type of transcription came to be known as *narrow transcription*, because it records very fine distinctions between speech sounds.

It quickly became apparent that a narrow transcription is not a practical way to transcribe particular languages. For example, consider English *t* in words like *stop*, *top*, *hat*, *not you*, *trap*, and *writer*. When looked at closely, these [t]s are all different: unaspirated [t] in *stop*; aspirated [t<sup>h</sup>] in *top*; unreleased [t̚] in *hat* (optionally: it may also be released and aspirated); palatalized [tʲ] in *not you*; retroflexed [ɖ] in *trap*; and flapped [ɾ] in *writer*. These are only some of the realizations of English *t*. Detailed examination of other sounds of English reveals

---

that they, too, are not unitary sounds but groups of sounds, distinguishable by separate phonetic symbols. The IPA has a way of distinguishing all these sounds, and this is desirable if we wish to give an accurate transcription of what each variant actually is. But it would be very cumbersome and quite impractical to actually attempt to use this type of transcription as a way of writing English.

More important, a narrow transcription fails to do justice to some basic facts about the sound system of English. For there is something correct about the intuitions of speakers that the sounds listed above are all 'variants of *t*'. A transcription system that treats [t] and [t<sup>h</sup>] as being as different from each other as each is to [p] or [n] is missing something important: the English spelling system, for all its faults and quirks, does a better job at capturing the way sounds actually pattern in English. Thus, alongside narrow transcription there developed the notion of a *broad* transcription, which is designed to abstract away from predictable variations and alternations in sounds.

In the above example, [t] and [t<sup>h</sup>] are allophones of the same phoneme /t/, whereas [n] in *nip* is an allophone of a different phoneme, /n/. We know that /t/ and /n/ are different phonemes in English because they are in *contrast*: *tap* and *nap* are different words in English, as are *fit* and *fin*. So one function of a broad transcription is that it abstracts away from allophonic variation and represents only contrastive differences of sounds.

Rule-governed behaviour in sound systems is not limited to allophonic variation. Consider the English plural -s. Following a voiced sound it is pronounced [z], as in *dogs*, *beds*, *bees*, *sins*, and *dolls*. Following a voiceless sound, it is pronounced [s], as in *cats*, *ropes*, and *sticks*.<sup>5</sup> Since this alternation between [s] and [z] is rule-governed and entirely predictable, it is plausible to suppose that

---

the regular plural morpheme has a single lexical representation, say /z/, and that English speakers apply a rule devoicing /z/ to [s] following a voiceless sound.

English s and z, however, are not merely allophonic variants of a single phoneme in English. When these sounds do not immediately follow a consonant, they contrast, as in *sip* vs. *zip*, and *bus* vs. *buzz*. Therefore, the phonological rules in (6) also include rules that change one underlying phoneme into another, in addition to rules that create allophones of a phoneme (without changing phonemic identity). In the case of the plural, the English spelling system again more closely approximates a broad than a narrow transcription, consistently writing the regular plural as <s> even when it is pronounced [z]. Similarly, the final segment in *electric* is consistently written with a <c>, whether it is pronounced [k], or [s] (as in *electricity*), or [ʃ] (as in *electrician*).

Whereas a narrow transcription should ideally be universal, a broad transcription is language particular, reflecting the patterning of sounds in particular languages.

### *1.2. Narrow transcription in phonological theory*

Students of phonology brought up in the tradition of generative grammar will readily identify broad transcription with Chomsky and Halle's systematic phonemic level and narrow transcription with the systematic phonetic level. Indeed, the model in (6) appears to be a natural translation into phonological theory of the two types of transcription. However, some difficulties had to be

---

overcome in arriving at the model in (6). The first of these concerns the nature of the phonetic level: to what extent is it truly a 'systematic' level of representation?

Leonard Bloomfield (1887–1949), and the American linguists who followed him, known as the post-Bloomfieldians, maintained that a phonetic level corresponding to a narrow transcription cannot be supported as a legitimate linguistic representation because it is not systematic, but arbitrary. According to Bloomfield, such a transcription is dependent on the background and perception of the transcriber: some transcribers will notice and note down certain subphonemic distinctions, but others that are less familiar to them will go unrecorded, particularly as they are not crucial to marking contrastive sounds in the language.

For example, an English-speaking transcriber might record that the *t* in the English word *two* is aspirated, because the distribution of aspirated and unaspirated /t/ in English is systematic. But there are many other aspects of this sound that may or may not be noted: whether the sound is dental (made with the tongue against the teeth) or alveolar (tongue against the alveolar ridge); whether the lips are rounded (as they are in *two*) and, if so, how much; whether the tongue is released quickly and simultaneously with the puff of air, or whether the tongue lags a bit, creating an affricated (tending to [tʰ]) or palatalized [tʰ] sound; and so on.

Since a linguistic representation must be based on more than just the whims of individual transcribers, Bloomfield concluded that there is no principled level of phonetic representation corresponding to a narrow transcription.

---

As pointed out by Chomsky (1964), this argument rests on the assumption that there is no universal theory of phonetic representation. Lacking such a theory, it would appear that a phonetic representation has no principled basis. However, a universal feature theory, of the sort initiated by Prague School linguists and developed in works such as Jakobson, Fant, and Halle 1952, Jakobson and Halle 1956, and subsequently revised by Chomsky and Halle (1968), can serve as the basis for a phonetic transcription. The universal set of distinctive features is designed to discriminate all and only those aspects of sounds that are contrastive in the languages of the world. SPE, for example, uses twelve distinctive features to represent the consonant sounds of English.<sup>6</sup> The existence of a universal set of phonetic features constrains what can go into a phonetic representation. No such theory existed in American linguistics, so there was no basis for a systematic phonetic level.

How, then, are the sounds of a language to be represented in a linguistic description? We are left with broad transcription, or, in terms of (6), the systematic phonemic level. However, the systematic phonemic level is too remote from the surface phonetics, that is, too abstract, to serve as the only level of phonological representation.

Consider the English vowel system. In English, unstressed vowels tend to reduce to schwa [ə] in many contexts:

## (7) English vowel - schwa alternations

Tense vowel	Reduced	Lax vowel	Reduced
Canáđian [ej]	Cánada [ə]	Asiátic [æ]	Ásia [ə]
managériál [ij]	mánager [ə]	telégraphý [ɛ]	télegraph [ə]
horízón [aj]	horizóntal [ə]	medícinal [ɪ]	médicine [ə]
custóđian [ow]	cústody [ə]	photógraphý [a]	phótograph [ə]
sulfúric [juw]	súlfur [ə]	prodúction [ʌ]	próduct [ə]

Thus, [ə] is an allophone of every English vowel phoneme. It follows that a phonemic representation of the above words should include unreduced vowels only; reduction to schwa would then be a rule-governed allophonic variation.

Bloch (1941) argued that while such a system is indeed elegant, it poses problems for a learner (as well as a linguist unfamiliar with the language). What happens when learners come across a schwa whose unreduced version is unknown to them, as in words like *sofa* or *of*? Or even *manager*, if they haven't heard a related form such as *manageriál*? If there were a 'lower' phonetic level of representation, a learner could at least represent the phonetic form of such words with a schwa, while deferring a decision as to which underlying phoneme to assign it to. But, having rejected a phonetic level, post-Bloomfieldian theory had no recourse to such a level of representation. The consequence is that learners (and linguists) would be unable to assign any phonological representation to such utterances.

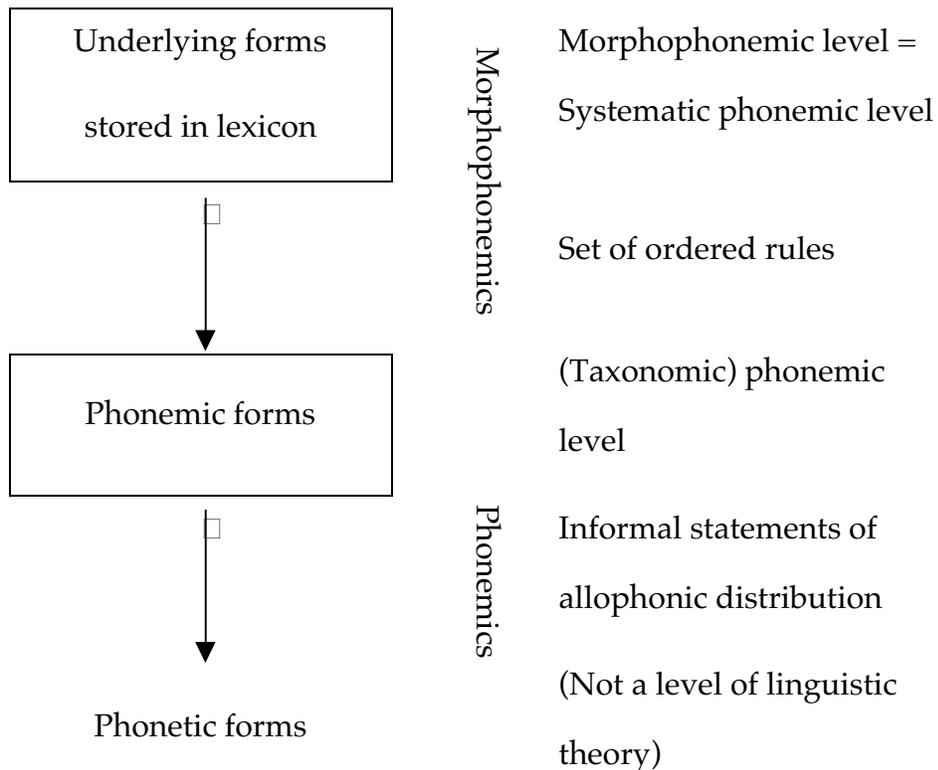
Moreover, according to Bloch, the only data relevant to phonemic analysis are "the facts of pronunciation," that is, the distribution of surface allophones,

---

and not, for example, the existence of morphologically related forms. This assumption severely limits the evidence one can use in arriving at a phonological analysis. It presupposes an analyst who has no access to the fact that the word *manager* is related to *managerial*. Such an analyst would not be in a position to know that the final schwa of the former is related to the stressed vowel of the latter.

Thus, without a systematic phonetic level, the post-Bloomfieldians needed a new level of representation that was much less abstract than the systematic phonemic level, and that did not suffer from the arbitrariness they attributed to phonetic representation. In structuralist terminology, this level was simply called the phonemic representation, and the more abstract systematic phonemic level was called the morphophonemic representation. In the terminology of Chomsky (1964), the new level is called the taxonomic phonemic level. The post-Bloomfieldian conception of the phonological component was thus as in (8).

## (8) Levels in Post-Bloomfieldian American structuralist phonology

1.3. *The new (taxonomic) phonemic level*

Bloch's argument assumes that language learners must be able to encode utterances into phonemes based only on the distribution of surface sounds, or phones. Of course, learners of a language must acquire not only the phonological system, but the rest of the grammar as well, including the morphology and syntax. However, it became an entrenched assumption of American structuralist linguistics that acquisition of language went 'bottom-up', from phones to phonemes, from phonemes to morphemes, from morphemes to syntax, and so on. Though this assumption had no empirical support whatsoever, it had important consequences for the development of phonological theory.

One consequence was the dictum that it is impermissible to ‘mix levels’ in developing a phonemic analysis. That is, a phonemic analysis must be justifiable solely on the basis of allophonic distribution, making no appeal to ‘higher’ levels such as morpheme identity. As Chomsky (1964) showed, this assumption had disastrous consequences for the generality and simplicity of the phonological analysis.<sup>7</sup>

An example is Hockett’s (1951) discussion of a hypothetical language with no underlying contrast between voiced and voiceless consonants: all consonants are voiceless except word-medially between vowels, where they are voiced. This is a fairly common situation, and the natural assumption is that a phonemic representation should indicate only voiceless consonants, since voicing is predictable. Hockett considers the case of two sequences of words in such a language, *pat adak* and *padat ak*. In the standard analysis, these words would have the phonemic representations shown in (9), where # represents a word boundary separating the words.

(9) Consonants voiced only word-medially between vowels

	a.	b.
Phonemic	/#pat#atak# /	/#patat#ak /
Voicing	#pat#adak#	#padat#ak#
Phonetic	[patadak]	[padatak]

Hockett argues that there is a problem with this analysis. The word boundaries do not correspond to any sound, or even to a regular absence of sound or pause. If one hears (9b) [padatak] one would not know, on phonetic grounds alone, whether it derives from /patat#ak/ or /pata#tak/. Of course, one could determine this if one knew something about the lexicon: one might find, for

---

example, that there is a stem /pata-/ with the appropriate meaning but no stem /patat/. But performing such a look up is to 'mix levels'. In Hockett's interpretation of phonemic theory, the phonemic representation must not rely on the proper positioning of word boundaries. If boundaries are omitted from the phonemic representation, then both utterances in (9) are represented as /patatak/. But now we cannot account for the different distribution of voiceless and voiced consonants in (9a) and (9b); the only solution, according to Hockett, is to represent voiced consonants as such in phonemic representation. Rather than the phonemic forms in (9), we would posit /patadak/ for (9a) and /padatak/ for (9b). This result appears to be incompatible with the original concept of a phoneme. Moreover, the generalization that the voicing of consonants is predictable and therefore need not be learned on a case-by-case basis is lost.<sup>8</sup>

The attempt to constrain the phonemic level so as to keep it closer to the surface phonetics led to the requirement that the phonemic level meet a number of further conditions. Their effect was to ensure that there be a one-to-one relation (*biuniqueness*) between allophones and phonemes: given an allophone, it should be possible to unambiguously assign it to a phoneme; and given a phoneme, it must be clear what allophone instantiates it in any given context. The # boundaries in (9) violate biuniqueness, because phonetic [padatak] can derive either from /patat#ak/ or /pata#tak/. Like the prohibition on mixing of levels, these conditions resulted in a loss of generalizations, with no compensatory gain in descriptive or explanatory force.

Consider again the interaction of Flapping and rules affecting the /aj/ diphthong shown above in the *writer/rider* example (4), repeated here.

---

(4)	A simple derivation	
	<i>writer</i>	<i>rider</i>
Underlying	/rajtər/	/rajdər/
Lengthening	–	ra:jdər
Raising	rʌjtər	
Flapping	rʌjɾər	ra:jdər
Phonetic	[rʌjɾər]	[ra:jdər]

It is clear in (4) that [ɾ] is a predictable allophone of both /t/ and /d/, and there is no difficulty in formulating rules to account for its distribution. However, this simple derivation fails a number of conditions that the post-Bloomfieldian linguists placed on phonemic representations.

First, the phonemes /t/ and /d/ have a common allophone, [ɾ]. This amounts to a *partial overlapping* of the two phonemes, which violates the biuniqueness condition. The problem with overlapping is that it is not possible, upon inspection, to decide which phoneme an allophone belongs to. Of course, if we could appeal to morpheme identity we would know that the [ɾ] of *writer* belongs to /t/, because of *write*, and that the [ɾ] of *rider* belongs to /d/, because of *ride*. But this again violates the constraint against mixing of levels.

Another problem with this analysis from the point of view of post-Bloomfieldian theory is that there is a mismatch between the location of the phonemic and the phonetic contrast in *writer* and *rider*: forms that are phonemically different only in their fourth member (/t/ vs. /d/) are

phonetically different only in their second member ([ʌj] vs. [a:j]). According to Chomsky (1964), this mismatch is a violation of the condition of *linearity*.

This example shows also that the notion of *minimal pair* is not a self-evident one. A minimal pair is a pair of words that differ in a single phoneme. Minimal pairs are often used to show that two sounds contrast in a language. For example, we can demonstrate that [s] and [z] contrast in English by adducing minimal pairs such as *sip* and *zip*, or *bus* and *buzz*. Since the only difference in these words is the [s] vs. [z], we conclude that they belong to distinct phonemes. However, a similar test would show that [a:j] and [ʌj] are distinct phonemes in English, since *writer* and *rider* appear to be minimal pairs distinguished in their second elements, not their fourth.

As Chomsky (1964) points out, minimal pairs are thus not evident from the surface, but require that we take into account various kinds of information. In the case of *sip* and *zip* there are no further facts that contradict the conclusion that the distinction is simply between /s/ and /z/.<sup>9</sup>

According to the logic of post-Bloomfieldian phonemics, then, we would have to transcribe *writer* as /rʌjrɚ/ and *rider* as /ra:jrɚ/. In effect, the phonemic level would fail to capture any of the generalizations about English sound patterns discussed above. It would fail to note that [r] is a predictable allophone of /t/ and /d/, and that [ʌj] and [a:j] are predictable allophones of /aj/.

Finally, we have seen that the rules in (4) have to be ordered; ordering was not permissible in American structuralist phonemics. The relationship between a

---

taxonomic phoneme and its allophones had to be stated as a set of unordered distributional statements.

What, then, of the generalizations about sound patterning that are thereby excluded from the phonemics? Where in the grammar, for example, do we represent the fact that there is a single regular English plural, or that the sounds of *write* are systematically related to the sounds of *writer*, or that the stressed vowel in *managerial* is related to the final schwa in *manager*? American structuralist theory had a place for all these generalizations: the morphophonemic component.

#### 1.4. Morphophonemics

In *Menomini Morphophonemics*, Bloomfield presented an analysis that resembles a generative derivation: starting from underlying representations, a series of rules apply in order to yield phonemic representations. This type of analysis, and morphophonemics itself, had a marginal status in structuralist theory. There was very little theorizing done in this area, as opposed to the attention devoted to phonemic theory. In contrast to the latter, morphophonemics had a freewheeling, anything-goes character, which led, as it turns out, to interesting and insightful analyses.<sup>10</sup> Why was there such a contrast between the two components? The answer lies in the degree of 'reality' attributed to each of these levels.

As we have seen, the taxonomic phonemic level was the lowest linguistic level recognized in the theory. By assumption, it had to be a level that could plausibly be attributed to speakers, including those just learning the language. In keeping with the very restricted conception of psychology and learning that prevailed at the time, learners were credited with only the most basic ability to

---

perform operations of grouping and classification. The various constraints placed on the phonemic level were designed to allow a phonemic representation to be easily discovered from the phonetic input available to such a learner.

By contrast, morphophonemics was not given a psychological interpretation. Morphophonemic representations were not necessarily considered to be things that speakers had. According to Anderson (1985:276), Bloomfield considered morphophonemic description to be “an elegant artifact, providing a uniform and concise account of a complex set of facts, but not to be confused with the actual language capacity of speakers. Only the phonemic forms, and the morphological fact of relations between them, could be considered to have that status.”<sup>11</sup>

It is significant that Chomsky’s first work in linguistics, his MA thesis (a later version of his BA thesis), is titled *Morphophonemics of Modern Hebrew*. His later contributions to generative phonology essentially adapted the techniques used in morphophonemics - rules and derivations - and placed them at the centre of phonological theory (and other components of grammar). To make this move, however, Chomsky and Halle had to overcome the arguments in favour of the taxonomic phoneme. Having shown that the ‘elegant fiction’ of the morphophonemic component was actually real, they now had to show that the taxonomic phoneme, the rigorous core of phonological theory, was a fiction.

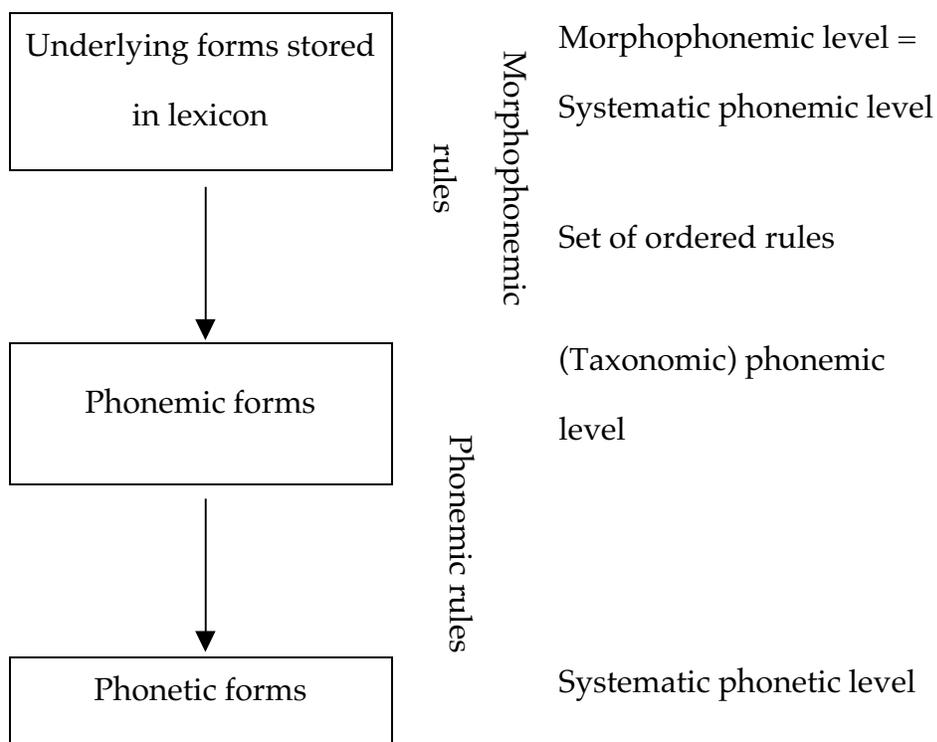
### *1.5. Against the taxonomic phoneme*

As we saw above, there is no empirical support for a taxonomic phonemic level that adheres to the various conditions and restrictions imposed by the post-Bloomfieldians, suggesting that such a level is *unnecessary*. In a famous

argument, Halle (1959) demonstrated that such a level is also *undesirable*, because it leads to a loss of generalizations.

Imagine, then, a phonology with three significant levels: a morphophonemic (systematic phonemic) level. that was not in dispute; a systematic phonetic level based on a universal distinctive feature theory; and mediating between them, the taxonomic phonemic level (10).

(10) Three-level phonological component



In Russian, voicing is a contrastive feature that distinguishes pairs of obstruent phonemes. Thus, phoneme /t/ is distinct from /d/, /k/ is distinct from /g/, /s/ contrasts with /z/, and so on. There is a rule that voices word-final obstruents if a voiced obstruent follows in the next word. Thus, we find [m'ok l,i] 'was (he) getting wet?', with a *k* preceding the sonorant *l*, but [m'og bi] 'were (he) getting wet', where *k* voices to *g* before voiced obstruent *b*. The voicing

rule that changes *k* to *g* changes one phoneme to another, and so it must be a morphophonemic rule, applying as in (11a).

(11) Russian voicing applying twice

a. Morphophonemic voicing

Systematic phonemic	/ /m'ok bi/ /	/ /ž'eč bi/ /
---------------------	---------------	---------------

Voicing	m'og bi	---
---------	---------	-----

Taxonomic phonemic	/m'og bi/	/ ž'eč bi/
--------------------	-----------	------------

b. Allophonic voicing

Taxonomic phonemic	/m'og bi/	/ž'eč bi/
--------------------	-----------	-----------

Voicing	---	ž'ež bi
---------	-----	---------

Systematic phonetic form	[m'og bi]	[ž'ež bi]
--------------------------	-----------	-----------

Three obstruents, /c/, /č/, and /x/, do not have corresponding voiced consonants. However, voicing also applies to these segments as well. We have [ž'eč l,i] 'should one burn?', with voiceless č before the sonorant *l*, but [ž'ež bi] 'were one to burn', where ž is the voiced counterpart to č. Because [ž] is not a phoneme in its own right, but exists only as an allophone of /č/, this application of voicing is an allophonic rule, and must be assigned to the component that maps phonemic forms into phonetic forms (11b)

Halle argued that the derivation in (11) needlessly splits the voicing rule into two (or, alternatively, applies the same rule twice). However, there is no evidence that voicing applies differently in these cases, or that the change occurs in two stages rather than just once. Having a taxonomic phonemic level makes it

---

impossible to capture the generalization that there is one voicing rule at work here, applying equally to all the segments in its purview.<sup>12</sup>

Without the taxonomic phonemic level, the grammar takes on the form of (6), with only two significant levels of representation: the systematic phonemic, or lexical, level, and the systematic phonetic level.

## **2. Grammar as a system of knowledge**

One of Chomsky's most fundamental contributions was to reposition linguistics as a field with implications for the nature of mind and learning. This reorientation required a new way of looking at linguistic description, one that was diametrically opposed to that prevailing in American linguistics up to that time.

### *2.1. Bloomfieldian philosophy of science and psychology*

Leonard Bloomfield introduced a particularly radical form of behaviorism and scientific empiricism to linguistics. This view included an approach to science in general, and to psychology in particular, that together had a great influence on the development of phonological theory.

With respect to science, Bloomfield and his followers took a view that was influential in the 1920s and 1930s, known as Operationalism, that all science must be framed in terms of statements that describe basic operations, such as reports of how long it takes an object to travel a certain distance. Putting the focus on operations seemed to make the 'content' of science out to be observable matters (measurements and experimental techniques) rather than obscure-looking

---

'hidden' entities and principles. Theoretical terms and theory itself, it was thought, could be treated as codes or shorthand for observations and techniques. Such an approach is incapable of characterizing most important scientific theories, and the main body of philosophy of science soon abandoned Operationalism.

In American linguistics, however, this general orientation remained influential, and provided a theoretical underpinning to the bottom-up approach to linguistic analysis. The scientist (here, the linguist) begins with basic data, that is, a set of utterances in a language (a *corpus*). These utterances appear to observers (linguists) as a stream of speech. Observers then perform basic operations to analyze the speech stream: they can *segment* an utterance into individual sounds, or phones; and they can *classify* the phones into phonemes. These operations can then be repeated at a higher level of analysis, segmenting and classifying strings of phonemes into morphophonemes, then into words and phrases, and so on.

American linguistics in fact took this approach even farther than other fields, making it an aim of linguistic theory to devise *discovery procedures* that would automatically apply the techniques of segmentation and classification to any corpus and produce an analysis. As Chomsky (1957a) argued, no other science has hoped to arrive at such procedures, which in effect, would be an algorithm for arriving at the correct theory in a particular domain, using only the restricted 'data' that operationalist ideology allows.<sup>13</sup>

These general views about science were complemented by a set of assumptions about psychology, which amounted to a radical form of behaviorism. Bloomfield and his circle believed that there was no point to

---

attributing a 'mind' to any organisms, including humans.<sup>14</sup> Rather, behaviour is a set of responses to stimuli. Language, then, is simply another form of behaviour, verbal behaviour, a set of learned responses, or habits.

It followed, on this view, that linguistics had nothing to contribute to psychology proper. The big questions of psychology - how learning takes place, for example - were held to be the province of psychologists. The special mandate of linguistics was to study verbal behaviour proper - to describe (and merely that) the way utterances are put together and used (primarily, it was assumed) in communication. Thus, linguists in this framework were given to say that linguistics should be free of psychology, and should make no assumptions about psychology. This statement, given the above, was quite obviously disingenuous. What was meant, however, is that linguists need not get involved in psychological speculation, beyond the background assumptions sketched above. That is, internal to a linguistic description there is no need to invoke psychology. Nevertheless, the analysis is couched in a framework that is heavily indebted to a specific model of psychology.

The combination of scientific and psychological assumptions that formed the background of Neo-Bloomfieldian linguistics placed narrow limits on the types of linguistic theory that could be entertained. Moreover, they tended to draw the focus of inquiry away from evaluating how successful particular theories were in accounting for the facts of language - capturing significant generalizations, accounting for how language is acquired, and so on - and emphasized instead conformity to what were in effect a priori and unmotivated restrictions on the form of a linguistic theory, whatever the consequences might be for particular analyses.

## 2.2. Chomsky's philosophy of science and psychology

Chomsky's approach represented a radical break with Bloomfieldian thought both with respect to science and psychology. With respect to science, Chomsky began with the premise that a linguistic description is a hypothesis about how the data is organized, and thus is a scientific theory of the grammar. He pointed out that no other science has developed, or even seeks to develop, discovery procedures – a set of automatic procedures applied to basic data that results in a theory of the data. Rather, scientists arrive at hypotheses about nature however they can. What can be expected of a scientific theory is not an account, let alone justification, of how it was arrived at, but criteria for assessing how good a theory it is. That is, rather than discovery procedures, linguists should be concerned with developing evaluation procedures, means by which to compare competing theories (grammatical descriptions) with a view to determining which is the better theory.

With respect to psychology, Chomsky argued that behaviorism chose the wrong target. There is not, and likely will not be, a theory of human behaviour.<sup>15</sup> But it may be possible to characterize aspects of cognition, or *knowledge*, that contribute to behaviour. He therefore proposed that the proper object of linguistics is not verbal behaviour, which is the product of diverse systems, but rather knowledge of language. Chomsky proposed that a grammar is actually a theory of the knowledge that native speakers have of their language.

---

### 2.3. *Learnability and Universal Grammar*

If a grammar is an account of knowledge, the question immediately arises as to how native speakers come to have this knowledge. As he has argued at length (notably in Chomsky 1975), one cannot consider learning in a domain apart from the cognitive principles that the learner brings to it. In the case of language, where many considerations point to a specialized ability shared by all humans, there is no general learning theory we can appeal to. Rather, linguistic inquiry itself must determine, first, what is actually acquired, and second, what cognitive principles learners employ. The answer to the first question will be a grammar of the language that has been acquired. The answer to the second question is a set of universal principles collectively called Universal Grammar (UG). Whether a particular grammar is easy to learn or difficult depends on the combination of accessibility to relevant data and the nature of UG.

*SPE* begins with the sentence, "The goal of the descriptive study of a language is the construction of a grammar." Much of *SPE* is devoted to constructing the phonological part of the grammar of English. Freed from restrictions on the relation between phonemes and their surface allophones and from the requirement that morphological relations may not enter into a phonological analysis, the phonology of English that emerges in *SPE* is radically different from previous accounts.

As the opening sentence indicates, *SPE* is much more concerned with the particular grammar of English than with how this grammar could be acquired. Nevertheless, the phonology of English presented in *SPE* is set within a general theory of phonology which is intended to provide a first attempt at a theory of UG in the domain of phonology.

---

The main mechanism proposed by *SPE* for guiding a learner to the correct grammar is the *evaluation measure*. The evaluation measure, assumed to be part of UG, assigns a higher value to a rule that uses fewer features over a rule that uses more features. This measure reflects the fact that features capture natural classes: therefore, the fewer features, the more general the class. A rule that applies to a more general class is generally preferred to one that delimits a narrower class. That is, a simpler rule is preferred over a more complex one.

While such an evaluation measure can guide a learner in making local decisions between similar rules that differ in their formal simplicity, it is of little help in allowing a learner to choose between sets of rules, or even entire competing grammars. With respect to this problem, there are parallels in the early development of phonological and syntactic theory. In both cases, the theories inherited from structuralism were too descriptively limited: they did not possess the resources to provide a descriptively adequate account of their subject matter. In both cases, therefore, the emphasis at first was on expanding the descriptive resources of linguistic theory. In the case of syntax, this meant exploring the power of transformations; in the case of phonology, the power of derivations and the rule formalism introduced by *SPE*.

In syntax, a concern with constraining the theory arose in the late 1960s and eventually came to dominate thinking in the field, at least in the part of the field associated with Chomsky. By this time, however, Chomsky had largely given up working on phonology, and was not involved with subsequent developments.<sup>16</sup> Many of these developments took as their starting point the theory of *SPE*. Here I will only highlight certain issues that particularly pertain to learnability.<sup>17</sup>

---

#### 2.4. *Abstractness of phonology*

The underlying forms posited by *SPE* are in general rather abstract with respect to their phonetic surface forms. Much ink has been shed over the extent to which such abstractness is learnable, or whether there should be constraints on abstractness.<sup>18</sup> It is an empirical question whether phonological systems obey constraints limiting the 'distance' between underlying and surface forms. If they do, we expect to find evidence for this from the patterning of the phonology itself. However, in some cases proposals to constraint abstractness have tended to be based not on empirical evidence, but, rather, on a priori assumptions concerning what is learnable.

There is no particular reason to suppose that there is a simple relation between abstractness and learnability. Consider again, for example, the derivation in (4). The rule of Flapping contributes to create a considerable mismatch between the underlying and surface forms, as we have seen. At the same time, Flapping obscures the contexts of Raising and Lengthening, by making it appear that Raising applies before a voiced consonant (the flap [ɾ]) in *writer* and that Lengthening fails before a voiced consonant in the same form. This derivation is ruled out by theories that place restrictions on abstractness. Kiparsky (1973), while not ruling out such derivations, proposed that rules that are *opaque* by his definition (i.e., contradicted by surface forms) are difficult to learn. Intuitively, one might expect this to be so: a learner who has not yet acquired the rule that raises /aj/ to [ʌj] before voiceless consonants might be

---

misled by hearing forms like *writer* ([rʌjɾər]), where the raising has taken place, though the following consonant is voiced on the surface.

Nevertheless, it is not obvious that this grammar is difficult to learn. The rules of Raising and Centralization apply without complicating factors in the words *write* and *ride*, as well as many other words of English. Similarly, the rule of Flapping applies transparently in many words, such as *sit* ~ *sitter* and *dad* ~ *daddy*. A learner who has learned this much of the grammar already has the main ingredients of the derivation except for the relative ordering of Flapping and the other rules, which must be acquired on the basis of forms like *writer* and *rider*.<sup>19</sup>

Just as it is not the case that abstractness necessarily leads to learnability problems, neither is it the case that reducing abstractness necessarily makes learning easier. For example, in Seri (Marlett 1981) there is a set of vowel-initial words that behave as if they begin with a consonant. Marlett proposes that such words have an underlying initial consonant, designated C, which is deleted toward the end of the derivation. This is an abstract analysis, because it posits a segment that is present in underlying and intermediate representations that is never audible at the surface.

Marlett and Stemberger (1983) present a different analysis. They adopt a nonlinear theory of phonology in which phonological representations consist of different tiers. One of the tiers is a CV-tier, or 'skeleton', which relates segmental features to syllable structure. Rather than posit an abstract underlying consonant, they propose that the exceptional verbs have an initial C position on

the CV-tier that has no counterpart on the feature tiers. The two analyses are presented schematically in (12).

(12) Two analyses of Seri exceptional verbs

a. Abstract C

Phonemic /CaX/

Rules

C-deletion aX

Phonetic [aX]

b. Empty C

CV-tier	C	V	C
---------	---	---	---

--	--	--	--

Feature tiers		a	X
---------------	--	---	---

Marlett and Stemberger comment that the Empty C analysis is superior to the Abstract C analysis because it is less abstract, since the empty C is present at the surface. However, if we focus on learnability and ask what would lead a learner to posit the empty C in (12b), we find that it is exactly the same evidence that would lead the learner to an abstract C in (12a). Since abstractness is a relation between levels of a derivation, it is no surprise that one can reduce abstractness by enriching representations. In (12), the ‘vertical’ relation between a phonemic C and its null phonetic realization is replaced by a ‘horizontal’ relation between a C on one tier and a null representation on the feature tier(s). The analysis in (12b) raises interesting issues about the nature of representations, but does little to advance the problem of learnability.<sup>20</sup>

---

### 2.5. *Markedness and other substantive principles*

While abstractness in itself is not necessarily a special problem for learnability, when combined with the otherwise rather loosely constrained apparatus of *SPE* a real problem for learnability does arise. In short, the theory of *SPE* does not provide an adequate answer as to how learners are able to arrive at the actual grammar of their language and not any number of other grammars.

One shortcoming of the *SPE* theory, recognized within *SPE* itself, is that the theory is overly formal: the evaluation measure counts only the number of symbols in a rule, without taking account of intrinsic properties of features or of phonological processes. Thus, a three-vowel system with vowels /i a u/ is very common, but one with the vowels /ü ε æ/ is unheard of. Similarly, a rule changing *i* to *u* is much more common than one changing *i* to *i*, even though *u* differs from *i* by two features ([back] and [round]) and *i* differs from *i* by only one feature ([back]). To remedy this fundamental theoretical inadequacy Chomsky and Halle introduce a theory of markedness. Taking up and extending ideas from the Prague School, they propose that certain feature values and combinations are *unmarked*, or default, whereas others are *marked*, and entail a greater cost to the grammar. By making phonological rules and inventories sensitive to markedness considerations, one can explain why certain segments are more common than others, and why certain rules are more highly valued than other rules with the same number of symbols. Markedness theory contributes to learnability by giving further structure to the hypothesis space of possible grammars.<sup>21</sup>

---

Asymmetries in intrinsic content also arise in the relationship between phonological changes and their contexts. For example, rules assimilating one or more features of a segment to features in the environment are very common. An extension of markedness theory would also take account of the *naturalness* of rules.

I have suggested that various proposals to limit abstractness have not improved the learnability of grammar. On the contrary, it can be argued that abstractness of the right sort actually improves learnability, by working together with markedness and naturalness to limit the choices available to a learner.

Dresher (1981b) adduces an argument along these lines from Old English. The verb *eotan* 'to eat' can be shown to have the underlying stem vowel /e/. This vowel is raised to [i] in the present indicative second and third person singular forms *ites* and *iteθ*. Dresher proposes that the suffixes in these forms derive from underlying /+is/ and /+iθ/, respectively, and the forms in question are derived by means of two rules, given informally in (13) and (14). The derivations are shown in (15).

(13) *e*-Raising

Stressed *e* raises to *i* when an *i* follows in the next syllable.

(14) *i*-Lowering

Unstressed *i* lowers to [e] when it follows a light syllable (a syllable containing a short vowel and followed by a single consonant).

---

(15) Derivation of forms with stem vowel *i*

Underlying	/et+is/	/et+iθ/
<i>e</i> -Raising	itis	itiθ
<i>i</i> -Lowering	ites	iteθ
Phonetic	[ites]	[iteθ]

This analysis is supported by a web of other evidence (see Dresher 1985 for details). Assuming here the correctness of the analysis, the question we wish to answer is how a learner of Old English could arrive at it. In particular, what would lead a learner to suppose that the suffixes are /+is/ and /+iθ/ when they always appear as [+es] and [+iθ]?

Learners would know that unstressed *e* does not in general cause raising of a preceding *i*: there are many words like the present subjunctive *eten* 'we/you (pl.)/they eat'. It is possible that raising in the present singular is simply an exception, conditioned by morphology just as the English vowel change from *foot* to *feet* must be attributed to an irregular rule that operates in the plural. While such an analysis cannot be excluded, the assumption of *SPE* is that it is not preferred by the learning theory. In the formulation of Postal (1968), the *SPE* theory incorporates the Naturalness Condition.

(16) The Naturalness Condition (Postal 1968)

Phonological classifications are preferred to morphological or arbitrary classifications at all levels of the phonology.

The Naturalness Condition instructs the learner to seek a phonological solution before falling back on a morphological one.

---

In a learning theory which values simple and natural rules learners could proceed from the change to make hypotheses about the context. The raising of /e/ to [i] involves a change of one feature, [-high], to [+high]. The most favoured context for such a change is in the vicinity of another [+high] feature. In our case, there is no such segment visible on the surface; however, the vowel -e in the suffixes -es and -eθ is only one feature away from a high vowel, namely -i. Thus, even lacking other evidence, learners equipped with a learning theory of the kind sketched above may suspect at an early stage that the suffix vowel in -es and -eθ may be a disguised /i/. In this case, the suspicion will be supported by evidence from other parts of the phonology.

Notice that this solution would not be possible without abstractness. By allowing a certain distance between phonemic and phonetic forms, we can take advantage of rule ordering to formulate rules that are maximally simple and natural. In this way, abstractness contributes to learnability.

---

**References**

- Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago: University of Chicago Press.
- Avery, Peter, and Keren Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6, 179-200.
- Bloch, Bernard. 1941. Phonemic overlapping. *American Speech* 16: 278-284.  
Reprinted in Joos (1957), 93-96. Also in Makkai (1972), 66-70.
- Calabrese, Andrea. 1995. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26, 373-463.
- Chomsky, Noam. 1957a. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1957b. Review of R. Jakobson and M. Halle, *Fundamentals of language*, *International Journal of American Linguistics* 23: 234-241.
- Chomsky, Noam. 1959. Review of Skinner, *Verbal Behavior*, *Language* 35, 26-58.  
Reprinted in . In Jerry A. Fodor and Jerrold J. Katz (eds.), *The structure of language*, 547-578. Englewood Cliffs, NJ: Prentice-Hall.
- Chomsky, Noam. 1964. Current issues in linguistic theory. In Jerry A. Fodor and Jerrold J. Katz (eds.), *The structure of language*, 50-118. Englewood Cliffs, NJ: Prentice-Hall.
- Chomsky, Noam. 1975. *Reflections on language*. New York: Pantheon.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English [SPE]*. New York: Harper & Row.
- Chomsky, Noam, Morris Halle, and Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle et al. (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday*, 65-80. The Hague: Mouton.

- 
- Dresher, B. Elan. 1981a. Abstractness and explanation in phonology. In Norbert Hornstein & David Lightfoot (eds.) *Explanation in linguistics*, 76-115. London: Longman.
- Dresher, B. Elan. 1981b. On the learnability of abstract phonology. In C. L. Baker and John J. McCarthy, eds., *The logical problem of language acquisition*, 188-210. Cambridge, MA: MIT Press.
- Dresher, B. Elan. 1985. *Old English and the theory of phonology*. New York: Garland.
- Dresher, B. Elan. 1996. Learnability and phonological theory. In Jacques Durand and Bernard Laks (eds.), *Current trends in phonology: Models and methods*, Vol. 1, 245-266. Manchester: European Studies Research Institute, University of Salford Publications.
- Dresher, B. Elan, Glyne L. Piggott and Keren Rice. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13.1, iii-xvii.
- Goldsmith, John A. (ed.). 1995. *The handbook of phonological theory*. Oxford: Blackwell.
- Gussenhoven, Carlos and Haike Jacobs. 1998. *Understanding phonology*. London: Arnold.
- Gussmann, Edmund. 1980. *Studies in abstract phonology*. Cambridge, MA: MIT Press.
- Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton. Second printing, 1971.
- Halle, Morris. 2002. *From memory to speech and back: Papers on phonetics and phonology 1954-2002*. Berlin: Mouton de Gruyter.
- Harris, Zellig. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.

- 
- Hockett, Charles, F. 1951. Review of *Phonology as functional phonetics* by A. Martinet, *Language* 27: 333-342. Reprinted in Makkai (1972), 310-317.
- Hooper [Bybee], Joan. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
- Jakobson, Roman, and Morris Halle. 1956. *Fundamentals of language*. The Hague: Mouton.
- Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. 1952. *Preliminaries to speech analysis*. MIT Acoustics Laboratory, Technical Report, No. 13. Reissued by MIT Press, Cambridge, Mass., Eleventh Printing, 1976.
- Joos, Martin (ed.) (1957) *Readings in linguistics I*. Chicago: University of Chicago Press. Second edition New York: American Council of Learned Societies, 1958.
- Kaye, Jonathan D. 1974. Opacity and recoverability in phonology. *Canadian Journal of Linguistics* 19, 134–149.
- Kaye, Jonathan D., Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2, 305-28.
- Kean, Mary-Louise. 1980. The theory of markedness in generative grammar. Ph.D. dissertation, MIT (1975). Bloomington: Indiana University Linguistics Club.
- Kenstowicz,, Michael. 1994. *Phonology in generative grammar*. Oxford: Blackwell.
- Kenstowicz, Michael, and Charles Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.
- Kiparsky, Paul. 1968. How abstract is phonology? Bloomington: Indiana University Linguistics Club. Reprinted in Kiparsky (1982b), 119–164.

- 
- Kiparsky, Paul. 1973. Abstractness, opacity, and global rules. Bloomington: Indiana University Linguistics Club. Also in O. Fujimura (ed.), *Three dimensions of linguistic theory*. Tokyo: TEC. 57-86.
- Kiparsky, Paul. 1982a. From cyclic phonology to lexical phonology. In Harry van der Hulst and Norval Smith (eds.), *The structure of phonological representations (Part I)*, 131-176. Foris: Dordrecht.
- Kiparsky, Paul. 1982b. *Explanation in phonology*. Dordrecht: Foris.
- Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2. 85-138.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lasnik, Howard. this volume. Grammar, levels, and biology.
- Linell, Per. 1979. *Psychological reality in phonology: a theoretical study*. Cambridge: Cambridge University Press.
- Makkai, Valerie Becker (ed.). 1972. *Phonological theory: Evolution and current practice*. New York: Holt, Rinehart and Winston.
- Marlett, Stephen A. 1981. The abstract consonant in Seri. In *Proceedings of the Berkeley Linguistics Society* 7. Berkeley: University of California at Berkeley. 154-65.
- Marlett, Stephen A., and Joseph P. Stemberger. 1983. Empty consonants in Seri. *Linguistic Inquiry* 14, 617-39.
- Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper & Row.
- Rice, Keren. 2002. Vowel place contrasts. In Mengistu Amberber and Peter Collins (eds.), *Language universals and variation*. Westport, Connecticut; London: Praeger, 239-270.

---

Roca, Iggy, and Wyn Johnson. 1999. *A course in phonology*. Oxford: Blackwell.

Sapir, Edward. 1933. La réalité psychologique des phonèmes, *Journal de Psychologie Normale et Pathologique* 30, 247–265. Reprinted as The psychological reality of phonemes in *Selected writings of Edward Sapir in language, culture, and personality*, ed. by David Mandelbaum, 46–60. Berkeley and Los Angeles: University of California Press (1949). Also in Makkai (1972), 22–31.

Skinner, B. F. 1938. *Behavior of Organisms*. Appleton-Century-Crofts, New York.

Skinner, B. F. 1957. *Verbal Behavior*. Appleton-Century-Crofts, New York.

Swadesh, Morris, and Charles F. Voegelin. 1939. A problem in phonological alternation. *Language* 15: 1-10. Reprinted in Joos (1957), 88-92.

Tranel, Bernard. 1981. *Concreteness in generative phonology: Evidence from French*. Berkeley, CA: University of California Press.

Twaddell, F. 1935. On defining the phoneme. *Language Monograph*, no. 16. Reprinted in Joos (1957), 55-80.

---

**Notes**

\* I would like to thank Morris Halle for discussing some of these issues with me, and Jim McGilvray for his detailed comments on an earlier version of this chapter. All errors of fact or interpretation are mine. I am grateful for the support of grants 410-99-1309 and 410-2003-0913 from the Social Sciences and Humanities Research Council of Canada.

<sup>1</sup> For a detailed historical account of context of twentieth century phonological theory see Anderson (1985).

<sup>2</sup> A rule of the form  $A \rightarrow B / C \_\_\_\_\_\_ D$  is to be read, "A becomes B in the context C  $\_\_\_\_\_\_$  D." Parentheses indicate optional material. Thus, rule (1) indicates that a vowel becomes [+long] before a glide followed by a voiced consonant or directly before a voiced consonant.

<sup>3</sup> More precisely, Flapping may not precede the other two rules. In this case, the correct results would obtain if all three rules applied simultaneously to the underlying form. There are many cases, however, in which simultaneous application does not succeed.

<sup>4</sup> I abstract away from details and refinements such as the cycle.

<sup>5</sup> After a class of sounds called sibilants the plural has a third pronunciation, [əz], as in *busses*, *churches*, *bushes*, and *lounges*.

<sup>6</sup> The relatively small number of features is a legacy of the work of Jakobson and Halle in the 1950s, who sought to arrive at the smallest possible number of features required to distinguish contrastive sounds. For a more recent survey of the phonetic parameters that can be used distinctively in the languages of the world see Ladefoged and Maddieson (1996). Though the range of variation seen

---

across the spectrum of the world's languages may be greater than was known in the 1950s, there still appear to be significant restrictions on how many of the possible contrasts may actually be employed in a single language. See Rice (2002) for a discussion of vowel systems.

<sup>7</sup> In one of the earliest works in generative grammar, Chomsky, Halle and Lukoff (1956) show that an elegant analysis of English stress can be achieved by rules that are sensitive to aspects of the syntax, in sharp contrast to previous treatments in the American structuralist tradition. This work constitutes an early argument against the notion that phonological analyses must be constructed without reference to other aspects of grammar.

<sup>8</sup> Hockett (1951) argues that the generalization concerning the distribution of voiceless and voiced consonants can be regained at the morphophonemic level, where word boundaries and other morphological relations come into play. Thus, in his system the underlying morphophonemic representation of (9b) would be // #patat#ak // . A morphophonemic rule of word-medial voicing would derive the (taxonomic) phonemic form /padatak/. The effect of this reshuffling of levels is to turn the taxonomic phonemic level into a surrogate phonetic level, with the morphophonemic level doing much of the work formerly (and subsequently) assigned to the phonemic level. Once a systematic phonetic level is added to the model of grammar, there is no need for another level (the taxonomic phoneme) to play a similar role.

<sup>9</sup> Of course, there could have been: we might have found, for example, that *sip* derives from /ssip/ and *zip* derives from /sip/, and that the phonemic difference is between /s/ and /ss/, not /s/ and /z/. The analysis of a phonetic

---

string is thus not possible in isolation, but must take into account other parts of the grammar.

<sup>10</sup> Another well-known example is Swadesh and Voegelin's analysis of the morphophonemics of Tübatulabal (Swadesh and Voegelin 1939). This work done under the influence of Edward Sapir (1884–1939), is a testament to a more liberal period, before the neo-Bloomfieldian framework had considerably narrowed the range of what was an acceptable analysis.

<sup>11</sup> This interpretation of the status of morphophonemics was evidently not accepted by Swadesh and Voegelin (1939): "If it has been possible...to reduce the apparent irregularity of Tübatulabal phonology to system, this very fact guarantees the truth of our theory."

The question of what sort of reality was to be attributed even to phonemic analysis did not receive a clear answer in American linguistic thought. Sapir argued unambiguously that phonological analyses were 'psychologically real' (Sapir 1933), but this 'mentalist' interpretation was rejected by most American linguists (cf. Twaddell 1935). Psychological realism was replaced by a recurring debate in American Structuralist linguistics as to whether an analysis should be thought of as 'God's truth' or as 'hocus-pocus' (see Joos 1957:80). Chomsky (1957b) found these discussions to be "quite empty and sterile."

<sup>12</sup> See the contribution by Lasnik in this volume for further discussion of this argument and the notion of levels in grammar. The notion has nevertheless persisted within generative phonology that there may be some fundamental difference between rules that deal only in contrastive feature values and rules that deal in redundant values. One expression of this is the theory of Lexical

---

Phonology and Morphology (Kiparsky 1982a, 1985), which divides the phonology into lexical phonology (roughly, the old morphophonemic rules) and the postlexical phonology (mainly allophonic rules). The dividing line between lexical and postlexical phonology thus occupies a place somewhat like that of the taxonomic phoneme, without, however, having to observe the old constraints on this level. A further difference is that proposals for a lexical - postlexical distinction in the phonological component are supported by empirical evidence, by showing, for example, that the components have different properties.

<sup>13</sup> Discovery procedures are not to be confused with the generative quest for *explanatory adequacy*. The latter aims to account for how first language learners arrive at the grammar of their language, given that the available evidence appears to greatly underdetermine the choice of grammar. Chomsky has proposed that learners are endowed with a rich theory of Universal Grammar that guides and limits their acquisition of grammar. A theory of Universal Grammar differs from discovery procedures in that the latter are primarily intended to guide the *linguist*. Generative grammar posits no procedures to guide the linguist in the construction of a theory of Universal Grammar, nor are any conditions placed on how a theory of Universal Grammar can be related to the facts of language. Of course, an important criterion in choosing between different linguistic theories is how well they fare in explaining how language can be acquired.

<sup>14</sup> This form of behaviorism is thus quite different from the classical empiricism of Hume and Locke. Though he believed in the priority of sense impressions

---

over innate ideas in the sense of Descartes, Hume saw his task as discovering the principles of the mind, including innate properties such as the imagination.

<sup>15</sup> Chomsky (1959) demonstrated this in his famous review of B. F. Skinner's *Verbal behavior* (Skinner 1957). Chomsky showed that concepts that had concrete meanings in Skinner's (1938) *Behavior of Organisms*, where they were applied to classical but limited problems involving the relationship between stimuli and responses in rats and pigeons, became either trivial or false when applied to human verbal behavior.

<sup>16</sup> Morris Halle continued, of course, to be at or near the centre of developments in phonological theory for many years. In this essay, however, I focus only on his early work with Chomsky. See Halle (2002) for a collection of representative papers from 1954 to 2002.

<sup>17</sup> For recent introductions to contemporary phonology, see Gussenhoven and Jacobs (1998) and Roca and Johnson (1999). For more detailed treatments of phonological theory up to the mid-1990s, see Kenstowicz (1994) and Goldsmith (1995).

<sup>18</sup> The debate on abstractness was initiated by Kiparsky (1968). For a sampling of the extensive literature on the abstractness controversy from various points of view, see Dresher (1981a), Gussmann (1980), and Kenstowicz and Kisseberth (1977, Chap.1) on one side, and Hooper (1976), Linell (1979), and Tranel (1981) on the other.

<sup>19</sup> In assessing learnability, it is relevant to ask what the learner already knows about the grammar. Kaye (1974) comments that derivations that are opaque in Kiparsky's sense may nevertheless aid a listener (and a learner) in recovering

---

underlying forms, and hence the lexical identity of a morpheme. For example, the short raised diphthong in [rʌjɾər] (*writer*) in the dialects under discussion makes the Raising rule opaque, but signals to a listener who already knows the rule that the flap derives from a /t/. In a dialect where the rules apply in a different order so that *writer* sounds the same as *rider*, the rules are less opaque, but the identity of the lexical items is more difficult to recover.

<sup>20</sup> For further discussion of this case see Dresher (1996).

<sup>21</sup> The *SPE* markedness theory was developed further by Kean (1980), but was otherwise not much pursued in the years immediately after *SPE*. Some version of markedness is found in most current approaches to phonology, albeit in different forms. Calabrese (1995) presents a version that is much in the spirit of Kean and *SPE*. The theory of Government Phonology (Kaye, Lowenstamm and Vergnaud 1985) builds markedness into phonological representations. The same is true of Modified Contrastive Specification (Avery and Rice 1989, Dresher, Piggott and Rice 1994).