

# 18 Developments Leading Toward Generative Phonology

B. Elan Dresher & Daniel Currie Hall

## 18.1 Introduction

This chapter discusses key figures in the 1950s and 1960s whose work led to the theory of generative phonology. In section 18.2, we show how generative phonology developed on one side from the collaboration of Roman Jakobson and Morris Halle, continuing the Prague School approach by elaborating and modifying earlier ideas concerning distinctive features and relating them to mathematical models from the then-new field of information theory. Another source was the formalization of American structuralist phonology by Zellig Harris and the critique of that theory by Noam Chomsky. These sources merged in the collaboration of Chomsky and Halle and their critiques of prevailing notions of the phoneme and the strict separation of levels (section 18.3). We show how the emerging synthesis both built on and diverged from earlier ideas, and discuss some of the controversies of the years leading up to the appearance of *The sound pattern of English*.

## 18.2 Sources of generative phonology

### 18.2.1 Prague School sources of generative phonology

Many of the themes of Prague School phonology were carried over into the generative framework, thanks in no small part to the influence of Roman Jakobson. Having left Prague at the beginning of the Second World War, Jakobson settled in the United States in 1941 (passing through Copenhagen and Stockholm along the way; see Battistella, this volume). Jakobson's work in the United States, notably his collaboration with his student Morris Halle (e.g., Jakobson, Fant, & Halle 1952; Cherry,

Halle, & Jakobson 1953; Jakobson & Halle 1956), contributed much to the formative influence of the Prague School on generative phonology.

### 18.2.1.1 Distinctive features

One area in which the Praguian influence is particularly salient is that of distinctive features (as discussed in some detail by Fischer-Jørgensen 1975, among others). Though the ‘Projet de terminologie phonologique standardisée’ (Cercle linguistique de Prague 1931) defined the phoneme as a ‘phonological unit that cannot be decomposed into smaller and simpler phonological units’,<sup>1</sup> Vachek ([1936] 1976) pointed out that phonological oppositions could involve units smaller than the phoneme. For example, in the English minimal pair *bad* : *pad*, the contrast between /b/ and /p/ can be reduced to the contrast between voicing and voicelessness (or, as Vachek [1936] 1976: 17 expresses it, between sonority and 0). Some phonemes might still be irreducible: for instance, Vachek ([1936] 1976: 17) suggests that the phoneme /p/ consists of just the archiphoneme *p*, while /b/ is *p* plus sonority, and /bʲ/ is *p* plus sonority and ‘palatal character’. Thus, although Vachek separated out properties such as voicing and palatalization as distinctive units smaller than the phoneme, he did not propose that phonemes should be fully decomposable into features, as has been assumed in much subsequent work.

Trubetzkoy’s (1939) *Grundzüge der Phonologie* (‘Principles of phonology,’ translated into English as Trubetzkoy 1969) discussed oppositions between pairs of segments in terms that have influenced many subsequent authors’ conception of phonological features.<sup>2</sup> Trubetzkoy (1939: 67) divides these oppositions into three types: privative (*privativ*), gradual (*graduell*), and equipollent (*äquipollent*). In a privative opposition, one member has a property that the other lacks, such as voicing, nasalization, or rounding; more generally, the opposition is between marked (*merkmaltragend*) and unmarked (*merkmallos*), as in Vachek’s description of the opposition between /b/ and /p/ as an opposition between sonority (voicing) and 0 (its absence). Members of a gradual opposition

---

<sup>1</sup> ‘*unité phonologique non susceptible d’être dissociée en unités phonologiques plus petites et plus simples*’

<sup>2</sup> See Battistella (this volume) for a wider discussion of the *Grundzüge* in its Prague School context.

exhibit the same phonetic property to different degrees; Trubetzkoy's examples here are of pairs of vowels that differ in height. Equipollent oppositions, for Trubetzkoy, are essentially anything else; his examples include the pair /p/–/t/, which differ along the single articulatory dimension of place, but also /f/–/k/, which differ in both place and manner. Members of an equipollent opposition are equally marked, rather than involving a contrast between the presence or absence of a mark, or between differing degrees of some continuous property.

For Trubetzkoy, the categorization of an opposition must be consistent with the phonetic facts, but is ultimately to be determined by its place and function in the phonological system. Viewed phonetically and in isolation, any opposition is both equipollent and gradual (Trubetzkoy 1939: 67). For example, the opposition between /d/ and /t/ in many languages is phonetically equipollent, in that each member involves a marked articulation: voicing in the case of /d/ and tension in the case of /t/; it is also phonetically gradual, in that it is generally possible to produce a range of different degrees of voicing and of tension. But the opposition may be phonologically privative if one member patterns as unmarked in the phonemic system: if /t/ is unmarked, then the opposition is between voicing and its absence, with no phonological role for tension; if /d/ is unmarked, then the opposition is between tension and its absence, with no role for voicing. Similarly, the characterization of /u/–/o/ as a gradual opposition in German is due to the existence of other segments at other degrees along the same height continuum (/ɔ/ and /ɑ/; Trubetzkoy 1939: 68). If there were no low vowels in the system, there would be no phonological justification for treating the contrast between /u/ and /o/ as gradual rather than as either privative (close versus non-close, or non-open versus open) or equipollent.

In Trubetzkoy's system, any opposition can be phonologically equipollent; it is always possible for two contrasting phones to be equally marked. A phonologically gradual opposition, on the other hand, must also be phonetically gradual (not a difficult criterion to meet, given that Trubetzkoy claims that all oppositions are phonetically both equipollent and gradual when considered outside the context of the phonemic system), and a phonologically privative opposition must be characterizable as phonetically privative (that is, there must be some phonetic property that one

member has and the other lacks, even if the converse is also true). Trubetzkoy (1939: 69) claims that some oppositions can only ever be equipollent, giving /k/-/l/ as an example, though one might argue that in the sparse phonemic inventory of Hawaiian, where /k/ and /l/ are the only two oral consonants with lingual constriction, the opposition between them could be treated as a privative contrast between non-lateral and lateral. In any case, Trubetzkoy (1939: 69) schematizes the possible correspondences between ‘potential or logical’ (*potentiell oder logisch*, here meaning, in effect, phonetic) types of oppositions and their ‘actual’ (*faktisch*) or phonological character with the diagram reproduced in Figure 18.1.

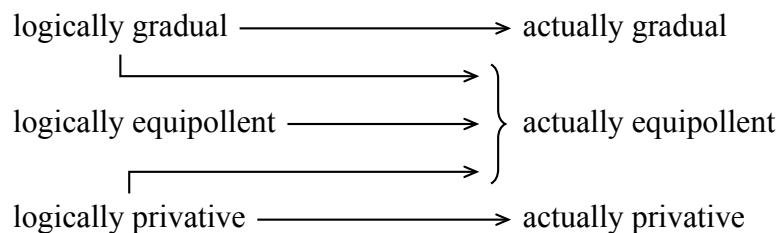


Figure 18.1: Correspondences between ‘logical’ (phonetic) and ‘actual’ (phonological) oppositions (Trubetzkoy 1939: 69)

The idea that any kind of phonetic opposition can be phonologically equipollent leads naturally to the possibility of representing all phonological oppositions in terms of binary distinctive features, in which positive and negative values have equal status. As Halle (1976: 86–87) notes, it was Jakobson ([1939] 1962) who took this step, both proclaiming ‘the primacy of the feature over the phoneme’ and proposing that features could all be of a single formal type, even when they represent different kinds of phonetic contrasts. Jakobson ([1939] 1962: 273) writes that ‘all vowel systems in general obey the principle of *dichotomy* and can be reduced to a restricted number of phonological features forming binary oppositions.’<sup>3</sup> Although Jakobson preserves something like Trubetzkoy’s notion of gradual and privative oppositions, neither his choice of examples nor his subsequent treatment of these oppositions shows this to be a clear-cut distinction, or a particularly

---

<sup>3</sup> ‘*Tout système vocalique en général obéit au principe de la dichotomie et se laisse réduire à un nombre restreint de qualités phonologiques formant des oppositions binaires.*’

useful one. For Jakobson ([1939] 1962: 273), a ‘contrary’ opposition (*opposition des termes contraires*) involves two terms that differ in degree along some dimension such as grave–acute,<sup>4</sup> and a ‘contradictory’ opposition (*opposition des termes contradictoires*) involves the presence or absence of an element. But his example of a contradictory opposition, namely that between short vowels (as lacking length) and long vowels (as having it), could just as easily be interpreted as a gradual or contrary one, particularly in languages such as Estonian, which has a three-way length contrast at least on the surface (Lehiste 1965).

Jakobson’s features also represented a move toward unifying the representations of consonants and vowels, and toward explaining the phonological affinities between the articulatorily distant labial and dorsal consonants. Jakobson ([1939] 1962: 274–5) proposes a system for representing consonantal place of articulation in which the labial and dorsal consonants are grave, in opposition to the acute dentals and palatals, thus identifying this contrast in the consonants with the vocalic opposition between grave /u/ and acute /i/. The grave–acute opposition cross-classifies with the anterior–posterior opposition to distinguish four major places of articulation as shown in Figure 18.2.<sup>5</sup>

<i>grave</i>	<i>acute</i>		<i>grave</i>
labials	dentals	palatals	velars
/p, b, .../	/t, d, .../	/c, ʃ, .../	/k, g, .../
<i>anterior</i>		<i>posterior</i>	

Figure 18.2: Consonant place features proposed by Jakobson ([1939] 1962)

This approach to features was further developed in Jakobson, Fant, & Halle’s (1952) influential

---

<sup>4</sup> Jakobson, Fant, & Halle (1952: 30) write that in a grave phoneme the second formant is closer to the first formant, whereas in an acute phoneme the second formant is closer to the third formant; in terms of production, gravity ‘is generated by a larger and less comparted mouth cavity, while acuteness originates in a smaller and more divided cavity’. See Hyman (1975: Chapter 2) for an accessible discussion of these and other features.

<sup>5</sup> Jakobson’s ([1939] 1962) opposition between anterior (*antérieur*) and posterior (*postérieur*) place of articulation is replaced in Jakobson, Fant, & Halle (1952: 27) by the opposition between diffuse and compact (as in (1g)), in keeping with their general practice of using acoustic names for features.

*Preliminaries to speech analysis*, which proposed a universal set of twelve binary oppositions, listed in (1).

(1) Features used in *Preliminaries* (Jakobson, Fant, & Halle 1952)

- |                                    |                        |
|------------------------------------|------------------------|
| a. vocalic vs. non-vocalic         | g. compact vs. diffuse |
| b. consonantal vs. non-consonantal | h. grave vs. acute     |
| c. interrupted vs. continuant      | i. flat vs. plain      |
| d. checked vs. unchecked           | j. sharp vs. plain     |
| e. strident vs. mellow             | k. tense vs. lax       |
| f. voiced vs. unvoiced             | l. nasal vs. oral      |

Each opposition is presented as equipollent, in the sense that each value has its own name, although some of the names are formed using privative prefixes (*non-* or *un-*); in the tables they provide in an appendix, the first member in each opposition is designated by + and the second by –. The term *plain* is opposed to both *flat* and *sharp*, suggesting that (1i) and (1j) could be conflated into a single ternary opposition of flat (F2 lowered by lip-rounding or pharyngealization) vs. plain (unaltered F2) vs. sharp (F2 raised by palatalization). Although Jakobson, Fant, & Halle do not make this conflation, they do contemplate another way of combining two binary features into one ternary one in their appendix. Jakobson, Fant, & Halle (1952: 44) write that if the value  $\pm$  is allowed, then the features interrupted/continuant and strident/mellow can be combined in the phonology of English into a contrast between the class ‘optimal constrictive’ (which is both strident and continuant) and its opposite ‘optimal stop’ (which is both mellow and interrupted). As shown in Table 18.1, the  $\pm$  value of this feature identifies obstruents that are either strident but not continuant or continuant but not strident, thus identifying a questionably natural class comprising affricates and dental fricatives.

Jakobson, Fant, & Halle’s tables of feature values further depart from pure binarity by leaving some specifications blank (e.g., nasal/oral on English vowels, whose value for this feature is determined contextually), and by parenthesizing others to indicate that they are redundant (e.g., [–nasal])

	<i>strident</i>	<i>mellow</i>
<i>continuant</i>	/f v s z ʃ ʒ/ [+optimal constrictive]	/θ ð/ [±optimal constrictive]
<i>interrupted</i>	/tʃ dʒ/ [±optimal constrictive]	/p b t d k g/ [−optimal constrictive]

Table 18.1: Combining two binary features into one ternary one (Jakobson, Fant, & Halle 1952: 44)

on obstruents). These compressions (i.e., reduction of redundancy) of feature values are intended primarily to make it easier ‘to determine the amount of significant information the phonemes actually carry in linguistic communication’ (Jakobson, Fant, & Halle 1952: 44), rather than necessarily to capture any phonologically significant generalizations; this interest in measuring information comes up again in Cherry, Halle, & Jakobson (1953), discussed below in §18.2.3.

The features of Jakobson, Fant, & Halle (1952) and Cherry, Halle, & Jakobson (1953) formed the basis, with some revisions, for the feature system used in Halle’s (1959) *The sound pattern of Russian (SPR)*. Chomsky & Halle (1968) made more extensive revisions in *The sound pattern of English (SPE)*, although the differences between *SPR* and *SPE* are somewhat artificially amplified by the fact that *SPE* uses articulatory names for the features rather than acoustic ones. As Halle (1983: 94) points out, the distinctive features of Jakobson, Fant, & Halle (1952) are formal phonological abstractions that have both acoustic and articulatory ‘correlates’ or ‘implementations’; their acoustic names, like the articulatory names used in *SPE*, are mnemonic rather than definitional. Still, Chomsky & Halle’s decision to focus on articulation in the names and descriptions they give their features, though explicitly attributed only to a desire for brevity,<sup>6</sup> has had a lasting effect on how features are conceived of in generative phonology.

---

<sup>6</sup> Chomsky & Halle (1968: 299) write: ‘We shall speak of the acoustical and perceptual correlates of a feature only occasionally, not because we regard these aspects as either less interesting or less important, but rather because such discussions would make this section, which is itself a digression from the main theme of our book, much too long.’ See further Kenstowicz (this volume).

To a large extent, the questions about features that interested Vachek, Trubetzkoy, and Jakobson continue to engage generative phonologists. Efforts to unify consonantal and vocalic representations (and to account for the affinity between labials and velars) have continued in Dependency Phonology (Anderson & Ewen 1980, 1987), Radical cv Phonology (Hulst 1994, 2005, 2020), Unified Feature Theory (Clements & Hume 1995), Element Theory (Harris & Lindsey 1995), Rice's (2002) Coronal-Peripheral theory of vowel place, Lahiri & Reetz's (2002; 2010) Featurally Underspecified Lexicon (FUL) model, and Morén's (2006) Parallel Structures Model, among many others. Jakobson, Fant, & Halle's (1952) proposal that there is a small set of universal features has been repeatedly revisited (e.g., *SPR*; *SPE*; Lahiri & Reetz 2010; Duanmu 2016) and challenged (e.g., Pulleyblank 2003; Mielke 2008; Samuels 2009). And questions of markedness have occupied generative phonologists from chapter 9 of *SPE* onward; see Steriade (1995), Rice (2007), and Hume (2011) for discussion.

### **18.2.1.2 The scope and goals of the field**

The Prague School took a more inclusive view of the scope of phonological inquiry than many American structuralists, and this, too, was a source of inspiration for generative phonologists. Jakobson's (1941) ambitious *Kindersprache, Aphasie und allgemeine Lautgesetze* ('Child language, aphasia, and phonological universals,' translated as Jakobson 1968) drew connections between typological patterns of markedness and the order of acquisition and attrition of linguistic structures. This approach treats language as a cognitive phenomenon, and therefore looks for insights into how language works not only in the minds of normally fluent adult speakers, but also in minds in the process of learning or losing languages. It thus stands in stark contrast with the anti-mentalism of Bloomfield (1933) and Hockett (1955), but prefigures generative approaches to acquisition, aphasia, and neurolinguistics (e.g., Stemberger & Stoel-Gammon 1991; Béland & Favreau 1991; Fikkert & Levelt 2008; Cornell, Lahiri, & Eulitz 2011, among many others). Jakobson's interest in universals likewise contrasts with the American structuralists' emphasis on differences among languages (see Silverstein and Ladd, both this volume), but resonates with generative



conceptions of Universal Grammar.<sup>7</sup>

In generative phonology, the relevance of acquisition is an underlying assumption even in work that does not deal with acquisition data directly. Chomsky & Halle (1965: 100) define explanatory adequacy, for them the ultimate goal of theoretical linguistics, as follows:<sup>8</sup>

A linguistic theory meets the level of explanatory adequacy insofar as it succeeds in describing the internal structure of AM [the acquisition model] and thus shows how the descriptively adequate grammar arises from the primary linguistic data.

The acquisition model is a representation of Universal Grammar: whatever is present in the minds of all (more or less cognitively normal) human beings that makes it possible for them to acquire natural languages. While the immediate concern of a given linguist might be the formulation of a descriptively adequate grammar of a particular language (i.e., one that accurately models a fluent speaker's implicit knowledge), the choice between competing descriptively adequate grammars ultimately comes down to explanatory adequacy, and thus to acquisition. This interest in the ability of acquisition to explain linguistic universals is very much in sympathy with the spirit of Jakobson (1941), and, as Chomsky & Halle (1965: 102) point out, very different from the taxonomic approach of American structuralists such as Harris (1951), who aspired primarily to what Chomsky & Halle call observational adequacy—the accurate reporting of directly observable linguistic data.

In a similar vein, Halle (1976: 95) remarks that generative linguistics has embraced Jakobson's use of explanation as a criterion in argumentation, in contrast to American structuralists' reliance on discovery procedures as a standard for guaranteeing the validity of a given analysis:

---

<sup>7</sup> Another stream of research into language universals was initiated by Joseph H. Greenberg and Charles A. Ferguson, the directors of the Stanford Language Universals Project (Greenberg 1966; Greenberg, Ferguson, & Moravcsik 1978).

<sup>8</sup> For a skeptical contemporary reaction to this objective, see Householder (1965), who describes observational adequacy—the most basic level in Chomsky & Halle's set of three—as the only one whose definition is intelligible to him.

The abstract structure of these arguments is usually of the following form: given a particular theoretical proposal the configuration of facts under discussion is highly probable, whereas in the absence of the theoretical proposal that observed configuration of facts would have to be regarded as a mere accident. Since it is *prima facie* implausible that the configuration is accidental, its existence provides evidence in support of the theoretical proposal. This form of argumentation, though standard in most science, was long regarded as questionable by linguists, many of whom felt that theoretical proposals and constructs could be justified only if it were shown that they were discoverable in the data by following certain procedures laid down in advance. In using the type of circumstantial argument just outlined, Jakobson was thus going against a very powerful current in the field. He had the good fortune to see the current reverse direction and his own practice fully vindicated.

### **18.2.2 American sources of generative phonology**

As has often been recounted (Anderson 1985: 314), Chomsky was a student of Zellig Harris and first learned linguistics by reading the proofs of Harris's *Methods in structural linguistics*; Harris (1951: v) thanks him for giving 'much-needed assistance with the manuscript'. Close study of Harris's text gave Chomsky a close-up view of the strengths and weaknesses of post-Bloomfieldian American phonology. In the words of Newman's (1952: 404) review, the book was

the most important contribution to descriptive linguistics since the publication of Bloomfield's *Language* in 1933. Its major significance lies in the fact that it makes explicit the direction in which linguistics has been moving... Its main value to linguists, however, will consist in its coherently reasoned and clearly formulated sequence of operational procedures for analyzing linguistic data.

Though Chomsky and Halle ultimately took issue with most of the main tenets of contemporary American phonology, generative phonology was influenced by Harris's 'rigorous and consistent

logic' (Newman 1952: 404), and by the general emphasis in American phonology on making its procedures explicit. Chomsky (1964*a*: 75) provides the following assessment of what he calls 'taxonomic phonemic theory'; in this he includes to some extent the approaches of Jakobson and Trubetzkoy,<sup>9</sup> as well as other European phonologists, but his discussion is mainly concerned with the work of American phonologists such as Bloch, Harris, Hockett,<sup>10</sup> Joos, and Twaddell:

Taxonomic phonemic theory constitutes the first attempt to formulate a linguistic theory with sufficient clarity and care so that questions of theoretical adequacy can seriously be raised. The development of taxonomic phonemics has led to standards of explicitness and precision that had rarely been attained in previous linguistic description, and to many insights into sound structure.

As we will see in the following sections, Chomsky and Halle did raise questions about the theoretical adequacy of taxonomic phonemics, and found it to be inadequate in fundamental ways. However, generative phonology continued to aspire to the high standards of explicitness and precision of the older theory. First, however, we will look at how information theory influenced phonologists of differing theoretical orientations.

### **18.2.3 Information theory and phonological representations**

The emerging field of information theory (or communication theory), as represented in particular by Shannon & Weaver (1949),<sup>11</sup> had an affinity with, and an early influence on, the aims and methods of theoretical linguistics, including phonology. Shannon & Weaver's work offered a mathematical

---

<sup>9</sup> Vachek (1964), responding to a preprint version of Chomsky (1964*b*), objects to the apparent subsumption of the Prague School under the label of 'taxonomic' phonology, and identifies instances in which Prague School phonologists anticipated aspects of the generative approach.

<sup>10</sup> See Ladd (this volume) for Hockett's generous assessment of generative grammar in his 1964 presidential address to the Linguistic Society of America.

<sup>11</sup> The theory was proposed in Shannon (1948), but Shannon & Weaver (1949) is the publication that is cited by linguists such as Jakobson, Fant, & Halle (1952) and Hockett (1953); as Hockett (1953: 69) remarks, 'the reviewer

basis for describing the information-carrying capacity of a system of communication, and for quantifying notions such as contrast and redundancy. In addition to the prospect of formally rigorous description of properties of natural languages, information theory suggested ways for linguistics to make connections with new technological developments in telecommunications and computer science.

The fundamental ingredients of Shannon & Weaver's theory of communication were essentially compatible with the existing apparatus of phonology. They begin by considering the case of a discrete channel—that is, one that transmits sequences composed by combining 'a finite set of elementary symbols  $S_1 \dots S_n$ ' (Shannon & Weaver 1949: 7). Although their examples take these symbols to be orthographic units, as in telegraphy, the extension to phonemes is a natural one. In the simplest case, the symbols are equally probable, and the likelihood of any given symbol appearing as the next one in the sequence is independent of what has come before it. This, however, is unrealistic as a model of either orthographic or spoken natural language. First, some letters or phonemes are more frequent than others, and thus more likely to occur at any given point in a sequence. Second, the existence of gaps in the lexicon, both accidental and systematic, means that the likelihood of a given phoneme or letter will vary according to context. For instance, about 2.73% of the letters in English writing are tokens of U (disregarding case); however, if the previous character is known to be a Q, then the likelihood that a letter will be a U rises to about 99.1% (Norvig n.d.). Shannon & Weaver show that the information conveyed by a symbol in a given context can be measured in terms of the recipient's uncertainty as to which symbol will appear. For example, if the symbol is entirely predictable from the context, then it conveys no new information; if there are two different symbols that could appear with equal likelihood, then the one that is actually transmitted conveys one bit of information—a single binary choice. The degree of uncertainty, and thus the rate of transmission of information, is quantified as entropy. Although greater unpredictability conveys more information at a time, redundancy increases reliability: the ability to check transmissions

---

found Weaver's discursive treatment of great value in grasping Shannon's often highly technical presentation'. See Pierrehumbert (this volume) for discussion of Shannon's influence on probabilistic approaches to phonology.

against expectations makes it possible to correct errors arising from noise in the channel. Increasing redundancy can yield arbitrarily high degrees of reliability, though with concomitant reductions in the speed of transmission.

The information-theoretic notion of entropy resonated with ongoing debates about contrast and redundancy in phonology. Hockett wrote a long and appreciative review of Shannon & Weaver (1949), published in *Language* in 1953, in which he gives a fairly detailed summary of some of the main points of the book, and explores how they might be applied to problems in linguistics, such as modelling how listeners transduce a continuous acoustic signal into a sequence of discrete phones. His discussion of differences between structuralist phonemics and a purely information-theoretic view of contrast reflects contemporary questions about the nature of the phoneme and prefigures debates about opacity and abstractness in generative phonology.

Hockett (1953: 81 et seq.) invites the reader first to consider a binary code using two equiprobable symbols ‘1’ and ‘2’; the average entropy of this code is one bit per symbol. He observes that randomly replacing some instances of ‘1’ with a third symbol produces free variation, and has no effect on the rate of transmission of information. Hockett considers two possibilities for the third symbol: ‘1’, which is graphically similar to ‘1’, and ‘3’, which is not. Using visual similarity as a proxy for phonetic similarity, he says that a linguist would readily treat ‘1’ and ‘1’ as constituting a single phoneme, but not ‘1’ and ‘3’; despite their information-theoretic equivalence, ‘the linguist at this point would balk’ at grouping together two such dissimilar allophones. The same logic applies if the replacement of ‘1’ by ‘1’ or ‘3’ is not random, but instead systematically occurs every time the ‘1’ is immediately followed by a ‘2’, yielding complementary distribution rather than free variation.

Hockett then imagines a further transformation of the original signal. Suppose that after ‘1’ has been changed to ‘1’ wherever it immediately precedes ‘2’, ‘1’ is deleted wherever it immediately follows ‘1’. Hockett’s thought experiment can be presented in *SPE*-style ordered rules as follows:<sup>12</sup>

---

<sup>12</sup> The rightmost column of (2) shows the names Hockett gives to each of the resulting encodings; code (b), omitted in (2), is a variant in which 1 and 1 are in free variation.

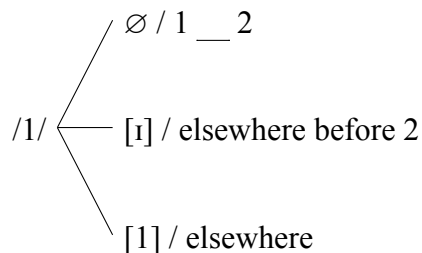
- (2) Original message 1121211121 Hockett's code (a)  
 1 → i / \_\_\_ 2      1i2i211i21 Hockett's code (c)  
 i → ∅ / 1 \_\_\_      12i21121 Hockett's code (d)

The original message can reliably be recovered from the output: every sequence '12' in code (d) can be expanded to the equivalent '1i2' in code (c) or '112' in code (a). From the perspective of information theory, this illustrates the trade-off between the number of distinct symbol-types in the code and the number of symbol-tokens needed to encode a given message. But 'from the linguistic point of view,' Hockett (1953: 83) writes, "'i'" and "'1'" are in contrast in (d), because there are such initial sequences as 12... versus i2... and such medial sequences as ...212... versus ...2i2....'<sup>13</sup> He continues (Hockett 1953: 84):

Some linguists would be tempted to 'phonemicize' code (d) by saying that where overtly one has the allophonic sequence '12' there is 'really' a variety—a zero alternant—of 'i' between them; to base a writing-system on this consideration would clearly be feasible in an unambiguous way, but within phonemics such a step is not valid.

In other words, the mapping from code (a) to code (d) could be represented by a distributional statement describing the 'allophones' in (d) of the 'phoneme' /i/ in (a) as in (3):

- (3) 'Allophones' of /i/ in code (d)



However, Hockett excludes analyses of this type in phonology, because the existence of a zero allophone is inconsistent with biuniqueness: not every instance of ∅ in code (d) is an allophone

---

<sup>13</sup> Parenthetical material in the original has been elided without marking so as to avoid adding editorial ellipses to the authorial ones already present in the quotation.

of /l/.<sup>14</sup> Likewise, Hockett would reject derivations like the one in (2) because they are inconsistent with the taxonomic conception of phonemes as strictly inferable from contrasting and non-contrasting sequences of allophones. Halle and Chomsky famously disagreed (Halle 1957, 1959, 1962; Chomsky 1964*a*; Chomsky & Halle 1965, 1968), as discussed below in sections 18.3.2 and 18.3.3.

Hockett (1953: 84) concludes this section of his review as follows:

The communications engineer is right in not understanding fully what linguists mean by phonemics, for we linguists have been fairly muddy in our phonemic thinking. The establishment of phonemic units can be rendered relatively non-arbitrary by accepting the criteria of phonetic similarity and of contrast versus no contrast, and by preferring that otherwise valid phonemicization which maximizes average entropy per symbol. But the selection of these criteria is itself arbitrary. A redefinition of the aims and procedures of phonemic analysis along the lines suggested above, and a clearer segregation of purely orthographic considerations, is a desideratum.

The preference for maximizing average entropy per symbol is an information-theoretic way of describing a preference for elegance and efficiency in phonological analyses—for capturing generalizations about patterns rather than simply reciting facts observable at the phonetic surface. Unchecked by other considerations, such a preference might, for example, lead one to treat English [h] and [ŋ] as allophones of a single phoneme, on the grounds that they are in complementary distribution. In Shannon & Weaver's terms, this is analogous to using a single symbol to transmit both [h] and [ŋ] because the recipient of the transmission would always be able to infer from the context which one was intended. But in phonology, phonetic dissimilarity militates against the conflation of [h] and [ŋ] (as pointed out by Chao 1934), and for Hockett the criterion of 'contrast

---

<sup>14</sup> See Roberts (1999) for further discussion of zeros in structuralist phonology, and in particular for their relation to juncture symbols, which are discussed in section 18.3.1 of this chapter.

versus no contrast,' evaluated strictly at the surface, likewise excludes some analyses that might be preferred if efficiency were the only consideration.

Cherry, Halle, & Jakobson (1953) voice a similar note of caution against the uncritical pursuit of information-theoretic efficiency in their discussion of the phonological inventory of Russian. They observe that each phoneme can be uniquely identified by the answers to a set of yes–no questions, with each question corresponding to one binary phonological feature, encoding one bit of information. In an inventory of 42 phonemes, the logical minimum number of questions/features/bits per phoneme would be  $\log_2 42$ , or 5.38. On the assumption that phonological features have phonetic content and group sounds into phonologically relevant classes, though, asymmetries in the inventory may necessitate a greater average number of specified features per phoneme. Cherry, Halle, & Jakobson (1953: 37) write:

In our analysis of language we are concerned, however, not only with questions of logic but also with matters of fact; hence the answers yes or no...are provided for us by considerations of the natural process of speaking.

An entirely abstract set of features could, in principle, function as a more efficient binary encoding of the inventory. But this option has not generally been seen as offering any useful phonological insights, and most work in generative phonology has explicitly or implicitly adopted some version of Postal's (1968) Naturalness Condition, positing that phonologically relevant classes of segments can and should be identified in phonetic terms (cf. Ladd 2014 for discussion and criticism and Kingston, this volume). Even in 'substance-free' approaches to phonology, features are based on something other than efficiency of encoding. Hale & Reiss (2000, 2003) and Hale, Kisoock, & Reiss (2007) posit that while phonological rules need not be phonetically natural in any way, the features they operate on have consistent phonetic definitions; for them, substance-freedom lies in the ability of rules to refer to arbitrary combinations of feature values. The radically substance-free phonology of Odden (2006) and Blaho (2008) takes the converse approach, claiming that features express classes of segments that pattern together phonologically, but may have arbitrary phonetic



interpretations, or in some cases no phonetic interpretation at all (see also Fudge 1967 and the emergent feature theory of Mielke 2008 and Samuels 2009).

Cherry, Halle, & Jakobson (1953), who use the features of Jakobson, Fant, & Halle (1952), propose that some redundancy can be removed from phonological representations by organizing the features into a binary decision tree, otherwise known in phonology as a contrastive hierarchy (Dresher 2009). If each feature represents a question about a phoneme ('Is it vocalic? Is it consonantal? Is it compact?' and so on), then the questions can be ordered so as to reduce the average number of questions needed to uniquely identify each phoneme. Although they do not present their ordering of features in tree form, their Table B (Cherry, Halle, & Jakobson 1953: 40) can be translated into the contrastive hierarchy in Figure 18.3.<sup>15</sup>

The twenty-first century has seen a renewed interest in entropy as a phonologically relevant metric (Hume 2004; Hume & Mailhot 2013; Hall 2009, 2013) and as a basis for acquiring constraint rankings in Optimality Theory (Goldwater & Johnson 2003; Riggle 2006; Hayes & Wilson 2008; Rasin & Katzir 2016). More broadly and pervasively, though, Shannon & Weaver's (1949) approach to communication helped to inspire the emphasis on mathematical formalization that can be found (with varying degrees of explicitness and rigour) throughout generative linguistics.

---

<sup>15</sup> The one feature whose place in the hierarchy is unclear is [ $\pm$ strident], whose positive value is assigned only to /ts/. Cherry, Halle, & Jakobson's (1953: 40) Table B shows /ts/ as having all the same specifications as /t/ except for this feature, which implies that [ $\pm$ strident] should take the lowest possible scope, below [ $\pm$ sharp] in the hierarchy. However, they show [ $-$ strident] specified not just on /t/ but also on /tʰ d dʰ/, which suggests that [ $\pm$ strident] takes scope over [ $\pm$ voice], but this ordering should leave [ $\pm$ voice] and [ $\pm$ sharp] unspecified on /ts/. Cherry, Halle, & Jakobson's (1953: 38) Table A indeed shows /ts/ with a zero for [ $\pm$ sharp], but still has it specified as [ $-$ voice]; this implies that [ $\pm$ strident] is below [ $\pm$ voice] and above [ $\pm$ sharp] in the hierarchy, as we have shown in our Figure 18.3, leaving [ $-$ strident] specified only on /t tʰ/.

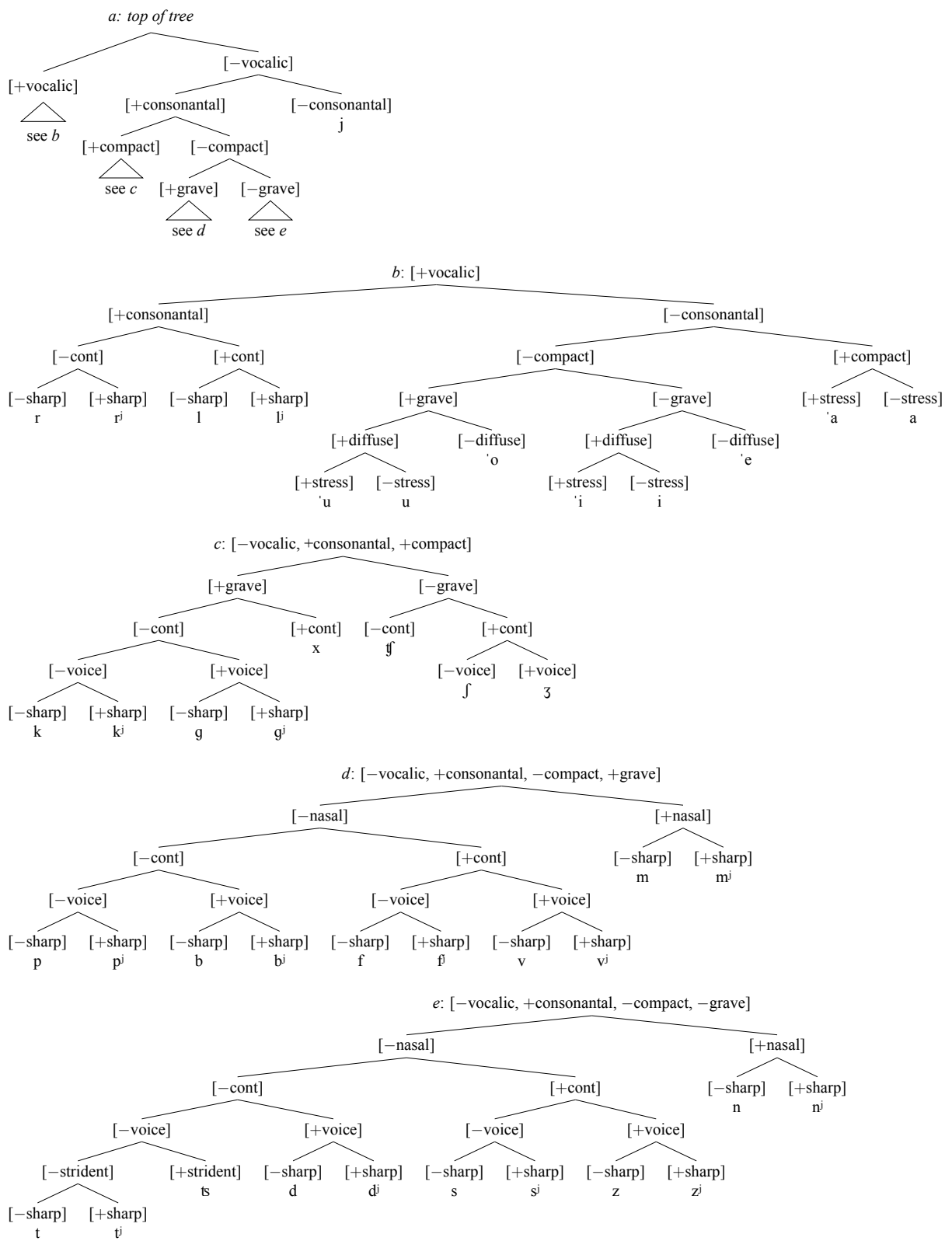


Figure 18.3: Cherry, Halle, & Jakobson's (1953) contrastive hierarchy for Russian

## 18.3 Chomsky and Halle's early revisions of phonological theory

### 18.3.1 Juncture and English stress

With the notable exception of Pike (1947), the dominant view among American structuralists was that a phonemic analysis could not admit of any information from higher (morphological or syntactic) levels of description. It was thought that such information is not available to a learner at the point where a phonemic analysis is being constructed out of lower-level phonetic data.<sup>16</sup> This has consequences for the analysis of what was referred to as 'juncture' (internal and external boundaries). As Aronoff (1980) remarks in his review of the issue, a phonemic analysis could not refer to the fact that *nightrate* is a compound consisting of two words, whereas *nitrate* is a single morphological word. Rather, the presence of an internal boundary in the former word, but not the latter one, must be inferable entirely from phonetic facts; in this case, that the unreleased [t'] in *nightrate* is also found word finally (for example, in *night*), whereas the slightly affricated [tʃ] in *nitrate* is not (see Ladd, this volume, for further discussion).

Chomsky, Halle, & Lukoff (1956) took direct aim at this tenet of American phonology by showing that a more adequate account of English stress can be obtained only by abandoning it, and recognizing not only that phonological junctures correspond to morphological and syntactic boundaries, but that they are hierarchically nested in a manner that reflects morphosyntactic structure.

Chomsky, Halle, & Lukoff (1956: 67) state a condition that a phonemic transcription must meet that contradicted the common understanding of the notion of the independence of levels:

---

<sup>16</sup> Michael Kenstowicz suggests to us that part of the motivation for this 'bottom-up' proceduralism may have derived from fieldwork practices that many American linguists were engaged in, where the first step was to transcribe the words and sentences of the informant without knowing much, if anything, about the morphology and syntax.

(4) Significance condition for junctures

Junctures should be distributed in a manner that is significant on higher levels. Specifically, junctures should appear only at morpheme boundaries, and different junctures should correspond, by and large, to different morphological and syntactical processes.

Chomsky, Halle, & Lukoff (1956) state that the status of this principle needs to be made clear ‘in view of the many recent discussions on the interdependence of levels’, without citing any in particular. Scheer (2011: 43) identifies Hockett (1942) as expressing ‘the base line of structuralist orthodoxy’ on this matter. Hockett (1942: 15) presents a clear statement of this position (5a); one of Hockett’s six requirements for a correct phonological system (Hockett 1942: 21) is given in (5b). Compare Moulton (1947), who expresses a similar view (5c); he holds this view despite the fact that ‘/+/ occurs almost exclusively at syntactic and morphological boundaries’.

(5) The ‘orthodox’ American structuralist position on juncture

- a. Hockett (1942: 15): Junctural phones are not matters of grammatical segmentation, though a junctural situation may define phonological segments which are of grammatical significance.
- b. Hockett (1942: 21): There must be no circularity; phonological analysis is assumed for grammatical analysis, and so must not assume any part of the latter. The line of demarcation between the two must be sharp.
- c. Moulton (1947: 225 n.14): I believe that the phonemes of a language should be analyzed without reference to syntax or morphology (as I have tried to do in this paper).

Chomsky, Halle, & Lukoff (1956) present positive and negative arguments for adopting their condition given in (4). Negatively, they state a number of times that ‘there is no circularity’ in the statement of this condition (Chomsky, Halle, & Lukoff 1956: 67), if one distinguishes between the linguist who prepares a transcription and the user of that transcription:

To read a phonemic representation properly one need know nothing of the morphological and syntactic structure of the language; one must only know the values of the

symbols in the phonemic transcription (including the junctures) and the rules governing their combinations. On the other hand, to evaluate a phonemic transcription, or to prepare one, the linguist must know the morphology and syntax, as well as the phonemics of the language.

Positively, Chomsky, Halle, & Lukoff (1956) argue that without imposing condition (4) as ‘a requirement of significance for the transcription T’, one can misuse junctures in ways that result in absurd analyses. Thus, a common structuralist argument in favour of positing junctures is that they enable a reduction in the number of phonemes that would be required without them. For example, without juncture, the unreleased [t̚] in *nightrate* would be in contrast with the affricated [tʃ] in *nirate*, and two phonemes, /t̚/ and /tʃ/, would have to be posited for English. By adding a phoneme of juncture, say /+/, we can capture the generalization that both of these phones are allophones of a single phoneme /t/: [t̚] occurs before /+/ in *night+rate* as well as word finally in *night+*, and [tʃ] occurs directly before /r/ when no juncture intervenes.

While this argument makes sense, Chomsky, Halle, & Lukoff (1956) point out that it can be abused if junctures are not required to have significance at higher levels of the grammar. For example, they observe that the number of nasal consonant phonemes in English can be reduced to one, /N/, by introducing a juncture /-/ and marking [n] as /N/, [m] as /-N/, and [ŋ] as /N-/; thus, the triad *ran*, *ram*, *rang* would be phonemicized as /ræN/, /ræ-N/, and /ræN-/, respectively.

One might try to rule out the above analysis, which no phonologist would accept, by requiring that if two phones contrast in medial position, like English [t̚] and [tʃ], then one of them can be represented as the phoneme /Z/ (here, /t/), and the other as /Z#/, if and only if the phone represented as /Z#/ (here, [t̚]) occurs finally and the other phone ([tʃ]) does not occur finally. This criterion rules out a juncture analysis of the English contrast between *m* and *n*, which contrast in every position: compare *might~night*, *spammer~spanner*, *rum~run*. But Chomsky, Halle, & Lukoff (1956) show that it does not prevent us from trading in a voicing contrast in a language like German for a junctural contrast: since voiceless and voiced obstruents both occur medially but only voiceless obstruents occur finally, we can write [b] as /p/ and [p] as /p-/ everywhere, with unacceptable results. They

conclude that no criterion based purely on phonetic distribution is likely to be able to rule out analyses that employ junctures in arbitrary ways.

The new approach to juncture is essential to Chomsky, Halle, & Lukoff's (1956) main purpose, which is to show that an adequate account of English stress can be obtained using only an opposition between stressed and unstressed, rather than four stress phonemes, as had been proposed by Trager & Smith (1951). Trager & Smith (1951) posit three stress phonemes in items of one, two, three, or more syllables: primary (´), tertiary (˘), and weak (˘).<sup>17</sup> A fourth stress phoneme, secondary stress (ˆ), occurs only in the presence of juncture (that is, in compounds and phrases). Examples of words with these stress phonemes according to Trager & Smith (1951) are given in (6).<sup>18</sup>

(6) Words with the four stress phonemes (Trager & Smith 1951)

- a. One stressed syllable: *ánimāl, tĕrrífíc, éffīgý*
- b. Main stress on the right: *ànimátiōn, hètĕrōgénĕōus, òpĕrátiōn*
- c. Main stress on the left: *ánimàte* (verb), *rĕfūgĕe, ópĕràte, sýntàx*
- d. Secondary stress: *bláckbôard, élĕvátōr òpĕrátōr, òld máid* ('former servant')

Chomsky, Halle, & Lukoff (1956) propose that the various levels of stress are derived by rules from a binary distinction between accented and unaccented syllables. The four different stress phonemes posited by Trager & Smith (1951) are thus allophones of one of these two phonemes; a tertiary stress, for example, may derive from an underlying accented or unaccented vowel, depending on the derivation. We will not attempt here to present their rules in detail, but will give a general idea of how they work by showing how they apply to the examples in (6) and to a more elaborate phrase that illustrates how the hierarchy of junctures influences the outcome. The various

---

<sup>17</sup> Trager & Smith (1951) observe that the last syllable in *animal* is a bit stronger than the middle one, though both are phonemically weak. The fact that weak stress has allophones suggests to them that it is a phoneme, and not simply the absence of a stress phoneme.

<sup>18</sup> The division of words into those with main stress on the right and those with main stress on the left is not made by Trager & Smith (1951), but it will be useful when discussing Chomsky, Halle, & Lukoff's (1956) reanalysis of the data.

levels of stress are designated with numbers; in simple cases the correspondences are as follows: primary or main stress (ˈ) is level 1, secondary stress (ˌ) is 2, tertiary stress (˙) is 3, and weak stress (˘) is 4 or higher.

The words in (6a) have a single accent on the stressed syllable. This syllable receives stress level 1. Unaccented syllables with unreduced vowels that follow the main stress receive level 3; other unstressed syllables (that precede the main stress or consist of the reduced vowel [ɪ]) receive level 4. Therefore, *animal* has stresses 1 4 3 and *terrific* receives 4 1 3. To account for Trager & Smith's (1951: 37) observation that the final syllable of *effigy* (1 4 4) has lower stress than the final syllable of *refugee* (1 4 3), Chomsky, Halle, & Lukoff (1956: 74) propose that the final [i] in the former derives from /iy/, in contrast to /i/ in *refugee*.<sup>19</sup>

The words in (6b) have two underlying accents, one on the initial syllable and the other on the syllable that surfaces with main stress. Both of these syllables are initially assigned level 1. Then the stress of the initial syllable is weakened by two to 3. This rule derives the tertiary stresses in the initial syllables of these words. In words with accented syllables between the first and last accents, the medial accents are weakened so they are one less than the stress of the preceding accent. Chomsky, Halle, & Lukoff (1956: 72) give the example *counter-revolutionary* (when represented with no internal junctures): whereas the initial syllable of *revolutionary* receives level 3, in *counter-revolutionary* it is demoted to 4 to be one less than the 3 stress of *coun-*.

Unlike Trager & Smith (1951), Chomsky, Halle, & Lukoff (1956) analyze the words in (6c) as having only one accent, on the initial syllable. This is because their rules require the rightmost accented syllable in a word to receive the main stress.<sup>20</sup> The tertiary stresses in these words are derived by the rules that assign stress levels to unaccented vowels that follow the main stress: 3 to unreduced vowels, and 4 to reduced vowels.

The words in (6d) contain junctures: internal juncture (–) corresponds to word formation, and

---

<sup>19</sup> Trager & Smith (1951: 36) specify '*refugee* (with primary on the first syllable)'.

<sup>20</sup> *SPE* would later attempt to predict the location of stresses in a word, and would add rules to account for words in which main stress is not the rightmost stress.

external juncture (=) corresponds to phrase formation. The basic rule is that the main stress in the domain to the left of an internal juncture causes the weakening of stresses to the right of the juncture; for external juncture, left and right are reversed in the statement of the rules. In examples with more than one juncture, the constituent structure influences the outcome of the derivation, which applies the stress rules first to the lowest-order constituents, and then successively to higher-order ones. Chomsky, Halle, & Lukoff (1956: 76) illustrate by giving the derivations of the phrases *excess-profits tax* (7) and *excessive profits-tax* (8).

(7) Derivation of *excess-profits tax*

	é	k	s	ε	s	= <sub>2</sub>	p	r	á	f	i	t	s	- <sub>1</sub>	t	á	k	s	<i>Rule applied</i>
1.	1					= <sub>2</sub>	1							- <sub>1</sub>	1				Assign 1 to accents
2.	2						1							- <sub>1</sub>	1				Weaken stress to left of =
3.	2						1										2		Weaken stress to right of –
4.	3						1										2		Weaken 2 in left domain of –
5.	3	4					1	5									2		Rules for unaccented vowels

(8) Derivation of *excessive profits-tax*

	ε	k	s	é	s	i	v	= <sub>1</sub>	p	r	á	f	i	t	s	- <sub>2</sub>	t	á	k	s	<i>Rule applied</i>
1.				1				= <sub>1</sub>	1							- <sub>2</sub>	1				Assign 1 to accents
2.				1				= <sub>1</sub>	1									2			Weaken stress to right of –
3.				2					1									2			Weaken stress to left of =
4.				2					1									3			Weaken 2 in right domain of =
5.	4	2	5						1	5								3			Rules for unaccented vowels

The derivations first assign 1 to accented syllables, treating each string containing no junctures in isolation. In the next step, the rules apply to the constituents separated by the most embedded juncture (with index number 2). In (7) this is the phrase *excess = profits*; because this has a = juncture, the stress on the left is weakened to 2. In (8), the more embedded phrase is *profits – tax*. This is a compound involving a – juncture, so the stress on the right is weakened to 2. The derivation continues in steps (3) and (4), now applying the rules that pertain to the juncture with



index number 1. After all the rules that apply to accented vowels have applied, stress levels are assigned by rule to the unaccented vowels.

Chomsky, Halle, & Lukoff (1956: 63) observe that they ‘have traded a transcription in which various levels of stress are marked, for one in which the constituent organization of the utterance on the phonological level is marked.’ They argue that this approach has a number of advantages: it leads to a simplification of the grammar, because the constituent structure imposed on phonological considerations correlates closely with the one required for English on other levels. Secondly, by reducing stress to a binary accented/non-accented opposition, it is possible to treat stress like other binary features, and to dispense with the whole category of suprasegmental phonemes.<sup>21</sup> Thirdly (Chomsky, Halle, & Lukoff 1956: 79), ‘the simplicity and the symmetry of the rules is further support for the essential correctness of the proposed treatment’. They conclude (Chomsky, Halle, & Lukoff 1956: 79–80):

Finally, when we interpret the phonological data in this way we are able to suggest a natural and simple explanation for the fact that native speakers can assign stress patterns to new utterances in a fairly consistent and uniform manner. A linguistic grammar must provide an explanation for this fact, just as it must provide an explanation for the fact that speakers form new sequences of words in an orderly and relatively consistent manner.

### **18.3.2 Halle’s critique of the taxonomic phoneme in *SPR***

Halle’s (1959) *The sound pattern of Russian* presents an analysis of the phonology of Russian in the same spirit as the theory of grammar, applied to syntax, of Chomsky (1957). This book takes further Chomsky, Halle, & Lukoff’s (1956) critique of American Structuralist practices, such as the separation of levels and the strict constraints imposed upon the linguist by theorists such as

---

<sup>21</sup> Readers familiar with later developments will be struck by the fact that one of the main arguments in favour of the metrical theory of stress (Lieberman & Prince 1977) is that stress does *not* behave like a binary feature; see Kisseberth (this volume).

Bloch (1941) and Hockett (1948). Halle (1959: 12) writes that a linguistic description, like other scientific theories, should be evaluated on its empirical adequacy, not on how the linguist arrived at it; those who propose a phonological transcription should not be subject to ‘a rigorous deductive procedure which would lead from the physical data’ to the proposed phonemes and boundaries. Though his practice departs from that of the American phonologists of the time, Halle (1959: 13) suggests that he follows methods characteristic of the work of Edward Sapir, to whose famous 1925 paper the title of *SPR* alludes. Like Sapir, Halle’s description does not make use of a phonemic transcription in addition to a morphophonemic one; rather, he posits a single significant level of lexical representation consisting of units he calls ‘morphonemes’, designated by curly braces {}.

The book is perhaps best known for its celebrated argument<sup>22</sup> against what Chomsky (1964a: 86) calls the ‘taxonomic phoneme’, a conception of the phoneme that was then influential in American linguistics.<sup>23</sup>

Halle (1959: 20–1) begins this argument by stating what he calls Condition (3) (shown in (9a)), ‘the weakest condition that can be imposed’ on the relation between the constructs of a phonological description (the segments and boundaries) and the observable data (actual speech events). He writes that while this condition is accepted by all, many contemporary linguists impose a stronger constraint, Condition (3a) (our (9b)).

(9) Two conditions on phonological descriptions (Halle 1959)

- a. Condition (3): A phonological description must provide a method for inferring (deriving) from every phonological representation the utterance symbolized, without recourse to information not contained in the phonological representation.

---

<sup>22</sup> See Anderson (2000) for extensive discussion of the reception of this argument. Anderson (2000: 16) writes that it is hard to say ‘why Halle’s argument should have been so earth-shaking’, given that the Russian facts were not novel and the problem he raised had been discussed before to no great effect. For our view see section 18.4.

<sup>23</sup> Thus, Halle’s argument is not against *any* conception of the phoneme, just the taxonomic one; Halle (1959)’s morphoneme (the ‘systematic phoneme’ of Chomsky 1964a) became the phoneme of generative phonology.

- b. Condition (3a): A phonological description must include instructions for inferring (deriving) the proper phonological representation of any speech event, without recourse to information not contained in the physical signal.

If one is to maintain Condition (3a), one must avoid a situation whereby one sound (a phone) corresponds to two different phonemes in the same context; an example is a sound [k] that in word-final position derives from the phoneme /k/, as well as from the phoneme /g/ by word-final devoicing. Such a transcription satisfies Condition (3), because a word-final phoneme /g/ will unambiguously become surface [k], just as phoneme /k/ in the same position becomes [k]. However, this transcription does not meet Condition (3a), because given an instance of word-final [k], one cannot know purely from phonetic information whether this [k] is an instance of /k/ or of /g/. One would need to have additional information from higher levels of grammar; for example, one would need to know if the final sound of the morpheme of which [k] is the last segment is realized as [k] or [g] when not in a devoicing context. To rule out this kind of phonemic overlapping (Bloch 1941), many American linguists adopted the requirement that the mapping between sound and phoneme be bi-unique (Hockett 1951).

In practical terms, this means that the phonemic representation of the Russian morpheme for ‘wet’ would have to be /mók/ in the phrase *mók li* ‘was (he) getting wet?’, but /móg/ in the phrase *móg bi* ‘were (he) getting wet’. A unique representation for the morpheme ‘wet’ would occur not at the phonemic level but at the morphophonemic level (/mók/), where morphological information is permitted to be available to the linguist. One might think that this is simply a definitional matter of apportioning the various representations to particular levels: the information that *mók* and *móg* are variants of the same morpheme (allomorphs) is encoded in the grammar at the morphophonological level by /mók/, and the information that the final consonant is voiced before a following obstruent is encoded at the phonemic level by /mók/ and /móg/.

Halle (1959: 22–23) argues, however, that there is a problem with this transcription. Russian has three ‘unpaired’ voiceless obstruents, /ts/, /tʃ/, and /x/, that have no contrastive voiced counterparts; nevertheless, just like /k/, which must be contrastively voiceless because it has a voiced

counterpart //g//, they are voiced before a voiced obstruent to *dʒ*, *dʒ*, and *ɣ*, respectively. Since the latter alternations create allophones from a single phoneme, rather than phonemes from a single morphophoneme, Halle (1959) argues that Condition (3a) and the requirement of bi-uniqueness that follows from it lead to a grammar that is more complicated than it needs to be, because the voicing rule must now appear in both the morphophonemic and phonemic components of the grammar, as shown in (10).

(10) Derivation with voicing applying in two different components

Morphophonemic level	//mók bi//	//mók li//	//ʒéʃ bi//	//ʒéʃ li//
Voicing rule	g			
Phonemic level	/móg bi/	/mók li/	/ʒéʃ bi/	/ʒéʃ li/
Voicing rule			dʒ	
Phonetics	[móg bi]	[mók li]	[ʒédʒ bi]	[ʒéʃ li]
Gloss	‘were (he) getting wet’	‘was (he) getting wet?’	‘were one to burn’	‘should one burn?’

If one were to do away with Condition (3a), one could posit a simpler grammar (11) with a single morphonemic level and a single application of the voicing rule. Thus, criteria of simplicity and generality favour the grammar in (11) over the one in (10), and show that there is no need for a ‘phonemic’ level of representation. Halle (1959: 24) sums up the argument: ‘The conclusion, therefore, imposes itself that Condition (3a) is an unwarranted complication which has no place in a scientific description of language.’

(11) Derivation with voicing applying once

Morphonemic level	{{mók bi}}	{{mók li}}	{{ʒéʃ bi}}	{{ʒéʃ li}}
Voicing rule	g		dʒ	
Phonetics	[móg bi]	[mók li]	[ʒédʒ bi]	[ʒéʃ li]

Halle (1959: 46) explicitly presents the features in a tree diagram, using a slightly different ordering of features than Cherry, Halle, & Jakobson (1953), but still with the primary aim of making the underlying representations of phonemes efficient. Like Cherry, Halle, & Jakobson (1953),

Halle (1959) orders  $[\pm\text{continuant}]$  over  $[\pm\text{voiced}]$ , with the result that the ‘unpaired’ obstruents are unspecified for  $[\pm\text{voiced}]$ . This ordering is illustrated for  $/tʃ/$  in Figure 18.4a; the other unpaired obstruents work the same way. As Hall & Dresher (2016) point out, though, Halle (1959) missed an opportunity to make the contrastive hierarchy do more phonological work: placing  $[\pm\text{voiced}]$  higher in the tree, as shown in Figure 18.4b, would cause it to be specified on the unpaired voiceless obstruents  $/ts\ tʃ\ x/$ , which would be consistent with their behaviour in voicing assimilation.<sup>24</sup> Instead, Halle simply has this feature filled in by a redundancy rule before voicing assimilation applies, making its underspecification in the tree irrelevant to its phonological patterning.

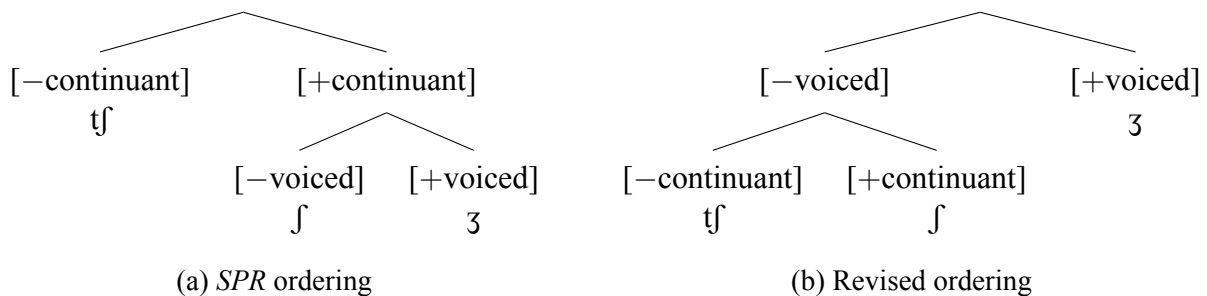


Figure 18.4: Two orderings of  $[\pm\text{continuant}]$  and  $[\pm\text{voiced}]$

Because Halle’s (1959) contrastive hierarchy had no real phonological consequences, it is not surprising that Chomsky & Halle (1968) turned away from contrastive hierarchies altogether, appealing instead to a theory of markedness (Chomsky & Halle 1968: ch. 9) as the basis for underspecification. The contrastive hierarchy went largely unemployed in generative phonology until being revived in Dresher, Piggott, & Rice (1994) and subsequent work; see Dresher (2009, 2015) for discussion and additional citations.

---

<sup>24</sup> Iosad (this volume) points out that essentially this analysis was proposed by Kasevich (1983). Notice that in the revised order  $/ʒ/$  is not contrastively specified for  $[\pm\text{continuant}]$ ; Hall & Dresher (2016) argue that this is a desirable consequence of the reanalysis.

### 18.3.3 Chomsky's critique of the taxonomic phoneme

Chomsky (1964*a*) incorporates and extends the critiques of Chomsky, Halle, & Lukoff (1956), Halle (1959), and other earlier publications to mount a comprehensive attack on the foundations and major tenets of phonological theory as then practiced. The brunt of his critique was aimed at the system of levels that came to prevail in post-Bloomfieldian phonology, and the distinction drawn between the morphophonemic and phonemic levels of representation. We have alluded to this organization above in (10); a more general diagram is given in Figure 18.5; see Dresher (2005) for further discussion.

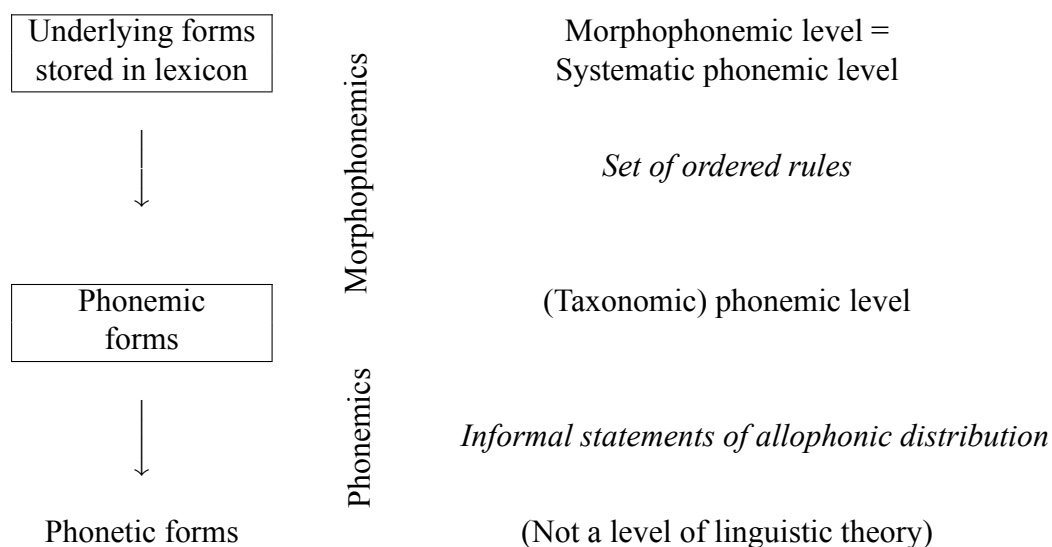


Figure 18.5: Levels in post-Bloomfieldian American structuralist phonology

Chomsky (1964*a*) argues that a key to the development of the phonemic level in Figure 18.5 is Bloomfield's rejection of a phonetic level as a systematic level of representation. According to Bloomfield (1933: 85), there are no principles that would govern how much detail should go into such a level, what Chomsky (1964*a*) calls the level of 'systematic phonetics'. Chomsky (1964*a*: 66–7) proposes that such principles are provided by a universal phonetic theory, to which the feature theory of Jakobson, Fant, & Halle (1952) is a contribution, building on earlier work by the classical British phoneticians and Saussure. Chomsky (1964*a*: 67) proposes that a linguistic theory must meet the condition of *phonetic specificity* (12). The 'fixed alphabet' of the universal phonetic

theory sets the limits as to what can be represented at the systematic phonetic level.

(12) Phonetic specifiability

A general linguistic theory must incorporate a universal phonetic theory, with a fixed alphabet.

Chomsky (1964a: 109) observes that the lack of a systematic phonetic level had the consequence of making the phonemic level the lowest systematic level of representation.<sup>25</sup> The morphophonemic level, what Chomsky (1964a) calls the ‘systematic phonemic’ level, was too remote from the phonetics to play this role; hence, over time a new level of representation was developed, what Chomsky (1964a) calls the ‘taxonomic phonemic’ level. While the properties of such a level remained ill-defined, and not all phonologists had the same conception of it, Chomsky (1964a: 78) identified a number of conditions, given in (13), that characterized this level of representation.

(13) Four conditions on the taxonomic phoneme (Chomsky 1964a: 78)

- a. Linearity
- b. Invariance
- c. Biuniqueness
- d. Local determinacy

Chomsky (1964a) argues that these conditions are untenable, no matter how they may be revised, because in many cases they have the effect of ruling out what simplicity and descriptive adequacy would identify as the best analysis. Since these conditions come in different versions and are difficult to make completely explicit, we will briefly illustrate them here with some examples adduced by Chomsky (1964a) that show why they fail. An example from English is supplied

---

<sup>25</sup> Chomsky (1964a: 77–8) writes that Bloomfield’s explicit rejection notwithstanding, in *practice* all varieties of taxonomic phonemics can be regarded as ‘resting squarely on assumptions concerning a universal phonetic theory’ that imply a systematic phonetic level (see Ladd 2014: 31–7, this volume). Ladd (2014) goes on to argue that despite the near-universal assumption that such a level exists, its characterization remains problematic.

by Malécot (1960), who observed that before an unvoiced stop, a sequence of a lax vowel followed by a nasal consonant is often realized as a nasalized vowel, as shown in (14).<sup>26</sup>

(14) Nasalization and nasal deletion in English

Phonemic level	/kænt/	/hænd/	/kæt/
Nasalization	kãnt	hãnd	—
Nasal Deletion	kæt	—	—
Phonetics	[kãnt]	[hãnd]	[kæt]
	<i>can't</i>	<i>hand</i>	<i>cat</i>

In (14), a rule of Nasalization applies to the vowels in *can't* and *hand*. In the latter case, this is a simple case of allophony, whereby the phoneme /æ/ has the allophone [ã] before a nasal consonant. In the case of *can't*, however, the /n/ deletes before the voiceless stop /t/, leaving the nasalized [ã] bereft of its conditioning context, and thus seemingly in contrast with [æ] in *cat*. This analysis of *can't* violates linearity, which requires that if a phoneme A (/æ/) precedes B (/n/) at the phonemic level, then the phonetic realization of A ([ã]) must precede the phonetic realization of B. Invariance, broadly stated, requires that every phoneme must be associated with a set of defining features that must appear in the phonetic representation whenever the phoneme occurs in the phonemic representation. This condition is arguably violated by the derivation of *can't* in (14), if we take the defining features of /n/ to be minimally [consonantal], [coronal], and [nasal].

We have alluded to the biuniqueness condition above: it requires that every sequence of phonemes corresponds to a unique sequence of phones (not counting free variation), and every sequence of phones corresponds to a unique sequence of phonemes. One might argue that the derivation of *can't* in (14) satisfies biuniqueness, in that the sequence [ãnt] uniquely corresponds to the phoneme sequence /ænt/. However, Chomsky (1964a) observes that many phonologists appear to interpret biuniqueness in conjunction with local determinacy, which requires that the unique phonemic sequence corresponding to the phonetic one must be determinable by purely phonetic considerations

---

<sup>26</sup> The phonetic transcription in (14) shows only what is relevant to this example, and thus omits aspiration, consonant release, and other such details.



involving neighbouring sounds. In this case, without being able to consider the morphology of *can't* so that one can find in it the same morpheme as occurs in *can*, one would have to take at face value the apparent contrast between *can't* and *cat*, and conclude that the phoneme /æ/ is in contrast with another phoneme /æ̃/, notwithstanding the straightforward analysis in (14).

All the conditions in (13) are aimed at regulating the relation between a phonemic representation and its phonetic realization.<sup>27</sup> Chomsky and Halle's core argument against taxonomic phonemics turns on the claim that this relation cannot be subject to such constraints, because the determination of the most descriptively adequate grammar depends on 'general systemic considerations' (Chomsky 1964a: 86), and not on local decisions arrived at in isolation from the grammar as a whole.

## 18.4 Conclusions

It is important to note that most of the examples cited by Chomsky (1964a) were first discussed by American structuralist phonologists, and their consequences for the taxonomic phoneme were also recognized. However, the fact that taxonomic phonemics could result in analyses that might appear to be uneconomical and even counter-intuitive was not considered to be a problem for the theory. On the contrary, the results were viewed as a sign of the maturity of the theory—as Ladd (this volume) puts it, 'a feature, not a bug'. And indeed earlier critiques, such as those of Pike (1947), did not gain much traction.

What was novel about Chomsky and Halle's critique was that they took issue with the scientific and psychological assumptions that underpinned taxonomic phonemics. With respect to science, they argued for applying standard criteria used to evaluate scientific theories, rather than focusing on the development of procedures for arriving at a correct analysis. With respect to psychology, Chomsky (1964a) added to his well known critique of Behaviorism (Chomsky 1959) by identifying

---

<sup>27</sup> This issue arose again in generative phonology as the controversy over 'abstractness'; see Kenstowicz & Kisseberth (this volume).

and attacking the cognitive assumptions that might underpin the conditions in (13) as well as the restriction against ‘mixing of levels’. These cognitive principles involved both perception and acquisition. Chomsky (1964a: 101) sums them up as follows:

[The conditions in (13)] pertain to the “perceptual model”; they assert that the phonemic correspondent to a given phonetic sequence must be determinable by operations involving only neighbouring sounds, *once the phonemic system is fixed*. But the condition of separation of levels is not a formal condition on a phonemic system and the rules that relate it to sound; it is a methodological condition on information relevant to determining the correct choice of a phonemic system. It thus pertains to an acquisition model...rather than to a perceptual model....

Chomsky goes on to observe that the conditions are all connected in that they exclude higher-level information from the determination of the taxonomic system. Chomsky and Halle’s central arguments against this approach were, first, that there is no empirical support for the assumptions about speech perception and phonological acquisition that underpin these assumptions; and second, that in the absence of such support, the modest procedural goals of phonological theory should be replaced by more ambitious scientific goals that ‘go far beyond the restricted framework of modern taxonomic linguistics and the narrowly-conceived empiricism from which it springs’ (Chomsky 1964a: 113).

## References

- Anderson, John M. and Colin J. Ewen (eds.). 1980. *Studies in dependency phonology*. Ludwigsburg: R.O.U. Strauch.
- Anderson, John M. and Colin J. Ewen. 1987. *Principles of dependency phonology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago, IL: University of Chicago Press.
- Anderson, Stephen R. 2000. Reflections on ‘On the phonetic rules of Russian’. *Folia Linguistica* 34(1—2): 11—27.
- Aronoff, Mark. 1980. The treatment of juncture in American linguistics. In Mark Aronoff and Mary-Louise Kean (eds.), *Juncture: A collection of original papers*, no. 7 in *Studia linguistica et philologica*, 29–36. Saratoga, CA: Anma Libri.
- Béland, Renée and Yves Favreau. 1991. On the special status of coronals in aphasia. In Carole Paradis and Jean-François Prunet (eds.), *The special status of coronals: Internal and external evidence*, no. 2 in *Phonetics and Phonology*, 201–221. San Diego: Academic Press.
- Blaho, Sylvia. 2008. *The syntax of phonology: A radically substance-free approach*. Ph.D. thesis, Universitetet i Tromsø.
- Bloch, Bernard. 1941. Phonemic overlapping. *American Speech* 16(4): 278–284.
- Bloomfield, Leonard. 1933. *Language*. London: Allen & Unwin.
- Cercle linguistique de Prague. 1931. Projet de terminologie phonologique standardisée. *Travaux du Cercle linguistique de Prague* 4: 309–323.
- Chao, Yuen-Ren. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4(4): 363–397.

- Cherry, E. Colin, Morris Halle, and Roman Jakobson. 1953. Toward the logical description of languages in their phonemic aspect. *Language* 29(1): 34–46.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1959. A review of B. F. Skinner's *Verbal behavior*. *Language* 35(1): 26–58.
- Chomsky, Noam. 1964a. *Current issues in linguistic theory*. The Hague: Mouton.
- Chomsky, Noam. 1964b. The logical basis of linguistic theory. In Horace G. Lunt (ed.), *Proceedings of the ninth International Congress of Linguistics*, 914–978. The Hague: Mouton.
- Chomsky, Noam and Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1(2): 97–138.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Chomsky, Noam, Morris Halle, and Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle, Horace G. Lunt, Hugh McLean, and Cornelis H. van Schooneveld (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday, 11 October 1956*, 65–80. The Hague: Mouton.
- Clements, G. N. and Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Oxford: Blackwell.
- Cornell, Sonia A., Aditi Lahiri, and Carsten Eulitz. 2011. What you encode is not necessarily what you store: Evidence for sparse feature representations from mismatch negativity. *Brain Research* 1394: 79–89.
- Dresher, B. Elan. 2005. Chomsky and Halle's revolution in phonology. In James McGilvray (ed.), *The Cambridge companion to Chomsky*, 102–122. Cambridge: Cambridge University Press.

- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- Dresher, B. Elan. 2015. The motivation for contrastive feature hierarchies in phonology. *Linguistic Variation* 15(1): 1–40.
- Dresher, B. Elan, Glyne L. Piggott, and Keren D. Rice. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13: iii–xvii.
- Duanmu, San. 2016. *A theory of phonological features*. Oxford: Oxford University Press.
- Fikkert, Paula and Clara Levelt. 2008. How does Place fall into place? The lexicon and emergent constraints in children’s developing phonological grammar. In Peter Avery, B. Elan Dresher, and Keren Rice (eds.), *Contrast in phonology: Theory, perception, acquisition*, 231–270. The Hague: Mouton de Gruyter.
- Fischer-Jørgensen, Eli. 1975. *Trends in phonological theory: A historical introduction*. Copenhagen: Akademisk Forlag.
- Fudge, Erik C. 1967. The nature of phonological primes. *Journal of Linguistics* 3(1): 1–36.
- Goldwater, Sharon and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson, and Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. Stockholm: Stockholm University Department of Linguistics.
- Greenberg, Joseph H. (ed.). 1966. *Universals of language*. Cambridge, MA: MIT Press.
- Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik (eds.). 1978. *Universals of human language*, vol. I–IV. Stanford, CA: Stanford University Press.
- Hale, Mark, Madelyn Kissonck, and Charles Reiss. 2007. Microvariation, variation, and the features of universal grammar. *Lingua* 117: 645–665.

- Hale, Mark and Charles Reiss. 2000. 'Substance abuse' and 'dysfunctionalism': Current trends in phonology. *Linguistic Inquiry* 31(1): 157–169.
- Hale, Mark and Charles Reiss. 2003. The subset principle in phonology: Why the *tabula* can't be *rasa*. *Journal of Linguistics* 39(2): 219–244.
- Hall, Daniel Currie and B. Elan Dresher. 2016. Trade-offs in the contrastive hierarchy: Voicing versus continuancy in Slavic. In Christopher Hammerly and Brandon Prickett (eds.), *NELS 46: Proceedings of the forty-sixth annual meeting of the North East Linguistic Society*, vol. 2, 43–56. Amherst, MA: GLSA.
- Hall, Kathleen Currie. 2009. *A probabilistic model of phonological relationships from contrast to allophony*. Ph.D. thesis, The Ohio State University.
- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2): 215–275.
- Halle, Morris. 1957. On the phonetic rules of Russian. Presented to the Linguistic Society of America, Chicago, 1957.
- Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton.
- Halle, Morris. 1962. Phonology in generative grammar. *Word* 18: 54–72.
- Halle, Morris. 1976. Roman Jakobson's contribution to the modern study of speech sounds. In Ladislav Matejka (ed.), *Sound, sign and meaning: Quinquagenary of the Prague Linguistic Circle*, no. 6 in Michigan Slavic Contributions, 79–100. Ann Arbor, MI: University of Michigan.
- Halle, Morris. 1983. On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory* 1(1): 91–105.

- Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In Jacques Durand and Francis Katamba (eds.), *Frontiers of phonology: Atoms, structures, and derivations*, 34–79. Harlow, Essex: Longman.
- Harris, Zellig S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.
- Hayes, Bruce and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39(3): 379–440.
- Hockett, Charles F. 1942. A system of descriptive phonology. *Language* 18(1): 3–21.
- Hockett, Charles F. 1948. A note on ‘structure’ [review of de Goeje by W. D. Preston]. *International Journal of American Linguistics* 14(4): 269–271.
- Hockett, Charles F. 1951. Review of *Phonology as functional phonetics: Three lectures delivered before the University of London in 1946* by André Martinet. *Language* 27(3): 333–342.
- Hockett, Charles F. 1953. Review of Shannon & Weaver (1949). *Language* 29(1): 69–93.
- Hockett, Charles F. 1955. *A manual of phonology*. Baltimore: Waverly Press.
- Householder, F. W. 1965. On some recent claims in phonological theory. *Journal of Linguistics* 1: 13–34.
- Hulst, Harry van der. 1994. An introduction to radical CV phonology. In Susanna Shore and Maria Vilkuna (eds.), *1994 yearbook of the Linguistic Association of Finland*. Helsinki: Suomen Kielitieteellinen Yhdistys.
- Hulst, Harry van der. 2005. The molecular structure of phonological segments. In Philip Carr, Jacques Durand, and Colin J. Ewen (eds.), *Headhood, elements, specification and contrastivity*, 193–234. Amsterdam: John Benjamins Publishing Company.
- Hulst, Harry van der. 2020. *Radical cv Phonology – a theory of segmental and syllabic structure*. Edinburgh: Edinburgh University Press.

- Hume, Elizabeth. 2004. Deconstructing markedness: A predictability-based approach. *BLS* 13: 182–198.
- Hume, Elizabeth. 2011. Markedness. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 1, chap. 4, 79–106. Oxford: Wiley Blackwell.
- Hume, Elizabeth and Frédéric Mailhot. 2013. The role of entropy and surprisal in phonologization and language change. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization*, 29–47. Oxford: Oxford University Press.
- Hyman, Larry M. 1975. *Phonology: Theory and analysis*. New York, NY: Holt, Rinehart & Winston.
- Jakobson, Roman. [1939] 1962. Observations sur le classement phonologique des consonnes. In *Selected writings, volume I: Phonological studies*, 272–279. The Hague: Mouton. Originally published in *Proceedings of the Third International Congress of Phonetic Sciences*.
- Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Språkvetenskapliga Sällskapet i Uppsala Förhandlingar. Uppsala: Uppsala Universitetet.
- Jakobson, Roman. 1968. *Child language aphasia and phonological universals*. The Hague: Mouton. Translated by Allan R. Keiler from Jakobson (1941).
- Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. 1952. *Preliminaries to speech analysis: The distinctive features and their correlates*. Tech. Rep. 13, MIT Acoustics Laboratory.
- Jakobson, Roman and Morris Halle. 1956. *Fundamentals of language*. No. 1 in *Janua Linguarum*. The Hague: Mouton.
- Kasevich, Vadim Borisovich. 1983. *Fonologicheskie problemy obshchego i vostochnogo yazykoznaniya*. Moscow: Nauka & Shkola ‘JaRK’.
- Ladd, D. Robert. 2014. *Simultaneous structure in phonology*. Oxford: Oxford University Press.



- Lahiri, Aditi and Henning Reetz. 2002. Underspecified recognition. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology 7*, 637–676. Berlin: Mouton.
- Lahiri, Aditi and Henning Reetz. 2010. Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics* 38(1): 44–59.
- Lehiste, Ilse. 1965. The function of quantity in Finnish and Estonian. *Language* 41(3): 447–456.
- Liberman, Mark and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2): 249–336.
- Malécot, André. 1960. Vowel nasality as a distinctive feature in american english vowel nasality as a distinctive feature in american english vowel nasality as a distinctive feature in american english vowel nasality as a distinctive feature in american english. *Language* 36(2, Part 1): 222–229.
- Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.
- Morén, Bruce. 2006. Consonant–vowel interactions in Serbian: Features, representations and constraint interactions. *Lingua* 116: 1198–1244.
- Moulton, William G. 1947. Juncture in modern standard German. *Language* 23(3): 212–226.
- Newman, Stanley S. 1952. Review of Harris (1951). *American Anthropologist* 54(3): 404–405.
- Norvig, Peter. n.d. English letter frequency counts: Mayzner revisited, or ETAOIN SRHLDCU. Published online at <http://norvig.com/mayzner.html>. Accessed 21 July 2017.
- Odden, David. 2006. Phonology ex nihilo. Presented at the Phonology Get-Together, Universitetet i Tromsø, December 2006. Handout available online at [http://www.ling.ohio-state.edu/~odden/Phonology\\_ex\\_Nihilo.pdf](http://www.ling.ohio-state.edu/~odden/Phonology_ex_Nihilo.pdf).
- Pike, Kenneth L. 1947. Grammatical prerequisites to phonemic analysis. *Word* 3(3): 155–172.
- Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.

- Pulleyblank, Douglas. 2003. Covert feature effects. *WCCFL* 22: 398–422.
- Rasin, Ezer and Roni Katzir. 2016. On evaluation metrics in Optimality Theory. *Linguistic Inquiry* 47(2): 235–282.
- Rice, Keren D. 2002. Vowel place contrasts. In Mengistu Amberber and Peter Collins (eds.), *Language universals and variation*, 239–270. Westport, CT: Praeger.
- Rice, Keren D. 2007. Markedness in phonology. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 79–98. Cambridge: Cambridge University Press.
- Riggle, Jason. 2006. Using entropy to learn OT grammars from surface forms alone. In Donald Baumer, David Montero, and Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 346–353. Somerville, MA: Cascadilla Proceedings Project.
- Roberts, E. Wyn. 1999. On the notion of zero. In Sheila Embleton, John E. Joseph, and Hans-Josef Niederehe (eds.), *The emergence of the modern language sciences: Studies on the transition from historical-comparative to structural linguistics in honour of E.F.K. Koerner*, vol. 2, 41–66. Amsterdam: Benjamins.
- Samuels, Bridget D. 2009. *The structure of phonological theory*. Ph.D. thesis, Harvard University.
- Sapir, Edward. 1925. Sound patterns in language. *Language* 1(1): 37–51.
- Scheer, Tobias. 2011. *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin/New York: De Gruyter Mouton.
- Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423, 623–656.
- Shannon, Claude E. and Warren Weaver. 1949. *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

- Stemberger, Joseph Paul and Carol Stoel-Gammon. 1991. The underspecification of coronals: Evidence from language acquisition and performance errors. In Carole Paradis and Jean-François Prunet (eds.), *The special status of coronals: Internal and external evidence*, no. 2 in *Phonetics and Phonology*, 181–199. San Diego: Academic Press.
- Steriade, Donca. 1995. Underspecification and markedness. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 114–174. Oxford: Blackwell.
- Trager, George L. and Henry Lee Smith, Jr. 1951. *An outline of English structure*. Studies in Linguistics: Occasional Papers 3. Norman, OK: Battenburg Press.
- Trubetzkoy, Nikolai S. 1939. Grundzüge der Phonologie. *Travaux du Cercle linguistique de Prague* 7.
- Trubetzkoy, Nikolai S. 1969. *Principles of phonology*. Berkeley and Los Angeles: University of California Press. Translated by Christiane A. M. Baltaxe from Trubetzkoy (1939).
- Vachek, Josef. [1936] 1976. Phonemes and phonological units. In *Selected writings in English and general linguistics*, 15–19. Prague: Academia. Originally published in *Travaux du Cercle linguistique de Prague* 6: 235–239.
- Vachek, Josef. 1964. On some basic principles of ‘classical’ phonology. *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung* 17(1): 409–432.