LOOKING BACK, MOVING FORWARD

The road not taken: *The Sound Pattern of Russian* and the history of contrast in phonology[1]

B. ELAN DRESHER ⓘ

University of Toronto

DANIEL CURRIE HALL ⓘ

Saint Mary's University

This article examines a turning point in the history of the theory of phonological distinctive features. In Morris Halle's (1959) *The Sound Pattern of Russian*, features are organized into a contrastive hierarchy designed to minimize the number of specified features. Redundancy rules, however, ensure that the resulting underspecification has no real phonological consequences and, in subsequent generative approaches to phonology, contrastive hierarchies were largely abandoned. We explore how Halle's hierarchy would have been different if it had been based on phonological patterns such as voicing assimilation, and show that this reorganization makes plausible predictions about other aspects of Russian phonology. We conclude by pointing to recent work in which the concept of a contrastive hierarchy has been revived, illustrating the range of phenomena that this theoretical device can account for if minimizing specifications is not the primary concern.

KEYWORDS: contrast, distinctive features, phonological representations, Russian

---

405

B. ELAN DRESHER & DANIEL CURRIE HALL

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth

(Robert Frost, 'The road not taken', 1916)

1. INTRODUCTION

Morris Halle's (1959) *The Sound Pattern of Russian* (*SPR*) sits at a significant fork in the road in the development of phonological theory. It can be viewed from one perspective as the culmination of a tradition of phonological analysis associated with the Prague School, the last in a line of works that include Roman Jakobson's (1941) *Kindersprache, Aphasie und allgemeine Lautgesetze*, Jakobson, Fant & Halle's (1952) *Preliminaries to Speech Analysis*, and Jakobson & Halle's (1956) *Fundamentals of Language*. From another perspective, it is the first major work in the new framework of generative phonology, a precursor to Chomsky & Halle's (1968) *The Sound Pattern of English*. We will show that from his earlier work with Jakobson, Halle retained the notion of a branching tree that generates all and only the contrastive features of each Russian underlying segment. In this approach, lexical representations are underspecified. *SPR* contains a novel argument for why it is necessary to generate feature specifications by means of such trees.

Halle's *SPR* is perhaps best known in the phonological literature for his argument against a certain formulation of the structuralist phoneme, based on his analysis of Russian regressive voicing assimilation (RVA). Halle showed that a unified treatment of this phenomenon required relinquishing the rigid division between morphophonemic and allophonic rules then favoured by many structuralists. We will suggest that this analysis, somewhat ironically, had the effect of devaluing the importance of the branching tree, and of contrastive features more generally. For while Halle's hierarchical tree structure assigned contrastively underspecified features to Russian phonemes, *SPR* also posited a system of rules that made underspecification effectively irrelevant. As a consequence, Chomsky & Halle's (1968) *The Sound Pattern of English* (*SPE*) abandoned contrastive feature specification and feature hierarchies. The result, in our view, is that generative phonology gave up some major insights of the Prague School phonologists N. S. Trubetzkoy and Roman Jakobson.

However, Halle could have made essentially the same argument by taking a different path. In this article, we explore the road not taken in *SPR,* showing how Halle could have made more substantive use of a contrastive feature hierarchy in his treatment of voicing assimilation and what consequences that would have had for the rest of his analysis of Russian in the short term, and, in the longer term, for the place of contrastive specification in the theory of generative phonology.

We begin in Section 2 by summarizing Halle's feature analysis of the segments of Russian, and his argument for the necessity of representing contrastive feature

406

specifications hierarchically, by means of branching trees. We also discuss his criteria for how to order features into a hierarchy; these criteria privileged the MINIMALITY PRINCIPLE, which requires that features be ordered so as to minimize the number of feature specifications. In Section 3 we review Halle's famous argument against the 'taxonomic' phoneme and its perhaps unintended by-product, which was to hasten the demise of contrastive specification in the developing theory of generative grammar.[1]

In Section 4 we begin our exploration of the road not taken by arguing that the original rationale for contrastive specification in the work of Trubetzkoy, Jakobson, and even Halle himself was not Minimality, but rather Activity, the principle that features are ordered so as to account for patterns of phonological activity in a language. We then consider what Halle's analysis of Russian and his argument against the taxonomic phoneme would have looked like had he adopted the ACTIVITY PRINCIPLE as the primary criterion for ordering features. In Section 5 we survey some recent synchronic and diachronic analyses that are based on contrastive feature hierarchies governed by the Activity Principle, and suggest that this road, which is still there to be taken, is a promising direction for phonological theory.

## 2. *SPR* : THE BRANCHING TREE AND THE BRANCHING ROAD

2.1 *The branching tree*

On page 46 of *SPR* is Figure I-1, a tree diagram that shows the contrastive feature specifications of every phoneme in Russian.[2] Each branch in the tree subdivides the inventory according to some distinctive feature. This tree is reproduced in our Figure 1. Since the tree is very large, we have divided it into separate subtrees for vocalic and non-vocalic segments. We have also replaced Halle's transcription symbols by their IPA equivalents (as we have done throughout this article).

Halle's numbering of features, listed in the key in Figure 1, is preserved; feature 1 in Figure 1a is [±vocalic], which forms the top branch in his tree. Because this is the first feature, every segment in the inventory must receive a value for this feature. In Halle's feature system, the [+vocalic] class comprises vowels and liquids; non-liquid consonants, including the glide /j/, are [−vocalic]. From each numbered node in the tree, the branch to the left contains phonemes with the negative value of the corresponding feature, and phonemes with the positive value are in the branch to the right.

---

[1] On the term 'taxonomic' please see Section 3.2. 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 without causing a major shift in the direction of phonological theory. This is an interesting question but it is not our focus here; see Anderson (2000) and Dresher & Hall (to appear) for discussion. Here, we accept that Halle's argument was influential and correct in its main conclusions; our question is, what would it have looked like had he adjusted a few assumptions about contrast and underspecification?

[2] Halle refers to the underlying segments that are the terminals of the feature tree as 'morphonemes' and encloses them in braces { }. He reserves the term 'phoneme', enclosed in diagonals / /, for the structuralist phoneme that he will argue against (Halle 1959: 23n13). In keeping with later usage, we refer to the underlying segments of *SPR* as 'phonemes' and enclose them in / /.
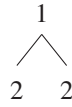
(a) Top of the tree
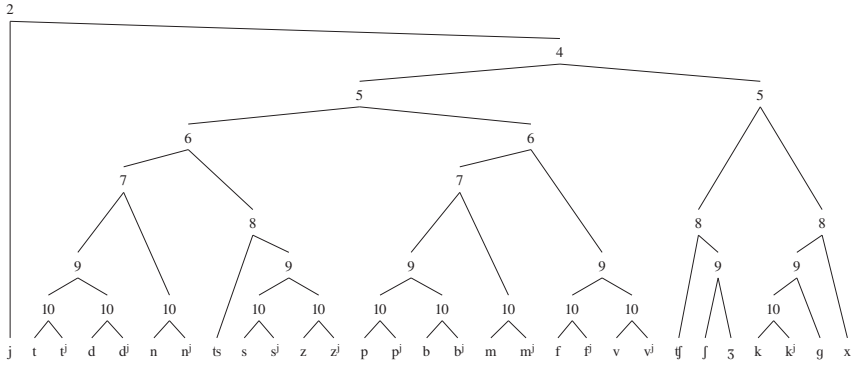
(b) [−vocalic] phonemes
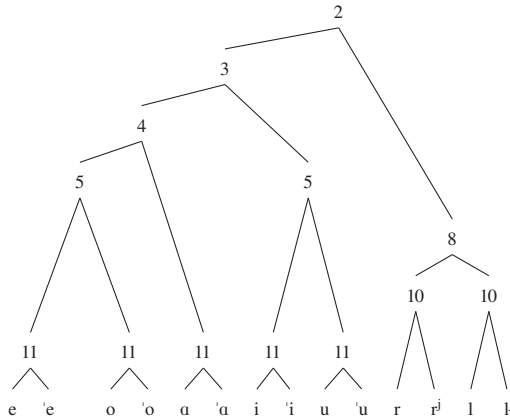
(c) [+vocalic] phonemes

Key: 1=[±vocalic], 2=[±consonantal], 3=[±diffuse], 4=[±compact], 5=[±low tonality], 6=[±strident], 7=[±nasal], 8=[±continuant], 9=[±voiced], 10=[±sharped], 11=[±accented]

Figure 1
Contrastive hierarchy for Russian phonemes, adapted from Halle (1959: 46).

408

The next feature, numbered 2, is [±consonantal], which is contrastive in both major branches of the tree. Looking first at the [−vocalic] branch of the tree in Figure 1b, only the glide /j/ is contrastively [−consonantal]. Because the tree has no non-branching non-terminal nodes, no feature is assigned to any segment unless it serves to distinguish that segment from at least one other. As /j/ is now unique, no further features are assigned to it. All the other segments in Figure 1b are [+consonantal], and therefore need to be distinguished from each other by additional features. On the [+vocalic] side of the tree in Figure 1c, vowels are [−consonantal] and liquids are [+consonantal].

The third feature in the ordered list is [±diffuse], which is distinctive only for the vowels (Hale 1959: 126). As it does not make any contrasts in the consonants in Figure 1b, it is skipped in that tree and the next feature is 4, [±compact]. The liquids in Figure 1c are not distinguished by any other feature until number 8, [±continuant] (perhaps unexpectedly, /r, rʲ/ are [−continuant] and /l, lʲ/ are [+continuant]).

The division of the tree by contrastive features continues until every phoneme has been uniquely distinguished.

2.2 *An argument for specification by branching trees*

Before continuing, we should ask why the branching tree occurs at all: why is it necessary to specify features in this way? In *SPR*, Halle makes an argument on behalf of branching trees – the first such argument we have found in the literature. He argues that phonological features must be ordered into a hierarchy because this is the only way to ensure that segments are kept properly distinct, as defined in the DISTINCTNESS CONDITION (Halle 1959: 32) in (1):

(1) *The Distinctness Condition*
     Segment-type /A/ will be said to be different from segment-type /B/, if and only if at least one feature which is phonemic in both, has a different value in /A/ than in /B/; i.e. plus in the former and minus in the latter, or vice versa.

This formulation is designed to disallow contrasts involving a zero value of a feature. Consider, for example, the typical sub-inventory /p, b, m/ shown in (2), and suppose we characterize it in terms of two binary features, [±voiced] and [±nasal]. In terms of full specifications, /p/ is [−voiced, −nasal], /b/ is [+voiced, −nasal], and /m/ is [+voiced, +nasal].

(2) /p b m/: Full specification for [±voiced] and [±nasal]

	/p/	/b/	/m/
[±voiced]	−	+	+
[±nasal]	−	−	+

Which of these specifications are contrastive? Many phonologists reason as follows. We observe that /p/ and /b/ are distinguished only by [±voiced], so these

409

specifications must be contrastive. Similarly, /b/ and /m/ are distinguished only by [±nasal]; these specifications must also be contrastive. Specifying only these features yields the partial specification in (3).

(3)   /p b m/: Partial specification for [±voiced] and [±nasal]

	/p/	/b/	/m/
[±voiced]	−	+	
[±nasal]		−	+

The remaining specifications are predictable from the ones in (3). Since /p/ is the only [−voiced] phoneme in this inventory, its specification for [±nasal] is predictable, hence redundant. We can write a rule or constraint: if [−voiced], then [−nasal]. Similarly, /m/ is the only [+nasal] phoneme, so its specification for [±voiced] is redundant: if [+nasal], then [+voiced]. This is a still-popular way of thinking about contrastive specifications. Dresher (2009) refers to it as 'pairwise comparison'; we will call it the 'Minimal Difference' approach. By whatever name, essentially this approach is adopted explicitly by Padgett (2003), Calabrese (2005), Campos Astorkiza (2007), and Nevins (2010), and by many others implicitly.

According to the definition proposed by Nevins (2010: 98), a segment S with specification [αF] is contrastive for F if there is another segment S′ in the inventory that is featurally identical to S, except that it is [−αF]. The condition is met by the specifications of [±voiced] on /p/ and /b/. It is not met by [−nasal] on /p/, because there is no other segment in this inventory that differs from /p/ only by this feature. For this feature to be contrastive on /p/ there would have to be a voiceless nasal /m̥/ in this inventory, as in (4). The same would be required by Minimal Difference to make [+voiced] contrastive on /m/. In the absence of voiceless nasal /m̥/, Minimal Difference gives us the specifications in (3).

(4)

	/p/	/b/	/m/	/m̥/
[±voiced]	−	+	+	−
[±nasal]	−	−	+	+

According to the Distinctness Condition, however, the set of specifications in (3) is not properly contrastive. /p/ is 'different from' /b/, because /p/ is [−voiced] and /b/ is [+voiced]. Similarly, /b/ is 'different from' /m/, because /b/ is [−nasal] and /m/ is [+nasal]. But /p/ is not 'different from' /m/: where one has a feature, the other has no specification. Therefore, these specifications are not properly contrastive. They violate the Distinctness Condition because no feature hierarchy yields this result. If we order [±voiced] ahead of [±nasal] (notated [±voiced] ≫ [±nasal]), we generate an 'extra' specification on /m/ (5a, 6a). If we order [±nasal] ≫ [±voiced], we generate an 'extra' specification on /p/ (5b, 6b).

(5) (a) [±voiced] ≫ [±nasal]   (b) [±nasal] ≫ [±voiced]
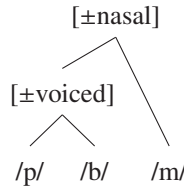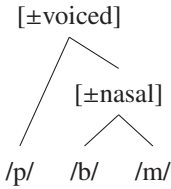
(6) (a) [±voiced] ≫ [±nasal]   (b) [±nasal] ≫ [±voiced]

	/p/	/b/	/m/
[±voiced]	−	+	+
[±nasal]		−	+

	/p/	/b/	/m/
[±voiced]	−	+	
[±nasal]	−	−	+

That is, according to *SPR*, the correct theory of contrast is a hierarchical one, and not Minimal Difference. To be properly contrastive, a set of feature specifications must be derivable by a branching tree. This view of contrastive specifications is reflected in the first generative phonology textbook by Harms (1968), but the matter became dormant after the critique of underspecification by Lightner (1963) and Stanley (1967) (see further Section 3.3). We believe, however, that Halle was correct in requiring contrastive specifications to be hierarchical. As demonstrated by Archangeli (1988) and Dresher (2009), the Minimal Difference approach often fails to yield any intelligible set of specifications. Conceptually, the main flaw of Minimal Difference is its failure to recognize that contrastive relations in an inventory exist not just between pairs of segments, but also between groups of segments at different levels of the hierarchy.

For example, [+voiced] on /m/ in (6a) is predictable from the other specifications, but it is required to distinguish /b, m/ as a group from /p/; at this point in the tree, this feature is contrastive for the entire set /p, b, m/. The same is true for [−nasal] on /p/ in (6b). These contrasts involve 'minimal differences' at the point where they are made – in (6a) [±voiced] makes a minimal difference between /p/ and /b, m/ and in (6b) [±nasal] makes a minimal difference between /p, b/ and /m/. The layered character of hierarchically derived contrastive features is what results in contrastive features that are not necessarily unpredictable.[3]

2.3 *Rationale for feature hierarchies: Minimality*

In many cases, there is more than one feature that could potentially distinguish two segments (or sets of segments), and so the ordering of the features has consequences for which of them are actually specified. In the simple example of /p, b, m/ in the previous section, we have seen that different feature orderings result in different

---

[3] For further discussion of the logic of contrast see Dresher (2009: 11–36).

411

contrastive specifications. Returning to *SPR*, consider, for example, the class of labial consonants /p, pʲ, b, bʲ, m, mʲ, f, fʲ, v, vʲ/, which is identified within the non-vocalic inventory in Figure 1b by the features [+consonantal, −compact, +low tonality]. Within this set, feature 6, [±strident], distinguishes the stops and nasals /p, pʲ, b, bʲ, m, mʲ/ from the fricatives /f, fʲ, v, vʲ/. This means that feature 8, [±continuant], is not marked on any of the labials, because the distinction that it would make has already been made by [±strident]. If the order of these two features were reversed, then [±continuant] would mark the division, and [±strident] would be redundant.

Although [+strident] is specified on the labiodental fricatives, it is not marked on the prototypically strident postalveolar fricatives and affricate /ʃ, ʒ, tʃ/. The set of posterior consonantal segments /tʃ, ʃ, ʒ, k, kʲ, g, x/ identified by the features [+consonantal, +compact] in Figure 1b (see also Figure 2b in Section 3.1), is divided into postalveolars /tʃ, ʃ, ʒ/ and velars /k, kʲ, g, x/ by feature 5, [±low tonality]. Because there are no non-strident postalveolar consonants and no strident velar ones, neither subset can be further divided by [±strident] (feature 6). The later-ordered [±continuant] (feature 8), on the other hand, does have a role to play here, separating the affricate /tʃ/ from the fricatives /ʃ, ʒ/ and the fricative /x/ from the stops /k, kʲ, g/. All these specifications could be altered if the features were ordered differently.

What, then, is the basis for the ordering? Halle (1959: 29–30) provides the rationale for it in his Condition (5):

> *Condition (5): In phonological representations, the number of specified features is consistently reduced to a minimum compatible with satisfying Conditions (3) and (4).*

Conditions (3) and (4) set out basic requirements of adequacy: Condition (3), which we discuss further in Section 3.2, requires that phonological representations contain enough information to derive the corresponding phonetic utterances without referring to other (e.g. semantic or syntactic) sources, and Condition (4) requires the phonological description to be 'appropriately integrated into the grammar of the language' (see further Section 4.2). Once those conditions are met, the primary desideratum in ordering the features is to minimize the number of feature specifications.

This concern for minimality stems at least in part from Halle's interest in information theory; Cherry, Halle & Jakobson (1953) show how binary features, parsimoniously assigned in a hierarchical order, can provide a metric for how much information phonemes can encode, and with how much redundancy. We will refer to this emphasis on efficiency of encoding as the Minimality Principle, which can be stated as in (7):

(7) *Minimality Principle for feature ordering*
     The object of ordering features into a hierarchy is to minimize redundancy in phonological representations and thereby to maximize the amount of information conveyed by each feature.

412

A concern with Minimality is reflected in Halle's (1959: 44–45) observation that his analysis of Russian contains 43 phonemes specified by 271 feature specifications, or 6.3 distinctive feature statements per phoneme. He compares 6.3 with the lower limit of $\log_2 43$ ($\approx$5.43) specifications, which would represent the most efficiently branching tree for 43 phonemes.[4] By the criterion of Minimality, this improves upon the feature hierarchy for Russian posited by Cherry et al. (1953), which uses an average of 6.5 features per phoneme for an inventory of 42 phonemes (compared to a logical minimum of $\log_2 42 \approx 5.39$).[5] Halle (1959: 45) cautions against reading too much into the precise numbers involved, concluding only that 'the minimization process has achieved results of the type that might reasonably be expected'.

Minimality governs (at most) only the hierarchical ordering of features, and not the choice of features themselves; an underlying assumption for both Cherry et al. (1953) and Halle (1959) is that features have phonetic content, and so the inventory can only be divided along lines of identifiable phonetic difference. 'One might ask,' Cherry et al. (1953: 37–38) write, 'Why cannot a type of feature pattern be invented which employs only' the logical minimum average number of specifications per phoneme? Their answer is that '[t]his could perhaps be done; but the distinctive features used at present… serve other purposes and are intimately related to the physical production of speech' (Cherry et al. 1953: 38). The 'other purposes' that distinctive features serve include not only phonetic description, but also accounting for phonological patterns; in the sections that follow, we examine the consequences of giving Minimality priority over phonological patterning in the ordering of features and show how a different approach to building feature hierarchies could have changed the agenda that *SPR* set for generative phonology.

3. Halle's argument against the 'taxonomic' phoneme

3.1 *The unpaired obstruents*

The ordering of two features in the part of the tree we looked at in Section 2.3 had momentous consequences for the development of phonological theory. These are features 8, [±continuant], and 9, [±voiced]. The relevant portions of the hierarchy are repeated in Figure 2 in a more easily legible form and size.

In these subtrees, [±continuant] is ordered above [±voiced]. Each segment in these diagrams has a specification for [±continuant], but the same is not the case for [±voiced]. The three circled consonants in Figure 2, /ts, ʧ, x/, have no phonemic

---

[4] Actually, by our count his analysis has 272 specifications, a discrepancy discussed below in §7. Also, Halle mistakenly says that $\log_2 43 = 5.26$; as far as we can see, this is just a typo.

[5] The difference in total numbers of phonemes is due to the fact that Cherry et al.'s inventory omits unstressed /e o/, while *SPR* omits palatalized /gʲ/. See Chew (1999: 62–63) for a summary of why the phonemic status of palatalized velars in Russian has been controversial, and Iosad (2012) for a description and analysis of the reduction of unstressed mid vowels.
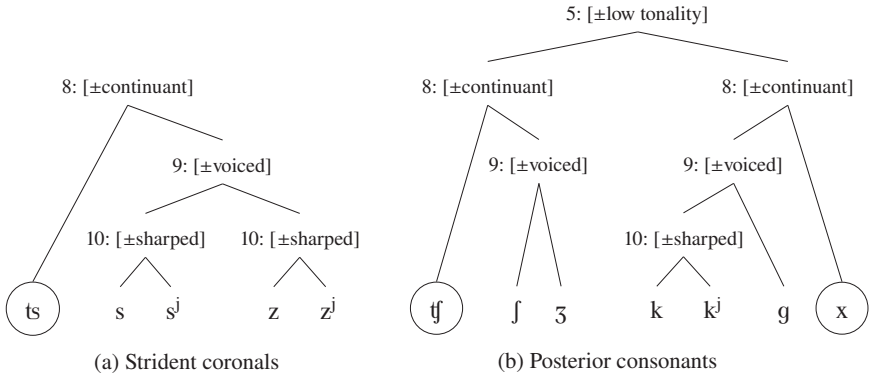
Figure 2
Subtrees resulting in underspecification of [±voiced] in Halle (1959).

voiced counterparts; that is, there are no underlying */dz, dʒ, ɣ/ in Russian. These are the 'unpaired' voiceless obstruents that formed the focus of Halle's (1957, 1959) famous argument against a certain conception of the phoneme. The unpaired obstruents receive no specification for [±voiced], because they have already been fully distinguished from the rest of the inventory by features higher in the hierarchy. But, as Halle (1959: 22–23) pointed out, these segments behave phonologically like other voiceless obstruents with respect to voicing assimilation. This fact is accounted for in *SPR* by rule ordering: Rule P1b, which fills in redundant [−voiced] on the unpaired obstruents, precedes the assimilation Rule P3a:

(8)　(a)　*Rule P1b*
　　　　Unless followed by an obstruent, /ts/, /tʃ/ and /x/ are voiceless (Halle 1959: 63).
　　(b)　*Rule P3a*
　　　　If an obstruent cluster is followed by a – (dash) boundary or by a sonorant, then, with regard to voicing, the cluster conforms to the last segment; if it is voiced, so are all the other segments in the cluster; if it is voiceless, so is the entire cluster (Halle 1959: 64).

The effect of this ordering is that the unpaired obstruents end up triggering assimilatory devoicing of obstruents to their left, despite being underlyingly underspecified. This can be seen, for example, in the derivation of [safxós] 'state farm' from /sovxoz/, the relevant portion of which is schematized in (9):

(9)　　　　　　　　　P1b　　　　P3a　　　　other rules

　　　　sovxoz　→　sovxoz　→　sofxoz　　　→　　[safxós]

[±voiced]:　　+∅　　　　+−　　　　−−

The underlying form is /sovxoz/, where /v/ is specified as [+voiced] and /x/ has no specification for [±voiced], indicated in (9) by ∅. By Rule P1b this zero specification is immediately changed to [−voiced]. After this point in the derivation, /x/ is like any other contrastively voiceless segment, and it acts as such to condition regressive voicing assimilation (Rule P3a) on /v/, which devoices to [f]. In sum, although 'unpaired' /ts, tʃ, x/ are not specified for [±voiced] by the branching tree, they are assigned [−voiced] early in the derivation, and subsequently behave like other voiceless segments.

With respect to contrast, then, this derivation demonstrates that the contrastive status of an obstruent with respect to voicing has no bearing on its behaviour with respect to regressive voicing assimilation. Though Halle did not focus on this point, we will see that it will have the effect of making contrastive underspecification a further, perhaps unintended, target of Halle's argument.

3.2 *The taxonomic phoneme*

The analysis reviewed in the previous section formed the basis of Halle's argument against the 'taxonomic' phoneme as employed in American structuralist phonology (Halle 1957, 1959; continued in Chomsky 1964 and Chomsky & Halle 1965; and see also Postal 1964, 1968).[6] From the 1930s to the 1950s, American linguists who were followers of Leonard Bloomfield attempted to make the definition of the phoneme more precise than it had been. In practice, this amounted to placing an increasing number of restrictions on the relationship between the phonemic level and the phonetic signal. One such restriction concerned 'phonemic overlapping' (Bloch 1941): one sound (a phone) may not correspond to two different phonemes in the same context. In Russian, for example, an underlying /ts/ can surface as [ts], or, when subject to regressive voicing assimilation, as [dz]. Because there is no underlying */dz/ in Russian, both [ts] and [dz] are unambiguous allophones of a single phoneme /ts/. This situation is permitted.

Consider now an underlying /t/: just like /ts/, it can surface as voiceless [t], or as [d] when voiced by regressive voicing assimilation. In this case, however, there is another source for [d] in Russian, namely underlying /d/. In some positions /d/ is in contrast with /t/, but the contrast is neutralized in assimilation environments.[7] Consequently, without additional information from the meaning or morphology, we would not know which phoneme any given [t] or [d] belongs to in neutralization

---

[6] The term 'taxonomic phoneme' does not appear in *SPR*; rather, Halle aims his argument against a representation that adheres to his Condition (3a), what he calls (Halle 1959: 23) 'phonemic' (in quotes), citing the usage of Joos (1957: 92). The term 'taxonomic phonemics' was coined by Chomsky (1964: 75) to refer to 'a body of doctrine… to all or part of which a great many linguists would subscribe'.

[7] The overlap is even worse than this, because regressive voicing assimilation can devoice /d/ to [t] if the final obstruent in the cluster is voiceless; so both phonetic [t] and [d] could be allophones of either /t/ or /d/.

415

contexts. This type of overlapping was not allowed by most American structuralist phonologists in the 1950s. The condition that rules it out came to be known as 'bi-uniqueness'; Halle (1959: 21n8) cites Hockett's (1951: 340) formulation: 'For a notation to be phonemic we require a bi-unique, one–one relation rather than a many–one relation' between representation and utterance.

Halle (1959: 24) observes that bi-uniqueness did not play a role in the practice of Sapir or even of Bloomfield, but was a more recent addition to phonological theory.[8] As Halle (1959: 20–21) puts it, a relatively weak condition on phonological descriptions that is 'accepted by all' is Condition (3), which requires that one be able to derive an utterance from a phonological representation. To enforce bi-uniqueness, an additional condition is needed that requires that one be able to derive the (underlying) phonological representation from the utterance. Halle's (1959: 21) version of this constraint is Condition (3a):

> *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.*

> *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.*

The only solution, in a theory that adheres to Condition (3a) in addition to Condition (3), is to require that every phonetic [t] must be an allophone only of the phoneme /t/, and every [d] must be an allophone only of the phoneme /d/. The fact that *t* alternates with *d* must be expressed in a mapping between the morphophonemic (lexical) level and the phonemic level; overlapping is permitted between morphophonemes. We will represent morphophonemes between double slashes, // //.

In a theory that adheres to Condition (3a) in addition to Condition (3), then, rules that change one phoneme into another must belong to the morphophonemic system, while phonemic rules that change phonemes into their various allophones in the appropriate environments apply later. Russian regressive voicing assimilation, though, does both of these: it can change some voiceless phonemes into their voiced counterparts or vice versa (e.g. mapping //t// to /d/ or //d// to /t/), but it can also change an unpaired voiceless phoneme into its voiced allophone (e.g. mapping /ʦ/ to [ʣ]). Adherence to Condition (3a) would preclude a unified account of assimilation, requiring it to be treated as two separate rules operating at two separate levels, as schematized in (10).

---

[8] On the distance between the practice of Bloomfield and the 'post-Bloomfieldians' who claimed to be following in his footsteps, see Ladd (2014, to appear).

416

(10)   Morphophonemic representations: //t// //ts//

Morphophonemic RVA rule: ↓

Phonemic representations: /d/   /ts/

Phonemic RVA rule: ↓

Phonetic representations: [d]   [ʤ]

Halle (1959) argues that rather than conforming to Condition (3a), eliminating the taxonomic phonemic level, and with it the distinction between morphophonemic and phonemic rules, allows a unified treatment of assimilation and other processes and produces a simpler phonological grammar whose rules map directly from underlying lexical representations (the 'systematic phonemic level' of Chomsky 1964) to surface phonetic forms without any privileged intermediate level. This is schematized in (11).

(11)   Underlying lexical representations: /t/   /ts/

Regressive voicing assimilation: ↓   ↓

Phonetic representations: [d] [ʤ]

3.3 *The demise of underspecification and branching trees*

It is widely understood that Halle's rejection of the taxonomic phoneme played an important role in setting the course for generative phonology. But another aspect of his analysis of voicing assimilation also had far-reaching consequences that are less generally appreciated – and, we will argue, that were not necessary.

In the derivation in (9), the distinction between contrastive and redundant features is made irrelevant by the rule system: although /x/ is not assigned a value for [±voiced] underlyingly because this feature is not needed to distinguish it from the rest of the inventory, it nonetheless receives the value [−voiced] by Rule P1b in time to trigger assimilatory devoicing of /v/. More generally, any unspecified value could be filled in by rule in time to make its underlying absence phonologically inconsequential.

More than inconsequential, the derivations in (9) and (11) make it impossible to draw a line between the contrastive and non-contrastive phonology that would not incur the same problem as the taxonomic phoneme. Trying to limit the phonology to apply only to contrastively specified features – that is, reordering rules P1b and P3a in (8) – would either yield an incorrect result where regressive voicing assimilation does not apply to /v/ in (9) because /x/ has no voicing feature, or would again require regressive voicing assimilation to apply twice, first to contrastively specified features, and then again after non-contrastive features are filled in.

417

This made the whole notion of contrastive underspecification vulnerable to arguments such as those of Lightner (1963) and Stanley (1967), who pointed out that redundancy rules could be exploited to make underspecification for a given feature tantamount to a third value for that feature, increasing the expressive power of the grammar in a way that was argued to be improper (see further Section 6). Underspecification was thus abandoned altogether in *SPE*, along with the contrastive hierarchies that generated it.

The result was that language-particular feature contrasts did not play a role in the theory of generative grammar that developed from *SPE*. Rather, the underlying phonemes of a language came to be fully specified according to a universal feature theory; to a first approximation, these representations are not sensitive to language-particular phoneme inventories or contrasts. Where problems arose from this, *SPE* attempted to address them by introducing a theory of markedness (Chomsky & Halle 1968: ch. 9; see Dresher 2009: 104–117 for discussion). Contrastive hierarchies like the one in Figure 1 largely disappeared from phonological theory for a generation. With a few exceptions, they did not return even with the revival of interest in theories of underspecification in the 1980s (e.g. Kiparsky 1982; Broselow 1984; Archangeli 1984, 1985, 1988; Pulleyblank 1986; Clements 1987; Steriade 1987; Hargus 1989).

The consequences for underspecification of Halle's (1959) treatment of the unpaired obstruents have gone largely unremarked in the literature, particularly in comparison to their consequences for the definition of the phoneme. A recent exception is Reiss (2017: §2), who argues – in our view rashly and erroneously – that *SPR* showed contrast to be irrelevant to phonology. His contention is that, because a rule like regressive voicing assimilation (in its unified form) applies in a uniform way without regard for whether its effect is neutralizing or allophonic, phonological rules in general are never affected by whether there is an underlying contrast between two segments. Reiss (2017: 26) formulates the following as a principle called 'No contrast in phonology':

> *Rules are blind to the content of the lexicon. Rules are functions that map segments to segments depending on their local context (syllable structure, precedence relations, etc.). Global or systemic considerations such as the content of segment inventories are not relevant to rules.*

'In plain English,' Reiss (2017: 26) continues, 'the phonology doesn't "care" about contrast, because it has no mechanism by which to do so.' But *SPR* includes a mechanism by which the underlying contrasts in the lexicon can in principle affect phonological rules, namely the tree in Figure 1. If featural representations are determined by a contrastive hierarchy, then contrast can be relevant to phonological rules, because rules can refer to features. The *SPR* rule system happened to be rigged to make contrast irrelevant by filling in redundant features in time for them to behave identically to contrastive ones, but this was not a logically necessary property of the system – rather, it was made necessary by the fact that Halle's (1959) approach to feature ordering considered underspecification primarily in broad information-theoretic terms, rather than as a potentially explanatory

component of phonological theory. In the next section, we explore an alternative approach to feature ordering, and the consequences it could have had for Halle's analysis of Russian and the course of generative phonology.

4. ACTIVITY AS THE GUIDING PRINCIPLE FOR ORDERING FEATURES: THE ROAD NOT TAKEN

4.1 *Origins of contrastive feature hierarchies*

Although the main criterion for ordering features in *SPR* is Minimality, this was not the original rationale for the use of contrastive hierarchies in phonology. The notion of contrastive specification by hierarchical ordering of features can be traced back to the Prague School in work by Trubetzkoy (1929, 1939, 1969) and Jakobson ([1931] 1962). Although they do not use tree diagrams of the sort found in *SPR*, feature hierarchies are implicit in much of their work, as discussed in some detail by Dresher (2009, 2015, 2016).[9] A rare example of an explicit mention is Trubetzkoy (1969: 103), who describes the Polabian vowel inventory as involving 'a certain hierarchy… with respect to the correlation of tongue position and the correlation of lip rounding', which he diagrams as in (12).[10]

(12)   back vowels – front vowels $\begin{cases} \text{rounded} \\ \text{unrounded} \end{cases}$

In other words, the rounding contrast exists only within the category of front vowels. And this statement is an analytical claim, not a theory-neutral observation: given the triplets of non-low vowels /i, y, u/ and /e, ø, o/, it would be just as plausible from the inventory shape alone to say that the front–back contrast exists only within the category of rounded vowels. The difference between the two statements is the difference between the two contrastive hierarchies in (13):

(13)   (a) *Trubetzkoy's analysis*   (b) *The other order*

          [±back] ≫ [±round]   [±round] ≫ [±back]

---
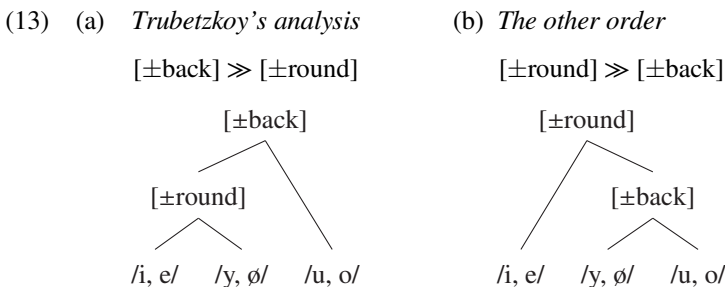
[9] Some examples adduced by Dresher include: Jakobson's ([1931] 1962) analysis of the Czech vowel system; Trubetzkoy's (1939) analyses of five-vowel (Latin, Archi, Japanese) and six-vowel (Bulgarian) systems, the contrastive status of German /h/ versus Czech /ɦ/, and the analysis of French and Greek consonant contrasts; and Jakobson & Lotz's (1949) rationale for their contrastive feature analysis of Standard French.

[10] The original German may be found on page 93 of Trubetzkoy (1939).

419

The hierarchy in (13a) groups front vowels together as a natural class, whereas the one in (13b) does not; it groups rounded vowels together instead. Trubetzkoy's (1969: 102) basis for choosing (13a) for Polabian is that '[t]he oppositions between the back and front vowels of the same degree of aperture were constant (non-neutralizable), while the oppositions between rounded and unrounded vowels of the same degree of aperture (*ü-i, ö-e*) were neutralizable after *v* and *j*'.[11] Given the hierarchy in (13a), /y, ø/ are readily identifiable as the rounded counterparts to /i, e/; in (13b), on the other hand, where /i, e/ have no specification for [±back], their 'counterparts' are all four of /y, ø, u, o/.[12]

Underpinning Trubetzkoy's use of features is the idea that only some properties of a segment are active, or relevant to the phonology, and these are the distinctive, or contrastive, properties. It is thus phonological activity that determines what the features are and how they are ordered, where activity can be defined as in (14), adapted from Clements (2001: 77):

(14)　*Phonological activity*
　　　A feature is phonologically **active** if it plays a role in the phonological computation; that is, if it is required for the expression of phonological regularities in a language, including both static phonotactic patterns and patterns of alternation.

Given this definition, we can define an alternative to the Minimality Principle that gives the feature hierarchy a more important role to play in accounting for phonological patterns, and which we believe makes explicit the original Prague School motivation for ordering features:

(15)　*Activity principle for feature ordering*
　　　The object of ordering features into a hierarchy is to account for patterns of phonological activity in a language.

4.2 *Contrast and activity*

The original intuition behind feature hierarchies is that there is a connection between phonological activity and contrast (Dresher 2009, 2015). In fact it can be shown that phonological activity still played an important role in choosing the ordering of features in *SPR*. Recall that Halle's (1959) analysis requires 272 feature specifications. There are other orderings that use the same features as Halle uses that result in far fewer specifications; however, one has to be willing to accept some

---

[11] We use IPA /y, ø/ for Trubetzkoy's *ü, ö*.

[12] The hierarchy in (13b), however, is appropriate for Finnish, which has vowel harmony in which /y/ alternates with /u/, /ø/ alternates with /o/, and /i/ and /e/ are neutral; see Hall (2017) for discussion.

420

strange-looking groupings of segments. Thus, the first feature in Halle's ordering is [±vocalic], which divides all phonemes into [−vocalic] glides and consonants and [+vocalic] vowels and liquids. However, [±vocalic] divides the inventory unequally: there are 29 [−vocalic] phonemes in Figure 1, and 14 [+vocalic] vowels and liquids. The next feature, [±consonantal], also makes very unequal divisions: in the [−vocalic] domain, it divides between /j/ and the other 28 consonants; in the [+vocalic] sub-tree, it opposes 10 vowels against 4 liquids. Unequal divisions make for relatively inefficient trees, and increase the number of contrastive specifications.[13]

Consider the tree in Figure 3. It has 243 specifications, 29 fewer than Halle's. The top feature, [±low tonality], divides the tree 25 against 18. The groupings are not very natural: only two vowels are [+vocalic], and only some consonants are [+consonantal]. The fact that Halle did not consider this ordering is evidence that he did not follow Minimality to the exclusion of Activity. Recall that Condition (5), Halle's version of Minimality, is restricted to feature orders that satisfy Conditions (3) and (4). Here is Condition (4) in full (Halle 1959: 24):

> *Condition (4): The phonological description must be appropriately integrated into the grammar of the language. Particularly, in selecting phonological representations of individual morphemes, these must be chosen so as to yield simple statements of all grammatical operations … in which they may be involved.*
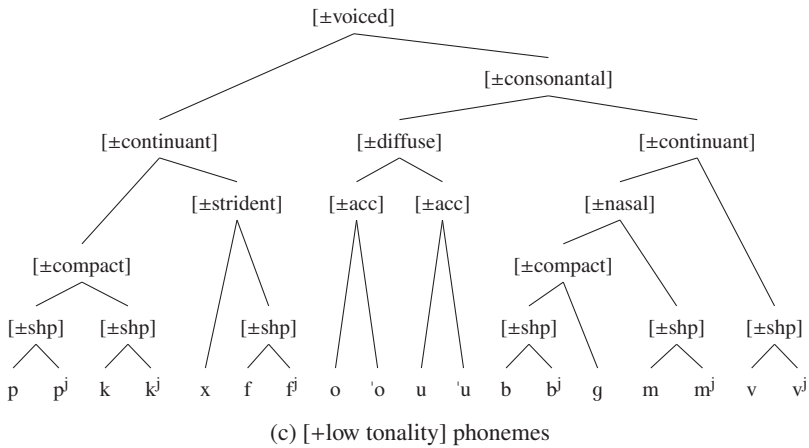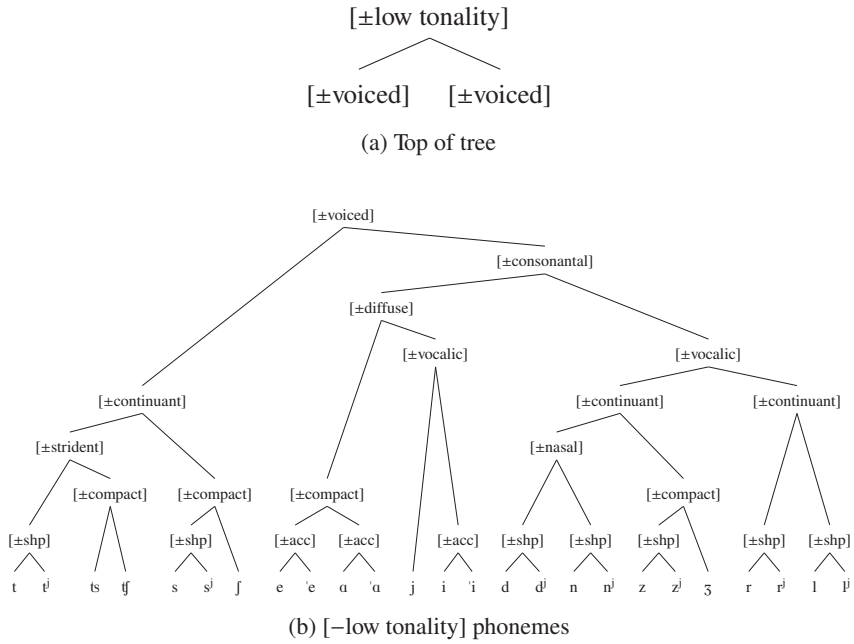
In other words, the number of feature specifications is reduced to a minimum as long as the resulting representations make sense in terms of the grammar of the language. That is, when assigning contrastive features to segments we are guided by Minimality as long as the result makes sense with respect to Activity.

The connection between contrast and activity is made explicit by what Hall (2007: 20) calls the Contrastivist Hypothesis:

(16) *Contrastivist Hypothesis*
  The phonological component of a language *L* operates only on those features which are necessary to distinguish the phonemes of *L* from one another.

---

[13] That *SPR* puts [±vocalic] and [±consonantal] at the top of the feature hierarchy, despite their lopsidedness, might suggest that Halle considered them major classes that ought to be highly placed, beyond the evidence of simple activity. We are grateful to a reviewer for pointing this out, and for suggesting that other conditions, such as the phonetic content of features, may have played a role in ordering. Indeed, Condition (4) could be said to include these criteria. Dresher (2015) cites Universality as another criterion for feature ordering that has been influential. For purposes of this article we can group these other non-Minimality criteria with Activity as empirical rationales for feature orderings. See Dresher (2015) for further discussion.

[±low tonality]

[±voiced]   [±voiced]

(a) Top of tree

[±voiced]

[±consonantal]

[±diffuse]

[±vocalic]

[±continuant]   [±vocalic]

[±continuant]   [±continuant]

[±strident]   [±nasal]

[±compact]   [±compact]   [±compact]

[±shp]   [±compact]   [±shp]   [±acc]   [±acc]   [±acc]   [±shp]   [±shp]   [±shp]   [±shp]   [±shp]

t tʲ ts tʃ s sʲ ʃ e ʲe ɑ ʲɑ j i ʲi d dʲ n nʲ z zʲ ʒ r rʲ l lʲ

(b) [−low tonality] phonemes

[±voiced]

[±consonantal]

[±continuant]   [±diffuse]   [±continuant]

[±strident]   [±acc]   [±acc]   [±nasal]

[±compact]   [±compact]

[±shp]   [±shp]   [±shp]   [±shp]   [±shp]   [±shp]

p pʲ k kʲ x f fʲ o ʲo u ʲu b bʲ g m mʲ v vʲ

(c) [+low tonality] phonemes

Key: [±low tonality] ≫ [±voiced] ≫ [±consonantal] ≫ [±diffuse] ≫ [±vocalic] ≫[±continuant] ≫ [±nasal] ≫ [±strident] ≫ [±compact] ≫ [±sharped] ≫ [±accented]

Figure 3
Contrastive hierarchy for Russian phonemes based on Minimality at the expense of Activity.

422

By identifying which features actually serve to mark contrasts in a given language, a contrastive hierarchy can limit which features can be phonologically active in that language. If, contra Halle (1959), redundant features are not filled in during the course of the phonological computation, then only the features identified as contrastive by the hierarchy can be active. This approach predicts that the size and shape of the underlying inventory place limits on how much information the phonological computation has access to; although it was not pursued in *SPR* or early generative phonology, it has been taken up in more recent work, some of whose results are discussed in Section 5.

Let us, then, revisit the Russian unpaired obstruents with (16) as our working hypothesis and Activity as the guiding principle for constructing feature hierarchies.

4.3 *The unpaired obstruents redux*

Regressive voicing assimilation shows that the feature [−voiced] is phonologically active on Russian /ts, ʧ, x/: these consonants can trigger assimilatory devoicing, even though they do not contrast with underlying voiced counterparts */dz, ʤ, ɣ/. Under the Contrastivist Hypothesis, this is only possible if [−voiced] serves to distinguish the unpaired obstruents from some other phonemes in the contrastive hierarchy. This is not the case in Halle's (1959) tree, as we have seen in Figure 2, but it can be the case if [±voiced] is given wider scope in the hierarchy. This can be accomplished with a minimal departure from Halle's ordering of features, by making [±voiced] take scope over [±continuant], which yields the specifications shown in Figure 4.[14]

The feature [±voiced] is now specified on the three circled segments, /ts, ʧ, x/, in each case because it marks some phonemic contrast: it distinguishes /ts, s, sʲ/ from /z, zʲ/, /ʧ, ʃ/ from /ʒ/, and /k, kʲ, x/ from /g/. But changing the order of features has other consequences for the specifications; a distinction that has already been marked by [±voiced] cannot also be marked by another feature. In this case, the trade-off is that the voiced consonants /z, zʲ, ʒ, g/, boxed in Figure 4, are now unspecified for [±continuant].

Is this new underspecification consistent with the facts of Russian phonology, or would the hierarchy in Figure 4 require redundancy rules to insert [±continuant] in the same way that the hierarchy in Figure 2 requires rules to insert [±voiced]? In fact, as discussed by Hall & Dresher (2016), this ordering of features conforms strikingly well to Russian sound patterns. Phonetically, /g/ is realized in some southern Russian dialects as a fricative [ɣ] or [ɦ]; this variability suggests that [−continuant] may not be a crucial part of the phonological specification of this segment.

---

[14] As Iosad (to appear) points out, essentially this analysis was proposed by Kasevich (1983). Kasevich (1983: 90–91) shows the necessary reordering of features in the hierarchy, but does not explore the consequences of narrowing the contrastive scope of [±continuant]. We are grateful to Iryna Osadcha for translating the relevant chapter of Kasevich (1983).
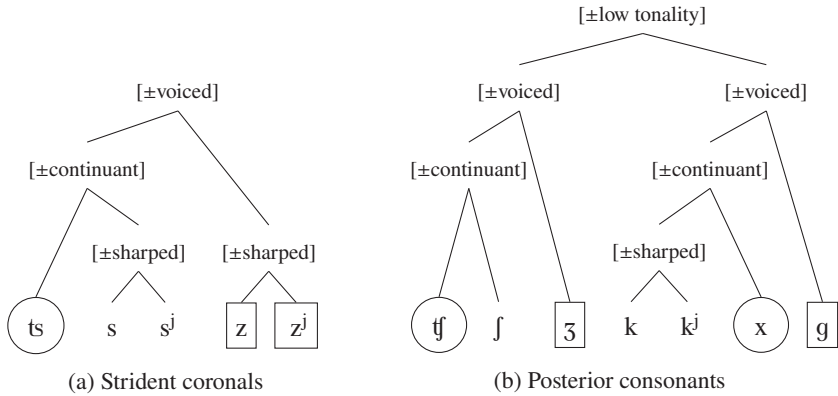
Figure 4
Revision of Figure 2 with [±voiced] taking scope over [±continuant].

There is also morphophonological evidence from alternations arising from the historical First Velar Palatalization.[15] In this pattern, velar consonants ([+low tonality] in Halle's (1959) feature system) alternate with their postalveolar ([−low tonality]) counterparts. As expected, these alternations pair /x/ with /ʃ/ and /k/ with /ʧ/; /g/ alternates with /ʒ/, which would be surprising if these two segments were specified for [±continuant], but can be elegantly characterized as part of the same pattern if they are not.[16] The alternating pairs are shown in (17) with their relevant feature specifications.

(17)

		[+low tonality]	[−low tonality]
[−voiced]	[+continuant]	x	ʃ
[−voiced]	[−continuant]	k	ʧ
[+voiced]	Ø	g	ʒ

Some examples of the alternations are shown in (18); the data are from Lightner (1965).

(18) (a) *Degree marking on adjectives*

	POSITIVE	COMPARATIVE	
(i)	tʲix-ij	tʲiʃ-e	'quiet(er)'
(ii)	ʒark-ij	ʒarʧ-e	'hot(ter)'
(iii)	dorog-oj	doroʒ-e	'dear(er)'

---

[15] It is not crucial to our argument that these alternations be synchronically productive in the modern language.

[16] A similar analysis of the corresponding pattern in Serbian is proposed by Radišić (2009).

424

(b) *Number agreement on verbs*

         3PL.   3SG.

   (i) max-ut   maʃ-et   'wave(s), wag(s)'

  (ii) pek-ut   petʃ-et   'bake(s)'

 (iii) strig-ut striʒ-et   'shear(s)'

(c) *Deriving denominal adjectives*

         NOUN   ADJECTIVE

   (i) ʧerepax-a ʧerepaʃ-ij   'turtle'/'chelonian'

  (ii) volk   volʧ-ij   'wolf'/'lupine'

 (iii) vrag̊   vraʒ-ij   'enemy'/'hostile'

Other alternations, historically arising from the Second Velar Palatalization, pair /k/ with /ts/ and /g/ with /z/.[17] Again, these pairings are unremarkable given the specifications in Figure 4, but would be harder to account for in a unified way if /g/ and /z/ were specified for [±continuant]. In Halle's (1959) feature system, the Second Velar Palatalization alternations involve changing more features than the First Velar Palatalization ones do, but they can still readily be captured using the specifications in Figure 4. The alternating pairs are shown in (19) with their relevant features.

(19)

$$
\begin{bmatrix} +\text{compact} \\ +\text{low tonality} \end{bmatrix}
\qquad
\begin{bmatrix} -\text{compact} \\ -\text{low tonality} \\ +\text{strident} \end{bmatrix}
$$

[−voiced]   [−continuant]   k   ts

[+voiced]   ∅   g   z

Examples of these alternations, again from Lightner (1965), are shown in (20).

(20) (a) brjakatʲ   'let fall with a clang' brjatsatʲ   'clang'

      (b) voskliknutʲ 'exclaim' (PFV.)   vosklitsatʲ 'exclaim' (IPFV.)

      (c) tjagatʲsja   'sue'   sostjazatʲsja 'contend with'

---

[17] Additionally, /gʲ/, which Halle (1959) does not treat as phonemic because of its restricted distribution, alternates with /zʲ/ as part of the same pattern. Regardless of the phonemic status of /gʲ/, the low-ranking feature [±sharped] can distinguish [gʲ] from [g] and appropriately match them with palatalized and non-palatalized /z/.

In sum, reordering [±voiced] and [±continuant] in the *SPR* feature hierarchy improves the empirical status of both features: in the former case by extending [−voiced] to the 'unpaired' /ʦ, ʧ, x/, and in the latter case by leaving /z, zʲ, ʒ, g/ unspecified for [±continuant].[18]

4.4 *Consequences of the revised analysis*

This analysis suggests a picture of phonological levels that is different both from the post-Bloomfieldian (10) and from the one in (11) advocated by Halle (1959) and Chomsky & Halle (1968). This model is shown in (21). Underlying lexical representations correspond to the systematic phonemic level of Chomsky and Halle, but with the proviso that they are made up only of contrastive features. These representations serve as input to the contrastive phonology, which, subject to the Contrastivist Hypothesis, computes only contrastive features. The representations which we designate here between straight brackets | | are the output of the contrastive phonology.[19] These representations serve as the input to additional levels not governed by the Contrastivist Hypothesis, which are the domain of enhancement and other phonetic processes that add or modify contrastive as well as non-contrastive features.[20]

(21) Underlying lexical representations: /t/ /ʦ/

      RVA, other rules: ↓ ↓ *Contrastive phonology*

      Contrastive representations: |d| |ʣ|

      Non-contrastive phonetic processes: ↓ ↓ *Enhancement, etc.*

      Phonetic representations: [d] [ʣ]

The Halle–Chomsky arguments against the taxonomic phoneme, defined by a series of conditions as discussed in Chomsky (1964), still go through, but the duplication problem raised by regressive voicing assimilation disappears. This is because the rule applies one time to segments bearing a contrastive specification of [±voiced], whether the result is an already existing phoneme or a new allophone of a phoneme. While the phonemic/allophonic distinction does not mark out a special level, the difference between contrastive and non-contrastive phonology does. Unlike the taxonomic phoneme, however, there are no constraints against over-lapping or any other constraints on the relation between this level and phonetic

---

[18] There is one further surprising consequence of the reordering that we reserve to the end.

[19] Although all the features at this level are ones that are underlyingly contrastive on the segments on which they originate, they may appear in combinations not found in the underlying inventory (as in Russian |ʣ|). This level of representation is thus not necessarily structure-preserving in the sense of Kiparsky (1982, 1985).

[20] See Section 5.4 for discussion of how phonetic enhancement interacts with the output of the contrastive phonology.

representations. Therefore, all the arguments against the taxonomic phoneme would have still been the same.

What is different is that between underlying lexical representations (now simply called phonemic) and the surface phonetic level (which has never been well defined) there is an intermediate level. This level is not the old taxonomic phoneme, but is the output of the contrastive phonology, where only contrastive features are in play, and the input to enhancement and other processes that add non-contrastive information (Dyck 1995; Hall 2011).[21]

5. THE ROAD AHEAD

Although the contrastive feature hierarchy was abandoned in *SPE*, it has been taken up again in recent work in generative phonology, which shows it to be a useful theoretical tool when features are ordered according to the Activity Principle and the role of redundant features is limited. In this section, we review some of the directions taken by this new line of research.

5.1 *Cross-linguistic variation*

Because the ordering of features is not universal, different languages with similar inventories, or different stages of the same language, may assign different featural representations to phonetically similar phonemes. Here we will briefly review two such cases: laryngeal harmony in two Chadic languages, and labial harmony in a Tungusic and a Mongolic language. In each case there are subtle differences as to which segments participate in the relevant harmonies that can be accounted for by positing a difference in the feature hierarchies of segment inventories that are very similar in the relevant respects.

5.1.1 *Laryngeal harmony in Chadic*

Mackenzie (2009, 2013) discusses the case of laryngeal harmony in Ngizim and Hausa, two Chadic languages. Each language has a three-way phonemic laryngeal contrast among obstruent stops, distinguishing plain voiced stops, plain voiceless stops, and implosives. Using the features [±voice] and [±constricted glottis], there are two possible contrastive hierarchies that can be applied to such an inventory. If [±constricted glottis] takes wider scope, as in (22a), then the implosives will be unspecified for [±voice]; if [±voice] takes wider scope, as in (22b), then the voiceless stops will have no specification for [±constricted glottis].

---

[21] This level might remind some readers of the division between lexical and post-lexical phonology, as it is commonly understood in the theory of Lexical Phonology and Morphology (Kiparsky 1982, 1985; Kaisse & Shaw 1985; Mohanan 1986). The equivalence may not be exact, however.

427

(22)   *Two hierarchical orderings of laryngeal features* (Mackenzie 2013: 300)
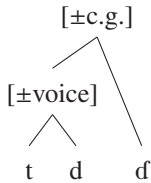
(a)   Ngizim   (b)   Hausa
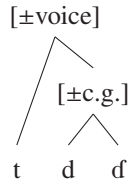
Mackenzie (2013) argues that (22a) is the correct hierarchy for Ngizim, which requires pulmonic obstruents to agree in voicing, but allows implosives to co-occur with either voiced or voiceless obstruents, as illustrated in (23):

(23)   *Ngizim voicing harmony* (Mackenzie 2013: 301, citing Schuh 1997)

(a) Harmonic forms with pulmonic consonants

(i)	gâːza	'chicken'	*k… z
(ii)	dɔ́ɓâ	'woven tray'	*t… b
(iii)	zɘ̀dú	'six'	*s… d
(iv)	kútɔ́r	'tail'	
(v)	tásáu	'find'	

(b) Implosives with voiced or voiceless pulmonic consonants

(i)	kìːɗú	'eat (meat)'
(ii)	fɔ́ɗû	'four'
(iii)	pɔ́ɗɔ́k	'morning'
(iv)	dɘ̀ɓú	'give water'

In Hausa, on the other hand, the hierarchy in (22b) appears to be the correct one. Hausa disallows a mismatch in [±constricted glottis] on homorganic voiced obstruents, as in (24a), but a voiceless pulmonic consonant can occur with an implosive, as in (24b).

(24)   *Hausa [±constricted glottis] harmony* (Mackenzie 2013: 302, citing Newman 2000)

(a)   ɗaɗa   'to strike a blow'   *ɗada

(b)   ɗata   'a small, bitter, green tomato'

As Mackenzie shows, the two different patterns of underspecification in (22) predict exactly the two different patterns of harmony and transparency found in Ngizim

428

and Hausa. In Ngizim, with the hierarchy in (22a), implosives are outside the scope of the harmonizing feature [±voice], while in Hausa voiceless stops are outside the scope of harmonizing [±constricted glottis].

5.1.2 *Labial harmony in Tungusic and Mongolic*

The Ngizim and Hausa obstruent stop systems are a kind of 'minimal pair': they have very similar phonemes and make use of similar features, but differ in how they pattern. In this case, the difference can be attributed to a difference in how the features are ordered. Van der Hulst & Smith (1988) called attention to another such 'minimal pair': Tungusic and Mongolic languages have similar vowel inventories and similar types of labial (rounding) harmony; however, in Tungusic /i/ blocks labial harmony, and in Mongolic /i/ is transparent to labial harmony. Ko (2013, 2018) proposes that this difference is also due to feature ordering. For Tungusic, here exemplified by Oroqen, he adopts the ordering of Zhang (1996), [low] ≫ [front] ≫ [round] ≫ [RTR], as in (25).[22] He proposes that Mongolic, represented by Khalkha, uses the same features, but ordered [front] ≫[low], as in (26).
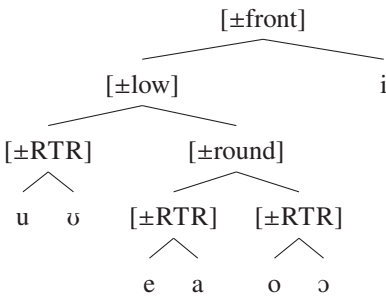
(25)   *Oroqen vowels* (Zhang 1996)

   [±low] ≫ [±front] ≫ [±round] ≫ [±RTR]
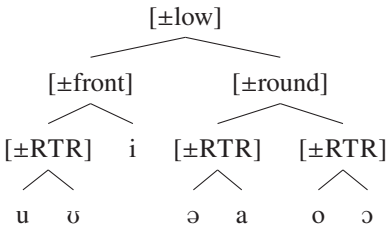
(26)   *Khalkha vowels* (Ko 2011)

   [±front] ≫ [±low] ≫ [±round] ≫ [±RTR]

---

[22] Solely for consistency with the features used in the rest of this section, we have changed Zhang and Ko's [coronal] to [front] and [labial] to [round]. We consider these to be functionally equivalent for present purposes, but make no theoretical claim about their status in general.

On this analysis, the key difference between Oroqen and Khalka /i/ is that Oroqen /i/ is specified [−low, +front] and Khalkha /i/ is specified as only [+front]. Ko (2013) observes that labial harmony in Tungusic and Mongolic is stratified by height, operating only in the domain of vowels that are [+low]. It is thus reasonable to suppose that adjacency in labial harmony is relativized to vowels that bear a feature for [±low]. Thus, labial harmony in both Oroqen and Khalkha is blocked by the [−low] specification of /ʊ/ and /u/. In Oroqen, it is also blocked by the [−low] specification of /i/; an example is shown in (27). Khalkha /i/ lacks a [±low] feature, and is thus transparent to labial harmony (28).

(27) *Labial harmony in Oroqen* (Tungusic) (Zhang 1995, 1996)

    (a)   bajta-wa   'affair DEF.OBJ'

    (b)   ɔlɔ-wɔ   'fish DEF.OBJ'   (*-wa)

    (c)   tɔrɔki-wa   'wild boar DEF.OBJ'   (*-wɔ)

(28) *Labial harmony in Khalkha* (Mongolic) (Svantesson 1985; Svantesson et al. 2005; Ko 2011)

    (a)   ʊd-aas   'willow-ABLATIVE'

    (b)   ɔd-ɔɔs   'star-ABLATIVE'   (*-aas)

    (c)   mɔrin-ɔɔs 'horse-ABLATIVE'   (*-aas)

5.2 Diachronic change

The same kind of variation in feature ordering can also be seen in diachrony, as different stages of the same language may have different contrastive hierarchies. A growing literature has made use of contrastive hierarchies in diachronic studies (see Dresher, Harvey & Oxford 2018 and the references therein).

A recurring idea that has been applied in a number of studies is that diachronic mergers tend to involve segments that have similar contrastive specifications. This idea has been formulated more specifically by Ko (2010, 2018) as the MINIMAL CONTRAST PRINCIPLE, given in (29a); Oxford (2015) reformulates it as the SISTERHOOD MERGER HYPOTHESIS (29b).

(29) *Diachronic phonological merger in contrastive hierarchy theory*
   (a)   The Minimal Contrast Principle (Ko 2010: 191; 2012: 35–37)
            Phonological merger operates on a minimal contrast – that is, on two segments that share a terminal branching node under a given contrastive hierarchy.
    (b)   The Sisterhood Merger Hypothesis (Oxford 2015: 314)
            Structural mergers apply to 'contrastive sisters'.

Here we will briefly review two such studies: Oxford's (2012, 2015) account of a contrast shift in the vowel systems of some Algonquian languages, and Krekoski's (2017) study of the evolution of Mandarin tone systems from Middle Chinese.

5.2.1 *Contrast shift in Algonquian vowel systems*

Based on what is known about the vowel system of Proto-Algonquian, Oxford (2012, 2015) posits that it had the contrastive feature hierarchy shown in (30).

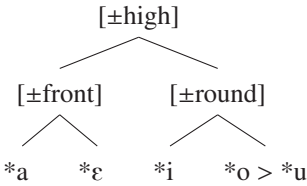(30)   *Contrastive hierarchy for Proto-Algonquian vowels*
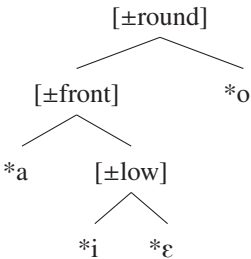
The hierarchy in (30) is motivated by feature activity that can be recovered as having been present in Proto-Algonquian. Thus, */o/ triggers rounding, an indication that it has an active, hence contrastive, [+round] feature. Similarly, */i/ triggers palatalization, indicating a contrastive feature we call [+front]. Patterns of partial neutralization relate */ɛ/ and */i/, suggesting that they are contrastive sisters by (29). Finally, */a/ does not trigger any processes, consistent with its being assigned no positive (marked) contrastive features.[23]

The Proto-Algonquian vowel feature hierarchy continues unchanged in the Central Algonquian languages and in Blackfoot (see Oxford 2015: 319 for details). It accounts for two recurring patterns: (a) palatalization always includes */i/ as a trigger; and (b) */ɛ/ regularly merges with */i/.

On the eastern and western edges of the Algonquian area, the vowels derived from Proto-Algonquian */o/ and */i/ began to pattern together as high vowels, contrary to the predictions of the Proto-Algonquian hierarchy (30), in which */o/ and */i/ share no vowel-specific features in common. Oxford posits that in these areas the feature hierarchy shifted: the feature [low] was reinterpreted as [high] and moved to the top of the vowel feature hierarchy, as shown in (31).

---

[23] See Oxford (2012, 2015) for detailed discussion of the evidence, and Dresher et al. (2018: 17–20) for a summary, from which this section is adapted.

(31)   *Contrastive hierarchy for eastern and western proto-languages*

```
                    [±high]
                   /       \
            [±front]         [±round]
             /    \           /     \
           *a     *ɛ        *i     *o > *u
```

Subsequent developments in the eastern and western daughter languages follow the predictions of the new hierarchy (see Oxford 2015: 319 for details). The patterns consistently differ from those of Central Algonquian: (a) palatalization in these languages is triggered by */ɛ/ but *excludes* */i/; and (b) */ɛ/ merges with or shifts to */a/ (not */i/).

Both of these patterns provide striking support for Oxford's analysis. In the hierarchy of (31), */ɛ/'s contrastive sister is */a/, which accounts for the new merger patterns. Perhaps more striking is the restriction of palatalization to */ɛ/, which is unexpected in light of the general cross-linguistic tendency for /i/ to be the preferred trigger of palatalization (Kochetov 2011), as is the case in Proto- and Central Algonquian. In those languages, both */i/ and */ɛ/ have the feature [+front] (or [+coronal]), which tends to be associated with vowels that trigger palatalization; */ɛ/ also has the feature [+low], which tends to inhibit its ability to trigger palatalization. However, in the shifted hierarchy of (31), only */ɛ/ is [+front], leaving it as the only potential trigger of palatalization. In sum, a single contrast shift accounts for the patterning of a large number of phonological changes across the Algonquian family.

5.2.2 *Contrastive diachronic stability and change in Mandarin tones*

Krekoski (2017) investigates the tone systems of a number of languages that descend from Middle Chinese, focusing on the patterns of activity that they display in the form of tone sandhi. He assumes that, where possible, tones related by a sandhi rule differ minimally, that is by only one feature (this is the synchronic counterpart of the diachronic principle in (29)). Paying attention to the contrastive relations among tones, rather than to their phonetics, reveals a surprising degree of stability in the contrastive structures of tone systems that descend from Middle Chinese, despite significant changes over time in the phonetic realizations of individual tones.

For example, Beijing Mandarin has the four contrastive tones in (32), which participate in two robust sandhi rules (33). Rule (33a), known as 'third tone sandhi', is obligatory, and states that a 214 tone becomes 35 before another 214 tone. This rule suggests that tone /214/ differs by one feature from tone /35/. The rule in (33b) is optional, and states that a lexical 35 tone becomes 55 when it immediately follows

432

either 35 or 55, and precedes any other tone. This rule suggests that /35/ differs by one feature from /55/.

(32) *Beijing tones*

    (a)   /55/   high level

    (b)   /35/   rising

    (c)   /214/   low concave

    (d)   /51/   high falling

(33) *Beijing tone sandhi*

    (a)   214 → 35 /__ 214

    (b)   35 → 55/ {35, 55}___ T   (T = any tone)

A tree satisfying these relations is given in (34). Krekoski (2017) then names the features ex post facto according to their phonetic correlates. The Roman numerals under each tone indicate the Middle Chinese tones from which the Beijing tones derive. (Middle Chinese tone I split into two tones designated Ia and Ib.)

(34) *Contrastive hierarchy for Beijing Mandarin*

$$[\pm\text{falling}]$$

$$[\pm\text{non-high}] \qquad [\pm\text{non-high}]$$

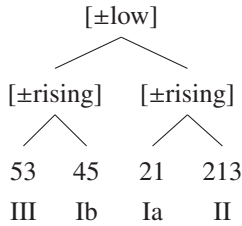| 55	35	51	214
Ia	Ib	III	II
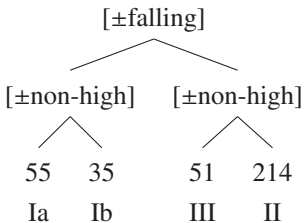
Following the same procedure for the Jin language Pingyao, Krekoski arrives at a hierarchy that is isomorphic to the one in (34), but whose features have different phonetic content. The tones in these languages have kept their places in the phonological system despite extensive changes to their phonetic realizations.

There is a limit, however, to the phonetic changes that the system of oppositions can withstand. Following the same methodology as for Beijing and Pingyao, Krekoski (2017) posits the tree in (35a) for the Mandarin language Tianjin. Surprisingly, these tones do not correspond as expected with their cognates in Beijing and Pingyao. Tones /21/ and /53/ are in the 'wrong place' relative to the other dialects that descend from Middle Chinese.

433

(35)   *Contrastive hierarchy for Tianjin*

    (a)   Tianjin (actual hierarchy)   (b) Mirroring Middle Chinese

```
          [±low]                              [±α]
         /      \                            /    \
   [±rising]  [±rising]                  [±β]      [±β]
    /   \      /   \                     /  \      /   \
  53    45   21   213                 21    45   53    213
  III   Ib   Ia    II                 Ia    Ib   III    II
```
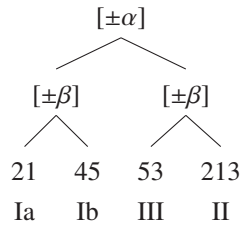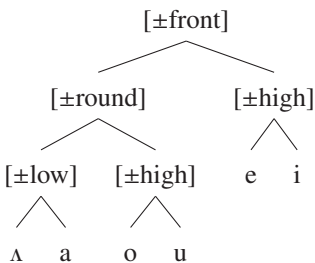
Why did a contrastive shift occur in the history of Tianjin? An answer can be found by considering what the contrastive feature tree of Tianjin would look like if the tones had maintained their relative positions from Middle Chinese (35b). Krekoski (2017: 66) observes that it is difficult to find plausible phonetic correlates for the features $[\pm\alpha]$ and $[\pm\beta]$ in this system, concluding that

> *Tonal drift likely accreted changes in height values until the system of contrasts reached some critical inflection point which precipitated the reanalysis of specifications. If one conceptualizes systemic change as occurring with subsequent generations of learners formulating new hypotheses about a phonological system from available data, a contrast shift such as this represents a new hypothesis of a system that is more phonetically natural and possesses a lower complexity than the preceding one.*

5.3 *Reduction and neutralization*

Contrastive hierarchies also offer insight into patterns of reduction (Spahr 2014; Nevins 2015). For example, Spahr (2014: §2.1) shows how the hierarchical organization of Bulgarian vowels in (36) identifies pairs neutralized in reduction.[24]

(36)   *Contrastive hierarchy for Bulgarian vowels* (based on Spahr 2014: 564)

```
              [±front]
             /        \
       [±round]      [±high]
        /    \         /  \
   [±low]  [±high]    e    i
    / \     / \
   ʌ   a   o   u
```

---

[24] We use /ʌ/ for the vowel that Spahr transcribes as /â/. Nevins (2015) independently proposes a similar analysis with a slightly different hierarchy (and does not include an underlying distinction between /a/ and /ʌ/).

In unstressed syllables, vowel reduction neutralizes contrasts at the lowest levels in the hierarchy. All dialects and registers reduce unstressed /ʌ/ and /a/ to [ə]; some combinations of dialect and register additionally reduce /o/ and /u/ to [ʊ]; and some also reduce /e/ and /i/ to [ɪ]. As Nevins (2015: 51) puts it, 'vowel reduction is a process of pruning the lowest node in the hierarchy'. Eliminating [±low] from /ʌ/ and /a/ produces a vowel [ə] that is [−front], [−round], and indeterminate in height. Similarly, pruning [±high] from /o/ and /u/ or /e/ and /i/ produces vowels that are no longer distinctly high or mid phonologically, and are intermediate in height between [o, e] and [u, i] phonetically, but which retain their other contrastive properties. Reducing all three pairs yields the truncated surface system in (37).[25]

(37) *Maximally reduced Bulgarian vowels*
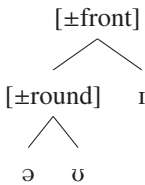
Nevins (2015) takes a similar approach to the reduction of vowel contrasts in whistled speech, and Rischel ([1966] 2009) uses a feature hierarchy to account for an otherwise puzzling reduction in distinctions connected with the Scandinavian runic reform.

This approach is consistent with the notion that reduction involves the elimination of formal structural complexity (not just articulatory effort), and offers an additional kind of evidence that can be brought to bear in identifying the feature hierarchy of any particular language. In doing so, it is important to keep in mind that phonetic neutralization does not always need to eliminate all surface evidence of the underlying featural contrasts to which it applies. For example, Compton & Dresher (2011) discuss Inuit varieties in which 'strong *i*' triggers palatalization and 'weak *i*', historically descended from an earlier schwa, does not. Although both vowels surface as [i] phonetically, the underlying distinction between them is seen in their different effects on adjacent consonants. This neutralization applies too late in the computation to bleed palatalization, whereas a reduction pruning the contrastive hierarchy earlier on would cause segments to behave alike rather than merely sound alike.

The pruned hierarchies in these cases of neutralization are epiphenomenal: the underlying segments are specified with all the features assigned to them by the full contrastive hierarchy for the language. Neutralizing the lowest features in the hierarchy produces a reduced surface inventory in certain contexts, but the underlying system of hierarchical contrasts remains constant. Another type of reduced

---

[25] As Iosad (to appear) points out, these pruned nodes are conceptually similar to Kuznetsov's (1941) 'hyperphonemes'.

a. Common three-vowel inventory   b. Unattested three-vowel inventory
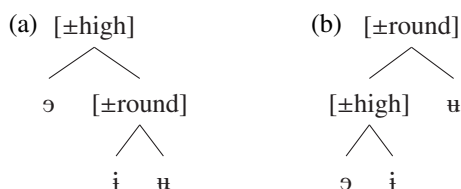
Figure 5
Three-vowel inventories.

inventory can be found in morphologically defined contexts, as in the case of many Romance languages, where inflectional suffixes (desinences) use a reduced inventory of vowels as compared with roots. Dyck (1995) and Frigeni (2002) have proposed analyses of such cases that involve separate contrastive hierarchies for the desinential domain. Here, different hierarchies are posited underlyingly, not just at the surface.

5.4 *Phonetic enhancement and the shapes of inventories*

Another benefit of the contrastive hierarchy approach, as Hall (2011) argues, is what it can contribute to our understanding of cross-linguistic patterns in inventories. It has long been known that phonemes tend to be spread out through the available phonetic space – for example, a vowel system is unlikely to have /y/ contrasting with either /i/ or /u/ unless it also has /i/ and /u/ in contrast with each other. From Liljencrants & Lindblom (1972) through more recent work such as Kirchner (1997), Flemming (2002), and Padgett (2003), Dispersion Theory has accounted for this pattern by positing mechanisms that explicitly enforce auditory distancing between sounds. What Hall (2011) shows is that the contrastive hierarchy, combined with the independently motivated process of phonetic enhancement (Stevens, Keyser & Kawasaki 1986 and subsequent work), generates dispersedness automatically, with no need for constraints that evaluate auditory distance.

For example, consider the vowel inventories in Figure 5. The widely attested three-vowel inventory in Figure 5a is spread out through the phonetic space. The unattested inventory in Figure 5b, on the other hand, is tightly packed into a small region of the available space, with very little acoustic difference between any two vowels.

If the vowels of Figure 5b are assigned contrastive feature specifications in a hierarchical ordering, the features can only mark how they differ; they cannot specify all the things these vowels have in common. The only two features that can be assigned here are [±high] and [±round] (or features with different names that mark the same distinctions, such as Jakobson et al.'s (1952) [compact/diffuse] and [flat/plain]). These may take scope in either order, producing either of the hierarchies in (38).

436

(38) *Ways of dividing Figure 5b*

(a) [±high]　　　　　　　(b) [±round]

　　ɘ　　[±round]　　　　　　[±high]　　ʉ

　　　　　ɨ　　ʉ　　　　　　　　ɘ　　ɨ

In either case, the features assigned to these vowels are ones that could equally represent the vowels of Figure 5a: /a/, /i/, and /u/. While it is possible to assign specifications to /a, i, u/ that could not correspond to /ɘ, ɨ, ʉ/, the reverse is not true.

Enhancement, in turn, means that a vowel inventory specified as in (38) is more likely to surface phonetically as [a, i, u] than as [ɘ, ɨ, ʉ]. As noted in Stevens et al. (1986) and subsequent work, phonetic properties that are not phonologically contrastive tend to be deployed in such a way as to reinforce the acoustic effects of contrastive features. For example, in (38) /ɘ/ is specified as [−high]. The elevated F1 frequency associated with [−high] can be enhanced by making the vowel [+low], so enhancement favours realizing /ɘ/ as [a]. Similarly, /ɨ/ is [−round], which causes a higher F2 that can be enhanced by realizing the vowel as front [i], while the lower F2 associated with the [+round] feature of /ʉ/ can be reinforced by making the vowel back [u].

The features of any individual segment can be enhanced without reference to the rest of the inventory. Because contrastive feature specifications encode differences, enhancing them makes the inventory as a whole more phonetically dispersed, with no need for constraints that explicitly evaluate acoustic distances between segments. As Hall (2011) shows, the combination of contrastive specification in the phonology and enhancement in the phonetics accounts for the typological patterns that Dispersion Theory was designed to explain.

6. A ROAD JUST AS FAIR?

> *Then took the other, as just as fair,*
> *And having perhaps the better claim,*
> *Because it was grassy and wanted wear;*
> *Though as for that the passing there*
> *Had worn them really about the same*

(Frost, 'The road not taken')

We do not want to leave the impression that generative phonology would have gone down the road that leads to the contrastive hierarchy theory if only Halle had ordered [±voiced] over [±continuant] in the tree in Figure 1. Unlike the choice

437

confronting Frost's traveller, the two roads we have identified would not have looked equally promising. In reality, a number of trends in pre- and early generative phonological theory were gaining momentum and conspiring against the notion of language-specific contrastive specification. In this section we briefly review some of these currents to show the wider context in which *SPR* was written and received.

6.1 *Universality*

The approach to contrastive feature hierarchies that we presented in Section 5 assumes that feature hierarchies are language particular. This was true of Trubetzkoy's (1939) theory, but through the 1950s and 1960s there was a movement, beginning with Jakobson (see Toman 1995: 119; Battistella to appear), to put as much as possible of feature theory on a universal basis, and *SPE* gave universalism a central place in generative phonology. By the time the notion of organizing features hierarchically returned with the introduction of autosegmental feature geometry, the hypothesis of universal features was well established, as reflected in works such as Clements (1985), Sagey (1990), Clements & Hume (1995), and Halle, Vaux & Wolfe (2000). In the context of this broader trend toward universalism, an attempt to entrench language-particular feature hierarchies might have had difficulty gaining traction in the first decades of generative phonology.

6.2 *Syntagmatic sources of underspecification in SPR*

The branching tree shown in Figure 1 is not the only source of unspecified feature values in *SPR*. All languages have constraints on sequences of phonemes, and these constraints give rise to additional feature redundancies that could serve as a further source of underspecification, particularly if Minimality is the guiding principle. Thus, the features left unspecified by the tree in Figure 1 (what we could call *paradigmatically* underspecified features) are supplemented in *SPR* by many more features that are lexically unspecified because of sequence structure constraints (*syntagmatically* underspecified features) that Halle (1959: 56–62) calls morpheme structure (MS) rules.

Whereas paradigmatic underspecification is entirely determined by the branching tree (that is, by the contrastive feature hierarchy), syntagmatic underspecification is much less well defined. As Halle (1959: 57) admits, 'It is not possible to give a *complete* procedure for discovering the most economical representation in every case. The best that can be done is to formulate the sequential constraints as rules specifying certain contexts' (emphasis in original). *SPR* lists a series of some 30 MS rules numbered 1a–11e. A sample of a few of these, shown in (39), illustrates how context-specific they are (in part because Halle made no recourse to syllable structure). In these rules, '&' represents a boundary, 'V' stands for any vowel, 'R' for a liquid, and 'C' for any consonant. These rules apply in a block prior to the phonological (P-) rules.

438

(39) (a) *Rule MS 1a*
           If the segment following the & marker is a glide, the next segment is
           vocalic and nonconsonantal, i.e. a vowel.
      (b) *Rule MS 3*
           If in an initial cluster a liquid is not in an initial position, the liquid is
           followed by a segment which is vocalic and nonconsonantal.
      (c) *Rule MS 10b*
           The sequence VCCR is followed by a vowel.
      (d) *Rule MS 11c*
           A consonant in position after a grave compact (velar) consonant is
           noncompact.

Morpheme structure conditions have been the subject of much discussion since
the writing of *SPR* (see Dresher & Nevins 2017 for a recent treatment); the issue
here is whether underspecification is the best way to account for them. The
arbitrariness and incompleteness of syntagmatic underspecification – unlike para-
digmatic underspecification, which is governed by an explicit and complete pro-
cedure (the branching tree) – suggests to us that it is not. However, the existence of
syntagmatic underspecification made it problematic to propose a model like the one
in (21), where non-contrastive features are filled in after the contrastive phonology.
Moreover, syntagmatic underspecification posed difficulties that made contrastive
underspecification more vulnerable to the criticisms of Lightner (1963) and partic-
ularly Stanley (1967), to which we now turn.

6.3 *Stanley and the critique of underspecification*

While Stanley (1967) devotes one section to a critique of 'branching diagrams'
(contrastive feature hierarchies) and another to the Distinctness Condition, much of
his discussion is aimed at Halle's (1959) morpheme structure rules, which he finds
(Stanley 1967: 435) have 'a certain kind of arbitrariness' that cannot be remedied.
With respect to the branching tree and paradigmatic underspecification, Stanley
makes a number of technical arguments that we cannot review here, but which are
discussed in detail by Dresher (2009: 96–102). Here we would like to highlight one
fundamental assumption that colours and vitiates Stanley's entire discussion of
underspecification, and this is his understanding of what an underspecified feature
(what he calls a 'blank' or '0' entry in a feature matrix) represents.

As should be clear from our discussion of the connection between contrast and
activity and from every case reviewed in Section 5, we understand an underspecified
feature to be phonologically different from both the + and − values of that feature,
irrespective of what its phonetic interpretation turns out to be. For example, the implosive
/ɗ/ in Ngizim is unspecified for [±voice], and is treated by the phonology as not having a
voiced feature; hence, it is exempt from the co-occurrence restriction that affects /t, d/.
The phonetically similar /ɗ/ in Hausa is specified differently and consequently cannot

439

co-occur with /d/. That is, contrastive underspecification has phonological significance because the logic of contrastive underspecification is a ternary logic.

This is not how Stanley understands underspecification, however. For him, a '0' in a feature matrix is merely an unexpressed + or −, as the case may be: 'we must take care that the feature values remain BINARY, and that "0" in a dictionary matrix is never allowed to function as a third value, distinct from "+" and "−"' (Stanley 1967: 410; emphasis in original). Therefore, the phonology must not be allowed to take advantage of a 0 value to achieve a simplification or generalization of a rule. For example, he would consider it improper to describe Ngizim as disallowing segments with differing values of [±voice] to co-occur in a word without further stipulating that laryngeal harmony is restricted to segments specified [−constricted glottis]; this would be 'a specious simplification, an improper use of blanks' (Stanley 1967: 416).

As Dresher (2009: 98) has remarked, 'What Stanley considered specious was later considered a result!' That Hausa implosives are unspecified for [±voice] does not just allow us to simplify the statement of voicing harmony, it *explains* why they do not participate in it. More generally, the Contrastivist Hypothesis does not make sense unless we assume that unspecified values have phonological significance and cannot be manipulated by the phonology. Nevertheless, Stanley's interpretation of blanks appears to have been widely shared, and was not contested until Kiparsky (1982) reintroduced underspecification into generative phonological theory (Dresher 2009: 117–121).

7. CONCLUSIONS

To sum up, if Halle had interpreted contrastively unspecified features as being phonologically significant, and if he had more consistently favoured Activity over Minimality as the principle governing feature ordering in Russian, his analysis of Russian would have been different in a number of respects. The feature [±voiced] would be ordered above [±continuant], and the 'unpaired segments' would be contrastively specified as [−voiced]. The famous argument against the taxonomic phoneme would still have shown that there is no distinct level between the underlying and surface representations that divides between rules that change phonemes into other phonemes and rules that turn phonemes into allophones. However, it would have introduced a new intermediate level that marks the dividing line between the contrastive and non-contrastive phonology.

Thus, the subsequent development of generative phonology might have taken a very different road. The connection between contrast and phonological activity would be maintained, and contrastive feature hierarchies (branching trees) would remain the way to generate contrastive representations. In sum, on this path, language-particular contrast remains an important means of accounting for phonological patterning. And, as the more recent work surveyed in Section 5 demonstrates, this road is still open and worth pursuing.

And here is a surprise: Recall that Halle's (1959: 44–45) report that his analysis of Russian requires 271 feature specifications appeared to be off by 1, and that he

440

actually has 272 specifications. It turns out that when we reorder [±voiced] and [±continuant] as we suggested, without changing anything else in Halle's (1959) tree, the number of feature specifications comes to precisely 271. Is this a coincidence? We cannot know. However, it turns out that ordering [±voiced] over [±continuant], the road not taken, would have been preferred even by Minimality.

REFERENCES

Anderson, Stephen R. 2000. Reflections on 'On the phonetic rules of Russian'. *Folia Linguistica* 34.1–2, 11–27.

Archangeli, Diana. 1984. *Underspecification in Yawelmani phonology and morphology*. Dissertation, Massachusetts Institute of Technology.

Archangeli, Diana. 1985. Yokuts harmony: Evidence for coplanar representations in nonlinear phonology. *Linguistic Inquiry* 16.3, 335–372.

Archangeli, Diana. 1988. Underspecification in phonology. *Phonology* 5.2, 183–207.

Battistella, Edwin. to appear. The Prague School: Nikolai Trubetzkoy, Roman Jakobson. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*. Oxford: Oxford University Press.

Bloch, Bernard. 1941. Phonemic overlapping. *American Speech* 16.4, 278–284.

Broselow, Ellen. 1984. Default consonants in Amharic morphology. *MIT Working Papers in Linguistics* 7, 15–32.

Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.

Campos Astorkiza, Judit Rebeka. 2007. *Minimal contrast and the phonology–phonetics interaction*. Dissertation, University of Southern California.

Cherry, E. Colin, Morris Halle & Roman Jakobson. 1953. Toward the logical description of languages in their phonemic aspect. *Language* 29.1, 34–46.

Chew, Peter. 1999. *A computational phonology of Russian*. Dissertation, University of Oxford.

Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, Noam & Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1.2, 97–138.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 225–252.

Clements, George N. 1987. Towards a substantive theory of feature specification. In James Blevins & Julie Carter (eds.), *NELS 18: Proceedings of the eighteenth annual meeting of the North Eastern Linguistic Society*, 79–93. Amherst, MA: GLSA.

Clements, George N. 2001. Representational economy in constraint-based phonology. In T. Alan Hall (ed.), *Distinctive feature theory*, 71–146. Berlin: Mouton de Gruyter.

Clements, George N. & Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Oxford: Blackwell.

Compton, Richard & B. Elan Dresher. 2011. Palatalization and 'strong *i*' across Inuit dialects. *Canadian Journal of Linguistics* 56.2, 203–228.

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. 2016. Contrast in phonology, 1867–1967: History and development. *Annual Review of Linguistics* 2, 53–73.

Dresher, B. Elan & Daniel Currie Hall. to appear. Developments leading toward generative phonology. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*. Oxford: Oxford University Press.

Dresher, B. Elan, Christopher Harvey & Will Oxford. 2018. Contrastive feature hierarchies as a new lens on typology. In Larry Hyman & Frans Plank (eds.), *Phonological typology*, 273–311. Berlin: De Gruyter Mouton.

Dresher, B. Elan & Andrew Nevins. 2017. Conditions on iterative rounding harmony in Oroqen. *Transactions of the Philological Society* 115.3, 365–394.

Dyck, Carrie. 1995. *Constraining the phonology–phonetics interface, with exemplification from Spanish and Italian dialects*. Dissertation, University of Toronto.

Flemming, Edward. 2002. *Auditory representations in phonology*. New York: Routledge.

Frigeni, Chiara. 2002. Metaphony in Campidanian Sardinian: A domain-based analysis. *Toronto Working Papers in Linguistics* (Special issue on *Contrast in Phonology*) 20, 63–91.

Hall, Daniel Currie. 2007. *The role and representation of contrast in phonological theory*. Dissertation, University of Toronto.

Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28.1, 1–54.

Hall, Daniel Currie. 2017. Contrastive specification in phonology. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. doi: 10.1093/acrefore/9780199384655.013.26.

Hall, Daniel Currie & B. Elan Dresher. 2016. Trade-offs in the contrastive hierarchy: Voicing versus continuancy in Slavic. In Christopher Hammerly & 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.

Halle, Morris. 1957. *On the phonetic rules of Russian*. Presented to the Linguistic Society of America, Chicago.

Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton.

Halle, Morris, Bert Vaux & Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31.3, 387–444.

Hargus, Sharon. 1989. Underspecification and derived-only rules in Sekani phonology. In Donna B. Gerdts & Karin Michelson (eds.), *Theoretical perspectives on Native American languages*, 70–103. Buffalo, NY: SUNY Press.

Harms, Robert T. 1968. *Introduction to phonological theory*. Englewood Cliffs, NJ: Prentice-Hall.

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.

Iosad, Pavel. 2012. Vowel reduction in Russian: No phonetics in phonology. *Journal of Linguistics* 48.3, 521–571.

Iosad, Pavel. to appear. Phonology in the Soviet Union. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*. Oxford: Oxford University Press.

Jakobson, Roman. [1931] 1962. Phonemic notes on Standard Slovak. In *Selected writings I: Phonological studies*, 221–230. The Hague: Mouton.

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze* (Språkvetenskapliga Sällskapets i Uppsala Förhandlingar). Uppsala: Uppsala Universitet.

Jakobson, Roman, C. Gunnar M. Fant & Morris Halle. 1952. Preliminaries to speech analysis: The distinctive features and their correlates. Technical report 13, MIT Acoustics Laboratory.

Jakobson, Roman & Morris Halle. 1956. *Fundamentals of language* (Janua Linguarum 1). The Hague: Mouton.

Jakobson, Roman & John Lotz. 1949. Notes on the French phonemic pattern. *Word* 5, 151–158.

Joos, Martin (ed.). 1957. *Readings in linguistics I*. Chicago: University of Chicago Press.

Kaisse, Ellen M. & Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook* 2, 1–30.

Kasevich, Vadim Borisovich. 1983. *Fonologicheskie problemȳ obshchego i vostochnogo yazȳkoznaniya* [Phonological problems in general and Oriental linguistics]. Moscow: Nauka & Shkola 'JaRK'.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm: Selected papers from SICOL-1981*, 3–91. Seoul: Hanshin.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2, 85–138.

Kirchner, Robert Martin. 1997. Contrastiveness and faithfulness. *Phonology* 14.1, 83–111.

Ko, Seongyeon. 2010. A contrastivist view on the evolution of the Korean vowel system. In Hiroki Maezawa & Azusa Yokogoshi (eds.), *Proceedings of the Sixth Workshop on Altaic Formal Linguistics (WAFL6)* (MIT Working Papers in Linguistics 61), 181–196. Cambridge, MA: MITWPL.

Ko, Seongyeon. 2011. Vowel contrast and vowel harmony shift in the Mongolic languages. In A. Simpson (ed.), *Proceedings of the seventh Workshop on Altaic Formal Linguistics (WAFL7)*, vol. 76, 187–202. Cambridge, MA: MITWPL.

Ko, Seongyeon. 2013. A contrastive hierarchy approach to Tungusic and Mongolic labial harmony. In Chundra Cathcart, I-Hsuan Chen, Greg Finley, Shinae Kang, Clare S. Sandy & Elise Stickles (eds.),

Proceedings of the 37th annual meeting of the Berkeley Linguistics Society (BLS 37), 185–198. Published online via eLanguage, https://escholarship.org/uc/item/47j9w9k5.

Ko, Seongyeon. 2018. *Tongue root harmony and vowel contrast in Northeast Asian languages* (Turcologica 112). Wiesbaden: Otto Harrassowitz Verlag.

Kochetov, Alexei. 2011. Palatalization. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 3, 1666–1690. Malden, MA: Wiley Blackwell. doi:10.1002/9781444335262.wbctp0071.

Krekoski, Ross. 2017. *Contrast and complexity in Chinese tonal systems*. Dissertation, University of Toronto.

Kuznetsov, Petr Savvich. 1941. K voprosu o fonematicheskoĭ sisteme sovremennogo frantsuzskogo yazȳka. *Kafedra russkogo yazȳka* 5.1, 140–174.

Ladd, D. Robert. 2014. *Simultaneous structure in phonology*. Oxford: Oxford University Press.

Ladd, D. Robert. to appear. Mid-century American phonology: The post-Bloomfieldians. In B. Elan Dresher & Harry van der Hulst (eds.), *The Oxford history of phonology*. Oxford: Oxford University Press.

Lightner, Theodore McGraw. 1963. A note on the formation of phonological rules. In *Quarterly progress report* 68, 187–189. Cambridge, MA: Research Laboratory of Electronics, Massachusetts Institute of Technology.

Lightner, Theodore McGraw. 1965. *Segmental phonology of Modern Standard Russian*. Dissertation, Massachusetts Institute of Technology.

Liljencrants, Johan & Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48.4, 839–862.

Mackenzie, Sara. 2009. *Contrast and similarity in consonant harmony processes*. Dissertation. University of Toronto.

Mackenzie, Sara. 2013. Laryngeal co-occurrence restrictions in Aymara: Contrastive representations and constraint interaction. *Phonology* 30.2, 297–345.

Mohanan, Karuvannur P. 1986. *Lexical Phonology*. Dordrecht: Reidel.

Nevins, Andrew. 2010. *Locality in vowel harmony*. Cambridge, MA: MIT Press.

Nevins, Andrew. 2015. Triumphs and limits of the contrastivity-only hypothesis. *Linguistic Variation* 15.1, 41–68.

Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. New Haven, CT: Yale University Press.

Oxford, Will. 2012. 'Contrast shift' in the Algonquian languages. *McGill Working Papers in Linguistics* 22.1. Published online at https://mcgill.ca/mcgwpl/files/mcgwpl/oxford2012_0.pdf.

Oxford, Will. 2015. Patterns of contrast in phonological change: Evidence from Algonquian vowel systems. *Language* 91.2, 308–357.

Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory* 21.1, 39–87.

Postal, Paul M. 1964. Boas and the development of phonology: Comments based on Iroquoian. *International Journal of American Linguistics* 30.3, 269–280.

Postal, Paul M. 1968. *Aspects of phonological theory*. New York: Harper and Row.

Pulleyblank, Douglas. 1986. Underspecification and low vowel harmony in Okpe. *Studies in African Linguistics* 17.2, 119–153.

Radišić, Milica. 2009. The double nature of the velar /g/ in Serbian. *Toronto Working Papers in Linguistics* 30, 91–103.

Reiss, Charles. 2017. Contrast is irrelevant in phonology: A simple account of Russian /v/ as /V/. In Bridget D. Samuels (ed.), *Beyond markedness in formal phonology*, 23–45. Amsterdam: John Benjamins.

Rischel, Jørgen. [1966] 2009. Phoneme, grapheme, and the 'importance' of distinctions: Functional aspects of the Scandinavian runic reform. In Nina Grønnum, Frans Gregersen & Hans Basbøll (eds.), *Sound structure in language*, 254–271. Oxford: Oxford University Press.

Sagey, Elizabeth C. 1990. *The representation of features and relations in nonlinear phonology*. New York: Garland.

Schuh, Russell G. 1997. Changes in obstruent voicing in Bade/Ngizim. Ms., University of California, Los Angeles.

Spahr, Christopher. 2014. A contrastive hierarchical account of positional neutralization. *The Linguistic Review* 31.3–4, 551–585.

Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43.2, 393–436.

443

Steriade, Donca. 1987. Redundant values. In *Papers from the 23rd annual regional meeting of the Chicago Linguistic Society*, vol. 2, 339–362. Chicago: Chicago Linguistic Society.

Stevens, Kenneth Noble, Samuel Jay Keyser & Haruko Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. In Joseph S. Perkell & Dennis H. Klatt (eds.), *Invariance and variability in speech processes*, 426–449. Hillsdale, NJ: Lawrence Erlbaum Associates.

Svantesson, Jan-Olof. 1985. Vowel harmony shift in Mongolian. *Lingua* 67.4, 283–327.

Svantesson, Jan-Olof, A. Tsendina, A. Karlsson & V. Franzén. 2005. *The phonology of Mongolian*. Oxford: Oxford University Press.

Toman, Jindřich. 1995. *The magic of a common language: Jakobson, Mathesius, Trubetzkoy, and the Prague Linguistic Circle*. Cambridge, MA: MIT Press.

Trubetzkoy, Nikolai S. 1929. Zur allgemeinen Theorie der phonologischen Vokalsysteme. *Travaux du cercle linguistique de Prague* 1, 36–67.

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).

Van der Hulst, Harry & Norval Smith. 1988. Tungusic and Mongolian vowel harmony: A minimal pair. In Peter Coopmans & Aafke Hulk (eds.), *Linguistics in the Netherlands 1988*, 79–88. Dordrecht: Foris.

Zhang, Xi. 1995. Vowel harmony in Oroqen (Tungus). *Toronto Working Papers in Linguistics* 14, 161–174.

Zhang, Xi. 1996. *Vowel systems of the Manchu-Tungus languages of China*. Dissertation, University of Toronto.

Authors' addresses: (Dresher)
University of Toronto, Department of Linguistics, 100 St. George Street, room 4073, Toronto, Ontario M5S 3G3, Canada
dresher@chass.utoronto.ca

(Hall)
Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia B3H 3C3, Canada
daniel.hall@utoronto.ca