Contrastive Feature Hierarchies in Phonology: Variation and Universality

B. Elan Dresher, University of Toronto

As a way of addressing the Georgetown University Round Table on Languages and Linguistics (GURT) 2017 conference theme, “Variable Properties: Their Nature and Acquisition,” I would like to ask the question: What is variable and what is fixed in phonology? In particular, I want to focus on phonological representations, and on the nature of features: are features innate and universal, or are they ‘emergent’ and language particular? The assumption that features are innate does not seem to leave enough room for the variability that we find; but the assumption that they are emergent could leave us with too much variation, with no account of why phonologies resemble each other as much as they do.*

Modifying a line of thought that can be traced back to Roman Jakobson, I propose that it is the concept of a contrastive feature hierarchy that is universal, not the features themselves or their ordering. I further adopt the Contrastivist Hypothesis, which holds that only contrastive features can be computed by the phonology. This hypothesis makes a connection between contrast and phonological activity that has implications for phonological theory as well as for language acquisition: it follows from this theory that learners are guided by phonological activity as well as by phonetics in acquiring the feature hierarchy for their language. I will argue that these principles suffice to account for many of the ways that phonological systems resemble each other. I will show how contrastive feature hierarchies contribute to accounts of synchronic and diachronic phonology, allowing for considerable variation, but governed by a uniform universal template.

I will begin by reviewing the groundbreaking contributions of Jakobson (1941). Building on Jakobson’s ideas about the emergence of phonological oppositions, I will then present a theory of phonological contrast. Next, I briefly rehearse the arguments against innate phonological features, and propose some prerequisites for a theory of emergent features. Finally, An example from Inuit vowel systems illustrates how contrastive feature hierarchies contribute to explanatory accounts of phonological patterning.

1. Jakobson’s Kindersprache: A reconsideration

Roman Jakobson’s Kindersprache, Aphasie und allgemeine Lautgesetze (Jakobson1941), translated into English as Child language, aphasia and phonological universals (Jakobson 1968), is important for its theory of phonological acquisition, as well as for how it connects acquisition to phonological theory more generally. Of the many influential ideas advanced in this book, the one that has attracted much discussion and criticism is the claim that acquisition proceeds in a fixed order. Jakobson does indeed emphasize this idea throughout the book. For example, he writes (Jakobson 1968: 20–28), “The fact that a fixed order must be inherent in language acquisition, and in phonological acquisition in particular, has repeatedly been noticed by observers … Again and again a number of constant features in the succession of acquired phonemes are observed…”

* This paper is a slightly revised version of portions of a talk presented at GURT 2017. I would like to thank participants and audience members for their comments and questions. For discussion, ideas, and analyses, I would like to thank Graziela Bohn, Elizabeth Cowper, Daniel Currie Hall, Paula Fikkert, Ross Godfrey, Christopher Harvey, Ross Krekoski, Will Oxford, Keren Rice, Christopher Spahr, and Zhang Xi.
In passages such as the above, Jakobson appears to be claiming that the fixed order of emergence refers to *phonemes*; for example, he writes that the acquisition of vowels is launched with a wide (low) vowel, *a*, and that the first consonant is generally a labial stop, *p* (hence, the first syllable is expected to be *pa*). In other places, however, he refers to the emergence of *oppositions*, that is *contrasts*, not individual phonemes. Thus, he proposes that the first vocalic opposition opposes the wide vowel, *a*, to a more narrow (high) vowel, *i*.

If the key notion, however, is contrasts, then the predictions about the order of emergence of individual sounds become much more obscure. This is because a contrast between a wider and a narrower vowel can be phonetically realized in a variety of ways: the phonemic labels ‘*a*’ and ‘*i*’ can each represent a broad range of phonetic vowels. Also, the boundary between two such phonemes can vary considerably from language to language. Hence the apocryphal tale recounted by Hyman (2008), in which Jakobson asserts in a lecture that in all languages the child’s first word is *pa*. An audience member objects that *his* child’s first utterance was *tʃɪk*. Jakobson replies, “phonetic [tʃɪk], yes, but phonologically /pa/!” This may be a joke, but there is truth to the notion that an emphasis on contrasts can overshadow the individual sounds that participate in a contrast.

This makes it harder than one might suppose to test Jakobson’s predictions about a fixed order of acquisition (Ingram 1988). Nevertheless, it appears that child phonology shows more variation, even within a single language, than Jakobson 1941 allows (Menn & Vihman 2011; Bohn 2017). But the claim that acquisition of phonology proceeds in a fixed order is not the only idea put forward in *Kindersprache*. More consequential, in my view, is the notion that contrasts are crucial and that they develop in a *hierarchical order*.

In particular, Jakobson proposes that learners begin with broad contrasts that are split by stages into progressively finer ones. He observes (1968: 65), “This system is by its very nature closely related to those stratified phenomena which modern psychology uncovers in the different areas of the realm of the mind. Development proceeds ‘from an undifferentiated original condition to a greater and greater differentiation and separation’” (citing E. Jaensch, *Zeitschr. f. Psychol.* 1928).

With this basic idea in mind, consider again the acquisition of vowel systems set out in Jakobson 1941 and its sequel, Jakobson & Halle 1956. At the first stage, there is only a single vowel. As there are no contrasts, we can simply designate it /V/ (1a). Jakobson & Halle write that this lone vowel is the maximally open vowel [a], the ‘optimal vowel’. But we do not need to be that specific: we can understand this to be a default value, or a typical but not obligatory instantiation. For contrastive purposes, any phonetic vowel will fit; for example, [ɪ]

(1) Early stages of vowel acquisition (Jakobson 1941; Jakobson & Halle 1956)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Stage 1</td>
<td>b. Stage 2</td>
<td>c. Stage 3</td>
</tr>
<tr>
<td>vowel</td>
<td>vowel</td>
<td>vowel</td>
</tr>
<tr>
<td>/V/</td>
<td>narrow /I/</td>
<td>narrow /i/</td>
</tr>
<tr>
<td></td>
<td>wide /A/</td>
<td>palatal /ɪ/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>velar /U/</td>
</tr>
</tbody>
</table>
In the next stage (1b), it is proposed that the single vowel splits into a narrow vowel, which is typically [i], and a wide vowel, typically [a]. I will continue to understand these values as defaults; I use capital letters to represent vowels that fit the contrastive labels that characterize them. Subsequently (1c), the narrow vowel splits into a palatal (front) vowel and a velar (back or round) vowel, typically [ʊ]. Jakobson (1968: 49) observes that this stage corresponds to the common three-vowel system /i, a, u/.

Of course, systems designated as /i, a, u/ vary considerably in their phonetic realizations. Dresher & Rice (2015) survey some three-vowel systems that are included in the online phonological database called PHOIBLE (Moran et al. 2014). It lists twelve Pama-Nyungan (Australia) three-vowel languages. Of these, eight are given as having the vowels /i, a, u/. The other four are listed as having different inventories: /i, a, u/, /ɪ, a, ʊ/, /ɪ, ɐ, ʊ/, and /ɪ, a, ə/. We found that there are no principled criteria for distinguishing between these systems: distinctions between /i/ ~ /ɪ/, /a/ ~ /ɑ/ ~ /ɐ/, and /u/ ~ /ʊ/ ~ /ə/ do not necessarily indicate significant differences between the languages. Conversely, the inventories designated /i, a, u/ exhibit considerable variation in the phonetic ranges covered by their three vowels.

Compare, for example, the vowel systems of two dialects of the Western Desert Language of central Australia: Pitjantjatjara (Figure 1, from Tabain & Butcher 2014) and Antakarinya (Figure 2, from Douglas 1955). The distributions of the vowels in the two languages are different, particularly that of the low vowel. They suggest that the languages may have different contrastive features, derived from different contrastive splits.

![Figure 1: Pitjantjatjara vowel ranges (Tabain & Butcher 2014: 194)](image)

The phonetic distribution of the vowels in Antakarinya is consistent with Jakobson & Halle's Stage 3 in (1c), in having a basic split between the low vowel and the other two. Updating the terminology, we can represent the contrasts as in (2a); I assume, for purposes of this example,
that the features [high] and [round] are the positive (marked) features, and their negative poles are defaults (on markedness see further §2).¹

Figure 2: Antakarinya vowel ranges (based on Douglas 1955: 221)

(2) Other types of contrasts in three-vowel systems

- **Antakarinya**

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>[high]</td>
</tr>
<tr>
<td>[ɪ]</td>
<td>[round]</td>
</tr>
<tr>
<td>[ʊ̣]</td>
<td>[back]</td>
</tr>
</tbody>
</table>

- **Pitjantjatjara**

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ɨ]</td>
<td>[high]</td>
</tr>
<tr>
<td>[ɪ]</td>
<td>[round]</td>
</tr>
<tr>
<td>[ʊ]</td>
<td>[back]</td>
</tr>
</tbody>
</table>

- **Western Arrarnta**

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a]</td>
<td>[low]</td>
</tr>
<tr>
<td>[ʌ̣]</td>
<td>[round]</td>
</tr>
<tr>
<td>[ʌ]</td>
<td>[front]</td>
</tr>
</tbody>
</table>

In Pitjantjatjara, however, it appears that /ʊ/ is not restricted to the low part of the vowel space, but ranges fairly high in the centre of the space, whereas /i/ and /ʊ/ have more restrictive ranges. These phonetic distributions might suggest a set of vertical contrasts, shown in (2b), whereby /ʊ/ is characterized by a feature such as [back] and /i/ is characterized by [front]; this leaves /ʊ̣/ as [non-back] and [non-front].²

The vowel ranges of another Pama-Nyungan language, Western Arrarnta (Anderson 2000: 36–40), are shown in Figure 3. Here, /a/ is restricted to a very small space; we infer it is [low]. /i/ “varies in quality from [ɛ] to [i].” We can assign it [front]. According to Anderson, /a/ is “extremely variable” in height and backness, with unrounded and rounded allophones (so it

---

¹ These feature assignments are made for the sake of concreteness, but they are underdetermined by the diagram. It is also possible that [low] is marked rather than [high], and [front] or [back] rather than [round]. Further study of these languages might reveal more phonological or phonetic facts to support or modify the assignments in (2).

² Again, there are other possibilities which are subject to disambiguation by further investigation.
could be written /u/, if we want to stick to /i, a, u/). It also appears to be the epenthetic vowel. This distribution is consistent with /a/ being [non-low] and [non-front]; in Jakobson’s terms, narrow and velar, that is, /U/ in (1c). Unlike /U/ in the other languages, however, this one appears to have only default feature values, as in (2c).

![Figure 3: Western Arrarnta vowel ranges (Anderson 2000: 37)](image)

We conclude, then, that the characterization of many three-vowel systems as /i, a, u/ may conceal the fact that they are very diverse. Similarly, the first stages of phonological acquisition may not be as unvarying as proposed by Jakobson (1941) and Jakobson & Halle (1956). On the other side, if Jakobson’s basic idea about the development of contrasts is correct, then all three-vowel systems are similar in being characterized by two features, even if these features are not the same in each case, or even universal.

After the first two stages, Jakobson & Halle allow variation in the order of acquisition of vowel contrasts. The wide branch of (1c) can be expanded to parallel the narrow one. Or the narrow vowels can develop a rounding contrast in one or both branches. Continuing in this fashion we will arrive at a complete inventory of the phonemes of a language, with each phoneme assigned a set of contrastive properties that distinguish it from every other one.

In a number of publications I have tried to reconstruct a history of ‘branching trees’ in phonology (Dresher 2009, 2015b, 2016, 2018). Early, though inexplicit, examples can be found in the work of Jakobson (1931) and Trubetzkoy (1939) in the 1930s, and continuing with Jakobson 1941 and Jakobson & Lotz 1949; then more explicitly in Jakobson et al. 1952, Cherry et al. 1953, Jakobson & Halle 1956, and Halle 1959. This approach was imported into early versions of the theory of generative phonology; it is featured prominently in the first generative phonology textbook (Harms 1968). Nevertheless, for reasons discussed by Dresher (2016: 70), branching trees were omitted from Chomsky & Halle’s Sound pattern of English (1968), and disappeared from mainstream phonological theory for the rest of the century.

In child language studies, however, branching trees continued to be used, for they are a natural way to describe developing phonological inventories (Pye et al. 1987; Ingram 1988; 1989; Leveilt 1989; Dinnsen et al. 1990; Dinnsen 1992; 1996; see Dresher 1998a for a review). Fikkert (1994) presents observed acquisition sequences in the development of Dutch onsets that follow this general scheme, shown schematically in (3).
(3) Development of Dutch onset consonants (Fikkert 1994)

a. Stage 1

consonant

/\P/

b. Stage 2

consonant

/\P/       \m
obstruent   sonorant

/\P/       /\N/

c. Stage 3

consonant

/\P/       /\F/       /\N/       /\L/\J/

Stage 3a  Stage 3b

\u  \m
obstruent sonorant

\u  \m
plosive  fricative nasal liquid/glide

At Stage 1 (3a), there are no contrasts; the value of the consonant defaults to the least marked onset, namely an obstruent plosive. The first contrast (Stage 2) is between obstruent and sonorant (3b). The former remains the unmarked option (u); the marked (m) sonorant defaults to nasal. After this stage, children differ (3c). Some expand the obstruent branch first, bringing in marked (m) fricatives in contrast with plosives (Stage 3a). Others (Stage 3b) expand the sonorant branch, introducing marked sonorants (either liquids or glides). And so on from there.

As a general theory of phonological representations, branching trees were revived, under other names, by Clements (2001; 2003; 2009), and independently at the University of Toronto, where they are called contrastive feature hierarchies (Dresher et al. 1994; Dyck 1995; Zhang 1996; Dresher 1998b; Dresher & Rice 2007; Hall 2007; Dresher 2009). It is the latter approach I will be presenting here. It has gone under various names: Modified Contrastive Specification (MCS), ‘Toronto School’ phonology, Contrast and Enhancement Theory, or Contrastive Hierarchy Theory. I will present the theory as I understand it.3

2. A theory of phonological contrast

The first major building block of our theory is that contrasts are computed hierarchically by ordered features that can be expressed as a branching tree. Branching trees are generated by what I call the Successive Division Algorithm (Dresher 1998b, 2003, 2009), given informally in (4).

(4) The Successive Division Algorithm

Assign contrastive features by successively dividing the inventory until every phoneme has been distinguished.

3 For a more complete summary of contrastive hierarchy theory, see Dresher (2015a).
Since feature hierarchies can vary from language to language, it is crucial to have criteria for selecting and ordering the features. Phonetics is clearly important, as we have seen in the discussion of three-vowel systems in (2), in that the selected features must be consistent with the phonetic properties of the phonemes. For example, a contrast between /i/ and /a/ would most likely involve a height feature like [low] or [high], though other choices are possible (e.g. [front] or [advanced/retracted tongue root]).

It should be noted that the contrastive specification of a phoneme can sometimes deviate from its surface phonetics. Proto-Eskimo, the ancestor of present-day Inuit dialects, has been reconstructed with four vowels: */a, i, u, i/ (see Compton & Dresher 2011 for references and further discussion). In modern dialects the fourth vowel */i/ has merged with */i/. In some dialects the merger is complete, and there is no synchronic trace of an original distinction between */i/ and */i/; in these dialects, there are only three underlying vowels, /i, a, u/.

Some dialects, however, still show evidence of an underlying distinction between */i/ and a fourth vowel. In Barrow North Alaskan Inupiaq, for example, some /i/, called ‘strong i’, cause palatalization of a following consonant (5b), but ‘weak i’ do not (5c). In this case, /i/ and the fourth vowel (designated /ɨ/) need to be distinguished by a contrastive feature, even though their surface realizations are identical (see further §3).

(5) Barrow palatalization after strong i in noun stems (Kaplan 1981: 81–82)

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Analysis</th>
<th>Gloss</th>
<th>Stem</th>
<th>‘and a N’</th>
<th>‘n OBL.PL’</th>
<th>‘like a N’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. u</td>
<td>/u/</td>
<td>‘house’</td>
<td>iyu</td>
<td>iyu-lu</td>
<td>iyu-nik</td>
<td>iyu-tun</td>
</tr>
<tr>
<td>b. Strong i</td>
<td>/i/</td>
<td>‘wound’</td>
<td>ikí</td>
<td>iki-ɻu</td>
<td>iki-nik</td>
<td>iki-sun</td>
</tr>
<tr>
<td>c. Weak i</td>
<td>/ɨ/</td>
<td>‘place’</td>
<td>ini</td>
<td>ini-lu</td>
<td>ini-nik</td>
<td>ini-tun</td>
</tr>
</tbody>
</table>

As this example shows, the way a sound patterns can override its phonetics (Sapir 1925). Thus, we consider as most fundamental that features should be selected and ordered so as to reflect the phonological activity in a language, where activity is defined as in (6) (adapted from Clements (2001: 77)):

(6) Phonological activity

A feature can be said to be 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.

The second major tenet has been formulated by Hall (2007) as the Contrastivist Hypothesis (7):

(7) The 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.

That is, only contrastive features can be phonologically active. If this hypothesis is correct, then (8) follows as a corollary:
Corollary to the Contrastivist Hypothesis

If a feature is phonologically active, then it must be contrastive.

One further assumption is that features are binary, and that every feature has a marked and unmarked value. I assume that markedness is language particular (Rice 2003, 2007) and accounts for asymmetries between the two values of a feature, where these exist. I will designate the marked value of a feature $F$ as $[F]$, and the unmarked value as $[\text{non-}F]$. I will refer to the two values together as $[\pm F]$.

Unless a vowel is further specified by other contrastive features (originating in another vowel or in the consonants), it is made more specific only in a post-phonological component. Stevens et al. (1986) propose that feature contrasts can be *enhanced* by other features with similar acoustic effects. Thus, /u/ in (2a) can enhance its [round] feature by adding [back], and /ʊ/ in (2b) can enhance [back] by adding [round]. These enhancements are not universal, however, as shown by the realizations of /ə/ in (2c).

### 3. Why do phonological features emerge?

There is a growing consensus that phonological features are not innate, but rather ‘emerge’ in the course of acquisition. Most of the contributions to a volume titled *Where do phonological features come from?* (Clements & Ridouane 2011) take an emergentist position; none argue for innate features. Mielke (2008) and Samuels (2011) summarize the arguments against innate features. From a biolinguistic perspective, phonological features are too specific, and exclude sign languages (van der Hulst 1993; Sandler 1993). Empirically, no one set of features have been discovered that ‘do all tricks’ (as Hyman 2011 writes with respect to tone features, but the remark applies more generally). Finally, since at least some features have to be acquired based on evidence of language-specific phonological activity, a prespecified list of features becomes less useful in learning than had once been thought.

But if features are not innate, what compels them to emerge at all? It is not enough to assert that features *may* emerge, or that they are a useful way to capture phonological generalizations. Assuming that phonological representations are composed of distinctive features, we need to explain why features *inevitably* emerge, and why they have the properties that they do. In particular, we have to explain why some learners do not simply posit segment-level representations. Further, are there limits to how broad or narrow features are, or how many features can be associated with a given phonological inventory?

Contrastive hierarchy theory provides an answer to these questions: learners must arrive at a set of hierarchically ordered features that distinguish between all the phonemes of their language. This requirement imposes strong constraints on the number of features that can be posited, and on what feature systems can look like. We have already seen that a three-vowel system allows for exactly two contrastive features. The features may vary, as well as their ordering, and either the marked (2a) or unmarked (2b, c) branches of the first feature may be expanded; but this is the extent of variation that is permitted.

---

4 See also Keyser and Stevens (2006) and the references therein.

5 See Dyck (1995) and Hall (2011) for further discussion and examples.
In general, the number of features required by an inventory of \( n \) elements falls in the following ranges: the minimum number of features is the smallest integer greater than or equal to \( \log_2 n \); and the maximum number of features is equal to \( n-1 \). Some sample values of feature minima and maxima for inventories of different sizes are shown in Table 1. By putting limits on the number and organization of feature systems, the contrastive hierarchy together with the Contrastivist Hypothesis account for why phonological systems resemble each other in terms of representations, without requiring individual features to be innate.

Table 1: Minimum and maximum number of features for various-sized phoneme inventories

<table>
<thead>
<tr>
<th>Phonemes</th>
<th>( \log_2 n )</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.58</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>2.32</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phonemes</th>
<th>( \log_2 n )</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>3.46</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>22</td>
<td>4.46</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>29</td>
<td>4.86</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>36</td>
<td>5.17</td>
<td>6</td>
<td>35</td>
</tr>
</tbody>
</table>

4. **Concluding example: Inuit vowel systems**

To illustrate some of the principles discussed above, consider again the Inuit dialects mentioned in §2. Dialects with four underlying vowels, such as Barrow Inupiaq, can support three contrastive features, as in (9a); the presence of [front] is what enables /i/ to cause palatalization.

(9) Contrastive feature hierarchies in Inuit dialects

a. Four-vowel dialects

\[
\begin{array}{c}
\text{vowel} \\
[\text{low}] & [\text{non-low}] \\
/\text{a/} & [\text{round}] & [\text{non-round}] \\
/\text{u/} & [\text{front}] & [\text{non-front}] \\
/\text{i/} & /\text{i/}
\end{array}
\]

b. Three-vowel dialects

\[
\begin{array}{c}
\text{vowel} \\
[\text{low}] & [\text{non-low}] \\
/\text{a/} & [\text{round}] & [\text{non-round}] \\
/\text{u/} & /\text{i/}
\end{array}
\]

In three-vowel dialects, however, where Proto-Eskimo */i/* has completely merged with */i/*, there is room for only two contrastive vowel features (9b). Compton & Dresher (2011) argue that there is evidence that [±low] and [±round] are active, hence, contrastive features, leaving /i/ with no marked feature that can trigger palatalization. The prediction that /i/ does not cause palatalization in three-vowel Inuit dialects is strikingly borne out: whereas there are four-vowel dialects with and without palatalization, no three-vowel dialects have palatalization. Contrastive hierarchy theory accounts for this conspicuous gap in the typology of Inuit dialects, demonstrating the close connection between contrast and phonological activity.
References


