# University of Toronto March 9–11, 2012

# Issues in the Theory of Contrast

B. Elan Dresher University of Toronto

## Introduction

I will consider some issues in the theory of phonological contrast, particularly issues surrounding the contrastive feature hierarchy.

I have argued (Dresher 2009) that the contrastive hierarchy is interesting mainly in connection with the Contrastivist Hypothesis (Hall 2007):

The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

### Introduction

The Contrastivist Hypothesis remains an important motivation for having a correct method for determining which specifications count as 'contrastive' in a language.

I will consider some empirical challenges to the Contrastivist Hypothesis (cf. Hall, this conference).

The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

## Introduction

However, the contrastive hierarchy has uses beyond the Contrastivist Hypothesis, and I will review some of them here:

- provide a measure of the amount of information conveyed by each phoneme in an inventory
- account for universal tendencies in inventories
- account for diachronic developments
- account for synchronic alternations
- applications outside phonology

Feature Hierarchies and Phonological Activity

### Active Features are Contrastive

An idea that can be traced to the beginnings of modern phonology is that only some properties of a segment are active, or relevant (Trubetzkoy), to the phonology, and these are the distinctive, or contrastive, properties.

# **Five-Vowel Systems**

Trubetzkoy (1939) reviews a number of five-vowel systems. He observes that in many such systems the low vowel does not participate in tonality contrasts.

He cites Latin as an example of this kind of system.

#### Latin

F	Front/unround	Back/round	-
	i	u	High
	e	0	Mid
	a		Low

### **Five-Vowel Systems**

This corresponds to a feature hierarchy with [low] at the top, removing /a/ from further participation in contrasts



However, he observes that other types of vowel systems exist.

In Artshi, a language of Central Daghestan, a consonantal rounding contrast is neutralized before and after the rounded vowels /u/and /o/. 'As a result, these vowels are placed in opposition with...unrounded *a*, *e*, and *i*'.

Artshi (East Caucasian)



'This means that all vowels are divided into rounded and unrounded vowels, while the back or front position of the tongue proves irrelevant...' (Trubetzkoy 1969: 100-101).

### Artshi (East Caucasian)

Unround	Round	
i	u	High
e	0	Mid
a		Low

'This means that all vowels are divided into rounded and unrounded vowels, while the back or front position of the tongue proves irrelevant...' (Trubetzkoy 1969: 100-101).

This analysis corresponds to ordering [round] first, followed by [high] and [low] (the latter only in the unrounded vowels).



Trubetzkoy argues that neutralization of the opposition between palatalized and non-palatalized consonants before iand e in Japanese shows that these vowels are put into opposition with the other vowels /a, o, u/.

### Artshi (East Caucasian)

Japanese

Unround	Round		Front	Back
i	u	High	i	u
e	0	Mid	e	0
a		Low		a

The governing opposition is that between front and back vowels, lip rounding being irrelevant.

#### Artshi (East Caucasian)

#### Japanese

Unround	Round		Front	Back
i	u	High	i	u
e	Ο	Mid	e	Ο
а		Low		a

The governing opposition is that between front and back vowels, lip rounding being irrelevant.

This analysis corresponds to ordering [front] first, followed by [high] and [low] (the latter only in the back vowels).



# The Contrastivist Hypothesis

These analyses effectively assume the Contrastivist Hypothesis, that is, that only contrastive features can be active in the phonology.

It follows from this that a heuristic principle for identifying contrastive features is: Assume that active features are contrastive.

## **Contrast via Feature Ordering**

Second, contrastive features are determined by ordering features into a *contrastive hierarchy*:

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

This method was called 'branching trees' in the literature, when referred to at all. I call it the *Successive Division Algorithm* (Dresher 1998, 2003, 2009).

# Variability of Feature Ordering

Third, we learn from the above examples that the contrastive hierarchy must allow for *variation*:

The contrastive feature hierarchy is not universal but may vary (within limits to be determined).

## Rationale for Feature Hierarchies

In keeping with the view that the contrastive hierarchy accounts for how sounds pattern, Jakobson and Lotz (1949) give empirical arguments for their choice of features for Standard French, based on two types of phonological activity:

the adaptation of foreign sounds

language internal alternations

They observe (1949: 153): 'the difference between velar and palatal is irrelevant in French phonemics...These contextual variations do not hinder French speakers from rendering the English velar  $\eta$  through the French palatal p... or the German 'ich-Laut' through  $\int$ .'



'The advanced articulation of k g before j or i, as well as the existence of n instead of n before w...illustrates the unity of the saturated consonants in French.'



# Feature Hierarchies and Information Theory

Despite these antecedents, this is not the approach taken by Halle (1959) in *Sound Pattern of Russian*.

The change in rationale for limiting specifications to contrastive features is hinted at by Jakobson and Halle (1956), when discussing Standard French.

Though their analysis is similar to that of Jakobson and Lotz (1949), their main justification is that theirs is 'the unique solution' on the grounds that it is optimal in terms of the number of binary decisions that have to be made.

In the 1950s, Jakobson and Halle became interested in the thennew field of information theory, and began to look at branching trees as a way of conveying information about phonemes in the most economical way (cf. Cherry, Halle and Jakobson 1953).

This criterion, the IT Principle, came to overshadow the earlier one, what we can call the Activity Principle, that is, to reflect the active features and account for phonological patterning.

Activity Principle

The purpose of a feature hierarchy is to identify the contrastive features that are relevant to the phonological computation.

### Activity Principle

The purpose of a feature hierarchy is to identify the contrastive features that are relevant to the phonological computation.

### IT Principle

The purpose of a feature hierarchy is to to minimize redundancy in phonological representations and to maximize the amount of information conveyed by each feature.

# Feature Hierarchies to Minimize Specifications

In *Sound Pattern of Russian* (29–30), Halle's version of the IT Principle is 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).

Roughly speaking, Conditions (3) and (4) require that the phonological description meet basic conditions of adequacy.

# Feature Hierarchies to Minimize Specifications

Halle observes (*SPR*: 44–5) 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  $log_243 = 5.26$  specifications, which would represent the most efficiently branching tree for 43 phonemes.

Feature hierarchies and phonological inventories (Clements 2001, 2003a, b, 2009)

# The Accessibility/Robustness Hierarchy

In a series of publications Clements posited a feature hierarchy, which he calls the Accessibility Hierarchy (2001), and later the Robustness Hierarchy (2009).

Clements (2001: 79): 'features can be ranked according to a universal hierarchy of accessibility. At the top of the hierarchy are features that are highly favored in the construction of phoneme systems, while at the bottom are features that are highly disfavored.'

# The Accessibility Hierarchy (Clements 2001)

# (8) Partial ranked scale of feature accessibility for consonants

	feature:	in:
a.	[coronal]	
b.	[sonorant]	
C.	[labial]	
d.	[dorsal]	[-sonorant)
e.	[strident)	
f.	[nasal]	
g.	[posterior]	[+sonorant, -nasal]
h.	[lateral]	[+sonorant]
1.	[voice]	[-sonorant]

This scale works *almost* like the contrastive hierarchy introduced earlier, but not exactly.

An important difference is that the ranking does not strictly dictate whether a feature will actually be specified.

## Universality of the Feature Hierarchy

The most important difference between Clements's approach and the one I argued for earlier is that Clements wishes to maintain a *universal* feature hierarchy.

Actually, his approach is quite nuanced (Clements 2001: 84–5); he does allow for some variability in the hierarchy, and he sometimes makes adjustments for particular languages.

The key question is how much relative weight should be given to the phonological patterning exhibited by a particular language, on the one hand, as compared to universal tendencies with respect to phonological inventories, on the other.

In general, Clements favours the latter, because of his interest in universals of feature economy.

### Feature Economy

Clements (2009: 34) observes that crosslinguistically inventories reflect the effects of Feature Economy working together with the Accessibility Scale, renamed now the Robustness Scale.



# The Robustness Hierarchy (Clements 2009)

The Robustness Scale is a somewhat revised version of the Accessibility Scale.

Rather than a strict ranking, features are placed in 5 groups of decreasing likelihood of occurring.

There are also some changes in the ordering.

Robustness scale: consonants

feature:

- a. [±sonorant] [labial] [coronal] [dorsal]
- b. [±continuant] [±posterior]
- c. [±voiced] [±nasal]
- d. [glottal]
- e. others

# Feature Economy

In order to maintain the proposed universal hierarchy, Clements (2009) is inclined to interpret the contrasts in inventories in accordance with the Robustness hierarchy, favouring it over other possible analyses.

He considers the typical consonant inventory shown below; capital letters indicate consonant types:

Р	Т	К	
	S		
Μ	Ν		
W	L~R J	H ~ ?	

34

# Feature Economy

For example, he considers that  $/T/\sim/S/$  are distinguished by [continuant], not [strident]; similarly, the  $/L/\sim/J/$  contrast could be based on [continuant] or [posterior], but not [lateral].

These may be the correct analyses in many, possibly most, maybe even all, inventories.

The crucial cases arise when phonological patterning diverges from the proposed universal ordering.

Р	Т	K	
	S		
Μ	Ν		
W	L~R J	H ~ ?	

35

# Weighting Rationales for Feature Hierarchies

To sum up, Clements does appeal to the Activity Principle: 'whether or not a given feature or feature value is specified in a given language can only be determined from an examination of its system of contrasts and sound patterns.'

### **Activity Principle**

The purpose of a feature hierarchy is to identify the contrastive features that are relevant to the phonological computation.
# Weighting Rationales for Feature Hierarchies

But in the end he gives preference to the Inventory Principle, which requires a universal feature hierarchy, to the extent possible:

**Inventory Principle** 

The purpose of a feature hierarchy is to express universal tendencies in the nature of phonological inventories.

# Weighting Rationales for Feature Hierarchies

#### Activity Principle

The purpose of a feature hierarchy is to identify the contrastive features that are relevant to the phonological computation.

#### **Inventory Principle**

The purpose of a feature hierarchy is to express universal tendencies in the nature of phonological inventories.

# Changing Rationales for Feature Hierarchies

To conclude this section, contrastive feature hierarchies may be useful in information theoretic computations, which may have some empirical applications.

Similarly, studying universal tendencies in inventories may suggest constraints or tendencies that apply to the ordering of features.

However, I believe that phonological activity should remain the main rationale for setting up a contrastive feature hierarchy for a particular language.

Empirical Challenges to the Contrastivist Hypothesis

### Challenges to the Contrastivist Hypothesis

As an empirical proposal, the Contrastivist Hypothesis is of course vulnerable to being challenged by data that may require that we modify it, or, in the worst case, abandon it.

I have argued that many apparent counterexamples arise because of incorrect assumptions about how to determine which features are contrastive.

Thus, phonologists who assume some version of Minimal Contrast (based on minimal pairs) or a fixed universal feature hierarchy, are likely to identify certain features as contrastive which are not actually contrastive in a hierarchical approach.

#### Weak Contrastivist Hypothesis

Some phonologists (e.g. Clements 2001) assume a version of what we could call the Weak Contrastivist Hypothesis:

The phonological component of a language L *may* operate only on those features which are necessary to distinguish the phonemes of L from one another.





## The "Oops, I Need That" Problem

Discarding cases where contrastive features have been wrongly identified, we still find genuine examples where noncontrastive features appear to be needed in the phonology.

In Andrew Nevins's colourful formulation, we can call this the "Oops, I Need That" Problem.

Given the number of cases in which the Contrastivist Hypothesis holds, a promising approach to this problem is to try to classify the types of cases in which OINT arises, and adjust the Contrastivist Hypothesis to accommodate them, if necessary.

One type of case, discussed by Hall (2007), arises in Yowlumne (Kuroda 1967, Archangeli 1984), which has 4 underlying long vowels as shown.



The features [round] and [high], which are motivated by vowel harmony, are contrastive. No other vowel feature can be contrastive.



High vowels lower: /u:/ lowers to [5:], merging with /5:/, as expected. We would similarly expect /i:/ to lower to [a:], but instead it lowers to [e:], a new allophone, which must be distinguished from /a:/ by a noncontrastive feature.



Hall (2007) proposes that /a:/ is specified with a redundant [low] feature.



A more serious violation of the Contrastivist Hypothesis would involve a case where a feature is active, hence presumably contrastive, with respect to one phonological process in a language, while being conspicuously inactive with respect to another process.

Such a case appears to arise in Finnish.

# Finnish /i/

Finnish /i/ and /e/ are neutral to front/back vowel harmony. These vowels have no [–round, +back] partner.

A common analysis posits that these vowels have no contrastive value for the feature [back].



#### Finnish Vowel Harmony

Vowels with contrastive [±back] can participate in vowel harmony. The vowels /i/ and /e/ are neutral, because they lack contrastive [back].

		[-round]	[+round]		
		[-back]	[-back]	[+back]	
[–low]	[+high]	i	ü	u	
	[—high]	е	ö	0	
[+low]		[–back]	[+back]		
		ä	a		

#### Finnish Vowels: Contrastive Hierarchy

A possible feature hierarchy for Finnish has the features in the order [low] > [round] > [back] > [high].

[round] is not contrastive under [+low], [back] is not contrastive under [–round]. Harmonic vowels have [back].



#### **Finnish Assibilation**

Finnish also has rule whereby /t/ becomes [s] before \_\_\_\_\_ *i* in certain contexts (Kiparsky 1973; Antilla 2003, 2006).

Assibilation:  $/t/ \rightarrow s / \____i$ 

a. halut–a 'to want'

b. halus-i 'wanted'

Evidently, /i/ has some feature or features that trigger the assibilation. Nevins (2010: 91) suggests that the relevant features are [+high, –back].

#### **Finnish Assibilation**

We decided with respect to vowel harmony that [back] is not a contrastive feature for /i/.

Assibilation:  $/t/ \rightarrow s / \_ [V, +high, -back]$ 

a. halut–a 'to want'b. halus–i 'wanted'

On this analysis, then, Assibilation is a counterexample to the Contrastivist Hypothesis.

I have no good solution to this case at the moment. Apparently there are similar problems in Votic.

Nevins (2012) observes that Vowel Harmony is a long-distance, vowel-to-vowel process, whereas Assibilation is a short-distance, consonant-consonant, process.

Perhaps these differences play a role with respect to the applicability of the Contrastivist Hypothesis.

Such a theory would still be more restrictive than most competitors.

#### Challenges to the Contrastivist Hypothesis

To conclude this section, the fact that there are empirical challenges to the Contrastivist Hypothesis is a good thing.

Such cases shed light on ways to refine the hypothesis, if needed, and point the way to discover new generalizations about the workings of phonology. Diachrony and the Contrastive Hierarchy

#### Applications to Diachrony

The contrastive hierarchy can help account for diachronic changes.

Zhang (1996) and Dresher and Zhang (2005) show how a change in the contrastive status of one vowel phoneme sets the stage for a series of further changes, leading to new patterns of phonological activity in the modern Manchu languages, Spoken Manchu and Xibe.







/ʊ/ had already merged with /u/ in many contexts in Classical Manchu. At some point, merger became complete.



















#### Evidence: /ə/ develops high allophones

Zhao 1989: Spoken Manchu /ə/ is a mid-high back unrounded vowel, with an allophone [ɤ]. Ji et al. 1989: [ə] is in free variation with a high back unrounded vowel [ɯ].



# Evidence: In Xibe /u/ participates in labial harmony, unlike Classical Manchu /u/.



## Applications to Diachrony

It is well known that Structuralism differs from the 19<sup>th</sup> century Neogrammarian approach in emphasizing the systemic nature of diachronic changes.

In Generative Phonology (Kiparsky 1965), this takes the form of analyzing phonological change with respect to the whole grammar, rather than just the phonetic surface.

The role of contrast in diachronic change was neglected in classical Generative Phonology. However, it was part of the original structuralist program.
#### **Contrast and Diachronic Change**



"Once a phonological change has taken place, the following questions must be asked:

What exactly has been modified within the phonological system?

...has the structure of individual oppositions [contrasts] been transformed? Or in other words, has the place of a specific opposition been changed...?"

Roman Jakobson, Principles of historical phonology, first published in German in *TCLP*, IV (Copenhagen, 1931).

## Applications to Diachrony

Recent papers by Oxford on Algonquian (2011, also this conference) and Harvey on Ob-Ugric vowel systems (2012 and this conference) show how Jakobson's questions can be answered via contrastive feature hierarchies.

They show that contrast shift is itself a type of diachronic change.

For example, consider the contrastive hierarchy of Proto-Algonquian posited by Oxford (2011): Contrastive Hierarchy for Proto-Algonquian Vowels (Oxford 2011) [labial] > [coronal] > [long] > [low]

Dialects with this hierarchy tend to merge  $/\epsilon / \sim /i/$  and have palatalization before one or more of the [+coronal] vowels.



Hierarchy for Proto-Eastern-Algonquian Vowels (Oxford 2011) [hi] >[lab] > [cor] > [long] > [low]

Here we find  $|\varepsilon|$  shifting to |a|, and palatalization before [+coronal]  $|\varepsilon|$ , but not |i|, which lacks [coronal].



## **Constraints on Mergers**

Oxford and Harvey also show that the feature tree constrains possible mergers and neutralizations.

This suggests that the tree itself plays a significant role in phonology beyond its function of specifying contrastive features. Synchrony and the Contrastive Hierarchy Applications to Synchrony

If the contrastive hierarchy can constrain diachronic processes then we might expect that it may play a role in synchronic phonology as well.

A number of papers at this conference address this issue, to varying extents: see (cf. Mackenzie, this conference; Motut, this conference; St-Amand, this conference).

Again, an interesting question is to what extent the hierarchy itself is relevant to phonology, apart from its role in distributing contrastive features.

Applications to Synchrony

For example, Motut (2012, also this conference) looks at how the feature hierarchy governs alternations in Oowekyala (Wakashan).

For example, what does  $/\chi^w/$ , which is both [+round] and [+RTR], do when it becomes [+voiced], given that no segment may be [+voiced, +round, +RTR]?

Again, an interesting question is to what extent the hierarchy itself is relevant to phonology, apart from its role in distributing contrastive features.

#### Oowekyala (Motut 2012)

The hierarchy has [round] > [RTR], so the former is retained, and the result is /w/:



# **Binary or Privative?**

Are Features Binary or Privative?

Are features binary (i.e., two-valued, [+F] versus [–F]), or are they privative (one-valued, [F] versus Ø)?

This issue is independent of the contrastive hierarchy itself, which can function in both modes. Though it should be noted that privative features do not distinguish between contrastive  $\emptyset$  (which would correspond to [–F]) and redundant  $\emptyset$  (which would correspond to no value for [F].

It appears that some processes function better assuming privative features (vowel coalescence, St-Amand 2012), while others, (co-occurrence restrictions, Mackenzie 2009), seem to require both + and – values of features. Are Features Binary or Privative?

Even if features are binary, it does not follow that the two values are symmetrical. Typically, one is marked and the other is unmarked.

One solution is to posit binary features, but to allow certain processes to refer only to marked values.

This is essentially the Calabrese-Nevins approach discussed earlier:

## **Two Modes of Application**

Recall on their view that some processes access only marked contrastive features (as if features are privative), and others access marked and unmarked contrastive features, that is, [+F] and [–F].

Marked contrastive features

Contrastive features (marked and unmarked)

# Theories of Phonology that do not have Features

## Phonology without Features

What about theories like Dependency Phonology, Government Phonology, and CV-phonology, which do not posit features of the sort I have been assuming?

Do considerations of contrast and the contrastive hierarchy apply to them?

Tobias Scheer (2011) suggests that they do not, that a contrastive hierarchy may be irrelevant to Government Phonology.

## Phonology without Features

I suspect that this may not be correct; it seems to me that notions of contrast are embedded in these theories, hence contrastive hierarchies apply just as much there as in feature theories.

However, this remains to be shown, and I can't pursue the issue here.

# The Contrastive Hierarchy Beyond Phonology

# Beyond Phonology

We might expect that morpho-syntactic features might also be subject to considerations of relative scope and contrast, and hence might be fruitfully viewed in terms of contrastive hierarchies.

Cowper and Hall (2011) apply this perspective to changes in the in the English Aspect-Voice system.

A key element of their analysis is that the expression of aspect and voice features can *vary* over time and across languages, and is not tied to a fixed universal structure.

## Way Beyond Phonology

Nevins (2012) observes that variable contrastive hierarchies can be fruitfully applied outside linguistics, to kinship systems.

The following two diagrams are from Nevins's presentation last month in Leiden: the first is a conventional diagram of the Seneca kinship system, and the second is contrastive hierarchy:

# SENECA KINSHIP





for these categories?

## Conclusion

## Conclusion

To conclude, I have argued that contrastive feature hierarchies are the correct way to assign contrastive feature values, and therefore are crucial to any attempt to assess the correctness of the Contrastivist Hypothesis.

But feature hierarchies have uses that do not depend on the Contrastivist Hypothesis, and are proving to be very useful in characterizing diachronic and synchronic phonological processes.

And we have seen that they have potential uses beyond phonology.

#### Conclusion

The variety of applications of feature hierarchies suggest that they are indeed, as stated by Jakobson, Fant and Halle (1952),



"the pivotal principle of the linguistic structure" and a fundamental principle of cognition.

#### **THANK YOU!**

I am grateful to members of the project on *Markedness and the Contrastive Hierarchy in Phonology* at the University of Toronto (Dresher and Rice 2007):

## http://homes.chass.utoronto.ca/ ~contrast/

This research was supported in part by grants 410-2003-0913 and 410-08-2645 from the Social Sciences and Humanities Research Council of Canada.