

Triumphs and Limits of the Contrastivity-Only Hypothesis

Andrew Nevins*
University College London
a.nevins@ucl.ac.uk

To appear in *Linguistic Variation* 15.1

Abstract: The Variable Hierarchy hypothesis of Dresher (2009) has a number of far-reaching consequences and applications, even beyond the domain for which it was originally developed, including overspecification, kinship terminology, vowel reduction and whistled languages. On the other hand, the Contrastivity-Only Hypothesis of Dresher (2009) is, in its present form, empirically too limited once one looks at a full system of phonological processes within a language, and some reference to noncontrastive features seems inevitable. However, I outline a possible set of diagnostics and distinctions that might allow one to localize and limit the reference to noncontrastive features.

Keywords: Contrastive hierarchies, vowel reduction, whistled languages, kinship, overspecification

1. Overview

The purpose of this paper is to explore the Contrastive Feature Hypothesis of Dresher (2009) (alternatively called the Contrastivist Hypothesis in closely related work, e.g. Hall (2011)), which in my view has at least two independently separable parts. One is the notion that different languages may promote the importance of certain features, and conversely relegate the unimportance of certain features to the point where they are absent from a given phonological system; let us call this the *Variable Hierarchy* Hypothesis. The other is the notion that only those features which are necessary in order to fully partition segments into contrastive leaves of a hierarchical tree can be active in a language's phonology; let us call this the *Contrastivity-Only* Hypothesis. In the present paper, I will explore a number of case studies that demonstrate areas in which the contrastive hierarchy works very well, even for empirical problems it was not originally 'designed' for, and then turn to case studies in which it seems that the contrastivity-only hypothesis is too strong. I conclude by outlining possible directions for further research to resolve this tension.

The Variable Hierarchy asserts that what matters to a language's phonology is determined by a hierarchy, and that this hierarchy assigns a set of contrastive features. Dresher (2009) makes explicit a practice that virtually all phonologists adopt (including myself, in certain prior work) only implicitly – that not all features in a language's feature-matrix are of equal importance to the phonology of a language. For example, in Nevins (2010a), I asserted that the Turkish 8-vowel system employed the three binary features $[\pm \text{back}]$, $[\pm \text{round}]$, $[\pm \text{high}]$, and in doing so, I tacitly grouped the four $[-\text{high}]$ vowels $/e, \ddot{o}, o, a/$ together as a natural class (e.g. in the fact that they do not undergo $[\pm \text{round}]$ harmony), even though phonetically it is obvious that $/a/$ is not of the same height as the former three. In asserting that Turkish was a two-height system based on $[\pm \text{high}]$, without saying so I implicitly relegated the feature $[\pm \text{low}]$ to the abyss for Turkish phonology, claiming that while $/a/$ might be phonetically distinct from $/e, \ddot{o}, o/$ in terms of height, phonologically it was not; indeed, as $/a/$ and $/o/$ alternate in suffixes for $[\pm \text{round}]$ harmony, both can be characterized as $[-\text{high}]$.¹ As such, a specific formalization in terms of the Variable Hierarchy hypothesis

*For very helpful discussion, I thank Marc van Oostendorp, Elan Dresher, Bert Botma, Daniel Currie Hall, Kenneth Dejong, Daniel Harbour, John Harris, Yuni Kim, Julien Meyer, Andreea Nicolae, Christopher Spahr, Bert Vaux, Joanna Zaleska, and anonymous reviewer for *Linguistic Variation*. A special thanks goes to the audience at the University of Leiden, where this productive debate was first staged in February 2012, my contributions to which are essentially reproduced in full in the text below.

¹When I say 'relegate to the abyss', I mean that $[\pm \text{low}]$ was assumed to not be present in phonological representations, without explicitly asserting its absence from the system of specifications. Such a move is implicit in a broad swath of approaches to vowel harmony in Turkish, even ones with very different assumptions about the harmony mechanism itself, e.g. van der Hulst and van de Weijer (1995); Krämer (2003); any model in which the (at least) three phonetic vowel heights of Turkish are treated using a single $[\pm \text{high}]$ contrast is thus implicitly stating that $[\pm \text{low}]$ is not

(under which [\pm low] is hierarchically so low as to be nonexistent for Turkish, but not for, say, Kalenjin) would have made this more explicit, and indeed Dresher's Contrastive Feature Hypothesis captures many insights and has a level of rigour that all future such theories should maintain. Indeed, he is right on the mark when he asserts "I will argue throughout that contrast is too central to be kept out of phonological theory for long, and I will show that it gradually leaked back into generative phonology in various forms. It is one of the aims of this work to reconnect phonology with its roots in this respect and to establish phonological contrast as a central principle of phonological theory." Dresher (2009, p.8)

In Section 2, I demonstrate the utility of the Variable Hierarchy Hypothesis for fricative patterning in Turkish vs. Russian, reviewing a study in Nicolae and Nevins (2015), and even for the specification of contrastive versus noncontrastive distinctions in kinship systems in Section 3, a domain that may seem completely distinct from phonology, but one which can be (and has been) modeled with distinctive features.

Now, while the fact that languages with the same phonetic inventory (e.g. 8 vowels, or 5 vowels, or a fortis/lenis contrast) may pattern quite differently phonologically from each other is perhaps not so novel (having its roots in Sapir 1925), I will attempt to show that the specific implementation in terms of a hierarchy enables the formalization of 'reduction as pruning' – in other words, when a language must reduce its inventory, the lowest (and by hypothesis, least important) contrasts in the tree are the first to be eliminated. I will demonstrate this for vowel reduction in Bulgarian versus Russian in Section 4, and for whistled Spanish versus whistled Greek in Section 5.

We then turn to the Contrastivity-Only hypothesis, viz. "Only contrastive features can be phonologically active". I will demonstrate that this hypothesis is too restrictive (and perhaps too simple) to model all of phonological activity (a point also raised in Clements (2001)), an issue I characterize as the 'Oops, I need that' Problem, demonstrated here for phenomena in Yoruba. However, all may not be lost – in Section 7, I explore the possibility of a typology of sensitivity to contrastive versus noncontrastive features. The overall conclusion is that the Contrastive Feature Hypothesis – jointly composed of the Variable Hierarchy Hypothesis and the Contrastivity-Only Hypothesis – works surprisingly well for certain phenomena, even ones it was not originally intended for.

2. Contrastivity and Overspecification in Turkish vs Russian fricatives

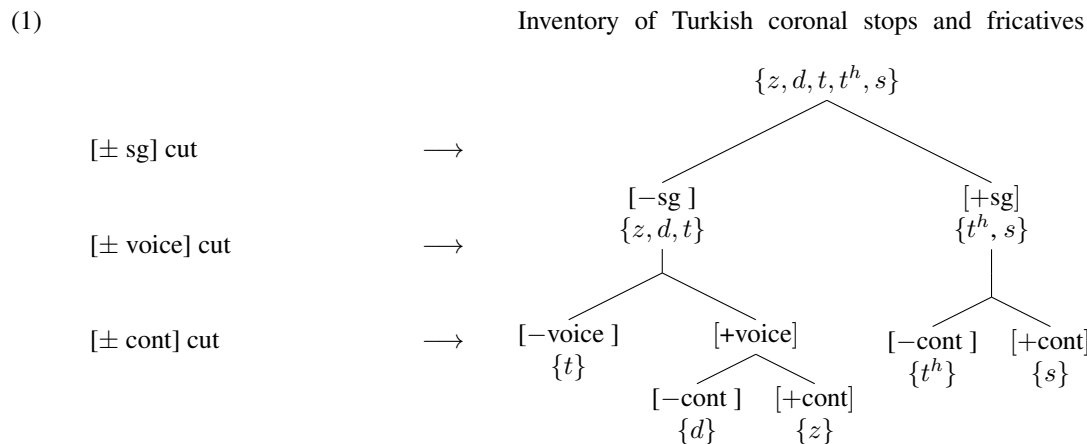
Dresher (2009) and related work within the 'Toronto School of Contrast' employ the use of the Variable Hierarchy and the Successive Division Algorithm, an approach through which the feature specifications of a language are governed by a hierarchical ordering of features. In this framework, only contrastive feature specifications are active in the phonology and contrast is taken to result from this ordering. Specifically, Dresher's account rests on the assumption that, initially, all sounds form one phoneme, and that contrastive feature specifications are determined by splitting the inventory by means of successive divisions, governed by an ordering of features.

It turns out that the Successive Division Algorithm (SDA) works very well for describing a few phonological differences in the patterning of fortis ('voiceless') fricatives in Turkish vs. Russian. In both languages, the fricatives show a two-way contrast (e.g. between /s/ and /z/). However, in Russian, fricatives undergo final devoicing, while in Turkish, they do not (see phonetic evidence in Nicolae and Nevins (2015)). A straightforward way of encoding the final devoicing that affects both stops and fricatives is to say that Russian has a [\pm voice] contrast in its stops and fricatives, and a process by which this feature is unlicensed and hence pruned in syllable-final positions. However, Turkish has a more complex system of laryngeal contrasts: there is final devoicing of some (but not all) stops, while there is no final devoicing of fricatives. Following the basic insight of Inkelas and Orgun (1995), the stops can be represented by a three-way contrast (although not necessarily with the particular set of features adopted in that work). Rather, adhering to Laryngeal Realism, the long-lag VOT of Turkish means the fortis stops are [+spread glottis], and the alternating vs non-alternating lenis stops vary for whether or not they are specified for [+voice]. What this means in terms of the laryngeal alternation processes themselves is that the alternating Turkish stops undergo intervocalic voicing, rather than final devoicing. Under this view, for the non-alternating lenis stops, nothing needs to be said, and for the non-alternating fortis stops, their [+spread glottis] specification protects them from intervocalic voicing; for further details, the reader is referred to Nicolae and Nevins (2015).

part of the phonology of the language. Note, however, that in some implementations this 'relegation to the abyss' such as in a universal constraint set version of OT, features like [\pm low] in Turkish might still be phonologically present, but the faithfulness constraints that refer to them would so low-ranked as to never demonstrate any activity in the system.

In light of the current discussion, what is most interesting about Turkish from a ‘systemic’ point of view is that there is a three-way contrast in stops, reduced to a two-way contrast in fricatives, but that the fricatives each ‘choose’ the marked pole. In other words, a reduction in the number of contrasts doesn’t necessarily mean that ‘one feature is removed from the system’. The fortis and lenis fricatives very well could have remained distinguished using a single feature, $[\pm \text{voice}]$. But they don’t, on the account of Nicolae and Nevins (2015): both are ‘marked’: one $[+\text{voice}]$, one $[+\text{spread glottis}]$, and no ‘plain’ version.² In this sense, the fricatives of Turkish are *overspecified*, as two features are being employed where only one is ‘needed’.³

The differing specifications of Turkish versus Russian fricatives can be understood as a consequence of the specifications needed for the obstruent system in a straightforward and insightful manner given the Successive Division Algorithm. Adopting this account for Turkish, where both $[\pm \text{spread glottis}]$ (henceforth $[\pm \text{sg}]$) and $[\pm \text{voice}]$ are phonologically active features, the hierarchical cuts would proceed as in (1), with stops and fricatives patterning together with respect to the $[\pm \text{sg}]$ and $[\pm \text{voice}]$ cuts. That is, we first partition the voiceless non-alternators from the alternators via $[\pm \text{sg}]$, followed by a cut via $[\pm \text{voice}]$. Importantly, it is not until the final cut via $[\pm \text{continuant}]$ that we are able to distinguish fricatives from stops.⁴ In other words, Dresher’s contrastive hierarchy allows us to impose an ordering on these features such that the stop/fricative split ($[\pm \text{cont}]$) follows the $[\pm \text{sg}]$ cut, therefore predicting that stops and fricatives will pattern similarly with respect to the $[\pm \text{sg}]$ feature. The fact that Turkish employs the use of $[\pm \text{sg}]$ in the feature specification of its stops has repercussions for the feature specification of its fricatives.



On the other hand, in a $[\pm \text{voice}]$ -only language such as Russian (Padgett, 2002; Halle, 1959), which has an uncontroversial two-way contrast in both its stop and fricative paradigms, $[\pm \text{sg}]$ is not an active feature in the language’s phonology since $[\pm \text{voice}]$ is the only feature needed to distinguish the two types of stops. Thus, since $[\pm \text{sg}]$ is not invoked in the stop inventory and the fricatives have only a two-way contrast, $[\pm \text{voice}]$ turns out to be the only laryngeal specification required. The result of a Successive Division hierarchy for Russian is illustrated in (2).

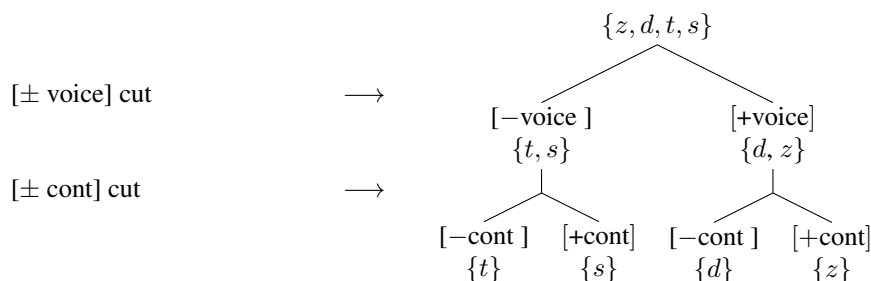
Observe the difference between Turkish and Russian in the specification of their fricatives. $[\pm \text{voice}]$ is phonologically active in both Russian and Turkish while $[\pm \text{sg}]$ is only active in Turkish, and is inert in Russian since it does not enter into alternation or neutralizations. Therefore Russian, unlike Turkish, does not belong to the subset of languages investigated by Vaux (1998) that phonologically mark their voiceless fricatives as $[\pm \text{sg}]$. While the fricative inventories may look similar, their contrastive specifications are very different. The Successive Division Algorithm and the Variable Hierarchy provide an effective way to model the different patterning of a two-way fricative opposition in these two languages in terms of the specifications that the phonology of a language provides based on the inventory as a whole.

²As such, they actually resemble the stop system of Swedish, on the analysis of Beckman et al. (2011).

³An interesting consequence of the present proposal is that the alternating lenis stops have no specification for $[\pm \text{continuant}]$. Whether or not the phonology of Turkish requires that they have one is an open question for further research, and connects to the ‘Oops, I Need That’ discussion in Section 6.

⁴It turns out that this is not the only possible hierarchy for this set of features; of the six possible orderings, either of the two in which $[\pm \text{cont}]$ is the last cut made would yield a feature assignment in concord with our phonological predictions.

(2) Inventory of Russian stops and fricatives



3. Variable Hierarchies and Contrast in Kinship Terminology

We now turn to a comparison of kinship systems, specifically looking at the difference between the Seneca kinship system (discussed in Jones (2010); Nevins (2010b)) and the English system. In the discussion that follows, parallel-cousins refer to cousins who are connected to the speaker ('ego') by a parent and a 'nuncle' (aunt/uncle) of the same gender. We might call nuncles linked by the same gender as the parent using the parallel vs cross terminology as well: the mother's sister is a parallel-aunt, while the father's sister is a cross-aunt. By the same token, the father's brother is a parallel-uncle, while the mother's brother is a cross-uncle. Extending this terminology further, ego's father's brother's children are parallel-cousins, and ego's mother's sister's children are parallel-cousins. By contrast, ego's father's *sister's* children are cross-cousins, and ego's mother's *brother's* children are cross-cousins. (Very often, but not always, cross-cousins can be married, while parallel-cousins cannot). Now, the fact is that 'cross-cousins' and 'parallel-cousins' logically exist in every genealogical tree, and every language. However, we can say that in English the distinction between cross-cousin and parallel-cousin is *syncretized* – the two use the same term 'cousin'. Similarly, the gender of cousins is syncretized in English – while siblings, nephews/nieces, and nuncles show a distinction for [\pm feminine], cousins do not (in English, although they do in French). Other languages may show even further syncretisms – for example, *nipote* in Italian is syncretic between 'nephew, niece, grandson, granddaughter'. And in fact, many languages, such as Mandarin Chinese, show distinct terms for older brother versus younger brother, while the [\pm elder] distinction is syncretized in English.

It is instructive to compare Seneca and English, which have the following clear differences, as characterized by Jones (2010):

(3) Comparative table of terminology:

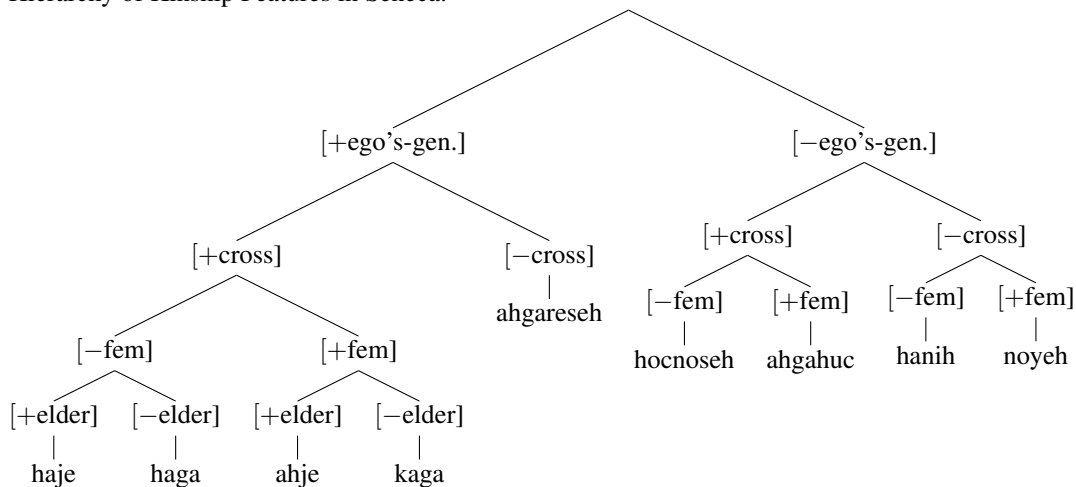
Seneca	English
hocnoseh	cross-uncle (mother's brother)
noyeh	parallel-aunt (mother's sister)
noyeh	mother
hanih	father
hanih	parallel-uncle (father's brother)
ahgahuc	cross-aunt (father's sister)
ahje	older sister
kaga	younger sister
haje	older brother
haga	younger brother
ahje	older parallel female cousin
kaga	younger parallel female cousin
haje	older parallel male cousin
haga	younger parallel male cousin
ahgareseh	cross-cousin (of any age or gender)

We may focus on the following three differences between Seneca and English

- (4) Three differences between Seneca and English kinship terminology:
- A distinction between cross-cousins and parallel-cousins, and between cross-nuncles and parallel-nuncles
 - An elder versus younger distinction in siblings (and a gender distinction in cousins)
 - Sibling/parallel-cousin syncretism, Parent/parallel-nuncle syncretism, and neutralization of all distinctions in cross-cousins

Since Goodenough (1967), differences in kinship systems have been modeled in terms of the Variable Hierarchy, namely in terms of the distinctions that are prioritized and de-prioritized (or, in the limit, never used) within a certain language’s inventory of terminology. Applying the Variable Hierarchy exactly as Drescher has developed it for phonology towards the Seneca kinship system would work as follows. A feature $[\pm \text{elder}]$ would differentiate older and younger siblings, a feature $[\pm \text{cross}]$ would essentially encode the cross/parallel distinction in terms of saying whether the relative in question is linked via the mother or not, and the feature $[\pm \text{ego’s-generation}]$ would state whether the relative is of the same generation as the ego.

- (5) Hierarchy of Kinship Features in Seneca:



A few remarks are in order. First, notice that $[\pm \text{elder}]$ distinctions are not made on the marked branch of the tree, namely $[\pm \text{ego’s generation}]$. The logic here is that ‘grade’ differences (relative age) are only relevant for those of the same generation as the ego. This kind of underspecification is easily achieved in the Successive Division Algorithm. Similarly, cross-cousins (*ahgareseh*) show no age or gender distinction. (It is an interesting question whether these features are completely underspecified for these relatives; in other words, is there anything else in the grammar of kinship that needs to refer to age/gender for these categories?). Finally, we conclude that the English kinship system has certain important differences as would be reflected by Variable Hierarchy, namely 1) it relegates features such as $[\pm \text{elder}]$ to the abyss; 2) it maintains a feature $[\pm \text{collateral}]$, which distinguishes parents from parallel-nuncles and siblings from parallel-cousins, unlike Seneca, which has relegated this feature to the abyss, and 3), it has no parallel/cross distinction, thereby dispensing with $[\pm \text{cross}]$. The one similarity that actually shines through is the word ‘cousin’, which bears no gender/age distinctions, just like the cross-cousin *ahgareseh*. Seneca and English are thus highly similar in this one corner of their kinship grammar. What about the differences? I adopt the point of view that kinship systems are, like phonological systems, a formal system of contrasts. Just as phonological contrast and classificatory categories are distinct from phonetic ‘veridicality’ (they are close, but distinct; see Nevins (2010a, 30-32, 70-76) for my views as applied to vowel harmony), so are kinship contrasts and classifications are distinct from social / taboo ‘veridicality’. Thus, while the Seneca kinship grammar adopts $[\pm \text{cross}]$ and indeed the society is matrilineal according to Jones (2010, 374), we should not leap to the conclusion that kinship terminology directly reflects societal practices, as revealingly epitomized in the quote from Kroeber from over a century ago: “In the Dakota language, according to Riggs, there is only one word for grandfather and father-in-law. Following the mode of reasoning sometimes employed, it might be deduced from this that these two relationships were once identical. Worked out to its implications, the absurd conclusion would be that marriage with the mother was once customary

among the Sioux.” (Kroeber, 1909)

Indeed, even Claude Lévi-Strauss at one point recognized that his Jakobson-inspired system of features for kinship was a purely formal system, and when asked whether it was anti-feminist to say that the mathematical character of the system involved tracking the exchange of women, he emphasized that the importance was the structure of contrasts within the system, quipping “One could just as well say that women exchange men; all you have to do is replace the plus sign with the minus sign and vice-versa – the structure of the system would not change.” (Wilcken, 2010, 163), paralleling arguments in phonology about the formal equivalence of notations such as [+ATR] vs [–RTR]. Stepping back towards the importance of the Successive Division Algorithm and the Variable Hierarchy, I contend that such tools are crucial in deciding what features/distinctions a system renders important/unimportant, and which have scope over others, and that they are extremely useful in modelling crosslinguistic variation given the same inventory of elements classified or grouped in different ways. Finally, pursuing them in detail leads to all kinds of important questions; for example, is there a kind of ‘Bayesian’ bias or set of priors that will guide the learner towards making certain cuts early, all else being equal? For example, it has been proposed that height/stricture is one of the first phonological distinctions a language tends to make (Steriade, 1987), even if this cannot be universalized as was attempted in the work by Jakobson and Halle in the 1950s. As we saw in the discussion above, certain features only ‘make sense’ in having scope over or under particular others, as in the case of [\pm elder], which kinship systems only seem to consistently employ within the ego’s generation. Finally, as the discussion above has revealed, the very fact that Drescher’s Successive Division Algorithm and Variable Hierarchy *can* be applied to kinship systems, a domain far beyond the empirical hearth for which they were originally developed, provides important confirmation for the applicability of this tool extremely generally within the linguistic sciences (kinship terminology, I would contend, being uncontroversially a part of human language, perhaps akin to the functional vocabulary within inflectional systems such as tense, mood and aspect).

4. Vowel Reduction as Pruning the Lower Branches

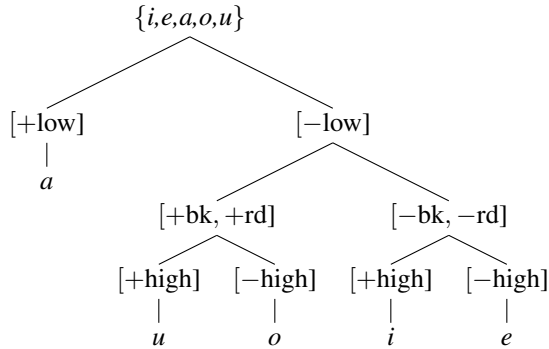
In this section, we turn to yet another novel use of the Variable Hierarchy, namely its ability to model crosslinguistic differences in vowel reduction. As a straightforward case study, I wish to focus on Bulgarian vs Russian vowel reduction, as both are essentially three-height systems that enact different patterns when their mid-vowels become unstressed. The term ‘vowel reduction’, after all, is best understood not as describing the process of *reducing a vowel* (for, as we will see, there is no sense in which an /o/ that becomes an [a] is ‘reduced’ in sonority), but rather describes *reducing the inventory of vowels*, i.e. turning a 5-vowel system of stressed vowels into a 3-vowel system of unstressed vowels. In what follows, I will contrast the ‘centrifugal’ system of vowel reduction in Russian, whereby the mid vowel /o/ becomes the more peripheralized and sonorous vowel [a], with the ‘centripetal’ system of vowel reduction in Bulgarian, whereby the mid vowel /o/ (and in fact all vowels) tend to reduce in sonority, in this case, raising to [u].⁵ These differences are characterized below, though see Crosswhite (2001) and Barnes (2006) for a much more thorough description of many of the additional complexities.

- (6) Bulgarian vs Russian Vowel Reduction:
- a. Bulgarian unstressed /o/ raises to [u]
 - b. Russian unstressed /o/ lowers to [a]

Intuitively, each language chooses whether to “keep” either color (i.e. [\pm round]) or sonority (i.e. [\pm high]). This choice can be formalized in terms of a hierarchy: the lowest cut on the stressed-vowel inventory is the first to be jettisoned. In order to make this clear, I will first demonstrate the hierarchies of the five-vowel systems in question (I’ve ignored schwa in Bulgarian and the unrounded allophone [u̠] in Russian here, which do not affect the overall argument).

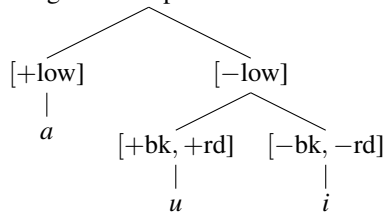
- (7) Bulgarian:

⁵I owe the insightful terms ‘centripetal’ and ‘centrifugal’ to Harris (2005).



The key idea I would like to introduce here, adopting the tools of the Variable Hierarchy, is that *vowel reduction is a process of pruning the lowest node in the hierarchy*. In fact, upon reading a first draft of this paper, Yuni Kim (pers. comm) made me aware that the use of pruning in conjunction with the Variable Hierarchy to explain vowel reduction phenomena has been independently proposed in Spahr (2012), roughly at the time of the initial presentation of this work in Leiden. I have found Spahr’s work (a later published version of which is Spahr (2014)) to be impressive and thorough development of the pruning hypothesis, and highly recommend the reader to consult it in conjunction with the present paper. The idea of vowel reduction as pruning is that there is ‘less space’ for a full set of vowels in unstressed position, and so the hierarchically least-important contrast is reduced/removed from the tree:

(8) Bulgarian VR prunes the lowest node:



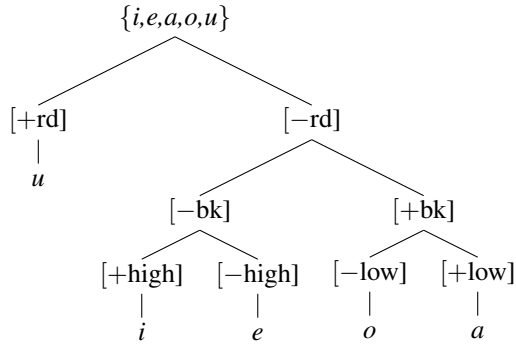
In other words, under the pressure of unstressedness, where not all contrasts can be maintained, the cheapest one to get rid of in Bulgarian is $[\pm \text{high}]$, as determined by the Variable Hierarchy. As a result, the $[\pm \text{high}]$ contrast is collapsed (and the two vowels that sit on this branch are neutralized to their unmarked value), and Bulgarian /o/ is neutralized to [u] under unstressedness. At this point, the pruning hypothesis alone does not determine in which direction the neutralization will go. At present, I take this to be a virtue, as neutralization *could* presumably be in either direction, or to some phonetically implemented value somewhere in the middle. Indeed, Spahr (2014), following Scatton (1984), points out that the reduced vowels can be intermediate [ɪ,ʊ] in height, which would be consistent with maintaining only a [-low] specification alongside their color.⁶

By contrast, Russian paints itself out of the corner in another way – again, by hypothesis, as determined by its hierarchy. As there are many different sequences of cuts which could be employed to contrastively characterize a 5-vowel system, the one below places the $[\pm \text{low}]$ contrast lowest for /o/ vs /a/, which are neutralized; it also places the $[\pm \text{high}]$ contrast lowest for /i/ vs /e/.⁷

(9) Russian tree:

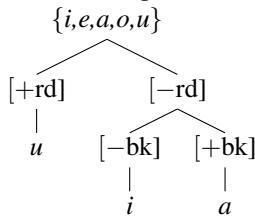
⁶Even more interestingly, Spahr (2014) suggests that the ‘non-terminal’ nodes that result from pruning, as for example the two resulting lower nodes in (8), may have their exact phonetic realization implemented by the addition of noncontrastive features in a postlexical block of computation – a point we will return to in Section 7.

⁷Many different hierarchies could account for the neutralization between /o/ and /a/, including $[\pm \text{high}] > [\pm \text{back}] > [\pm \text{round}]$. The one in the text allows both the /o ~ a/ neutralization and the /i ~ e/ neutralization to take place on the same level; I thank an anonymous reviewer for this suggestion. For discussion of the non-specification of /o/ as $[\pm \text{round}]$, see Spahr (2014), who follows Iosad (2012) in characterizing Russian /u/ and /o/ as phonologically orthogonal.



As a result, when the inventory must be reduced, and hence the tree must be pruned, it is the lowest one down that is the first to go:

- (10) Russian VR prunes the lowest node:



Thus, when the $[\pm \text{ low}]$ contrast is lost, Russian /o/ is neutralized to [a] under unstressedness. (Further reduction and allophony of all of these vowels is determined by the $[\pm \text{ back}]$ value of adjacent consonants). While the non-terminal nodes that remain from pruning do not specify the exact phonetic values of the neutralized result, I assume that this allows for intermediate realizations, and that some degree of the detail of phonetic implementation of these non-terminal nodes results from post-lexical specifications (as will be discussed in Section 7).⁸

Summing up more broadly the fact that these two languages enact vowel reduction ‘centripetally’ and ‘centrifugally’ in a way that one pushes /o/ upwards (neutralizing $[\pm \text{ high}]$) while the other pushes it downwards (neutralizing $[\pm \text{ low}]$) can be straightforwardly captured in terms of which feature specifications are least important to keep. In sum, a contrastive hierarchy may be useful not only in terms of order-of-acquisition, and determining the specification of items for phonological rules, but also for determining what contrasts are of lowest priority when vowel reduction forces a reduction in the inventory. In the next section, I would like to demonstrate the further applicability of this modelling tool in what could be called an even more massive type of ‘vowel reduction’.

5. Extreme Reduction: Whistled Languages

Whistled languages are some of the world’s most endangered language types (see Kouneli et al. (2013)) as they employ a ‘natural telecommunication system’, namely whistling, to send messages clearly at up to half a kilometer’s distance; nonetheless, with cellphone technology, this broadcast communication system is losing intergenerational transmission. Whistled languages, when based on non-tonal languages, involve mappings of the segmental system of

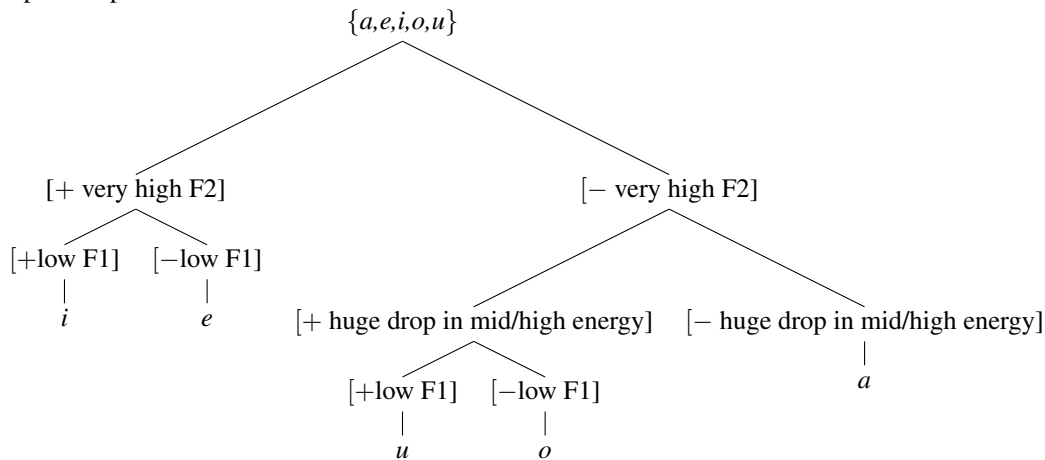
⁸I have employed binary features in the diagrams above and throughout this paper, and for the purposes of the present discussion will have little to say about unary-or privative-featured models. Of course the question arises as to how the Variable Hierarchy hypothesis could be modeled with unary features, and indeed with primes of the type in Element Theory (ET; Backley 2011). Assuming that Bulgarian and Russian /o/ would be represented by the combination $[A, U]$, then under vowel reduction, the former would remove $[A]$ and the latter would remove $[U]$, and indeed a Variable Hierarchy might be one way to implement this crosslinguistic difference in terms of pruning the ‘lower’ of the two. On the other hand, many of the questions of the language-specific phonological relevance of features present in a phonetic representation do not arise in ET, as ET’s abstraction away from language-particular phonetics is made before one even gets to specific languages, and as each prime is argued to pronounceable in isolation, its negative counterpart cannot be represented in terms of branching. Carvalho (2011) avers that “contrastive hierarchy theory runs the risk of being rejected as an artefact of binarism”. Artefact or not, in an ET system, the set of elements is so limited it may not need to decide what’s ‘unimportant’ to exclude in a given language, and as Drescher (2009, p.33) notes, “The effect of feature ordering is greatly reduced with privative features as opposed to equipollent features”.

the spoken language to a specific set of whistled frequencies, and in general the vowel system (and formant transitions) of the spoken language are mapped onto the whistled language. Thus, these whistled languages essentially have the same syntax, morphology, and syllable structure as their spoken equivalent, but the ‘channel’ for segmental contrasts is accomplished entirely via whistling, which acoustically does not have the source/filter distinction of spoken speech with formants that modulate the harmonic structure of f_0 , but rather only a single fundamental frequency and series of harmonics, call it H_0 , with no formants. As a result, all of the aspects of spoken speech encoded into formants are ‘competing’ for a single channel and a single acoustic dimension, and some decisions about what to throw away and what to keep must inevitably be made. As such, we might view whistled languages as ‘extreme vowel reduction’, as the reduction of the inventory takes place *in every syllable*, not just unstressed syllables.

What I would like to propose here is that the same mechanisms at work in unstressed vowel reduction are taking place when listeners must decide which aspects of their vowel system to cut.⁹ Namely, as whistled speech cannot reliably support 5 pitch distinctions (and indeed, few languages have 5 level tones), the whistled versions of 5-vowel languages must chuck out certain contrasts. The comparison I will make here is between Whistled Spanish, as found on the Canary Island of La Gomera, and Whistled Greek, as found in the village of Antia on Evia Island. Both of these are based on spoken languages with the same 5-vowel inventory, but nonetheless, their whistled versions collapse different aspects of the vowel system.

In particular, Whistled Greek neutralizes the vowels /e,u/ as a pair, as determined by analysis of the whistled tokens of each in Meyer (2008) (in addition to neutralizing /a/ and /o/). By contrast, Whistled Spanish collapses /o,u/, in the analysis of Rialland (2005). While the pattern of Spanish is understandable in terms of a number of feature systems, the neutralization of /e/ and /u/ in Greek is extremely hard to make sense of in terms of any articulatory features, as they differ in height, backness, and roundness. Nonetheless, there are reasons to think whistled speech strives for a kind of iconicity with the acoustic speech signal. I will thus assume acoustic features are at play in characterizing the phonological contrasts in whistled languages, and I will specifically adopt a set of both absolute values for acoustic landmarks in the spectrum and a characterization of the regions of overall concentration of energy (Jakobson et al. (1952), Backley (2011), Chistovitch and Lublinskaja (1979), Schwartz and Escudier (1989)). Specifically, I will assume that in Spanish, acoustic properties referring to the lowness of particular formants are crucial in characterizing the system, alongside a feature that characterizes a huge drop in mid-to-high energies in the spectrum (known as the ‘rump’ signature of roundness in Harris and Lindsey (1995)). In the Variable Hierarchy trees below, by convention I place the positive (+) value to the left and the negative (–) value to the right in each binary cut.

(11) Spanish spoken vowels tree

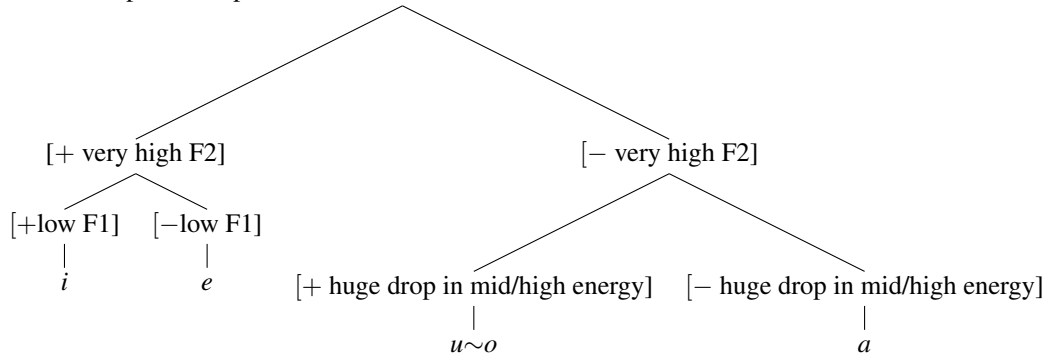


The idea here now becomes very straightforward: Whistled Spanish must deprioritize the lowest (or, in terms of acquisition, latest) cut made within the tree. As a result, the contrast between /u/ and /o/ of the spoken language is neutralized (a fact which indeed may reflect some of the phonology of the spoken Spanish of Gomera, in which,

⁹This decision is all the more complex once the correlates of stress are included in the equation, particularly in Modern Greek, where stress bears a very high functional load, and f_0 and formants must also compete; see Mussa (2010) for discussion.

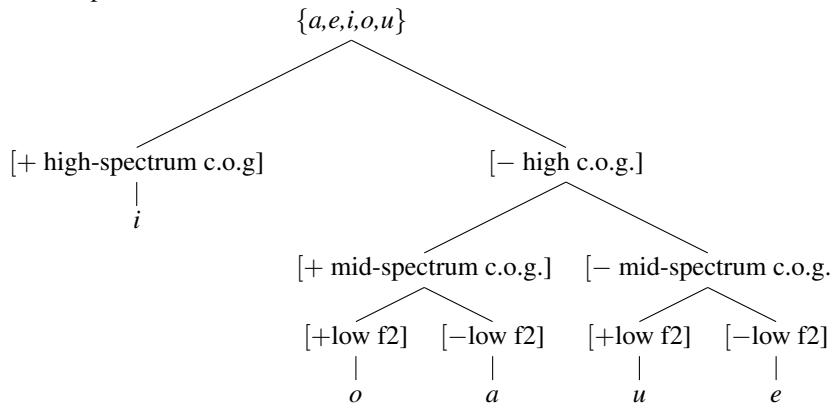
according to Julien Meyer, the /o~u/ contrast has a low functional load). Note, importantly, that it is the lowest (or latest-made) *cut* that is eliminated, and not the feature per se. Thus in the tree below, while [\pm low F1] is retained on a higher branch, it is removed on a lower branch. This emphasizes the importance of *hierarchical* contrast:

(12) Whistled Spanish Deprioritizes Lowest Cut:



While the Spanish hierarchy is implemented largely in terms of direct formant values themselves, the Greek hierarchy is based more on the notion of a spectral ‘centre of gravity’ (c.o.g.), i.e. the region in which the most energy is concentrated, sometimes resulting in ‘super formants’ that are a blurred combination of, say f2 and f3 (in the case of /i/). Adopting these types of features, we may construct the following hierarchy:

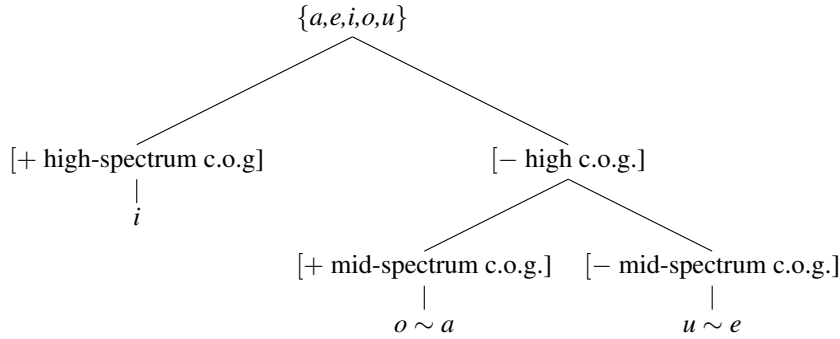
(13) Greek spoken vowels tree:



As a result, we can now understand the somewhat puzzling neutralization of /e,u/ in Whistled Greek: these two vowels are the strange bedfellows that share the property of neither having a very high f2 (which sets /i/ apart in a great many whistled systems) nor having a compressed f1 and f2 (characterizing the center of gravity shared by /o,a/).¹⁰ As a result, the low-ranking distinction between them, made in terms of raw f2 value in a system otherwise characterized by centres-of-gravity, is the first to go when the channel imposes a stricter throughput:

(14) Whistled Greek Deprioritizes Lowest Cut:

¹⁰Apparently, the mnemonics used in Korean solfeges employ similiar groupings to whistled Greek (Kenneth Dejong, pers. comm.), although I have not yet been able to explore this.



The comparison between these two 5-vowel systems is, I contend, revealing, in that there is no way other than a kind of Variable Hierarchy hypothesis of some sort in order to understand why the neutralization patterns are different across them, whether they be done with the explicit rigor of the hierarchies above or in terms of faithfulness constraints that choose to keep or discard certain aspects of the acoustic signal over others. Note, however, that as the neutralization, or co-patterning of /e/ and /u/ is extremely rare, if at all present in spoken languages, it may be that acoustic features of this type are specific to whistled language phonologies and the iconic mapping of the speech signal to a reduced channel with strikingly distinct articulation but an attempt to emulate the formant structure of the spoken variant (with a variable hierarchical way of doing so).¹¹

Interestingly enough, there are not many other 5-vowel whistled systems to compare. This is in part because whistled languages spring up in particular ecological niches with the right types of population activities and geography, which don't necessarily overlap with the distribution of 5-vowel systems. Turkish, for instance, also has a whistled variety in the village of Kuşköy, where the 8-vowels of Turkish must be massively reduced, but as Meyer (2008) points out, much of the potential confusion that would result from this neutralization is avoided given the predictable phonological restrictions imposed by vowel harmony. However, most dramatic is the fact that it is apparently a universal, according to Bagemihl (1988), that whenever the corresponding spoken language is a tone language, the whistled variant will always keep the tones at highest priority and throw away the segmental vowel features if needed. This is a striking decision, given that the functional load of tone, in Bagemihl's calculations, is often much lower than that of the consonants and/or vowels that are discarded instead. In a way, it points to one of the 'Bayesian' priors alluded to above, that apparently whenever possible, a whistled language's contrastive hierarchy will place tonal features above all segmental vowel features.¹²

In sum, vowel reduction and whistled transformations are irrevocably hierarchical in their decisions about 'what to keep', and such decisions are excellently modelled by hierarchical cuts with a notion of pruning the least important (or last-made) cut. As such, the Variable Hierarchy is an excellent model for making explicit the aspects of phonological organization that are often left implicit in many analyses, and have novel applications to understanding not only what the specification of certain segments is, but, given its explicitly hierarchical organization, also 'how important' each specification is in a relative sense.

6. The 'Oops, I Need That' Problem

We now turn to the second, and arguably logically independent part of the Contrastive Feature Hypothesis of Dresher (2009), namely the Contrastivity-Only hypothesis – that only the contrastive values that are expressed through a pass of the Successive Division Algorithm are active in the phonology of a language. In other words, the Contrastivity-Only hypothesis states that 'redundant' features (such as [\pm round] in a 3-vowel system where only [\pm back, \pm high] are needed to build the Variable Hierarchy) will never be active in the phonology of such a language. In particular, 3-vowel and 5-vowel systems are some of the most straightforward to look at in this sense, as the prediction is that [\pm round]

¹¹While the Greek whistlers with whom we conducted our fieldwork insisted that implementing the whistled version of another language – say, them attempting to whistle English or French – would work identically to whistling Greek, one cannot help but think that there might be such a thing as a 'foreign accent' in the whistled version of a language that has a different mapping (and hence a different hierarchy) from its acoustic speech signal to its respective reduced contrastive space of whistling.

¹²This seems to be largely true of drummed languages as well, although the Bâtá drum tradition, while using the right hand to drum tones from Yoruba, uses a whip in the left hand to encode [–high] vowels in its 7-vowel system (Debbie Klein, pers. comm.).

and [\pm back] can never *both* be active at the same time at the same vowel height in the same language with a 3-vowel or 5-vowel system.¹³ I have discussed this prediction for 5-vowel systems in Huave in Nevins (2010a, 214), and the point can be also presumably made for any 3-vowel or 5-vowel language that has both palatalization and labialization. In this sense, Japanese looks like a straightforward candidate, as it has the well-known palatalization of /s/ by /i/, and the allophonic process whereby /h/ becomes [ç] before /u/ (Vance, 1987; Tsujimura, 1996). However, given the fact that the realization of Japanese /u/ is often more like [u], it is a slightly trickier case to make that the feature in question behind the [h~ç] allophony is indeed [+round]. My focus, therefore, will be on cases where contrastivity-only has been employed in the analysis of processes that show supposed phonological invisibility, namely vowel harmony. In Nevins (2010a) I adopted the parametric approach to phonological visibility of Calabrese (1995) in order to deal with cases where a feature on a vowel is invisible for say, vowel harmony, but cannot be said to lack this feature altogether in the language once the phonology as a whole is considered.

In other words, just because a feature *F* is non-active for a process *P* doesn't mean it will be non-active for a process *Q*. Looking "only" at vowel harmony without looking at other processes in the whole phonology of a language might lead one to say that the vowel /i/ lacks the feature [–back] – indeed, Hall (2011, p.15) falls prey to exactly this temptation, but as I have discussed (Nevins, 2010a, p.81), Finnish needs to have the feature [–back] on the vowel /i/ in order to trigger assibilation.¹⁴ Similarly to Finnish, as Blumenfeld and Toivonen (2014) point out, in Votic harmony, the vowel /i/ is transparent for [\pm back] harmony across it (e.g. *ko:kkima* 'to dig'), suggesting it is not specified for this feature, while the same feature in the same vowel conditions the [–back] allophone of the lateral (e.g. *ma:lima* 'paint'), which has a dental vs velar allophone, the former conditioned exactly by the set of [–back] vowels, whether contrastive for this feature or not. Blumenfeld and Toivonen (2014) conclude that "a vowel's feature may have a double identity, active in one process and neutral in another".

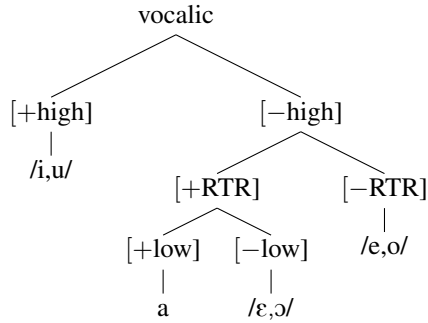
In order to put the nature of this pitfall of Contrastivity-Only into sharp relief, I call this the "Oops, I Need That" Problem – namely, while it is convenient to set up the contrastive hierarchy such that segment *S* lacks feature *F* for process *P*, it turns out that once the phonology of the language as a whole is looked at, indeed, one needs to have *F* on *S*. (As such, the issue is reminiscent of the problems for underspecification of coronals raised in McCarthy and Taub (1992), albeit in a different context, as Dresher's model is based on a very different kind of underspecification, namely one with a Variable Hierarchy – in which coronals are not necessarily the universally underspecified place of articulation.)

One of the cases that Dresher has focused on recently is Yoruba vowel harmony, precisely because in Nevins (2010a) I used the difference between Standard Yoruba and Ifẹ Yoruba as an argument for the parametric visibility mentioned above. Specifically, in Standard Yoruba, all vowels participate in RTR harmony (I will call it RTR, rather than ATR harmony, simply to maintain the same terminology as Dresher), while in Ifẹ Yoruba, the high vowels /i,u/ are invisible to ATR harmony. In a valiant attempt to re-analyze this case, Dresher (2013) argues that in Ifẹ Yoruba, the hierarchy is such that /i,u/ are underspecified for [RTR]: the cuts are such that they are separated off from the other 5 vowels by the highest cut, that of [\pm high].

(15) Dresher's Tree for Ifẹ Yoruba:

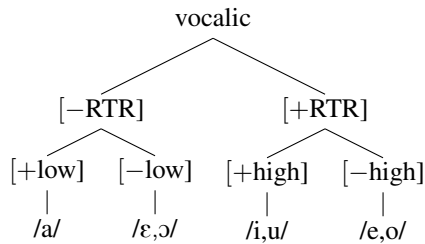
¹³There is, in principle, a way of specifying only [\pm round] and [\pm back] in a 3-vowel system, without specifying height at all, but I follow Jakobson et al. (1952), Steriade (1987), and in particular Harms (1987) in insisting that all vowel systems must distinguish at least one degree of height (and indeed, perhaps two-vowel systems, such as Kabardian, *only* distinguish height): "The burden of phonological distinctiveness is greater over those harmonics with the greatest energy . . . in other words, if God had wanted man to make greater use of higher frequencies, he would have (a) put more energy there or (b) made our basilar membrane significantly more sensitive to those frequencies" (Harms 1987:389).

¹⁴While [–back] alone is not the feature that distinguishes /t/ from /s/, I adopt the theory of Calabrese (2005, Ch. 4), under which universal correlational statements such as *Tongue Body* [+high, –back] ↔ *Tongue Blade* [–anterior, +distributed] are followed by language-particular, inventory based repairs that yield the varying outputs of palatalization such as [tʃ, s, ʃ, ç].



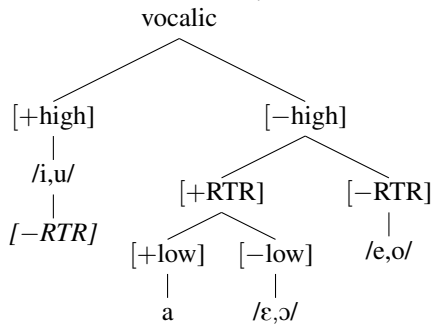
However, as the high vowels *are* visible for harmony in Standard Yoruba, its tree is as follows:

(16) Drescher’s Tree for Standard Yoruba:



Notably, the tree for Standard Yoruba fails to group the 5 minus high vowels /e,o,ε,ɔ, a/ as a natural class (even though we have seen that the Bâtá drum language groups these together). In fact, internal to the phonology of Yoruba, we can find evidence that [+high] is a natural class and [-high] is a natural class: Akinlabi (1993, p.144-145) discusses the process of *r*-deletion, which happens when adjacent to a [+high] vowel, and he also discusses the process of progressive vowel assimilation within the lexical (as opposed to postlexical) stratum, for which he must “refer to the class of nonhigh vowels” (p.151). In other words, even Standard Yoruba requires [-high] as a natural class for phenomena besides vowel harmony. There doesn’t seem to be any way out of the “Oops, I Need That” problem once we consider all three phonological processes. We certainly don’t want to say that Standard Yoruba has the hierarchical cut for Standard Yoruba in (16) when it does vowel harmony but that it has the hierarchical cut for Ifẹ Yoruba (15) when it does hiatus resolution and *r*-deletion. Learners who encounter vowel harmony in Yoruba alongside assimilation and *r*-deletion presumably have the same sort of tree for both Ifẹ and Standard Yoruba, but the additional, noncontrastive [± RTR] specification on the high vowels is still present to refer to in the latter. I notate the additional presence of noncontrastive feature specification below with italics, below the segments that are the leaves of the tree themselves.

(17) Shared Tree for both Ifẹ and Standard Yoruba:



On this approach, the Variable Hierarchy can indeed determine the relative importance of contrast and what is contrastive, but additionally noncontrastive features can still be present on segments, and referred to by particular phonological processes. Under the parametrized visibility hypothesis (see e.g. Calabrese (1995, 2005); Nevins (2010a)), phonological rules may parametrize whether they are sensitive to contrastive-values-only or to *all* values of a feature

(including the redundant or, as I prefer to call them, noncontrastive values, indicated in italics as above). This kind of parametrization allows for expressing the differential sensitivity of certain processes to contrastiveness or not:

- (18) Parametric settings of visibility:
- a. VH in S.Yoruba: *all-values* of [\pm RTR]
 - b. VH in I.Yoruba: *contrastive-only* of [\pm RTR]
 - c. r-deletion/vowel assimilation in IY and SY: *all-values* of [\pm high]
 - d. VH in Finnish: *contrastive-only* of [\pm back]
 - e. Assibilation in Finnish: *all-values* of [\pm back]

I am keen to point out that the identification of the “Oops, I Need That” problem is not specific to my attempt to preserve my analysis of vowel harmony. For example, Scheer (2010) points out that Germanic final devoicing requires sonorants to be cut apart from obstruents early. But Germanic vowel lengthening requires grouping voiced stops and sonorants as a natural class. In other words, prevoiced lengthening is an *all-values* sensitive process, even for noncontrastive instances of [\pm voice]. For this reason, it seems preferable to maintain the theory of parametric visibility discussed above: some phonological processes are sensitive to contrastivity-only, whereas others allow in noncontrastive values.

Now, all parameter-based theories in linguistics must first set up the descriptive adequacy of the parameters (*n* parameters exist; we have proof for microvariation along all and only these lines). The second step, only after doing so, is to then set up the explanatory adequacy: does a given parameter setting correlate with anything else? We are not yet at this step for the parametric visibility theory above, largely because there is not yet even agreement that this kind of parametric variation exists, under the Contrastivity-Only hypothesis. But once the cases of noncontrastive activity that pop up again and again in a language are admitted, we can move towards the descriptive adequacy step in (18), with the goal of then arriving at explanatory adequacy of why certain processes are parametrized as they are. It is to this speculative step that I wish to turn in the next section.

7. Emerging Correlational Landmarks

If any single one of the arguments from Huave, Japanese, Finnish, Yoruba, or Germanic above are immune to reanalysis, then the Contrastivity-Only hypothesis, valiant as it is, is empirically insufficient. Would that mean we have to discard all of the virtues of the Contrastive Feature Hypothesis explored in Dresher’s work and in the case studies above? I would argue that we can replace the approach to contrastive definition that I had in Nevins (2010a, p.70) in terms of pairwise contrast with Dresher’s Variable Hierarchy trees, as established by the Successive Division Algorithm, as the latter have the added virtue of being able to define *which contrasts are more important than others*, a distinction that I have argued becomes crucially relevant in understanding processes such as vowel reduction. Thus, the Successive Division Algorithm can be used to establish the specifications of contrast within an inventory, and then a second pass through the tree can take place in order to fill in the ‘redundant’, noncontrastive values – in fact, such a move has already been made in Hall (2011), who recognizes the importance of Enhancement (Stevens and Keyser, 1989) in phonological specification. Enhancement specifically refers to the fact that a noncontrastive feature is recruited in order to – for lack of a better word – *enhance* a phonological contrast. Thus, the [\pm round] nature of /u/ in a 5-vowel system may in fact be *secondary* to the [\pm high] distinction – and under this view, the cut for [\pm back] would be made first, but a subsequent traversal of the tree would add [+ round] to the [–back] vowels of the language. It is thus entirely possible to maintain the trees that express Contrastivity-Only, and then enhance them with noncontrastive features (such as the italicized [–RTR] under the Yoruba high vowels in (17) that can be referred to when a given phonological process is parametrically set to *all-values*. We then return to the question of a move towards explanatory adequacy: what does the setting of a given parameter for a given process correlate with?

By and large, it seems that long-distance, vowel-to-vowel processes are *contrastive-only*; it is for this reason that vowel harmony has been the bread-and-butter of many underspecification approaches (again, especially when they look only at this process in a given language). On the other hand, short-distance, consonant-consonant, vowel-vowel, and consonant-vowel processes are often *all-values*. Indeed, Votic, as mentioned above, seems to work precisely in this way, whereby the local CV interaction of lateral allophony is sensitive to noncontrastive features, while the non-local vowel harmony is relativized to contrastive features. As such, a research program emerges: taking parametrization

seriously and seeing whether it correlates with adjacency and trigger/target type. (Note that there is a third parametrization value in Calabrese (1995); Nevins (2010a), namely *marked-values*. As this is a *more restrictive* subset than even *contrastive-only*, it is orthogonal to the empirical limitations of Contrastivity-Only).

Other correlational landmarks are possible as well. Kiparsky (1985) provided an analysis of voicing contrasts and assimilation in Russian obstruents, focusing especially on the noncontrastive *ts, ch, x* famous from Halle (1959). In Kiparsky's analysis, *contrastive-values* of obstruent [\pm voice] were active in the lexical phonology, while *all-values* of obstruent [\pm voice] were active in the postlexical phonology.¹⁵ The way that feature-changing assimilation was implemented in Nevins (2010a, p.58) was to have a two-step process, namely postcyclic deletion of the [\pm voice] in *all* nonfinal obstruents, followed by all-value [\pm voice] copying. As such, yet another strand of research is possible, namely to investigate to what extent the departures from Contrastivity-Only occur in a later stratum of phonological computation, perhaps precisely after the step of adding noncontrastive values via Enhancement occurs. In fact, in the Yoruba discussion above, Akinlabi (1993) specifically appealed to underspecification in the lexical phonology with noncontrastive fill-in at the postlexical phonology for demonstrably postlexical processes. While certain processes (e.g. Votic lateral allophony) may not be demonstrably postlexical, the correlational 'landmarks' that I suggest may coincide with the restricted parameter settings above could take into account potentially orthogonal properties of the processes in question, including trigger/target type, distance, and morphological stratum.

We might therefore say that there are two classes of features: those that are contrastive, and those that are non-contrastive but still specified. The former are determined by the contrastive hierarchy, while the latter would have to be determined by a second 'pass' through the tree, assigning non-contrastive features to particular segments (perhaps in the way that Hall (2011) does for Enhancement). In the analysis of Yoruba I suggested above in (17), a noncontrastive specification of [-RTR] was added to the high vowels (in this case, not necessarily in a postlexical block). The two languages both had the same sets of specifications, and simply differ in whether noncontrastive features were visible or not to vowel harmony, a parametrization of visibility that seems independently needed for many phenomena.

In fact, precisely to deal with the fact that Japanese /u/ *does* require a color feature ([peripheral], in her work) distinct from backness, Hirayama (2003) proposes that postlexical processes may require the specification (and hence addition) of features beyond those needed for establishing contrast alone. I would contend that it is not clear that all of the case of noncontrastive feature activity can be argued to be postlexical – for example, the Votic case discussed above, dental vs velar allophony among the laterals does not have any characteristics that would make it seem derivationally 'later' than vowel harmony. To avoid the need for positing derivational stages whereby all noncontrastive-referring processes are necessarily later than contrastivity-only ones, I have argued that the visibility-parametrization hypothesis localizes this variability in one restrictive place, and have attempted to demonstrate that visibility parameters of exactly this type can explain microvariation in morphosyntax as well – in particular, different variants of the 'person case constraint' on pronominal clitic combinations (Nevins, 2007).

The Variable-Hierarchy hypothesis thus affords a way of describing the relative importance of particular contrasts, and indeed the fact that certain contrasts (e.g. [\pm low] in Turkish) are never employed at all, but an additional 'round' of specification of noncontrastive features to certain terminal nodes seems necessary (and even to non-terminal nodes, within the model of Spahr (2014), where it is argued that the precise value of the neutralized front vowel in Russian may be determined by a 'later pass' of noncontrastive feature assignment). Lest you say, if we have to go back through and assign non-contrastive features anyway, what is the advantage of using the contrastive hierarchy in the first place, let me remind you that in this paper I have presented another use for it, namely 'how contrastive' or 'how important' a certain feature is within a language, a status which becomes important for characterizing cross-linguistic differences in reduction processes.

The general conclusion I am to provide is to emphasize that hierarchical cuts are implicit in all phonological analyses, and that making them explicit can be crucial in deciding what features and distinctions a system renders important or unimportant, and which have scope over others, and as such, that the Variable Hierarchy is extremely useful in modelling crosslinguistic variation given the same inventory of elements that can be classified or grouped in different ways. Dresher (2015) points out that the Contrastive hypothesis is motivated in part by two principles – *Activity* and *Minimality*. The Variable Hierarchy is an excellent way of implementing Activity, and formalizing the way in which features such as [\pm low] are simply never active in the phonology of Turkish. The Contrastivity-Only hypothesis as

¹⁵Interestingly, however, Dresher (2015) presents an analysis in which [\pm voice] is indeed contrastive for Russian *ts, ch, x*, despite the absence of their voiced counterparts in the language.

an implementation of Minimality, however, arguably places too much emphasis on minimizing redundancy in phonological representations, as certain segments irrevocably bear feature specifications above and beyond what they need for contrast alone. A potential way forward is to see whether the limitations of the Contrastivity-Only hypothesis may be solved by embracing a parametrization of processes into contrastive-only vs all-values, where the latter enables the activity of certain noncontrastive specifications to be added to a segment (e.g. [–RTR] on Yoruba high vowels), and subsequently exploring how the setting of such parameters interact with distance parameters, the nature of the trigger and target, and the morphological stratum in which these phonological processes are localized.

References

- Akinlabi, A. (1993). Underspecification and the Phonology of Yoruba /t/. *Linguistic Inquiry*, 24:139–160.
- Backley, P. (2011). *An Introduction to Element Theory*. Edinburgh University Press.
- Bagemihl, B. (1988). *Alternative Phonologies and Morphologies*. PhD thesis, University of British Columbia.
- Barnes, J. (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Mouton de Gruyter, Berlin/New York.
- Beckman, J., Helgason, P., McMurray, B., and Ringen, C. (2011). Rate Effects on Swedish VOT: Evidence for Phonological Overspecification. *Journal of Phonetics*, 39:39–49.
- Blumenfeld, L. and Toivonen, I. (2014). A Featural Paradox in Votic Harmony. Ms., Carleton College.
- Calabrese, A. (1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*, 26:373–463.
- Calabrese, A. (2005). *Markedness and economy in a derivational model of phonology*. Mouton de Gruyter.
- Carvalho, J. B. (2011). Contrastive hierarchies, privative features, and Portuguese vowels. *Linguística : Revista de Estudos Linguísticos da Universidade do Porto*, 6:51–66.
- Chistovitch, L. and Lublinskaja, V. V. (1979). The center of gravity effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research*, 1:185–195.
- Clements, G. (2001). Representational economy in constraint-based phonology. In *Distinctive feature theory*, pages 71–146. Mouton de Gruyter.
- Crosswhite, K. (2001). *Vowel Reduction in Optimality Theory*. New York: Routledge.
- Dresher, B. E. (2009). *The Contrastive Hierarchy in Phonology*. Cambridge University Press.
- Dresher, B. E. (2013). Contrastive features and microvariation in vowel harmony. In *NELS 42: Proceedings of the Forty-Second Annual Meeting of the North East Linguistic Society, University of Toronto, Volume One*, pages 141–153. GLSA, University of Massachusetts, Amherst.
- Dresher, E. (2015). The Motivation for Contrastive Feature Hierarchies in Phonology. *Linguistic Variation*, (this volume).
- Goodenough, W. (1967). Componential Analysis. *Science*, 67:1203–1209.
- Hall, D. C. (2011). Phonological contrast and its phonetic enhancement: dispersedness without dispersion. *Phonology*, 28:1–54.
- Halle, M. (1959). *The Sound Pattern of Russian*. The Hague: Mouton.
- Harms, R. (1987). What Helmholtz Knew About Neutral Vowels. In *A festschrift for Ilse Lehiste*, pages 381–399.

- Harris, J. (2005). Vowel reduction as information loss. In *Headhood, elements, specification and contrastivity*, pages 119–132. John Benjamins.
- Harris, J. and Lindsey, G. (1995). The elements of phonological representation. In *Frontiers of phonology: atoms, structures, derivations*, pages 34–79. Harlow, Essex: Longman.
- Hirayama, M. (2003). Contrast in Japanese vowels. In *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology) 20*, pages 115–132. Department of Linguistics, University of Toronto.
- Inkelas, S. and Orgun, C. O. (1995). Level ordering and economy in the lexical phonology of Turkish. *Language*, 71(4):763–793.
- Iosad, P. (2012). Vowel reduction in Russian: No phonetics in phonology. *Journal of Linguistics*, 48:521–571.
- Jakobson, R., Fant, G., and Halle, M. (1952). *Preliminaries to Speech Analysis*. MIT Press.
- Jones, D. (2010). Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, 33.5:367–381.
- Kiparsky, P. (1985). Some Consequences of Lexical Phonology. *Phonology Yearbook*, 2:85–138.
- Kouneli, M., Meyer, J., and Nevins, A. (2013). Whistled languages: Including greek in the continuum of endangerment situations and revitalisation strategies. In *Keeping Languages Alive: Documentation, Pedagogy and Revitalisation*, pages 198–211. Cambridge University Press.
- Krämer, M. (2003). *Vowel Harmony and Correspondence Theory*. Mouton de Gruyter.
- Kroeber, A. (1909). Classificatory Systems of Relationship. *Journal of The Royal Anthropological Institute*, 39:77–84.
- McCarthy, J. and Taub, A. (1992). Review of Paradis & Prunet (eds.), *The Special Status of Coronals*. *Phonology*, 9:363–370.
- Meyer, J. (2008). Typology and acoustic strategies of whistled languages: Phonetic comparison and perceptual cues of whistled vowels. *Journal of the International Phonetic Association*, 38:69–94.
- Mussa, H. (2010). *An Analysis of the Greek Whistled Vowels of Evia*. Master's thesis, University College London.
- Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory*, 25(2):273–313.
- Nevins, A. (2010a). *Locality in Vowel Harmony*. MIT Press.
- Nevins, A. (2010b). The Applicability of Theories of Phonological Contrast to Kinship Systems. *Behavioral and Brain Sciences*, 33.5:398–399.
- Nicolae, A. and Nevins, A. (2015). Fricative patterning in aspirating versus true voice languages. *Journal of Linguistics*, to appear in 52.2:Preprint available, <http://ling.auf.net/lingbuzz/002156>.
- Padgett, J. (2002). Feature Classes in Phonology. *Language*, 78.2.
- Rialland, A. (2005). Phonological and phonetic aspects of whistled languages. *Phonology*, 22:237–271.
- Sapir, E. (1925). Sound patterns in language. *Language*, 1:37–51.
- Scatton, E. A. (1984). *A reference grammar of modern Bulgarian*. Columbus, OH: Slavica Publishers, Inc.
- Scheer, T. (2010). How to marry (structuralist) contrast and (generative) processing: Review of Drescher (2009). *Lingua*, 120:2522–2534.

- Schwartz, J.-L. and Escudier, P. (1989). A strong evidence for the existence of a large scale integrated spectral representation in vowel perception. *Speech Communication*, 8:235–259.
- Spahr, C. (2012). Positional neutralization in the Contrastive Hierarchy: The case of phonological vowel reduction. University of Toronto Generals Paper, available at <http://individual.utoronto.ca/spahr/>.
- Spahr, C. (2014). A contrastive hierarchical account of positional neutralization. *The Linguistic Review*, 31:551–585.
- Steriade, D. (1987). Redundant values. In Bosch, A., Need, B., and E.E.Schiller, editors, *Papers from the Parasession on Autosegmental and Metrical Phonology*, pages 339–362. Chicago: Chicago Linguistic Society.
- Stevens, K. and Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language*, 65.1:207–236.
- Tsujimura, N. (1996). *An Introduction to Japanese Linguistics*. Oxford: Blackwell.
- van der Hulst, H. and van de Weijer, J. (1995). Vowel Harmony. In Goldsmith, J., editor, *The Handbook of Phonological Theory*, pages 495–534. Blackwell.
- Vance, T. (1987). *An Introduction to Japanese Phonology*. Albany: SUNY Press.
- Vaux, B. (1998). The Laryngeal Specification of Fricatives. *Linguistic Inquiry*, 29:497–511.
- Wilcken, P. (2010). *Claude Lévi-Strauss: The Poet in the Laboratory*. London: Bloomsbury.

Author's address:
Department of Linguistics
University College London
2 Wakefield Street
London, WC1N 1PF, UK