

## RESEARCH

## Marginal contrasts and the Contrastivist Hypothesis

Daniel Currie Hall<sup>1</sup> and Kathleen Currie Hall<sup>2</sup><sup>1</sup> Saint Mary's University, 923 Robie Street, Halifax, NS B3H 3C3, Canada<sup>2</sup> University of British Columbia, 2613 West Mall, Vancouver, BC V6T 1Z4, CanadaCorresponding authors: Daniel Currie Hall ([daniel.hall@utoronto.ca](mailto:daniel.hall@utoronto.ca)), Kathleen Currie Hall ([kathleen.hall@ubc.ca](mailto:kathleen.hall@ubc.ca))

The Contrastivist Hypothesis (CH; Hall 2007; Dresher 2009) holds that the only features that can be phonologically active in any language are those that serve to distinguish phonemes, which presupposes that phonemic status is categorical. Many researchers, however, demonstrate the existence of gradient relations. For instance, Hall (2009) quantifies these using the information-theoretic measure of entropy (unpredictability of distribution) and shows that a pair of sounds may have an entropy between 0 (totally predictable) and 1 (totally unpredictable).

We argue that the existence of such intermediate degrees of contrastiveness does not make the CH untenable, but rather offers insight into contrastive hierarchies. The existence of a continuum does not preclude categorical distinctions: a categorical line can be drawn between zero entropy (entirely predictable, and thus by the CH phonologically inactive) and non-zero entropy (at least partially contrastive, and thus potentially phonologically active). But this does not mean that intermediate degrees of surface contrastiveness are entirely irrelevant to the CH; rather, we argue, they can shed light on how deeply ingrained a phonemic distinction is in the phonological system. As an example, we provide a case study from Pulaar [ATR] harmony, which has previously been claimed to be problematic for the CH.

**Keywords:** contrast; entropy; distinctive features; vowel harmony; Pulaar

## 1 Introduction

The Contrastivist Hypothesis (CH; Hall 2007; Dresher 2009) holds that the only features that can be phonologically active in any language are those that serve to distinguish the members of the underlying inventory—i.e., the phonemes—of that language from one another. Many phonological processes appear to ignore non-contrastive features (see, for example, Kiparsky 1985; Calabrese 2005; and Dresher 2009, among many others); the CH, in its strongest version, predicts that this should always be the case. In this context, the Successive Division Algorithm (SDA; Dresher et al. 1994; Dresher 2009) has been proposed as a means specifying contrastive (and hence potentially active) features. In essence, the SDA starts with all sounds assumed to be allophones of a single phoneme, and assigns contrastive feature specifications only when there is evidence that phones need to be distinguished underlyingly. In order to implement the SDA, then, one must be able to unambiguously determine what is a (separate) contrastive phoneme and what is not. This method thus presupposes that phonemic status is categorical.

A number of researchers, however, have pointed out the existence of phones in a language that are not easily categorized as either phonemic or allophonic (e.g., Gleason 1961; Crothers 1978; Goldsmith 1995; Hill 1998; Hualde 2005; Ladd 2006; Scobbie & Stuart-Smith 2008; Kager 2008; Hall 2009; 2012; Bye 2009; Dresher 2011; Ferragne et al. 2011; Boulenger et al. 2011). Hall (2013b) provides a comprehensive overview of such intermediate relationships, and in particular provides a typology illustrating the many

different ways in which contrasts can be ‘marginal.’ In this paper, we explore the question of whether the existence of such intermediate degrees of contrastiveness makes the Contrastivist Hypothesis untenable, or even meaningless. We propose that it does not, and that furthermore, marginal contrasts may offer insights into how contrastive hierarchies determined by the SDA change diachronically, and what it takes for learners to acquire them. In particular, we examine the case of Pulaar [ $\pm$ ATR] mid vowels, which have been claimed to be problematic for the CH (Archangeli & Pulleyblank 1994; Campos Astorkiza 2007), and show that while there is good evidence for the contrast between [+ATR] and [–ATR] to be marginal, it should nonetheless be considered a contrast, and is therefore not a counter-example to the CH.

In Section 2 we provide background, first on the Contrastivist Hypothesis and the Successive Division Algorithm, and then on marginal contrasts and how to measure them. Section 3 discusses the differences in how contrastiveness is treated by the SDA and by Hall’s (2009) entropy-based approach to marginal contrasts. In Section 4, we show what each of these two theories of contrast has to say about Pulaar; Section 5 concludes.

## 2 Background

### 2.1 The Contrastivist Hypothesis and the Successive Division Algorithm

The impetus behind the Contrastivist Hypothesis, and behind contrast-based theories of underspecification more generally, is the observation that redundant feature values often appear to be ignored by phonological patterns. For example, voicing is contrastive on obstruents but predictable on sonorants in Japanese (as in many other languages), and Japanese *rendaku*, which voices the initial consonant of the second member of a compound, is blocked by the presence of another voiced obstruent, but not by sonorants (Itô & Mester 1986). Accordingly, many phonologists have proposed that redundant features are omitted from some or all of the phonological computation, or that contrastive features have some special status (e.g., Kiparsky 1982a; 1985; Archangeli 1984; 1985; 1988; Itô & Mester 1986; Clements 1987; Pulleyblank 1988; Avery & Rice 1989; Piggott 1992; Dresher et al. 1994; Itô et al. 1995; Calabrese 1995; 2005; Dyck 1995; Ghini 2001; Hall 2007; 2011b; Dresher 2009; 2015; Mackenzie 2009; 2013; Nevins 2010; 2015; Iosad 2012; Oxford 2015); see Hall (2011a; forthcoming) for a concise overview and Dresher (2009) for a more detailed one.

Hall (2007) formulates the Contrastivist Hypothesis as in (1):

- (1) Contrastivist Hypothesis (Hall 2007: 20)  
The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

This is arguably the simplest and strongest possible hypothesis that assigns a special role to contrastive features. At a minimum, the phonological features included in lexical representations must suffice to distinguish all underlyingly contrastive segments. The hypothesis in (1) holds that these are the *only* features that can be phonologically active, not only in underlying representations, but also in the phonological computation.<sup>1</sup> In other approaches, redundant features may be filled in during the course of the derivation and thereby affect the application of subsequent rules (e.g., Archangeli 1985; 1988), or contrastive and redundant features may both be present from the start but be treated differently by the grammar (e.g., Calabrese 1995; 2005).

<sup>1</sup> In derivational, autosegmental terms, a feature is phonologically active if there are processes in which it is spread, or blocks spreading, or is a target for spreading or delinking. It is trickier to define phonological ‘activity’ in a non-derivational framework, but cf. Mackenzie & Dresher (2003) and Iosad (2012) for examples of how the CH has been applied in Optimality Theory.

Any theory along these lines requires a means of identifying which features are contrastive. One intuitive approach to this works analogously to the process of identifying contrasting phonemes by finding minimal pairs of words. First, all phonemes are fully specified; then, in every pair of phonemes that differ by only a single feature value, that feature is identified as contrastive on those phonemes. However, as Archangeli (1988) and Dresher (2009) have pointed out, the features identified as contrastive by this procedure do not always suffice to distinguish all the phonemes in the inventory. For example, consider the inventory /i ε a ɔ u/, which Paradis (1986; 1992) gives as the underlying vowel system of Pulaar. Full specifications of these vowels for five commonly used binary features are shown in (2):

(2)		i	ε	a	ɔ	u
	high	+	-	-	-	+
	low	-	-	+	-	-
	ATR	+	-	-	-	+
	back	-	-	+	+	+
	round	-	-	-	+	+

In (2), any two vowels differ by at least two feature specifications (e.g., /i/ and /ε/ are distinguished by both [ $\pm$ high] and [ $\pm$ ATR]; /ε/ and /a/ are distinguished by both [ $\pm$ low] and [ $\pm$ back]; and so on). The minimal pairs test will therefore fail to designate any feature values as contrastive on any of these vowels. In a framework such as that of Calabrese (1995; 2005), in which both contrastive and redundant features are specified, but some rules apply only to contrastive features, this is not necessarily problematic; it simply means that all rules in languages with inventories like the one in (2) must be able to see ‘redundant’ features. However, in theories that posit that redundant features are systematically absent from at least some levels of phonological representation, the minimal-pairs test is untenable; if all features that are non-contrastive according to this test are omitted, then it will not always be possible to differentiate all (or even any) of the underlying phonemes.

Accordingly, Dresher et al. (1994) propose an alternative procedure, inspired by Cherry et al. (1953), which Dresher (2009) gives in a more general form as the Successive Division Algorithm (SDA):

- (3) Successive Division Algorithm (Dresher 2009: 16)
- a. Begin with *no* feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.
  - b. If the set is found to consist of more than one contrasting member, select a feature and divide the set into as many subsets as the feature allows for.
  - c. Repeat step (b) in each subset: keep dividing up the inventory into sets, applying successive features in turn, until every set has only one member.

Unlike the minimal pairs test, the SDA will necessarily produce specifications that are sufficient to differentiate any phonemic inventory, because it does not terminate until all phonemes have been distinguished. Because it proceeds by adding specifications rather than by pruning them, it does not require a universal set of features to determine what the full specification for each phoneme would be, as the minimal pairs test does; it is thus equally compatible with universal feature theories (such as those of Jakobson et al. 1952; Chomsky & Halle 1968; or Clements & Hume 1995) or emergent features (as proposed by Mielke 2008).

The primary purpose of the SDA is to provide a formal definition of contrastiveness, and thus to make predictions about which features can be phonologically active in any given

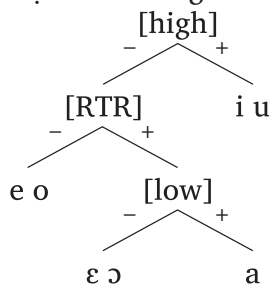
system. However, it is also framed as a procedure that could be part of the acquisition process, with learners assigning features as they discover the phonemic contrasts that define the underlying inventory. A complete learning algorithm based on the SDA would need to elaborate what it takes to identify the presence of a phonemic contrast and how the learner selects the features to assign, and it would also need to take into account how the acquisition of the phoneme inventory interacts with the acquisition of the system of rules or constraints operating on that inventory. (See Dresher (2009: §7.8) for discussion of the role of contrastive hierarchies in theories of acquisition.)

The SDA relies on a hierarchical ordering of features, capturing the intuition that features are contrastive or redundant not in isolation, but with reference to the other features whose values have already been determined. For example, in the inventory in (2), [±ATR] is predictable on any vowel whose value for [±high] is known, and [±high] is predictable on any vowel already specified for [±ATR]. For the minimal pairs test, this means that neither [±high] nor [±ATR] will be identified as contrastive on any vowel in this inventory, but the SDA resolves the problem by requiring that features be hierarchically ordered. Because the SDA itself does not stipulate any particular ordering of features, it allows for the possibility that similar inventories may have different contrastive feature specifications in different languages.<sup>2</sup>

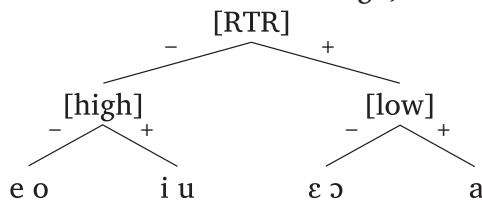
For example, Yoruba has a phonemic seven-vowel inventory /i e ε a ɔ o u/. In Ifẹ Yoruba, the (predictably tense) high vowels are ignored by tongue root harmony, while in Standard Yoruba, a high vowel requires a preceding mid vowel to be tense (Nevins 2010; Dresher 2013). Dresher (2013) proposes that in Ifẹ Yoruba, [±high] takes scope over [±RTR], whereas in Standard Yoruba, the opposite ordering applies, as shown in (4).<sup>3</sup>

(4) Partial contrastive hierarchies for vowels in two dialects of Yoruba (Dresher 2013: 149)

a. Ifẹ Yoruba: high » RTR » low



b. Standard Yoruba: RTR » high, low



In the hierarchy in (4a), all seven vowels are specified for [±high], the feature with the widest scope. [±RTR] is specified only on the [-high] vowels, because it is redundant on the [+high] vowels /i u/; high vowels are thus ignored by harmony. In (4b), on the other hand, [±RTR] takes scope over the whole inventory, and so it is specified, and thus

<sup>2</sup> Proposals for universal partial orderings of features have been made by, for example, Dyck (1995), Ghini (2001), and Clements (2009); cf. also Jakobson (1941). Such orderings could be implemented as restrictions on the SDA.

<sup>3</sup> Some additional feature not shown in Dresher’s hierarchies, such as [±back] or [±round], is needed to distinguish /u/ from /i/, /o/ from /e/, and /ɔ/ from /ε/.

potentially phonologically active, on /i/ and /u/, even though there are no [+RTR] high vowels. (Promoting [ $\pm$ RTR] in the hierarchy entails a tradeoff: in (4b), [+high] is contrastive only among the [−RTR] vowels. See Hall & Dresher (2016) for further discussion of tradeoffs of this kind.) In both hierarchies, the feature [ $\pm$ low] has narrower scope than [ $\pm$ RTR], and is contrastive only on [+RTR] vowels.

By positing different orderings of features in the two dialects, Dresher (2013) rebuts Nevins's (2010) claim that Standard Yoruba RTR harmony requires a redundant feature ([−RTR] on high vowels) to be phonologically active. For Nevins, who uses the minimal pairs test to establish contrastiveness, [ $\pm$ RTR] is contrastive only on the mid vowels /e o ε ɔ/. For Dresher, who uses the SDA, contrastiveness depends not on the inventory alone, but also on the language-specific hierarchical ordering of the features.<sup>4</sup> In this approach, there is greater room for variability in feature specifications, but no feature is assigned unless it serves to mark a phonemic contrast, and the total number of features predicted to be potentially phonologically active by the Contrastivist Hypothesis is constrained by the phonemic inventory.

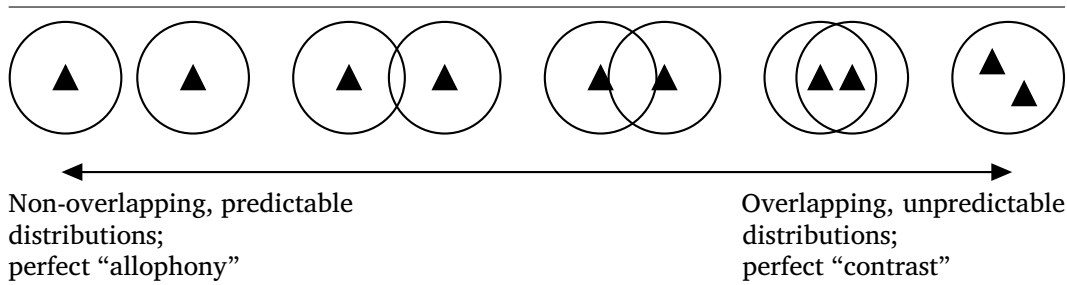
## 2.2 Marginal contrasts and entropy

The Successive Division Algorithm assigns contrastive features with reference to phonemic inventories, and thus presupposes that there is some reliable way of determining which segments need to be distinguished at the underlying level of representation—i.e., which segments are phonemic. At the same time, it has been widely noted that segments are not always clearly and categorically either in contrast with one another or in a relation of allophony; rather, there appears to be a continuum of intermediate possibilities (e.g., Gleason 1961; Crothers 1978; Goldsmith 1995; Hill 1998; Hualde 2005; Ladd 2006; Kager 2008; Scobbie & Stuart-Smith 2008; Bye 2009; Hall 2009; 2012; Boulenger et al. 2011; Dresher 2011; Ferragne et al. 2011).

There can be many reasons for relationships to be intermediate, as Hall (2013b) outlines. One particularly prevalent case is based on intermediate degrees of predictability of distribution (see, e.g., Hall 2009). In many cases, the phonological relation between two sounds can be determined by the extent to which they are predictably distributed in the language, with allophony defined as occurring if the sounds are complementarily distributed, and contrast occurring otherwise.<sup>5</sup> Trubetzkoy ([1939] 1969: 239) predicted, however, that partial predictability of distribution might also be important, noting that neutralizing a contrast can lead to a reduction in its 'distinctive force.' The psychological validity of this claim has been empirically verified by, e.g., Huang (2001) and Hume & Johnson (2003), who found that tone pairs in Mandarin that neutralize in certain contexts are perceived as being more similar to each other than non-neutralizing pairs, to an extent beyond what might be expected due to acoustic characteristics. Hall (2009; 2012) proposes that there is in fact a continuum of possible relations that hold between two phones, based on the degree to which they are predictably distributed. This continuum can be measured using the information-theoretic quantity of entropy (Shannon & Weaver 1949), a measure of uncertainty. Given a relation between exactly two sounds, entropy will range from 0 to 1, measured in bits. An entropy of 0 bits indicates a complete

<sup>4</sup> See Dresher (2015) and Nevins (2015) for further discussion.

<sup>5</sup> Note that we focus exclusively here on the issue of predictability of distribution. This means that we are ignoring issues of phonetic similarity as a filter on complementary distribution, or the possibility of free variation, among others. Our purpose here is not to catalogue all possible ways of defining phonological relations (cf. Hall 2013b) or to investigate the relationships among different measures (cf. Hall & Hume 2015), but rather to examine how one particular measure of marginality can be used to investigate the implications of potentially marginal contrasts for the application of the SDA.



**Figure 1:** A continuous set of phonological relationships defined on the predictability of their distributions (Hall 2009: 16).

absence of uncertainty, i.e., perfect predictability of distribution, and is analogous to perfect allophony. An entropy of 1 bit, on the other hand, indicates total uncertainty, i.e., complete unpredictability in all environments and hence total contrast. This is illustrated in Figure 1, where black triangles schematically represent individual phones, and circles indicate the set of environments each phone can occur in. At the left-hand side of the continuum, the circles are entirely non-overlapping; given a particular environment, there is no uncertainty as to which sound would occur, and so the entropy is 0. At the right-hand side of the continuum, the circles are entirely overlapping; both phones occur (with equal frequency) in all the same environments, and so there is perfect uncertainty about which might occur in any given environment—an entropy of 1. The two phones will have an entropy between 0 and 1 when they have partially overlapping distributions. As hinted at above, the frequency of occurrence of the phones also matters: if one phone is significantly more or less frequent than the other, that will decrease the uncertainty about which one will occur in a given environment. Thus, even if the two phones have entirely overlapping distributions, if one occurs more frequently than the other in those environments, the entropy of the pair will still be less than 1. This reflects, for example, the claim in Goldsmith (1995) that asymmetries in frequency of occurrence are relevant for understanding what he terms a ‘cline’ of contrast; see also discussion in Hall (2009).

The entropy of a pair of phones in a language (their systemic entropy) is calculated as in (5), as proposed by Hall (2009) and implemented in the *Phonological CorpusTools* software (Hall et al. 2015a). First, the entropy in each individual environment in which at least one of the two phones occurs is calculated, using the formula in (5a).<sup>6</sup> Here, the probability of each phone’s occurring in that environment is multiplied by the log of that same probability, and the products for both phones are summed. To calculate the systemic entropy, then, the entropy of each environment,  $H(e)$ , is multiplied by the probability of that environment’s occurring,  $p(e)$ , and these products are summed for all environments. The probabilities of the environments are calculated as in (5b), where  $N_e$  is the number of occurrences of a given environment and  $\Sigma N_{e \in E}$  is the number of occurrences of all environments in which at least one of the phones occurs.

$$(5) \quad \text{Systemic Entropy} = \Sigma(H(e) \times p(e))$$

$$a. \quad H(e) = -\Sigma p_i \log_2 p_i$$

$$b. \quad p(e) = N_e / \Sigma N_{e \in E}$$

As an example for how to implement this metric, we can examine the relation between [e] and [ɛ] in French, which are interesting from the perspective of the degree of complementary distribution they show. As with other tense–lax vowel pairs in French, [e] and [ɛ]

<sup>6</sup> See §4.3 for more about how to determine environments.

are subject to the *Loi de position* (Grammont 1913; Walker 2001; Féry 2003; Durand & Lyche 2004), such that the tense vowels tend to occur in open syllables while the lax vowels tend to occur in closed syllables. The vowels are generally still considered contrastive, though, because there are certain environments in which both vowels can occur.

Note that this same formula for calculating the degree of contrastiveness can be applied to different types of data. At one extreme, it can be applied to a corpus of running speech in a language, such that the probability of occurrence of an individual phone in a given environment is based on its token frequency in that environment in the corpus. The token frequency of the sound [e] occurring in word-final position in the *Lexique* corpus of French film subtitles (New et al. 2004) is 5,745,009. The token frequency of [ɛ] occurring in that same environment is 2,661,281. Thus, the probability of [e] is  $\frac{5,745,009}{5,745,009+2,661,281}=0.68$ ; the probability of [ɛ] is 0.32; and the entropy of [e, ɛ] in this environment is 0.9 bits. This number is relatively high (close to the maximum of 1), but not at 1; this suggests that while these sounds fall slightly short of complete unpredictability, they are more contrastive than not.

Alternatively, one could determine the probabilities of occurrence based on a lexicon of the language rather than a corpus, such that each word containing a phone in a particular environment is counted only once; i.e., using type frequency. The type frequency of word-final [e] in *Lexique* is 31,601, while that of [ɛ] is 15,981. The probabilities of each phone are thus 0.66 and 0.34, respectively, and the entropy in this environment is 0.92 bits. In this case, the choice between type and token frequency doesn't make much difference, but in cases where these two measurements are different, the entropy values will obviously follow suit.

A third way of applying this formula, however, would be in terms of type occurrences rather than type frequencies; that is, if there is at least one occurrence of a given phone in a given environment, it would be counted as having a type occurrence of 1; if there is not at least one occurrence, its type occurrence would be 0. In this case, both word-final [e] and word-final [ɛ] have type occurrences of 1 in *Lexique*; their probabilities would each then be 0.5, and the entropy would be that of perfect contrastivity, i.e., 1 bit. This third way of applying the entropy formula is of course most similar to traditional approaches to contrast; it allows for entropy values of only 0 and 1, i.e., allophony and contrast, with no intermediate relations possible. Under such an application, of course, the notion of 'marginal' contrasts—at least, those based on predictability of distribution—disappears, and what appears to be a gradient mathematical formula gives rise to binary, categorical distinctions.

In addition to being able to apply this formula to different kinds of frequency measures, one can, of course, apply it to different levels of representation. In the examples above, token frequency, type frequency, and type occurrences were counted at the level of surface representations, but it would be possible to count any of them on the basis of underlying representations instead, if one had independent evidence for such representations. Thus, the formula itself makes no inherent claim about what level of representation is relevant; it simply provides a means for objectively quantifying the extent to which two phones are predictably distributed in a language, given some measure of probability.

Recall, however, that predictability of distribution is not the only reason two phones might be considered to be intermediate between contrast and allophony. Other reasons include (1) cases where the sounds are predictably distributed, but where crucial reference is made to non-phonological information, such as morphological boundaries (e.g., the marginal contrast between [ai] and [ʌi] in *tied* vs. *tide* in Scottish English; see, e.g., Scobbie & Stuart-Smith 2008); (2) cases where one or both of the sounds is foreign, specialized, or belongs to some distinct stratum of the language (e.g., the marginally

contrastive status of [x] in English due to words like *Bach* and *loch*; see, e.g., Ladd 2006); or (3) cases where the sounds are highly variable in their phonetic realization (e.g., the marginal contrast between [e] and [ɛ] in Italian; see Ladd 2006).

Thus, as Steriade (2007: 140) points out, “the very existence of a clear cut between contrastive and non-contrastive categories” is “called into question.” If this is true, then what is to become of theories, such as the Contrastivist Hypothesis and the SDA, that crucially rely on such a clear cut being possible? As we will argue in the next section, the answer lies in understanding what kind of binary distinction the CH actually assumes, and at what level of representation.

### 3 What counts as contrastive, and where

As mentioned in §2.2, the use of a continuous quantification metric does not preclude the possibility of categorical distinctions. In the case of using entropy to define a continuum of contrastiveness, there is an obvious line to be drawn between pairs of phones with zero entropy and pairs of phones with non-zero entropy: if two phones are unpredictable in at least some contexts, then they will have an entropy greater than zero, and the system of phonological representations must have some means of distinguishing them. This is analogous to the case of applying the entropy formula to type occurrences; rather than counting the frequency of occurrence, one simply determines whether the phones occur at all in each environment.

In a theory of phonology that adopts the Contrastivist Hypothesis and the Successive Division Algorithm, the most immediately relevant form of contrastiveness is not the gradient degree to which two sounds contrast at the surface, but the categorical question of whether the phonological system needs to distinguish them in lexical representations. That is, if a phone is at least marginally contrastive—i.e., if there are at least some contexts in which it cannot be predictably derived from other information independently known to be present in the representation—then it must be treated as a separate phoneme for the purposes of the SDA. To take an extreme example for the sake of clarity, if a language adopts a single loanword containing a potentially contrastive new phone, then the novel segment must either be treated as an exception (and thus not integrated into the regular phonological system) or be regarded as phonemic and included in the SDA.

As an approximation, then, one could say that the SDA assigns distinct featural representations to any pair of phones with non-zero entropy. But this is only an approximation. Measuring the entropy of phones at the surface level gives an index of their functional contrastiveness in an information-theoretic sense,<sup>7</sup> but the SDA is concerned with formal, not functional, contrast, and at the level of underlying (lexical) representations, not surface forms. The set of phonemes to be distinguished by the SDA can be identified only by analysis of the phonological system, and is thus not trivially calculable from surface data.<sup>8</sup> In cases of absolute neutralization, two phonemes that appear identical on the surface may need to be distinguished underlyingly because they exhibit different phonological patterning. For instance, ‘strong *i*’ and ‘weak *i*’ in some Inuit dialects are both realized phonetically as [i], but one triggers palatalization and the other does not (Archangeli & Pulleyblank 1994; Compton & Drescher 2011).<sup>9</sup> Conversely, some surface

<sup>7</sup> ‘Functional contrastiveness’ as measured by entropy is an indication of the (un)predictability of a pair of phones, without reference to other pairs in the system; it should not be confused with the notion of functional load, which measures how much work a phonemic contrast does, relative to other contrasts, in distinguishing forms. See Hall (2009: §3.7.1) for further discussion.

<sup>8</sup> As noted above, the entropy metric can be applied to underlying representations as well as surface forms; as with the SDA, this is dependent on the underlying inventory posited by a particular analysis.

<sup>9</sup> Absolute neutralization has been a contentious issue in phonological theory; see, e.g., Hyman (1970), Brame (1972), Kiparsky (1982b), Hall (2011a), and Yu (2011) for discussion and further examples.



contrasts between phones can be analyzed as arising from differences in underlying structure that are not segmental. For example, European Portuguese has a marginal surface contrast between [a] and [ɐ], as in the minimal pair in (6) (Spahr 2016; see also Mateus & d'Andrade 2000; Carvalho 2004):

- (6) a. *falamos* [fɐˈlɐmuʃ] ‘we speak’  
 b. *falámos* [fɐˈlamuʃ] ‘we spoke’

However, Spahr (2016) argues that there is no underlying phonemic contrast between these two vowels; rather, the surface difference between (6a) and (6b) is derived from a morphological contrast in the presence or absence of a perfect suffix consisting of an empty V slot, as in (7). In (7b), the theme vowel /a/ at the end of the verb stem spreads to fill this empty position. Singly linked /a/ raises to [ɐ] in unstressed syllables and before nasals, while doubly linked /a/ is realized as [a] in all contexts.<sup>10</sup>

- (7) a. [fɐˈlɐmuʃ] = ‘speak’ + 1PL.  

f		l			m		ʃ
C	V	C	V	+	C	V	C
	a		a			u	

  
 b. [fɐˈlamuʃ] = ‘speak’ + PERFECT + 1PL.  

f		l			m		ʃ
C	V	C	V	+	V	+	C
	a		a	-----	u		

In Spahr’s analysis, then, there is no need to recognize two distinct underlying phonemes /a/ and /ɐ/; even though the two vowels contrast at the surface, they constitute a single phoneme for the purposes of the SDA. Similarly, in Northern Irish English (Harris 1990: §6.3) a surface contrast between [ɛ:] and [ɪə] in pairs such as *day* + s [dɛ:z] versus *daze* [dɪəz] arises from the interaction of morphological structure with rules that derive a long monophthong in open syllables and a diphthong in closed ones. The surface minimal pair need not be taken to imply an underlying phonological contrast, because there is another, independently motivated underlying difference from which it can be derived. Other cases in which marginal surface contrasts are predictable from underlying morphophonological information have also been discussed by Harris (1990; 1994), Carr (2008), Bye (2009), and Hall (2013b).

For the purposes of the SDA, then, contrastiveness is a categorical property, and one that depends on the analysis of the underlying phonological system. This does not mean, however, that the study of gradient surface contrastiveness as measured by type frequency or token frequency is untenable or meaningless. As mentioned above, there is empirical evidence that intermediate degrees of predictability affect the perception of phones (e.g., Huang 2001; Hume & Johnson 2003; Hall 2009; Hall & Hume 2015). Quantifying predictability of distribution can also be useful in tracking and interpreting phonological change (e.g., Hall 2013a). Furthermore, even within the context of the CH, the existence of a categorical split into ‘contrastive’ (assigned by the SDA, and therefore potentially

<sup>10</sup> Spahr (2016: 213), citing Gendrot & Adda-Decker (2007), notes that [a] also has longer duration than [ɐ], as the representations suggest. In Spahr’s representations, length is encoded in timing slots, and is thus orthogonal to the underlying inventory of contrasts in vowel quality.

phonologically active) and ‘predictable’ (and thus not available in the phonological computation) does not preclude further study into gradient contrastiveness and its possible relation to differing degrees of phonological activity. In the context of the SDA, a contrast might be considered ‘marginal’ if it is relatively low in the contrastive hierarchy, particularly if it appears on only a small number of branches. Contrasts with low scope might, in turn, be expected to be diachronically unstable—either emerging or disappearing (Dresher et al. 2014; Oxford 2015)—and synchronically more susceptible to neutralization (Spahr 2014). Similarly, but potentially independently, contrasts with low entropy may offer less evidence to the learner, and consequently also be diachronically precarious. Although the hierarchical and entropy-based notions of marginality will not necessarily always align, the study of marginal contrasts and the study of the Contrastivist Hypothesis can be entirely symbiotic.

## 4 The case of Pulaar

### 4.1 Pulaar [ATR] harmony patterns

Pulaar ATR harmony provides an example of the importance of marginal contrasts for the Contrastivist Hypothesis (see also Hall 2000; 2007). As mentioned in §2.1, the underlying vowel inventory according to Paradis (1986; 1992) is as in (8a); the surface inventory is shown in (8b). The surface inventory has a distinction between [−ATR] mid vowels /ε ɔ/ and [+ATR] /e o/, while the underlying inventory posited by Paradis has only [−ATR] mid vowels.

(8)	Pulaar vowel inventory			
	a. Underlying ( <i>per</i> Paradis)		b. Surface	
	i	u	i	u
			e	o
	ε	ɔ	ε	ɔ
	a		a	

The surface mid vowels are subject to a pattern of ATR harmony. Archangeli & Pulleyblank (1994) describe this pattern as follows: high vowels are [+ATR]; /a/ is [−ATR]; and mid vowels are [+ATR] when the next vowel to their right is [+ATR], and [−ATR] otherwise. In other words, mid vowels are normally [−ATR], and should surface as [+ATR] only when followed by a high vowel with no intervening /a/. An example of this is shown in (9), where a high vowel in the singular suffix triggers a [+ATR] mid vowel in the stem, while a mid vowel in the diminutive plural suffix yields a [−ATR] mid vowel in the stem.

(9)	SINGULAR	DIM. PL.	
a.	[sof-ru]	[cɔf-ɔn]	‘chick’
b.	[ser-du]	[cɛr-kɔn]	‘butt of a rifle’
c.	[ <sup>m</sup> bɛ:l-u]	[ <sup>m</sup> bɛ:l-ɔn]	‘shadow’
d.	[dog-o:ru]	[dɔg-ɔ-w-ɔn]	‘runner’

Similarly, (10) shows that the addition of the instrumental applicative suffix, containing [+ATR] /i/, triggers [+ATR] harmony to its left.

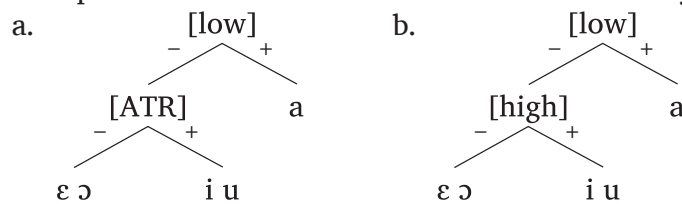
(10)	PLAIN INFINITIVE		INSTRUMENTAL INFINITIVE	
a.	[βɛt-dɛ]	‘to weigh’	[βɛt-ir-dɛ]	‘to weigh with’
b.	[hɛl-dɛ]	‘to break’	[hɛl-ir-dɛ]	‘to break with’

Finally, (11) illustrates the blocking of [ATR] harmony by /a/ when it intervenes between a mid vowel and a [+ATR] high vowel.

- (11) a. [bɔ:t-a:-ri] ‘lunch’  
 b. [pɔ:f-a:-li] ‘breaths’  
 c. [nɔdd-a:-li] ‘call’  
 d. [ʔgɔr-a:-gu] ‘courage’

These patterns would appear to be a problem for the Contrastivist Hypothesis. If [ $\pm$ ATR] is the feature that distinguishes /i u/ from / $\epsilon$  ɔ/, as in (12a), then spreading [+ATR] to the mid vowels would turn them into [i u], not into [e o]. On the other hand, if [ $\pm$ ATR] is not the feature that distinguishes /i u/ from / $\epsilon$  ɔ/, as in (12b), where [ $\pm$ high] is used instead, then [ $\pm$ ATR] has no contrastive role in the system; in that case, ATR harmony would entail that a non-contrastive feature is phonologically active, contra the Contrastivist Hypothesis.

- (12) Two partial contrastive hierarchies for the inventory in (8a)<sup>11</sup>



Indeed, the Pulaar facts have been cited in arguing against contrastive underspecification, although not with specific reference to a contrastive hierarchy. Archangeli & Pulleyblank (1994: 134) say that “although completely predictable, [ATR] values play an active role in the phonology of Pulaar,” and Campos Astorkiza (2007: 194) claims that “[a]ccording to underspecification theories, the [ATR] value for high and low vowels should not play an active role in the phonology of the language.”

#### 4.2 A solution

This apparent paradox can be solved, however, by recognizing that [e o] not only occur on the surface, but are marginally contrastive. According to Paradis (1992: 90), there are three morphemes that counterexemplify the surface generalization that [+ATR] mid vowels occur only to the left of high vowels; these are given in (13).

- (13) a. [fof] ‘all’  
 b. [-(g)el] diminutive singular  
 c. [-(g)ol] noun class marker

Two of these, [-(g)ol] and [-(g)el], are suffixes and therefore appear in a variety of longer words; they then trigger harmony in the stems that precede them, as illustrated in (14) and (15).

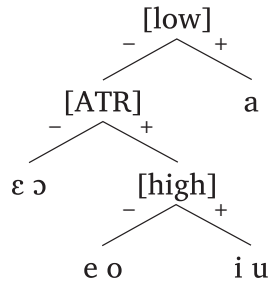
- |         |            |            |            |               |
|---------|------------|------------|------------|---------------|
| (14)    | SINGULAR   | DIM. SG.   | DIM. PL.   |               |
| Suffix: | -(g)ol     | -(g)el     | -ɔn        |               |
| a.      | [lef-ol]   | [lef-el]   | [lɛf-ɔn]   | ‘ribbon’      |
| b.      | [ke:r-ol]  | [ke:r-el]  | [kɛ:r-ɔn]  | ‘boundary’    |
| c.      | [ce:lt-ol] | [ce:lt-el] | [cɛ:lt-ɔn] | ‘cut’         |
| d.      | [cef-ol]   | [cef-el]   | [cɛf-ɔn]   | ‘incantation’ |

<sup>11</sup> Some additional feature such as [ $\pm$ round] or [ $\pm$ back] must also be present, to distinguish /u/ from /i/ and /ɔ/ from / $\epsilon$ /; we omit it from the hierarchies in (12) and (16) because it is not relevant here.

(15)	SINGULAR	DIM. SG.	
	Suffix: various	-(g)el	
	a. [m̥baro:-di]	[baro-gel]	‘lion’
	b. [p̥em̥bɔ:w-ɔ]	[p̥em̥bow-el]	‘hairdresser’
	c. [hɔr-dɛ]	[kor-el]	‘calabash’

Paradis (1992) avoids positing underlying /e o/ by analyzing these morphemes as /fɔuf/, /-geil/, and /-gɔul/, with the high vowels underlyingly lacking any association to a timing slot. The floating high vowels trigger harmony, but are not themselves pronounced. Archangeli & Pulleyblank (1994: 192) posit that these suffixes have (exceptional) underlying [+ATR] mid vowels, but maintain that [+ATR] is still redundant on the high vowels, because it is predictable from [+high]. However, if the marginal surface contrast is analyzed as a categorical underlying contrast, with /e o/ included in the phonemic inventory despite their apparently low type frequency, and [±high] is given relatively narrow scope, then [±ATR] can be characterized as contrastive on /e o i u/, as shown in (16).

(16) Partial contrastive hierarchy for Pulaar with phonemic /e o/



Given the representations in (16), harmony can be unproblematically represented as the leftward spreading of [+ATR] from one non-low vowel to the next.<sup>12</sup> But if /e o/ are indeed phonemic, then they might be expected to appear in more than three morphemes. A re-examination of the Pulaar system in §4.3 will provide evidence that this is indeed the case. Given that other related languages such as Wolof, Kisi, and Diola-Fogny have a more robust [ATR] contrast (Casali 2003), Pulaar as described by Paradis might represent a diachronic change from an earlier stage in which [+ATR] mid vowels were more prevalent. The Contrastivist Hypothesis predicts that if the marginal underlying contrast between /e o/ and /ε ɔ/ is lost, the harmony pattern could not remain productive in its present form.

A learner acquiring Pulaar phonology must acquire the feature [ATR] in order to generate the harmony pattern. According to the SDA, a feature can be assigned only if it serves to mark some contrast in the underlying inventory. In the hierarchy in (16), [ATR] serves to distinguish /ε ɔ/ not only from /e o/, but also from /i u/. If the acquisition of contrasts proceeds monotonically from the root of the contrastive hierarchy to its leaves, with no backtracking, and if a child acquires harmony before identifying /e o/ as distinct underlying segments—that is, before making the lowest of the three divisions in (16)—then we might expect the child to pass through a stage in which harmony raises /ε ɔ/ to [i u]. On the other hand, if phonemes are not acquired in perfect hierarchical order, or if the full inventory is acquired before the harmony rule, then no such stage is predicted. Note that the hierarchical rank of contrasts in (16) does not directly reflect their degree of prominence or marginality: while the marginal status of /e o/ relates to their relation

<sup>12</sup> This analysis assumes that it is the feature [+low] on /a/ that blocks spreading of [+ATR], either as a formal structure-preservation effect (Kiparsky 1985) or because of the articulatory antagonism between [+low] and [+ATR] (Archangeli & Pulleyblank 1994). Alternatively, giving [ATR] scope over [low] in the hierarchy would allow /a/ to be specified as [-ATR], as in Krämer’s (2003) analysis.

with / $\varepsilon \text{ } \text{ɔ}$ /, (16) groups them more closely with / $i \text{ } u$ /, suggesting that the acquisition path may not be wholly straightforward.

In any case, the complete set of underlying phonemes must be learned in order to arrive at the adult grammar's harmony patterns. If there are indeed only three morphemes that provide evidence for the phonemic status of / $e \text{ } o$ /, harmony would seem to rest on startlingly unsteady ground—as if, for example, there were some phonological rule of English that depended on the learner's exposure to *Bach* and *loch*, though in the Pulaar case the crucial morphemes are affixes that can combine with multiple stems, rather than low-frequency free roots. In §4.3, then, we turn to entropy as a means of quantifying the degree to which / $e \text{ } o$ / contrast with / $\varepsilon \text{ } \text{ɔ}$ / on the surface, and thus as a potential approximation of the robustness of the evidence available to the learner.

### 4.3 Quantifying the contrastiveness of [ATR] in Pulaar

As described in §2.2, the extent to which two phones are contrastive can be quantified using entropy, with an entropy of 0 indicating perfect predictability and an entropy of 1 indicating perfect unpredictability. To get a sense of the extent to which [ATR] distinctions might actually be contrastive in Pulaar, at least on the surface, the formulae in (5) were applied to the lexicon as represented by Niang (1997), a Pulaar–English and English–Pulaar dictionary based on “the Pulaar dialect spoken essentially in Mauritania, Senegal, and The Gambia” (Niang 1997: x), giving us a measure of type frequency.<sup>13</sup> All distinct forms appearing as headwords in the Pulaar–English section, a total of 6332 words, were taken as input. The entropy measure was calculated using *Phonological CorpusTools* (Hall et al. 2015a).

Interestingly, in Niang's orthography, [+ATR] mid vowels are represented as <é, ó>, and [–ATR] ones as <e, o>. The fact that Niang represents this distinction at all is itself suggestive, as entirely predictable variations are typically excluded from orthographic representations.

In order to calculate entropy, one must decide which environments are relevant. This is not a trivial decision, as the selection of environments can relatively drastically change the results, as shown in Hall (2012), where environments are chosen both to emphasize and minimize the apparent contrasts between vowels in Canadian English. In the limit, one can always simply assume that there are no conditioning environments and calculate entropy based on the relative frequencies of the sounds in question across all words in the lexicon (or tokens in the corpus). This move actively ignores the entire enterprise of phonological analysis by ignoring the visible patterns of conditioning that are at the heart of phonological rules or constraints. Using a mathematical formula does not make the analysis of a phonological pattern deterministic; one still must bring active meta-knowledge of the phonological patterns being analyzed to bear. What is crucial for the application of the formula is that (1) whatever environments are chosen, they exhaustively cover all the individual tokens of the sounds being analyzed and (2) they do so in a unique way, i.e., not double-counting any such tokens. For example, it would be meaningless to include both the environments ‘word-initial before a vowel’ and ‘word-initial before [i]’ because these environments overlap and would double-count any segment occurring word-initially before [i]. Thus, in order to calculate a sensible number, the environments must be exhaustive and unique, but they still must be selected by the analyst in a transparent way (see more discussion in Hall 2009).

<sup>13</sup> In principle, it would also be interesting to compare this calculation with one based on token frequency, which would of course require a corpus of transcribed Pulaar speech.

For the calculation of the entropy between [+ATR] mid vowels on the one hand and [−ATR] mid vowels on the other, three mutually exclusive and comprehensive environments were chosen. These environments were picked so as to maximize the possibility of perfect allophony, such that the expected entropy in each environment is 0 given the typical descriptions of Pulaar vowel harmony in the literature. The three environments were: (1) before [+ATR] vowels (where only [+ATR] mid vowels are predicted to occur under a harmony account); (2) before [−ATR] vowels (where only [−ATR] mid vowels should occur); and (3) where the vowel is the final vowel in the word (where again, only [−ATR] mid vowels are predicted to occur). A non-zero entropy in any one of these three environments would indicate at least a partial surface contrast in that environment.

As it turned out, all three environments had a non-zero entropy. Specific numbers will be reported shortly. However, this meant that there were a number of words that were ‘surprising’ from the perspective of a harmony account: there were 401 words in which a [+ATR] mid vowel occurred without a [+ATR] trigger to its right (or with such a trigger, but only to the right of an /a/), and there were 101 words in which a [−ATR] mid vowel occurred even though there was a [+ATR] trigger available to its right. Before continuing, it was important to verify that these were in fact real instances of non-harmonic [ATR] mid vowels rather than, say, typos in the dictionary. The 101 cases of [−ATR] vowels before a [+ATR] trigger, in particular, seemed likely candidates for typographical errors, as they could potentially be attributed to a simple lack of an acute accent mark on the first vowel. Therefore, all of these words were checked by a native speaker of Pulaar from Senegal.<sup>14</sup> Although not a trained linguist, this speaker had studied Pulaar at university and spontaneously used the terms “open” and “closed” to identify the [−ATR] and [+ATR] mid vowels, respectively; he had very little trouble in identifying which vowels occurred in any given word. While several typos were indeed discovered, there were still 397 words containing a [+ATR] mid vowel without an available [+ATR] trigger, and 88 words containing a [−ATR] mid vowel despite the presence of an available [+ATR] trigger. All entropy calculations given in the text are based on these verified words.<sup>15</sup>

Table 1 shows the total numbers of occurrences of the [±ATR] mid vowels in each of the three contexts: before [+ATR] vowels, before [−ATR] vowels, and word-finally. Numbers in bold are ones for which a non-zero value indicates a deviation from the expected pattern—viz., mid vowels followed by another vowel with the opposite value for [±ATR], or [+ATR] mid vowels in the final syllable of the word.

Table 2 shows the entropy calculations in the three environments for [±ATR] mid vowels, along with the systemic entropy (the weighted average entropy across the three environments). As can be seen, the entropy is non-zero in all three environments, and the overall weighted average entropy is 0.27. This is clearly in the realm of being contrastive, i.e., it is not 0, but it is also clearly a somewhat ‘marginal’ contrast, in that it has a relatively low entropy value considering the scale of 0–1.

To make this marginality clearer, consider the contrast between [±back] mid vowels, in the analogous conditioning environments of before [+back] vowels, before [−back] vowels, and word-finally. There has never been, as far as we know, any suggestion of vowel harmony or allophony in this domain. Thus, we would predict much higher entropy values in all three environments and systemically; Table 3 shows the actual calculations.

<sup>14</sup> Paradis’s data come from Futankoore Pulaar, also called Kaédi Pulaar, spoken in the Futa Toro region around the Senegal River in Mauritania and Senegal (Paradis 1992: 6).

<sup>15</sup> Two of the words that are represented distinctly in the dictionary were pronounced identically by our consultant, leaving a total of 6331 distinct word-forms.

	e	o	ɛ	ɔ
__[+ATR]	577	766	<b>40</b>	<b>48</b>
__[-ATR]	<b>0</b>	<b>24</b>	1378	1711
__#	<b>74</b>	<b>299</b>	3695	935

**Table 1:** Total instances of Pulaar mid vowels by environment.

Environment <i>e</i>	<i>H(e)</i>	<i>p(e)</i>
__[+ATR]	0.33	0.15
__[-ATR]	0.07	0.33
__#	0.38	0.52
Systemic	0.27	—

**Table 2:** Entropies for [+ATR] vs. [-ATR] mid vowels.

Environment <i>e</i>	<i>H(e)</i>	<i>p(e)</i>
__[+back]	0.83	0.26
__[-back]	0.93	0.22
__#	0.81	0.52
Systemic	0.84	—

**Table 3:** Entropies for [+back] vs. [-back] mid vowels.

As predicted, the [ $\pm$ back] mid vowels are much more clearly contrastive across all three environments, and the weighted average entropy is 0.84.

Thus, there is clearly a difference between the degree of contrastiveness for [ $\pm$ ATR] mid vowels on the one hand and [ $\pm$ back] mid vowels on the other, with a tendency for the [ $\pm$ ATR] mid vowels to be harmonically distributed. At the same time, both are indeed squarely in the ‘contrastive’ range of the continuum.<sup>16</sup>

As mentioned above, there are 397 words in the 6331-word lexicon that contain [+ATR] mid vowels that are not immediately followed by another [+ATR] vowel and are thus ‘contrastively’ (unpredictably) [+ATR]. While many of these do involve the suffixes in (13), they also include borrowings (largely from French) and a number of other apparently native words that are transcribed with [+ATR] mid vowels, some of which are

<sup>16</sup> Note that at present, there is no mathematical way of (1) categorically classifying contrasts as ‘marginal’ or ‘non-marginal,’ or (2) calculating whether two entropy values are even statistically significantly different from each other. The former is a question of phonological analysis; as we argue here, for example, there really is not a clear case to be made for making such a distinction in the application of the SDA, for the purposes of which the relevant question is simply whether sounds do or do not need to have distinct representations underlyingly. Whether there is a specific role for ‘marginal’ contrasts as a unified category remains to be seen; other work illustrating the relevance of gradience in phonological relationships has largely focused instead on showing a correlation between degree of contrastiveness and degree of perceived similarity (e.g., Hall 2009; Hall & Hume 2015; Hall et al. 2015b). The latter question, about whether two entropy values are statistically significantly different, is a statistical one. The difficulty with determining statistical difference in this case is that there is no variability in the measure—the entire corpus or lexicon has been searched, and the entropy values are the true values for that subsection of the language. Without reference to other corpora, or artificially creating and sampling sub-corpora of the given corpus, there is no easy way to establish whether the values are statistically different from each other because there is no variability about the calculated values.

shown in (17), lending credence to the hypothesis that such vowels are in fact underlyingly contrastive.<sup>17</sup>

(17)	SPELLING	IPA	GLOSS	PAGE
a.	<i>biyé</i>	bije	'ticket' (< French <i>billet</i> )	9
b.	<i>góó</i>	go:	'one'	xii, 37
c.	<i>jóy</i>	ɕoj	'five'	xii, 51
d.	<i>ñóndo</i>	ɲo <sup>n</sup> ɗo	'person whose speech is affected by nasalization'	78
e.	<i>pulóók</i>	pulo:k	'cassava; yam'	83
f.	<i>taartóyaade</i>	ta:rtɔja:de	'take a scenic route'	95
g.	<i>taggéé</i>	tagge:	'death message broadcast on radio'	95
h.	<i>téew</i>	te:w	'meat; flesh'	97

There were also 88 words that unexpectedly had [−ATR] mid vowels preceding high vowels; a few of these are listed in (18).

(18)	SPELLING	IPA	GLOSS	PAGE
a.	<i>fendude</i>	fɛ <sup>n</sup> dude	'set milk; condense'	31
b.	<i>kodijje</i>	kɔdi:ɕɛ	'silver coins'	54
c.	<i>meedfnude</i>	mɛ:nnude	'make someone taste sth.'	64
d.	<i>ñengi</i>	ɲɛŋ <sup>n</sup> gi	'magic'	77
e.	<i>yornude</i>	jɔrnude	'loosen'	110

These contribute to the surface unpredictability of [±ATR] on mid vowels, and suggest that this contrast extends beyond the suffixes, borrowings, and handful of native words containing unexpected [+ATR] mid vowels. They also suggest that the harmony pattern is less than fully productive, at least in the sense that it appears to have exceptions.

Interestingly, despite the clearly contrastive nature of [±ATR] mid vowels, the contrast has a nearly null functional load. There was exactly one minimal pair listed in Niang (1997), shown in (19), but our native-speaker consultant pronounced both of these as [doppude].

(19)	SPELLING	IPA	GLOSS	PAGE
a.	<i>doppude</i>	dɔppude	'put out a fire'	23
b.	<i>dóppude</i>	doppude	'repel'	23

Of course, even with a larger number of apparently unpredictable surface [+ATR] mid vowels, Paradis's (1992) floating-vowel analysis could still be tenable. The surface data demonstrate only that there is a phonologically necessary distinction in underlying representations; by themselves, they do not tell us whether that distinction is encoded paradigmatically, as an underlying [ATR] contrast among mid vowels, or syntagmatically, as a contrast between the presence and absence of a floating high vowel. Paradis (1992: 92) writes that "it is preferable to put the burden on a few underlying forms [...] than to expand the whole phonological system." The larger the number of 'exceptional' forms, then, the less reason there is to treat the unpredictable [+ATR] mid vowels as anything other than underlyingly contrastive phonemes. To the extent that there is any empirical

<sup>17</sup> Page references in (17), (18), and (19) are to Niang (1997). Niang lists source languages of loanwords where they are known; of the words in (17), only (17a) is marked as a borrowing. A linguistically informed native speaker of a related Fula dialect (A. Alkali, p.c.) confirms that (17b, c, d, f, g, h) are native words, but speculates that (17e) *pulóók* may not be, because the final plosive is phonotactically unexpected. Bah (2014) indicates that this form is specific to the Fouta Toro dialect, the more general word being *bantara*.



difference between positing marginal underlying /e o/ and positing floating high vowels, we might expect that there could be phonotactic principles or morpheme structure constraints restricting vowel sequences but not individual vowels, or that floating vowels might interact with other nearby vowels or glides. The forms in (17) are quite varied: some have long vowels, others short; some have a following glide /j/ or /w/, others not. If there are floating high vowels in the underlying representations of these words, it is not obvious that they have any consistent effect other than giving rise to the [ $\pm$ ATR] mid vowels.

## 5 Conclusion

In sum, we argue that there is no incompatibility between the Contrastivist Hypothesis and recognizing the existence of a continuum between pure contrast and pure allophony. To some extent, these two approaches to phonological contrast are simply orthogonal. The CH is about the information that may or must be present in underlying representations; Hall's (2009) entropy measure quantifies contrastiveness in terms of unpredictability of distribution at any level for which independent evidence is available for representations.

One might then conclude that the entropy measure could be applied within a framework embracing the CH only by using type occurrence measures on underlying representations, as this would result in the relevant categorical distinction between non-contrastive and contrastive segments. We argue, however, that this misses the potential for a deeper understanding of phonological phenomena. While there may not be a formal role for a distinction between fully contrastive and marginally contrastive phonemes under the CH, the distinction is certainly useful from an analytic perspective.

In the case of Pulaar, earlier researchers had noted the existence of morphemes containing [+ATR] mid vowels not followed by high vowels, but had discounted them as not reflective of an underlying [ $\pm$ ATR] contrast on these vowels. The Contrastivist Hypothesis, however, predicts that the Pulaar harmony pattern requires precisely such a contrast to exist. Both from the perspective of the learner and of the analyst, once a contrastive featural specification is posited, it becomes available to be used contrastively throughout the system, and might be expected to expand. For the learner, this might mean that loanwords containing underlying /e/ or /o/ are adoptable without adaptation, or that small acoustic differences are more easily interpreted as being a reflection of an actual vowel contrast (cf. Ohala 1981; 2012). For the analyst, the realization that a feature is unpredictable in one instance should lead one to examine whether the distinction is in fact contrastive elsewhere in the lexicon. Calculating the entropy of the sounds provides a way of exploring exactly how robust the contrast is at the surface level.

Our results also suggest how the measure of entropy can contribute to measuring the progress of changes in segmental inventories (see also Hall 2013a), though we do not yet have the historical data that would be needed to chart the diachronic course of the Pulaar vowels. Here, we went from the observation that the CH predicts a phonemic [ $\pm$ ATR] contrast to the verification of that contrast as being present but less robust than other contrasts in Pulaar. One could also go in the other direction, that is, from noting that there is a marginal surface contrast to exploring whether the feature specifications that might formally encode that contrast are in fact phonologically active.

It is of course important to keep in mind, though, that surface contrasts do not always correspond to underlying contrasts, as noted in Section 3. Even in the present case, as discussed in Section 4, we have no definitive proof that the distinction between [ $\pm$ ATR] vowels in Pulaar must be represented by adding underlying /e/ and /o/ to the inventory instead of something like Paradis's (1992) abstract high vowel analysis.

Under the strictest interpretation of the Contrastivist Hypothesis, any evidence that a feature is phonologically active—even in only a single form—implies that that feature should be categorically contrastive. Even in the contrastivist framework, though, isolated examples are more plausibly treated as exceptions rather than as reflective of fundamental properties of the system of representations; particularly from the perspective of learnability, the more robustly attested a contrast is, the more reasonable it is to posit that it is systematically encoded. It is therefore tremendously useful to have mathematical methods for quantifying gradient contrastiveness (Hall 2009) and software for applying those methods (Hall et al. 2015a), even in a theoretical framework that is primarily concerned with contrast as a categorical phenomenon (Hall 2007; Drescher 2009).

### Abbreviations

ATR = advanced tongue root, CH = Contrastivist Hypothesis, DIM = diminutive, IPA = International Phonetic Alphabet, PL = plural, RTR = retracted tongue root, SDA = Successive Division Algorithm, SG = singular, sth = something

### Acknowledgements

The authors wish to thank Henry Hane and Abubakar Atiku Alkali for help with the Pulaar data, John T. Hall and Doug Pulleyblank for assistance in marshalling and digitizing resources on Pulaar, Yuni Kim and Pavel Iosad for organising the themed session at the 2013 meeting of the LAGB that prompted the writing of this paper, and three anonymous reviewers for their helpful comments and suggestions.

### Competing Interests

The authors have no competing interests to declare.

### References

- Archangeli, Diana. 1984. *Underspecification in Yawelmani phonology and morphology*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA. Retrieved from <http://hdl.handle.net/1721.1/15440>.
- Archangeli, Diana. 1985. Yokuts harmony: Evidence for coplanar representation in non-linear phonology. *Linguistic Inquiry* 16(3). 335–372. Retrieved from <http://www.jstor.org/stable/4178441>.
- Archangeli, Diana. 1988. Underspecification in phonology. *Phonology* 5(2). 183–207. DOI: <http://dx.doi.org/10.1017/S0952675700002268>
- Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.
- Avery, J. Peter & Keren D. Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6(2). 179–200. DOI: <http://dx.doi.org/10.1017/S0952675700001007>
- Bah, Oumar. 2014. *Dictionnaire pular*. Dallas: SIL International. Published online at <http://pular.webonary.org/>.
- Boulenger, Véronique, Emmanuel Ferragne, Nathalie Bedoin & François Pellegrino. 2011. Derived contrasts in Scottish English. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*, 352–355.
- Brame, Michael K. 1972. On the abstractness of phonology: Maltese  $\text{ʔ}$ . In Michael K. Brame (ed.), *Contributions to generative phonology*, 22–61. Austin: University of Texas Press.
- Bye, Patrik. 2009. Three types of marginal contrast. Presented at the Toronto–Tromsø Phonology Workshop, University of Toronto, October 2009.

- Calabrese, Andrea. 1995. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26(3). 373–463. Retrieved from <http://www.jstor.org/stable/4178906>.
- Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter. DOI: <http://dx.doi.org/10.1515/9783110197600>
- Campos-Astorkiza, Judit Rebeka. 2007. *Minimal contrast and the phonology–phonetics interaction*. Ph.D. thesis, University of Southern California, Los Angeles, CA. Retrieved from <http://digitallibrary.usc.edu/cdm/ref/collection/p15799coll127/id/514484>.
- Carr, Philip. 2008. *A glossary of phonology*. Edinburgh: Edinburgh University Press.
- Carvalho, Joaquim Brandão de. 2004. Templatic morphology in the Portuguese verb. In Trudel Meisenburg & Maria Selig (eds.), *Nouveaux départs en phonologie*, 13–32. Tübingen: Gunter Narr.
- Casali, Roderic F. 2003. [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7(3). 307–382. DOI: <http://dx.doi.org/10.1515/lity.2003.018>
- Cherry, E. Colin, Morris Halle & Roman Jakobson. 1953. Toward the logical description of languages in their phonemic aspect. *Language* 29(1). 34–46. DOI: <http://dx.doi.org/10.2307/410451>
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Clements, G. N. 1987. Toward a substantive theory of feature specification. In James Blevins & Juli Carter (eds.), *Proceedings of NELS 18*, vol. 1, 79–93. Amherst, MA: GLSA.
- Clements, G. N. 2009. The role of features in phonological inventories. In Eric Raimy & Charles E. Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 19–68. Cambridge, MA: MIT Press. DOI: <http://dx.doi.org/10.7551/mitpress/9780262182706.003.0002>
- Clements, G. N. & Elizabeth V. Hume. 1995. Internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonology*, 245–306. Oxford: Blackwell.
- Compton, Richard & B. Elan Dresher. 2011. Palatalization and “strong *i*” across Inuit dialects. *Canadian Journal of Linguistics* 56(2). 203–228. DOI: <http://dx.doi.org/10.1353/cjl.2011.0020>
- Crothers, John. 1978. Typology and universals of vowel systems. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, vol. 2, 93–152. Stanford, CA: Stanford University Press.
- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press. DOI: <http://dx.doi.org/10.1017/CBO9780511642005>
- Dresher, B. Elan. 2011. The phoneme. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 1, 241–266. Oxford: Wiley-Blackwell.
- Dresher, B. Elan. 2013. Contrastive features and microvariation in vowel harmony. In Stefan Keine & Shayne Sloggett (eds.), *NELS 42: Proceedings of the forty-second annual meeting of the North East Linguistic Society*, 141–153. Amherst, MA: GLSA.
- Dresher, B. Elan. 2015. The motivation for contrastive feature hierarchies in phonology. *Linguistic Variation* 15(1). 1–40. DOI: <http://dx.doi.org/10.1075/lv.15.1.01dre>
- Dresher, B. Elan, Christopher Harvey & Will Oxford. 2014. Contrast shift as a type of diachronic change. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.), *NELS 43: Proceedings of the forty-third annual meeting of the North East Linguistic Society*, vol. 1, 103–116. Amherst, MA: GLSA.

- Dresher, B. Elan, Glyne L. Piggott & Keren D. Rice. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13. iii–xvii. Retrieved from <http://twpl.library.utoronto.ca/index.php/twpl/article/view/6331>.
- Durand, Jacques & Chantal Lyche. 2004. Structure et variation dans quelques systèmes vocaliques du français: L'enquête Phonologie du français contemporain. In Aidan Benedict Coveney, Marie-Anne Hintze & Carol Sanders (eds.), *Variation et francophonie*, 217–240. Paris: L'Harmattan.
- Dyck, Carrie Joan. 1995. *Constraining the phonology–phonetics interface, with evidence from Spanish and Italian dialects*. Ph.D. thesis, University of Toronto, Toronto. Retrieved from <http://twpl.library.utoronto.ca/index.php/twpl/article/view/652>.
- Ferragne, Emmanuel, Nathalie Bedoin, Véronique Boulenger & François Pellegrino. 2011. The perception of a derived contrast in Scottish English. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*, 667–670.
- Féry, Caroline. 2003. Markedness, faithfulness, vowel quality, and syllable structure in French. *Journal of French Language Studies* 13(2). 247–280. DOI: <http://dx.doi.org/10.1017/S0959269503001121>
- Gendrot, Cédric & Martine Adda-Decker. 2007. Impact and duration of vowel inventory size on formant values of oral vowels: An automated formant analysis from eight languages. *International Conference of Phonetic Sciences* 16. 1417–1420.
- Ghini, Mirco. 2001. Place of articulation first. In T. Alan Hall (ed.), *Distinctive feature theory*, vol. 2, Phonetics and Phonology, 147–176. Berlin: Mouton de Gruyter. DOI: <http://dx.doi.org/10.1515/9783110886672.147>
- Gleason, H. A. (Jr.). 1961. Review of *African language studies I* by Malcolm Guthrie; *The role of tone in the structure of Sùkúma* by I. Richardson. *Language* 37(2). 294–308. DOI: <http://dx.doi.org/10.2307/410866>
- Goldsmith, John A. 1995. Phonological theory. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 1–23. Cambridge, MA: Blackwell.
- Grammont, Maurice. 1913. *Le vers français, ses moyens d'expression, son harmonie*. Paris: H. Champion.
- Hall, Daniel Currie. 2000. Redundant features in Pulaar... and why they aren't. Presented at the Montréal–Ottawa–Toronto (MOT) Phonology Workshop, York University, Toronto, February 2000.
- Hall, Daniel Currie. 2007. *The role and representation of contrast in phonological theory*. Ph.D. thesis, University of Toronto, Toronto. Retrieved from <http://twpl.library.utoronto.ca/index.php/twpl/article/view/6497>.
- Hall, Daniel Currie. 2011a. Contrast. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 1, 27–53. Oxford: Wiley-Blackwell.
- Hall, Daniel Currie. 2011b. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28(1). 1–54. DOI: <http://dx.doi.org/10.1017/S0952675711000029>
- Hall, Daniel Currie. forthcoming. Contrastive specification in phonology. In Mark Aronoff (ed.), *The Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. DOI: <http://dx.doi.org/10.1093/acrefore/9780199384655.013.26>
- Hall, Daniel Currie & B. Elan Dresher. 2016. Trade-offs in the contrastive hierarchy: Voicing versus continuancy in Slavic. In Christopher Hammerly & Brandon Prickett (eds.), *NELS 46: Proceedings of the forty-sixth annual meeting of the North East Linguistic Society*, vol. 2, 43–56. Amherst, MA: GLSA.

- Hall, Kathleen Currie. 2009. *A probabilistic model of phonological relationships from contrast to allophony*. Ph.D. thesis, The Ohio State University, Columbus, OH. Retrieved from [http://rave.ohiolink.edu/etdc/view?acc\\_num=osu1250228987](http://rave.ohiolink.edu/etdc/view?acc_num=osu1250228987).
- Hall, Kathleen Currie. 2012. Phonological relationships. *McGill Working Papers in Linguistics* 22(1). 1–14. Retrieved from <http://www.mcgill.ca/mcgwpl/archives/volume-221-2012>.
- Hall, Kathleen Currie. 2013a. Documenting phonological change: A comparison of two Japanese phonemic splits. In Shan Luo (ed.), *Proceedings of the 2013 annual meeting of the Canadian Linguistic Association*, 1–15. Toronto: Canadian Linguistic Association. Published online at <http://cla-acl.ca/actes-2013-proceedings/>.
- Hall, Kathleen Currie. 2013b. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215–275. DOI: <http://dx.doi.org/10.1515/tlr-2013-0008>
- Hall, Kathleen Currie, Blake Allen, Michael Fry, Scott Mackie & Michael McAuliffe. 2015a. Phonological CorpusTools. Computer program. Retrieved from <https://github.com/PhonologicalCorpusTools/CorpusTools/releases>.
- Hall, Kathleen Currie & Elizabeth Hume. 2015. Modeling perceived similarity: The influence of phonetics, phonology and frequency on the perception of French vowels. Ms., University of British Columbia and University of Canterbury.
- Hall, Kathleen Currie, Veronica Letawsky, Alannah Turner, Claire Allen & Kevin McMullin. 2015b. Effects of predictability of distribution on within-language perception. In Santa Vinerte (ed.), *Proceedings of the 2015 annual conference of the Canadian Linguistic Association*, 1–14. Toronto: Canadian Linguistic Association. Published online at <http://cla-acl.ca/actes-2015-proceedings/>.
- Harris, John. 1990. Derived phonological contrasts. In Susan Ramsaran (ed.), *Studies in the pronunciation of English: A commemorative volume in honour of A. C. Gimson*, 87–105. London: Routledge.
- Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
- Hill, Kenneth C. 1998. Review of *Registro de la veración fonológica en el náhuatl moderno: Un estudio de caso* by Cristina Monzón. *International Journal of American Linguistics* 64(1). 70–73. DOI: <http://dx.doi.org/10.1086/466348>
- Hualde, José Ignacio. 2005. Quasi-phonemic contrasts in Spanish. *WCCFL* 23. 374–398.
- Huang, Tsan. 2001. The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. *Ohio State University Working Papers in Linguistics* 55. 23–42.
- Hume, Elizabeth & Keith Johnson. 2003. The impact of partial phonological contrast on speech perception. In Maria J. Solé, Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the fifteenth International Congress of Phonetic Sciences*, 2385–2388. Barcelona: Casual Productions.
- Hyman, Larry M. 1970. How concrete is phonology? *Language* 46(1). 58–76. DOI: <http://dx.doi.org/10.2307/412407>
- Iosad, Pavel. 2012. *Representation and variation in substance-free phonology: A case study in Celtic*. Ph.D. thesis, University of Tromsø, Tromsø. Retrieved from <http://hdl.handle.net/10037/4794>.
- Itô, Junko, Armin Mester & Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26(4). 571–613. Retrieved from <http://www.jstor.org/stable/4178915>.
- Itô, Junko & Ralf-Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17(1). 49–73. Retrieved from <http://www.jstor.org/stable/4178471>.


- Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*, Språkvetenskapliga Sällskapet i Uppsala Förhandlingar. Uppsala: Uppsala Universitetet.
- Jakobson, Roman, C. Gunnar M. Fant & Morris Halle. 1952. Preliminaries to speech analysis. *Tech. Rep.* 13, MIT Acoustics Laboratory.
- Kager, René. 2008. Lexical irregularity and the typology of contrast. In Kristin Hanson & Sharon Inkelas (eds.), *The nature of the word*, 397–432. Cambridge, MA: MIT Press. DOI: <http://dx.doi.org/10.7551/mitpress/9780262083799.003.0017>
- Kiparsky, Paul. 1982a. From cyclic to Lexical Phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations, part I*, 131–176. Dordrecht: Foris.
- Kiparsky, Paul. 1982b. How abstract is phonology? In Paul Kiparsky (ed.), *Explanation in phonology*, 119–163. Dordrecht: Foris. Originally published in 1968 by the Indiana University Linguistics Club.
- Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2. 85–138. DOI: <http://dx.doi.org/10.1017/S0952675700000397>
- Krämer, Martin. 2003. *Vowel harmony and correspondence theory*. Berlin: Mouton de Gruyter. DOI: <http://dx.doi.org/10.1515/9783110197310>
- Ladd, D. Robert. 2006. ‘Distinctive phones’ in surface representation. In Louis Goldstein, D. H. Whalen & Catherine T. Best (eds.), *Laboratory Phonology 8*, 3–26. Berlin: Mouton de Gruyter.
- Mackenzie, Sara. 2009. *Contrast and similarity in consonant harmony processes*. Ph.D. thesis, University of Toronto, Toronto. Retrieved from <http://hdl.handle.net/1807/17474>.
- Mackenzie, Sara. 2013. Laryngeal co-occurrence restrictions in Aymara: Contrastive representations and constraint interaction. *Phonology* 30(2). 297–345. DOI: <http://dx.doi.org/10.1017/S0952675713000146>
- Mackenzie, Sara & B. Elan Dresher. 2003. Contrast and phonological activity in the Nez Perce vowel system. In Pawel M. Nowak, Corey Yoquelet & David Mortensen (eds.), *Proceedings of BLS 29*, 283–294. Berkeley, CA: Berkeley Linguistics Society. DOI: <http://dx.doi.org/10.3765/bls.v29i1.979>
- Mateus, Maria Helena & Ernesto d’Andrade. 2000. *The phonology of Portuguese*. Oxford: Oxford University Press.
- Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.
- Nevins, Andrew Ira. 2010. *Locality in vowel harmony*. Cambridge, Mass.: MIT Press. DOI: <http://dx.doi.org/10.7551/mitpress/9780262140973.001.0001>
- Nevins, Andrew Ira. 2015. Triumphs and limits of the contrastivity-only hypothesis. *Linguistic Variation* 15(1). 41–68. DOI: <http://dx.doi.org/10.1075/lv.15.1.02nev>
- New, Boris, Christophe Pallier, Marc Brysbaert & Ludovic Ferrand. 2004. *Lexique 2: A new French lexical database*. *Behavior Research Methods, Instruments, & Computers* 36(3). 516–524. DOI: <http://dx.doi.org/10.3758/BF03195598>
- Niang, Mamadou Ousmane. 1997. *Pulaar–English English–Pulaar standard dictionary*. New York: Hippocrene.
- Ohala, John J. 1981. The listener as a source of sound change. In Carrie S. Masek, Roberta A. Hendrick & Mary F. Miller (eds.), *Papers from the parasession on language behavior*, 178–203. Chicago: Chicago Linguistic Society.
- Ohala, John J. 2012. The listener as a source of sound change: An update. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: Perception, production, and social factors*, 21–35. Amsterdam: John Benjamins. DOI: <http://dx.doi.org/10.1075/cilt.323.05oha>

- Oxford, Will. 2015. Patterns of contrast in phonological change: Evidence from Algonquian vowel systems. *Language* 91(2). 308–357. DOI: <http://dx.doi.org/10.1353/lan.2015.0028>
- Paradis, Carole. 1986. *Phonologie et morphologie lexicales*. Ph.D. thesis, Université de Montréal, Montréal.
- Paradis, Carole. 1992. *Lexical phonology and morphology*. New York: Garland. Translated by the author from Paradis (1986).
- Piggott, Glyne L. 1992. Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* 10(1). 33–77. DOI: <http://dx.doi.org/10.1007/BF00135358>
- Pulleyblank, Douglas. 1988. Underspecification, the feature hierarchy and Tiv vowels. *Phonology* 5(2). 299–326. DOI: <http://dx.doi.org/10.1017/S0952675700002323>
- Scobbie, James M. & Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the indeterminacy of the segmental inventory. In J. Peter Avery, B. Elan Dresher & Keren D. Rice (eds.), *Contrast in phonology*, 87–114. Berlin: Mouton.
- Shannon, Claude E. & Warren Weaver. 1949. *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- Spahr, Christopher. 2014. A contrastive hierarchical account of positional neutralization. *The Linguistic Review* 31(3–4). 551–585. DOI: <http://dx.doi.org/10.1515/tlr-2014-0008>
- Spahr, Christopher. 2016. Confronting the European Portuguese central vowel distinction. *Canadian Journal of Linguistics* 61(2). 211–217. DOI: <http://dx.doi.org/10.1017/cnj.2016.9>
- Steriade, Donca. 2007. Contrast. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 139–157. Cambridge: Cambridge University Press.
- Trubetzkoy, Nikolai S. 1939. Grundzüge der Phonologie. *Travaux du cercle linguistique de Prague* 7.
- Trubetzkoy, Nikolai S. [1939] 1969. *Principles of phonology*. Berkeley: University of California Press. Translated by Christiane A. M. Baltaxe from Trubetzkoy (1939).
- Walker, Douglas C. 2001. *French sound structure*. Calgary: University of Calgary Press.
- Yu, Alan C. L. 2011. Mergers and neutralization. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 3, 1892–1918. Oxford: Wiley-Blackwell.

**How to cite this article:** Hall, Daniel Currie and Kathleen Currie Hall. 2016. Marginal contrasts and the Contrastivist Hypothesis. *Glossa: a journal of general linguistics* 1(1): 50. 1–23, DOI: <http://dx.doi.org/10.5334/gjgl.245>

**Published:** 06 December 2016

**Copyright:** © 2016 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 