

Predicting Perceptually Weak and Strong Unmarked Patterns: A Message-based Approach

Elizabeth Hume,¹ Kathleen Currie Hall,² Andrew Wedel³
¹University of Canterbury, ²University of British Columbia, ³University of Arizona

1 Introduction

Perceptual factors have been drawn on to provide insight into sound patterns (see e.g., Ohala 1981; Lindblom 1990; Flemming 1995; Jun 1995; Hume 1998; Steriade 2008) and commonly serve as a diagnostic for markedness (e.g. Hamilton 1996; Boersma 1998). However, a puzzling situation has emerged: patterns associated with strong perceptual distinctiveness and those with weak distinctiveness are both described as unmarked. For example, it is widely assumed that the unmarked position for most consonants is in prevocalic/onset position (e.g. [ta]), a pattern commonly associated with multiple inherent and/or contextual phonetic cues (Wright 1996). Yet, unmarkedness is also associated with weak cues, as in positions of neutralization (Silverman 2012). In coda position, for example, Korean plain voiceless, aspirated and tense consonants (e.g. [t t^h t']) surface as the plain consonant (Cho 1990), the unmarked member of the opposition (Trubetzkoy 1939; Rice 1999). Context is important in understanding why strong and weak cues can both be unmarked: prevocalic/onset position hosts unmarked consonants with strong cues and postvocalic/coda position hosts unmarked consonants with weak cues. Yet *why* should the unmarked sound in coda position be commonly associated with weak cues while the unmarked consonant in onset position be associated with strong cues? Or, more fundamentally, what does it mean for a pattern to be unmarked?

In this paper we propose that insight into the unmarked nature of these patterns can be gained when we take seriously the view of language as a system of information transmission. In particular, we suggest that perceptually weak and strong unmarked patterns are those that effectively balance two competing properties of effective communication: (a) the contribution of the phonological unit in context to accurate message transmission, and (b) the resource cost of the phonological unit (see Hall, Hume, Jaeger & Wedel, forthcoming). In order to demonstrate this, we begin in §2 by describing key properties of language as a system of information transmission. In the following section, we turn to the predictions that follow from such a system for perceptually weak and strong sound patterns and show why it makes sense for them both to be described as unmarked.

2 Language as a system of information transmission

Central to this approach is the fact that language is a system of information transmission.¹ As such, it involves transmitting a message from one person/point to another. For spoken language, we can think of the message as a unit of meaning, e.g. a word, that is transmitted by an acoustic/auditory speech signal. When a system supports error-free message transmission in most contexts, the system is considered robust. Yet, a communication system must also be cost-effective, using as few resources as necessary to transmit messages. An *effective* system of information transmission is both robust and cost-effective.

Acknowledgements: The ideas in this paper build on a phonological framework that we have developed in conjunction with Florian Jaeger and thus we acknowledge his many contributions to the development of that approach (Hall et al., forthcoming). We are grateful to AMP 2015 participants for their constructive input on the ideas in this paper.

¹ The view that communicative functions shape language has a long history and includes, among others, Zipf 1932; Jakobson, Cherry & Halle 1952; Martinet 1972; Lindblom 1990; Silverman 1997; Boersma 1998; Bybee 2001; Kirby 2010.

© 2016 Elizabeth Hume, Kathleen Currie Hall, Andrew Wedel
Proceedings of AMP 2015
Completed April 29, 2016

One can approximate the ideal communication system by taking into account the amount of uncertainty associated with selecting the outcome of a transmitted message (Shannon 1948), e.g. choosing a word from a set of potential words. The greater the number of choices, the higher the uncertainty associated with choosing the outcome. Consider the two sentences in (1). For most readers, there is probably low uncertainty associated with predicting the outcome in (a). This is because the set of possible outcomes is small, most likely made up of *front* and/or *coronal*. Selecting the outcome in (b) has higher uncertainty since the set of potential outcomes includes all possible consonant types.

- (1) Uncertainty associated with selecting an outcome
- | | |
|---|------------------|
| a. [i e ε æ] belong to the class of _____ vowels. | LOW UNCERTAINTY |
| b. He pronounced a _____ consonant. | HIGH UNCERTAINTY |

In addition to choice, uncertainty is affected by noise, which can impact how identifiable a message's signal is. Noise can come from external sources, e.g. doors closing, cars passing. It can also be due to impediments in sending or receiving a message that relate to, e.g. planning, producing, processing a message. All combined, choice and noise can negatively impact the probability of transmitting a message without error.

3 Manipulating Redundancy

Depending on the amount of uncertainty associated with a message, manipulating redundancy can serve to enhance the robustness and cost-effectiveness of a communication system (Shannon 1948). When uncertainty is high, additional or enhanced cues to the identity of the message can lower the potential for errors, thereby increasing system robustness. When uncertainty is low, fewer cues are needed; consequently, redundancy can be reduced thus lowering resource costs.

Redundancy can be modified in spoken messages through the addition/enhancement or deletion/reduction of cues to the identity of phonological units in the message, and so to the message itself.² There are multiple dimensions along which redundancy can be enhanced or reduced, depending on the nature of the sounds involved and the context in which they occur. Modifying redundancy associated with a prevocalic stop, for example, could involve changes to the intensity of the burst or the length of the closure, while for a postvocalic stop, it could result in the addition or suppression of a burst, or enhancement/reduction of the closure. For our present purposes, a hypothetical *phonological unit in context* (PUC)³ can be represented in an abstract space defined by degrees of redundancy, as in the *Redundancy Continuum* for a word-final voiceless coronal stop <Vt#> in (2). We assume for simplicity that the symbols represent continuous phonetic properties associated with the stop consonant. Where a particular PUC is situated along the continuum depends in part on the rest of the linguistic system of which it is a part.

- (2) Hypothetical Redundancy Continuum for word-final <Vt#>
- | | |
|--|--------------------------|
| <i>lower redundancy</i> | <i>higher redundancy</i> |
|V_# V?# Vt# Vt ^h # Vt ^{hV} #.... | |

All else being equal, adding/enhancing cues to a message increases the probability of transmitting the message error-free, thus enhancing the robustness of the communication system. However, increasing redundancy generally imposes a cost on the system.⁴ As shown below, understanding that language systems

² A message's redundancy can also be modified at levels above the word through, for example, discourse markers.

³ A crucial assumption in this approach is that to understand asymmetrical sound patterns, the analysis of the sound in question cannot be divorced from the message, e.g. word, in which it occurs. This includes syntagmatic contexts within the word such as, for English /t/: <#tV> (top), <t_#> (hat), <VtV> (butter), etc., as well as the broader contexts within which a word occurs, e.g. preceding/following word, as these can also influence sound shapes (see below). We use the term *phonological unit in context* (PUC), as a general term referring to the phonological elements that make up words and larger units of meaning thereby including not only sounds but also units of smaller (e.g. features) or larger (e.g. syllables) size. The symbols <> are used to indicate PUCs.

⁴ As a simplifying assumption, we set aside considerations that cues may have different costs on the system and that some may come for “free” (for related discussion see e.g., Stevens & Keyser 1989).

balance both robustness and efficiency provides new insights into why asymmetrical patterns occur.

4 Message Modification

A diverse body of evidence suggests that a word that occurs, on average, in contexts with high probability will be prone to reduction, while one that occurs, on average, in contexts with low probability will be prone to enhancement (e.g. Zipf 1932; Aylett & Turk 2004; Cohen Priva 2015; Piantadosi et al. 2011; Seyfarth 2014). Building on this work, we can make predictions about where in a message change would be most effective. Specifically, given the trade-off between system robustness and resource cost, redundancy should be enhanced in phonological-units-in-context (PUCs) that have the potential to contribute most to accurate message transmission, while having the least negative impact on resource cost. Conversely, redundancy should be reduced in PUCs that have the potential to provide the least benefit to successful message transmission, thereby reducing resource cost. Phonological systems are predicted to evolve in ways that are consistent with these predictions, stated in (3).

(3) Message Modification Principles

- a. Enhance redundancy on phonological-units-in-context that will contribute the most to successful message transmission, while having the least negative impact on resource cost.
- b. Reduce redundancy in phonological-units-in-context that will provide the least benefit to successful message transmission, thereby reducing resource cost.

A key assumption in (3) is that greater redundancy is more cost-effective in some parts of a message than others, a view shared by phonological theories such as positional faithfulness (Beckman 1998, Smith 2002) and cue-based approaches (Steriade 2008). It is also well supported by phonetic and psycholinguistic evidence showing the importance of certain positions for accurate word identification (e.g., Cutler et al. 1985). In general terms, such PUCs are characterized as having acoustic/auditory prominence, and/or high informativity (Seyfarth 2014, Cohen Priva 2015), as stated in (4).

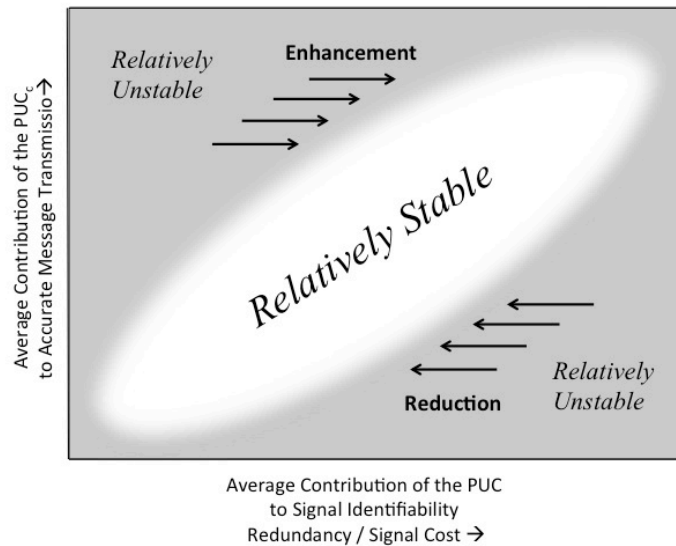
(4) Relative Importance of Phonological-Units-in-Context (PUCs)

The importance of a PUC in a message is correlated with the potential contribution that it could make to successful message transmission, given: (a) the physical prominence of the PUC; and/or (b) the informativity of the PUC.

PUCs with acoustic/auditory prominence can be prosodically or segmentally defined and include, for example, stressed (vs. unstressed) syllable, onset (vs. coda) position, or prevocalic (vs. preconsonantal) consonant. PUCs associated with greater prominence are those that have a low degree of predictability given the context. In general, segments at word beginnings typically have low predictability: because the speech stream is processed incrementally below the level of the word, cues at the word-beginnings contribute more to lexical identification than to cues later in the word. Take, for example, the English word *banana*. The initial 'b' reduces the set of possible messages from all words starting with any sound in English to the set of possible messages beginning with [#b]. The sequence [#bə] further constrains the set of outcomes to only those in which [ə] follows [b]. Segments toward the end of a word are typically predictable given the segments processed earlier, and thus contribute less to identifying the message (see van Son & Pols 2003 for related discussion and quantification along these lines).

Modifying redundancy in response to message uncertainty lays the foundation for understanding why some sound patterns are stable and others are unstable. The figure in (5) provides the basics of the model (see Hall et al., forthcoming). The delimited area in the figure can be thought of an abstract space that includes all potential PUCs associated with a given message. The y-axis represents the average contribution of the PUC to accurate message transmission. The x-axis is the average contribution of the PUC to the signal's identifiability; that is, the degree to which the signal uniquely identifies the message in question. Signal identifiability is modulated by the amount of redundancy in the signal with greater redundancy correlated with higher resource cost.

- (5) Relative stability of a message's phonological-units-in-context (PUC)
(based on Hall et al., forthcoming)



To the extent that a word-final PUC contributes little to the identity of the message in question, it would occur toward the lower part of the vertical axis; the redundancy continuum for word-final $\langle Vt \rangle$ given in (2) would be situated along the horizontal axis and toward the bottom end of the vertical axis. In contrast, a hypothetical redundancy continuum for a word-initial $\langle \#tV \rangle$, would be placed towards the top of the y-axis.

While the goal of this short paper is not to provide a detailed implementation of the proposed model, it is worth noting that *relative contribution to message identity* on the y-axis is quantifiable using mathematical tools such as informativity (see, e.g. van Son & Pols 2003, Seyfarth 2014, Cohen Priva 2015, for measures of segment and word predictability). See §3 for discussion regarding the quantification of resource cost along the x-axis.

Areas of relative stability and instability are shown in (5). Recall that stable patterns are those predicted to recur with regularity cross-linguistically while unstable patterns are underrepresented typologically and prone to change. Two regions of instability are indicated in the figure, toward the bottom right and the top left, indicating that the PUCs in these regions are unstable for difference reasons. A PUC in the bottom right has high redundancy, but since its potential contribution to error-free communication is low, having multiple cues to its identity is a costly use of resources. Consider as an example a voiced obstruent occurring word-finally in a word where the contrast between voiced and voiceless consonants distinguishes few, if any, words (i.e., has low functional load). In such a case, voicing in the word-final consonant plays a minor role in distinguishing the message from other potential messages thereby making it unstable. We refer to these patterns as *low contribution-high redundancy* (LC-HR) PUCs.

A PUC in the top left quadrant is also unstable although in this case because its potential contribution to communicating the message is high; however, signal identifiability is low given that it has few, if any, cues to its identity. Examples of *high contribution-low redundancy* (HC-LR) PUCs include word-initial sequences with poorly discriminated initial sounds, e.g. homorganic glide + vowel sequences.

Both HC-LR and LC-HR patterns are predicted to be relatively unstable and prone to change. In the case of HC-LR PUCs, an increased investment in resources is beneficial since word-initial segments have the potential to make a large contribution to message identity. The model thus predicts that it is advantageous to enhance HC-LR cases since increasing cues to the identity of the segments reduces uncertainty associated with predicting the outcome of the message. Patterns of fortition such as onset glide strengthening in Porteño Spanish (Lozano 1979) are consistent with this prediction. Conversely, LC-HR PUCs are prone to reduction given that the signal contributes little to message identity. For the above

example of a word-final voiced obstruent, reducing the voicing cue would be consistent with this approach. The common pattern of word-final devoicing is thus predicted by this model, as are cases of neutralization affecting high redundancy-low contribution PUCs.

Stable patterns are those that balance the trade-off between the contribution that the PUC can make to message identification, and the resources needed for the signal to be uniquely identifiable. Stable sound patterns correspond to those that we have traditionally referred to as unmarked. These involve *low contribution-low redundancy* (LC-LR) PUCs and *high contribution-high redundancy* (HC-HR) PUCs. Examples of LC-LR PUCs include word-final voiceless obstruents, reduced vowels in unstressed syllables, and a nasal consonant preceding a homorganic obstruent. Such patterns are stable since the PUCs contribute little to distinguishing words and thus to identifying the message in question; in other words, the PUC serves little contrastive value. Conversely, PUCs with *high contribution-high redundancy* (HC-HR) have the potential to contribute significantly to identifying the message and, as a result, the investment of additional cues is an effective use of resources. Stable patterns of this type are predicted to have strong perceptual cues and support contrasts along multiple dimensions. Examples include word-initial consonants, full vowels in stressed syllables, prevocalic plosives, intervocalic geminate consonants, etc.⁵

5 Predictions for perceptually weak and strong sound patterns

With this as a basis, we return to the issue of why perceptually weak and strong sound patterns can both pattern as unmarked. The solution to this problem is that in both patterns, there is an effective trade-off between a PUC's resource cost and its potential contribution to accurate transmission of the message. In an effective system of information transfer, stable sound patterns with weak perceptual cues are predicted to be associated with PUCs that contribute little to message transmission accuracy; that is, the PUC has high relative predictability within the message, by occurring for example, toward the end of the word, or in a high frequency affix or function word. Investing resources in a PUC that contributes little to message identity is not cost-effective. Conversely, it is worthwhile to invest in PUCs that make a strong contribution to message transmission accuracy. Stable patterns with perceptually strong sounds are thus predicted to occur in contexts that are informative since strong perceptual cues support the transmission of unpredictable information, e.g. beginning of words, content words (as opposed to affixes/function words). Stable PUCs with strong cues occur toward the upper right quadrant of the figure in (5), while those with weak cues are located in the lower left quadrant.

6 Conclusions

In this brief paper we have proposed a solution to the issue of why both perceptually weak patterns and perceptually strong sounds can pattern as “unmarked”. The reason, we suggest, is that both patterns effectively balance the potential cost associated with investing resources in the sound and its contribution to conveying the message accurately. In unifying these observations we have moved away from the traditional notions of marked/unmarked and reframed asymmetrical sound patterns as *stable* or *unstable*, defined in terms of the trade-off between resource cost and message transmission accuracy. In stable patterns the trade-off is effective and thus, the patterns recur with regularity cross-linguistically. The trade-off is ineffective in unstable patterns and, as a result, they are prone to change and are underrepresented typologically. This evidence suggests that language systems evolve to maximize the effective transfer of information.

We note in conclusion that our approach is message-based as opposed to sound-based and as such it generates strong predictions about the role of the message in shaping phonological systems. Significantly, given that the message typically takes the form of a morpheme, word, or higher unit of meaning, the

⁵ The discussion has focused on typologically common patterns yet there are also sound patterns within a given language that while not typologically common, are nonetheless “unmarked” in the language in question. For example, labial consonants, as opposed to coronals, bear the properties of stable PUCs in Sri Lankan Portuguese Creole (Hume & Tserdanelis 2002). As discussed in detail in Hall et al. (forthcoming), the current approach predicts that specific properties of the language's lexical system will predict such patterns.

approach predicts that the shape of phonological patterns occurs as a consequence of the trade-off between resource cost and accurate transmission of these messages. From this perspective, phonology does not exist for its own sake but rather in the service of transferring information.

7 References

- Aylett, Matthew & Alice Turk. 2004. The Smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47.31-56.
- Beckman, Jill. 1998. Positional faithfulness. Amherst, MA: University of Massachusetts Doctoral dissertation.
- Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph.D. dissertation, University of Amsterdam.
- Bybee, Joan L. 2001. *Phonology and Language use*. Cambridge: Cambridge UP.
- Cho, Yong-mee Yu. 1990. *Parameters of Consonantal Assimilation*. PhD dissertation. Stanford, CA.
- Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology* 6. 243-78.
- Cutler, A., J.A. Hawkins & G. Gilligan. 1985. The suffixing preference: a processing explanation. *Linguistics* 23. 723-58.
- Hall, Kathleen Currie, Elizabeth Hume, Florian Jaeger and Andrew Wedel. Forthcoming. The message shapes phonology. Ms.
- Hamilton, Philip. 1996. *Phonetic constraints and markedness in the phonotactics of Australian aboriginal languages*. Ph.D. dissertation, University of Toronto.
- Hume, Elizabeth. 1998. The Role of perceptibility in consonant/consonant metathesis. *WCCFL* 17. 293-307.
- Hume, Elizabeth & Frédéric Mailhot. 2013. The role of entropy and surprisal in phonologization and language change. *Origins of Sound Patterns: Approaches to Phonologization*, ed. by A. Yu, 29-50. Oxford: Oxford University Press.
- Hume, Elizabeth & Giorgios Tserdanelis. 2002. Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology* 19.441-58.
- Cherry, Colin, Morris Halle & Roman Jakobson. 1953. Toward the logical description of languages in their phonemic aspect. *Language* 29.34-46.
- Jun, Jongho. 1995. Place Assimilation as the Result of Conflicting Perceptual and Articulatory Constraints. *WCCFL* 14. 221-237.
- Kirby, James. 2010. Cue selection and category restructuring in sound change: University of Chicago Doctoral dissertation.
- Lindblom, Björn. 1990. *Explaining phonetic variation: A sketch of the H&H theory*. In Hardcastle & Marchal. Kluwer Academic. 403-439.
- Lozano, María del Carmen. 1979. *Stop and spirant alternations: Fortition and spirantization processes in Spanish phonology*. Bloomington, IN, Indiana University Linguistics Club: The Ohio State University Doctoral dissertation.
- Martinet, André. 1972. Function, structure, and sound change. *Reader in Historical and Comparative Linguistics*, ed. by A.R. Keiler, 139-74. New York: Holt, Reinhart, & Winston.
- Ohala, John. 1981. The Listener as a Source of Sound Change. In Masek, C.S., R.A. Hendrik, M. F. Miller (eds.), *Parasession on Language and Behavior: CLS*. Chicago: CLS. 178-203.
- Piantadosi, Steven T, Harry Tily & Edward Gibson. 2011. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* 108.3526-29.
- Rice, Keren. 1999. Featural markedness in phonology: variation. *Glott International* 4:7. 3-6, 4:8. 3-7.
- Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition* 133.140-55.
- Shannon, Claude E. 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.
- Silverman, Daniel. 192012. *Neutralization*. Cambridge, UK: Cambridge University Press.
- Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*. Amherst, MA: University of Massachusetts Doctoral dissertation.
- Steriade, Donca. 2008. The Phonology of perceptibility effects: The P-map and its consequences for constraint organization. In K. Hanson and S. Inkelas (eds.), *The Nature of the Word*, Cambridge: MIT Press. 151-179. (originally circulated in 2001)

- Stevens, Kenneth N. & Samuel Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65.81-106.
- Trubetzkoy, N. 1939. *Grundzüge der Phonologie*. Göttingen: van der Hoeck & Ruprecht. Translated 1969 by C. Baltaxe as *Principles of Phonology*. University of California Press.
- Van Son, R. J. J. H. & Louis C. W. Pols. 2003. How efficient is speech? *Proceedings of the Institute of Phonetic Sciences* 25.171-84.
- Wright, R. 1996. *Consonant Clusters and Cue Preservation in Tsou*. PhD dissertation. UCLA.
- Zipf, George Kingsley. 1932. *Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, MA.