SIMULATING THE INTERACTION OF FUNCTIONAL PRESSURES, REDUNDANCY AND CATEGORY VARIATION IN PHONETIC SYSTEMS

Bodo Winter[1] & Andy Wedel[2]

[1] University of California, Merced, Cognitive and Information Sciences
[2] University of Arizona, Department of Linguistics

ABSTRACT

Phonetic systems need to be able to signal communicatively relevant meaning distinctions. In this paper, we explore an evolutionary simulation which shows how the functional pressure to keep words perceptually distinct reduces variation at the phonetic level. Our simulation furthermore shows that adding redundancy to the system (e.g., through multiple phonetic cues or longer words) relaxes these functional pressures. Based on these results we argue that phonetic systems can be seen as finding a relative optimum: Efficient and unambiguous communication is maintained while at the same time, there is enough category variation to allow evolvability, the potential for future evolution.

Keywords: redundancy; category variation; evolutionary simulations; exemplar models; cultural evolution

1. INTRODUCTION

The sound systems of spoken languages constantly change, at both long and short time scales [1, 27, 8]. What remains constant amidst these changes is the ability of sound systems to subserve communication [39]. A broad range of work argues that this arises in part through functional pressures on language evolution to maintain sufficient contrast in phonetic categories [3, 13, 21, 29, 33, 34, 41]. In this paper, we look at how phonological systems evolve under such usage constraints.

We specifically investigate the role of a functional pressure towards keeping words acoustically distinguishably (henceforth "anti-ambiguity bias"). We suggest that this bias constrains variation at the phonetic level, i.e., different renderings of the same utterance vary less if lexical items need to be contrasted. Moreover, our simulations show that infusing redundancy into phonetic systems (e.g., via multiple phonetic cues or via longer words) relaxes this functional pressure, allowing systems to harbour more variation.

2. BACKGROUND

Phoneme contrasts such as /p~b/, /s~f/, or /ɑ~ɔ/ can be lost from a language when, for example, two phonemes merge with one another [16, Ch. 11]. As an example, the contrast between /ɑ~ɔ/, (exemplified by the words "cot" and "caught") has merged in many dialects of North America [17]. Wedel, Jackson, and Kaplan [35] demonstrated that the probability of such merger is cross-linguistically associated with how many lexical items are distinguished by the phonemic contrast: a greater number of such "minimal pairs" is significantly associated with lower merger probability (see also [36]).

The linguistic literature is rife with anecdotal reports of these kinds of effects as well. Blevins and Wedel [2] discuss attested cases of "inhibited" sound changes, where an otherwise regular sound change ignores sets of words that would lead to the breakdown of an entire morphological paradigm, such as the distinction between past tense and present tense. These kinds of observations suggests that biases toward communicative efficiency do influence the evolution of phonological systems. Moreover, they corroborate the statistical studies [35,36] which indicate that maintenance of contrast at the lexical level in particular, influences the phonetic level.

Wedel [34] proposes a multi-level exemplar model to account for these interactions (cf. [31]). In exemplar models of speech, phonological knowledge is characterized as being constituted by rich and detailed representations of experience rather than by abstract symbolic representations. Within this framework, a phonological category can be modeled as a collection of stored phonetic exemplars (an "exemplar cloud") that is acquired and continually enriched by experience.

Language evolution can then be modelled as resulting from a repeated cycle of production and perception events [26, 33]. To model interactions between the lexicon and sublexical structure, experiences need to contribute to two connected levels of representation: a lexical level, and a sublexical level (phonetics). In Wedel's computational model [32, 33, 34], an anti-ambiguity bias at the lexical level results in the evolution of a phoneme set that efficiently subserves lexical distinctions.

This paper extends this model to explore what has been called "cryptic variation" in biological systems [10, 30]. Cryptic variation refers to variation that is not selected for or against, that is, neutral variation that does not impact fitness. For language, it has been noted that sound systems harbour variation that is not consciously perceived by speakers [25], and therefore is not subject to overt communicative pressures.

We can use the concept of cryptic variation to understand an observation that has frequently been made by linguists: Sound systems that do not make use of certain distinctions tend to "allow for" or "afford" more variation. For example, Lavoie [19] shows that native speakers of English produce spirantized variants of /k/ more frequently than native speakers of Spanish, where /k/ and /x/ are contrastive. Thus, the English sound category of /k/ encroaches into "unfilled areas of the language's sound space" [19, p. 39]. However, in Spanish variation is more constrained, presumably because of the functional significance of /x/ in that particular language's system. This reduction of variation due to communicative significance is what we set out to model.

3. THE COMPUTATIONAL MODEL

This section briefly outlines the computational model, with more technical detail provided in [34]. In the model, two agents take turns talking to each other. Each agent has an internal lexicon. The speaker utters one token of each of the words, and the listener maps each token to its best fitting category, where it stores the input as a new exemplar.

Each word exemplar is further decomposed into a number of phonetic exemplars on one of two possible continuous dimensions, each with an arbitrary scale from 1-100. As a useful metaphor, we can think of one of the dimensions as voice-onset time (VOT), and the other dimension as vowel height on an /i-a/ continuum. Thus, each word exemplar maps onto a point in 2-dimensional space. For example, a token with the values [15 VOT, 25 TongueHeight] can be thought of as corresponding to [ba].

Each new exemplar is associated with an initial activation value that decreases over time, corresponding to the observation that memories decay [11, 15, 24, 26]. In production, exemplars are selected as a function of the activation level, with more strongly activated exemplars contributing more strongly to a production plan.

For each word production, a random exemplar is chosen from the word's exemplar cloud. Two types of changes apply to the target before it is passed to the listener for categorization: the addition of production noise, and the application of a similarity bias [26], reviewed in [33]. This similarity bias is implemented by biasing the phonetic values of the output target toward nearby values in memory both within the word category itself, and across the lexicon. The move towards nearby values of the lexicon leads the system to re-use phonetic features across words, which is a defining characteristic of human languages [12, 18, 20]. There is empirical support for such a cross-word similarity bias, reviewed in [34].

A final feature of the model is a bias against lexical confusability [34]. A bias with this effect is empirically motivated by the above-mentioned cross-linguistic studies of phoneme merger and inhibited sound change. We implement this bias computationally in a straightforward fashion: an output has a chance of not being stored as a new exemplar in the listener's memory in proportion to the degree to which it maps to multiple categories [32, 33]. In this way, unambiguous speaker outputs are more likely to be stored than ambiguous outputs, with the result that unambiguous exemplars contribute relatively more to the continuing evolution of the lexicon. The central result from this work is that contrast relationships between sounds may be constrained and maintained by contrast relationships between words.

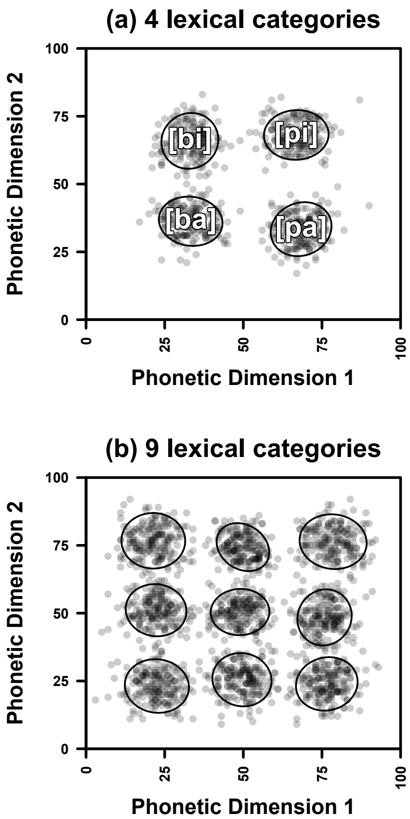
4. SIMULATION RESULTS

4.1. Category variation as a function of lexical density

We first explored how lexical density impacts cryptic variation at the phonetic level. Figure 1 shows a time slice of two representative simulations after 500 time steps, where agents either had 4 words or 9 words. Note that these are extremely small lexicons compared to natural languages because we are specifically interested in modelling elements of the lexicon that form a set of minimal pairs.

In the absence of other words, a lexeme's exemplar distribution is determined by the balance between noise, which promotes spread, and similarity bias, which promotes contraction [26]. When the exemplar clouds of two words get close enough such that some outputs become ambiguous in perception, the anti-ambiguity bias comes into play as well, which introduces an additional constraint on how broad a category can spread. This anti-ambiguity bias is stronger when there is higher lexical density. In other words, as we add more words in a given phonetic space, pronunciation variation at the boundaries between them becomes increasingly suppressed and the standard deviations of the exemplar clouds shrink. This can be seen in Figure 1, where the dashed lines indicate the standard deviations (SD) of all exemplar clouds for different numbers of lexical categories (after 500 simulation steps).

Figure 1: Simulation results with (a) 4 words and (b) 9 words, after 500 time steps. Each point represents an exemplar. Each cloud represents the totality of exemplars for a word. Labels are given for ease of interpretation. Ellipsoids represent confidence regions that cover 80% of the exemplars for each cloud.

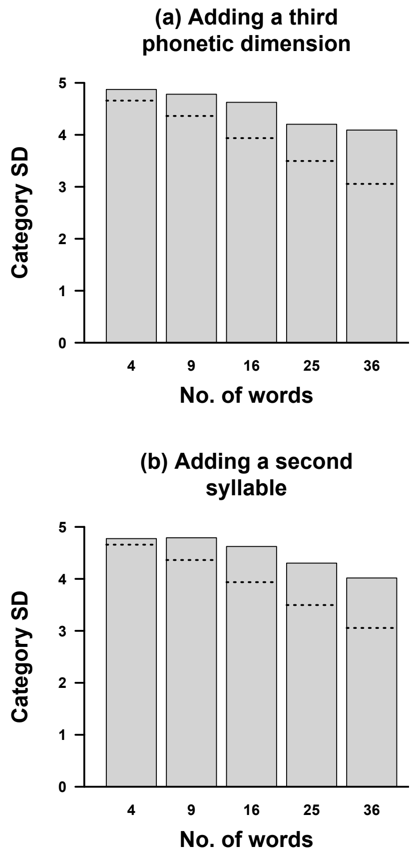


4.2. Increasing redundancy

It is well known that phonological categories are often distinguished by many different phonetic cues [9, 14, 28, 40], which has been argued to increase the robustness of speech communication [39]. What if we add another dimension, akin to having another phonetic cue?

We introduced an additional independent dimension. For ease of interpretation, we can imagine this to be an additional vowel contrast along the front-back dimension allowing an expanded vowel space of /i ~ ɯ ~ æ ~ ɑ/. The information provided by the additional phonetic cue allows for the maintenance of variation in the initial dimension of each lexical item even as more lexical items are added. In Fig 2a, each bar represents the SD with an added third dimension of phonetic contrast, while dashed lines indicate the SD of corresponding simulations without this third dimension. Thus, adding a third dimension increases standard deviations. This is because with this additional phonetic dimension, the role of each phonetic cue in reducing lexical confusability is reduced, which relaxes constraints on spread.

Figure 2: The relationship between category standard deviation and lexical density for systems with added redundancy through (a) a third phonetic dimension or (b) a second CV syllable. Dashed lines indicate SDs of simulation runs without such redundancy.



4.3. Making words longer

In this set of simulations, we limit the phonetic space to the two dimensions used in section 4.1, but double the length of each word by copying the values of the first two dimensions to a second "syllable" when initializing the simulation. After the start of the simulation, the numeric values of this second syllable are independent of those in the first syllable, except insofar as they belong to the same dimension and so are subject to the same phonetic similarity bias. Hence, the second syllable potentially adds just as much information about the output identity as the first. By increasing the amount of phonetic material transmitted (in terms of word length), but keeping the number of lexemes the same, we by definition increase redundancy. Again, similar to the case of adding a third dimension, the constraint on category standard deviation is relaxed (see Figure 2b) and SDs are higher.

5. DISCUSSION

The present simulations illustrate that lexical density directly affects variation within this model architecture, as expected. Exemplar clouds become more constrained when adding more possible lexical items. Within this model, the mechanism by which this happens is an anti-ambiguity bias, the same bias that has been proposed to explain patterns of phoneme merger [35, 36].

This anti-ambiguity bias directly relates to a listener's uncertainty about the incoming input. This is clearly demonstrated by adding redundancy to the system, which increases global phonetic distance between signals, rendering them less confusable. This is a concrete example of how redundancy serves to counteract noise [7].

Crucially, once redundancy has been added to the system —either via additional phonetic cues or via longer words— exemplar clouds are less constrained in their cryptic variation, and within-category variation is allowed to accumulate. This within-category variation is crucial for future change, as all evolution needs variation as "fodder". Hence, increasing redundancy increases variation and hence, assures the future evolvability of the system. Wedel [32, section 3.3] illustrates how variation provides a pathway for sound change in this architecture.

A comment should be made about the term "redundancy". From an information-theoretical perspective, adding a new syllable or adding an additional phonetic dimension are qualitatively similar changes; they both expand the phonetic channel capacity through which lexical contrasts can be distinguished, beyond what is strictly speaking necessary to distinguish lexemes. In the disyllabic case, redundancy is added in a sequential fashion. In the phonetic dimensionality case, redundancy is added in a simultaneous fashion. In the biological literature, each of these types of redundancy is technically called "degeneracy" (for review, see [23]), which refers to redundancy in which *different* structural components realize similar system functions [4, 22, 23, 37, 38]. In the linguistic case, this corresponds to different syllables and different phonetic cues signalling the same contrast. Even if the same syllable is repeated, this does strictly speaking not fall under the purview of redundancy, because the syllable conveys linguistic information at a different time point.

The present results are conceptually important because they show how evolution at one level (the lexicon) affects evolution at another level (the phoneme system). This deviates from standard exemplar models in the domain of speech, e.g. Goldinger [5], who models words as holistic acoustic traces—there are no separate sublexical and lexical levels in his model. However, a two-layer exemplar architecture is necessary in the present case to model lexicon/speech interactions.

This two-level architecture is furthermore illustrative because there are direct parallels to biological evolution, where evolution acts on phenotypes, and therefore selection only indirectly affects the frequency of genotypes within a population. Similarly, in the linguistic case modelled in the present paper, selection acts indirectly on phoneme inventories, via coupling relations from the lexical to the sublexical levels. This general picture is moreover in line with the idea that the "success" of phonetic categories is largely measured with respect to what they do at the communicative level [6, 35, 36]. The communicative "currency" in this model, so to say, is the word, not the phoneme. And this currency is ultimately the measure of success for different phonetic exemplars.

7. REFERENCES

[1] Blevins, J. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.

[2] Blevins, J., Wedel, A. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26, 143–183.

[3] de Boer, B. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28, 441–465.

[4] Edelman, G. M., Joseph A. G. 2001. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences* 98, 13763–13768.

[5] Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological review* 105, 251–279.

[6] Graff, P. 2012. Communicative efficiency in the lexicon. Doctoral dissertation, Massachusetts Institute of Technology.

[7] Hailman, J. P. 2008. *Coding and redundancy: Man-made and animal-evolved signals*. Cambridge: Harvard University Press.

[8] Harrington, J., Palethorpe, S., Watson. C. I. 2000. Does the Queen speak the Queen's English? *Nature* 408, 927–928.

[9] Hawkins, S. 2010. Phonological features, auditory objects, and illusions. *Journal of Phonetics* 38, 60–89.

[10] Hayden, E J., Ferrada, E., Wagner, A. 2011. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* 474, 92–95.

[11] Hintzman, D. L. 1986. 'Schema abstraction' in a multiple-trace memory model. *Psychological Review* 93, 328–338.

[12] Hockett, C. 1960. The origin of speech. *Scientific American* 203, 88–96

[13] Hockett, C. 1967. The quantification of functional load. *Word* 23, 320–339.

[14] Kingston, J., Diehl, R. D. 1994. Phonetic knowledge. *Language* 70, 419–454.

[15] Kruschke, J. K. 1992. ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review* 99, 22–44.

[16] Labov, W. 1994. *Principles of linguistic change: Internal factors*. Oxford: Wiley-Blackwell.

[17] Labov, W., Ash, S., Boberg, C. 2006. *The Atlas of North American English: Phonetics, phonology, and sound change*. Berlin: Mouton de Gruyter.

[18] Ladd, R. 2013. What is duality of patterning, anyway? *Language and Cognition* 4, 261–273.

[19] Lavoie, L. M. 2002. Subphonemic and suballophonic consonant variation: The role of the phoneme inventory. *ZAS Papers in Linguistics* 28, 39–54.

[20] Lindblom, B., Diehl, R. K., Park, S.-H., Salvi, G. 2011. Sound systems are shaped by their users: The recombination of phonetic substance. In: Nick, G. C., Ridouane, R. (eds.), *Where do phonological features come from?: Cognitive, physical and developmental bases of distinctive speech categories*. Amsterdam: John Benjamins, 67–98.

[21] Martinet, A. 1952. Function, structure, and sound change. *Word* 8, 1–32.

[22] Mason, P. 2010. Degeneracy at multiple levels of complexity. *Biological Theory* 5, 277–288.

[23] Mason, P, Domínguez D., J. F., Winter, B., Grignolio, A. 2015. Hidden in plain view: Degeneracy in complex systems. *BioSystems* 128, 1–8.

[24] Nosofsky, R. M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115, 39–57.

[25] Ohala, J. 1989. Sound change is drawn from a pool of synchronic variation. In: Breivik, Leiv E., Jahr, Ernst H. (eds.), *Language Change: Contributions to the study of its causes*. Berlin: Mouton de Gruyter.

[26] Pierrehumbert, J. B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In: Bybee, J. L., Hopper, P. J. (eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, 137–158.

[27] Sancier, M. L., Fowler, C. A. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25, 421–436.

[28] Stevens, K. N., Keyser, S. J. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38, 10–19.

[29] Trubetzkoy, N. 1939. *Grundzüge der Phonologie*. Prague, Czech Republic: Travaux du Cercle Linguistique de Prague.

[30] Wagner, A. 2005. *Robustness and evolvability in living systems*. Princeton: Princeton University Press.

[31] Walsh, M., Möbius, B., Wade, T., Schütze, H. 2010. Multilevel exemplar theory. *Cognitive Science* 34, 537–582.

[32] Wedel, A. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23, 247–274.

[33] Wedel, A. 2007. Feedback and regularity in the lexicon. *Phonology* 24, 147–185.

[34] Wedel, A. 2012. Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition* 4, 319–355.

[35] Wedel, A, Jackson, S., Kaplan, A. 2013. Functional load and the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech* 56, 395–417.

[36] Wedel, A., Kaplan, A., Jackson, J. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128, 179–186.

[37] Whitacre, J. M. 2010. Degeneracy: a link between evolvability, robustness and complexity in biological systems. *Theoretical Biology and Medical Modelling* 7, 1–17.

[38] Whitacre, J., Bender, A. 2010. Degeneracy: a design principle for achieving robustness and evolvability. *Journal of Theoretical Biology* 263, 143–153.

[39] Winter, B. 2014. Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. *BioEssays* 36, 960–967.

[40] Wright, R. 2004. A review of perceptual cues and robustness. In: Steriade, D., Kirchner, R., Hayes, B. (eds.), *Phonetically based phonology*. Cambridge, UK: Cambridge University Press, 34–57.

[41] Zuidema, W., de Boer, Bart. 2009. The evolution of combinatorial phonology. *Journal of Phonetics* 37, 125–144.