Chain Shifts and Transphonologizations are Driven by Homophony Avoidance

Stefon Flego[1] and Andrew Wedel[2]

Human languages balance pressures for effort reduction and communicative accuracy. This tension plays out in the structure of phoneme inventories and how they change over time. For example, merger of two phoneme categories (e.g., /p~b/) results in a simpler inventory, but also eliminates a phoneme contrast which can distinguish words (e.g. *pat~bat*). Previous work has shown that merger of a phoneme contrast is more likely when that contrast distinguishes few words. Within a diverse set of languages we extend this finding to two phoneme inventory changes which preserve word contrast. In chain shifts, two adjacent phonemes move in concert within the phonetic space. In transphonologizations, a primary phonetic cue to a phoneme contrast merges while another cue expands. A common feature of these two change types is that although the phonetics-phoneme category mapping changes, no new homophones are created. Here we show that the greater the contribution a phoneme contrast makes to word differentiation, the less likely that contrast is to merge, and conversely the more likely it is to undergo a chain shift/transphonologization. Traditional phonological theory assumes phonological systems are causally independent of actual words in the language. This work shows instead that change-type in inventories is strongly influenced by the particularities of the lexicon. These findings support a model in which the structure of phoneme inventories is shaped by usage-driven pressures to optimize effort reduction versus function in the transmission of meaning, paralleling broader cognitive work in efficient coding and in learning versus usage of category systems.

Keywords: sound change, merger, chain shift, transphonologization, functional load, category stability, hyperarticulation, phonological features

[1] *University of New Mexico (USA).* stefonflego@unm.edu
Department of Linguistics, MSC03 2130, 1 University of New Mexico, Albuquerque, NM 87131

[2] *University of Arizona (USA).* Corresponding author: wedel@arizona.edu, +1 415-217-9483
Department of Linguistics, The University of Arizona, P.O. Box 210028, Tucson, AZ 85721-0028

The order of authors is alphabetical.

1. Introduction

The notion that features of human language structure are influenced by competing pressures for simplicity/efficiency versus communicative function has a long history (Zipf 1949, Lindblom 1990, Cancho & Solé, 2003, Piantadosi, Tily, & Gibson, 2011, Gibson 2019, and many more). These shared competing pressures have been argued to underlie many typological universals of language (e.g., Liljencrantz & Lindblom 1972, Hawkins et al. 2009; Hahn, Jurafsky & Futrell 2020). For example crosslinguistically, word-lengths (Zipf 1949, Piantadosi, Tily & Gibson 2011) and phonological grammars (Wedel et al. 2019) appear to evolve under pressure to minimize effort for speakers while preserving accurate transmission of meaning categories for hearers. Here we provide evidence that phoneme inventories – the set of abstract sound categories in a language which function to distinguish words – similarly evolve to preserve functional complexity while simplifying through the elimination of category contrasts which are relatively redundant in the transmission of meaning.

All spoken languages build words from a finite inventory of individually meaningless sounds organized into an inventory of contrastive *phonemes*[1]. A phoneme contrast is a pair of sound-categories defined by their ability to distinguish so-called *minimal pairs*, which are pairs of words that are identical in form except for the contrasting sounds in question. For example, we know that the sound categories /p/ and /b/ in English are a phoneme contrast because they can distinguish minimal pairs like *pat* and *bat*. The inventories of phonemes in languages are not static: they can lose or gain phoneme contrasts over the course of language change. For example, the vowel inventory in English spoken in the eastern part of America includes two perceptually similar low-back vowels, /ɔ/ as in the word *caught*, and /ɑ/ as in the word *cot*. However in Canadian English and western dialects of American English, the distinction between these two vowels has collapsed, merging them to /ɑ/. For speakers of these dialects *caught* and *cot* are now homophonous, and as such their language has lost one of its phonemic tools to make words different from one another. On the other hand, perceptually similar English vowels like /ɪ/ as in *bit*, and /ɛ/ as in *bet* virtually never merge in dialects of English.

What prevents all perceptually similar contrasts from merging in this way? Why do /ɔ/ and /ɑ/ merge in English dialects, but not /ɪ/ and /ɛ/? Earlier work has theorized that the probability of loss of contrast through phoneme merger is inversely correlated with that contrast's *functional load*, i.e, the amount of work a phoneme contrast contributes to carrying information in a language (Gilliéron 1918; Trubetzkoy 1939; Martinet 1952; Hockett 1967; Surendran & Niyogi 2003). Functional load has been operationalized in many ways (Hockett, 1967; King, 1967; Surendran & Niyogi, 2006, Silverman, 2010, Kaplan, 2011), for example as the lexically local property of number of minimal pairs distinguished by a contrast, versus global properties such as the contribution of a contrast to overall entropy of the lexicon. Comparing these approaches, Wedel, Kaplan & Jackson (2013), showed that the simple count of minimal pairs did a better job predicting merger than either (i) global conceptions of the pressure for contrast exemplified by change in system entropy (Hockett 1967, Surendran & Niyogi 2003, 2006), or (ii) more diffuse conceptions of contrastive pressure exemplified by lexical neighborhood density (see also Nelson & Wedel 2017). Further, Wedel, Jackson & Kaplan (2013) showed that local morphosyntactic context appeared to modulate the minimal pair effect, because only within-syntactic category (e.g. noun-noun, verb-verb) minimal pairs statistically accounted for the effect. Word frequency was shown not to strongly modulate the minimal pair effect (Wedel, Jackson & Kaplan 2013).

Here we present new evidence consistent with the hypothesis that sound change is biased toward maintenance of phoneme contrasts which contribute more to distinguishing individual words (Martinet 1952, Wedel 2006, Blevins & Wedel, 2009; Kaplan 2011, Wedel, 2012; Flego 2022), and in particular to distinguishing words that appear in the same local syntactic context (Wedel, Jackson & Kaplan 2013). Continuing our example from American English vowels, the merged vowel contrast /ɔ~a/ distinguishes very few within-syntactic category minimal pairs in the lexicon of English, while the non-merging contrast /ɪ~ɛ/ distinguishes hundreds of such minimal pairs. These patterns suggest that maintenance of lexical expressiveness is a driving factor in phoneme inventory stability and change.

In previous studies, homophony avoidance was associated with lack of change, i.e. prevention of phoneme merger (Figure 1a; Kaplan 2011; Wedel, Kaplan & Jackson 2013; Wedel, Jackson & Kaplan 2013). Here, we show that homophony avoidance appears to also drive two superficially distinct, active sound changes, each of which gives the appearance of influencing the system of contrasts to actively *avoid* merger. *Chain shifts* occur when a set of phonemes move in concert within phonetic space, so that although the phonetic properties of each phoneme changes, the contrast between them is maintained (reviewed in Gordon 2002). For example, the front vowels in New Zealand English have undergone a chain shift upwards, such that the original vowel /æ/ in 'pat' has raised to /ɛ/, while the original /ɛ/ in 'pet' has raised to near /i/ (Figure 1a; Maclagan & Hay 2007; Hay, 2008).

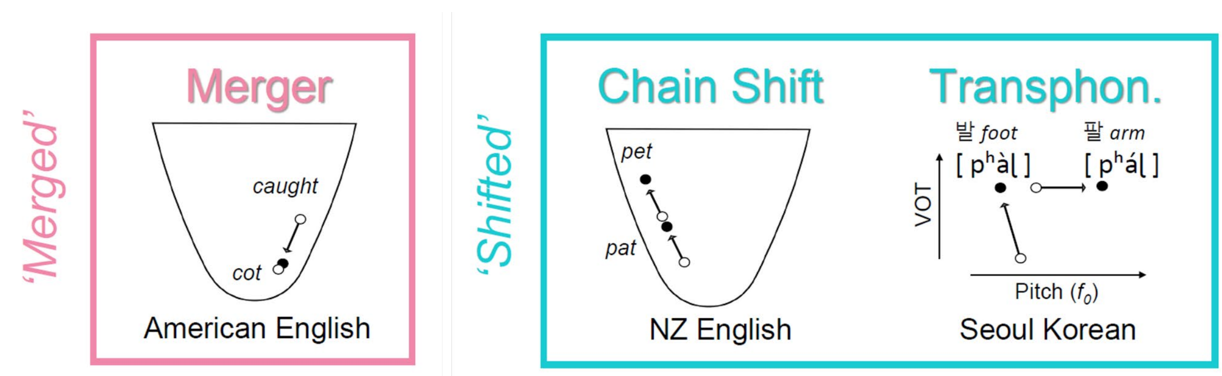

Figure 1. (a) Schematic of phoneme mergers. (b) Schematics of chain shifts and transphonologizations

Transphonologizations, on the other hand, occur when the primary cue distinguishing a contrast merges, while a minor cue expands to become the primary cue. For example, aspirated and lenis stops in Korean are historically distinguished by a voice-onset-time difference (VOT), with a minor distinction in pitch on the following vowel. In modern Seoul Korean, this voice onset time difference is collapsing, while the pitch difference has expanded to become the primary cue (Figure 1b; Silva, 2006). These two superficially distinct classes of sound change have in common that lexical contrast is maintained throughout the change: in a chain shift, two phonetically adjacent phonemes move in concert within the phonetic space, while in a transphonologization, one phonetic cue-distinction to a phoneme contrast merges, while at the same time another cue-distinction to the same contrast expands. Here we show that the greater the contribution a pair of phonemes makes to word differentiation in the lexicon (i.e., the higher its functional load as measured by within-category minimal pairs), the less likely those phonemes are to merge over the course of language change, and the *more likely* they are to participate in contrast-preserving changes such as chain shifts and transphonologizations.

Changes that involve the collapse or maintenance of contrast often involve parallel changes across multiple phonemes which are similar in some dimension, that is, phonemes which share *phonological features*. For example, the sounds /b/, /m/ and /z/ all share a *voicing* feature, because at some point in the sound the vocal cords vibrate. In the Korean example given above, aspirated and lenis stops undergo the same transphonologization from a laryngeal voice-onset-time (VOT) distinction to a pitch distinction, realized at three places of articulation in parallel, labial (p^h~p), alveolar (t^h~t) and velar (k^h~k). Many decades of work in phonology (e.g., Trubetzkoy 1939; Jakobsen et al. 1951; Jakobsen 1968; Halle & Jones 1971, Mielke 2008) suggest that this group of changes is cognitively represented as one change distributed over a class of similar phonemes, rather than as multiple independent, superficially similar changes. Here we show that treating parallel changes such as the Korean example as one change for the purpose of counting minimal pairs, rather than as multiple independent changes per phoneme pair, gives a significantly better fit to the data. This provides evidence from a novel process– minimal pair-driven homophony avoidance– that phonological pattern development and change proceeds over featural classes of sounds rather than individual sounds.

In the discussion below, we draw connections between these findings and broader themes in category evolution and stability, work in artificial language learning, and modeling of language change via acquisition by children versus usage by adults. Further, previous work on phoneme inventories has assumed that the causal locus for structure lies in the need for perceptual contrast between sound categories themselves, only indirectly referencing the need for contrast between words (Liljencrantz & Lindblom 1972, Padgett 2003, Flemming 2004, but see Lindblom 1990). Instead, we will argue that our results place the causal locus for phoneme inventory structure in the tension between effort reduction and accurate transmission of higher categories of meaning, such as words (Hall et al. 2018). In this sense, the results we show here suggest that phoneme inventory structure is parasitic on lexicon structure. From this perspective, a phoneme inventory can be thought of as a system that carves up a semi-continuous perceptual space to map efficiently, yet expressively to a categorical informational space used to transmit meaning.

1. Although this work focuses on sounds, we anticipate that the phenomena studied here apply as well to other modalities such as sign languages.

2. Methods

2.1 Database

Mergers, chain shifts and transphonologizations are rare-enough events in a language's history that only a small number of historically recent changes of these types are usually attested in any particular language. To address this issue and to allow us assess the generalizability of our results, we pooled data from 12 languages- Cantonese, Danish, Dutch, English (RP and American), French, German, Icelandic, Korean, Slovak, Spanish, Turkish, and Vietnamese. These languages were chosen because one or more of their dialects have recently undergone, or are currently undergoing a change that is well documented, and can be characterized as a merger, chain shift, or transphonologization. In addition, large phonemically transcribed, lemmatized word lists were available for all of these languages, including information about lexical frequency and syntactic category. Grammatical and function words were excluded from the dataset. A summary of our data sources and the sound changes are presented in Appendix A.

To be able to relate our results more directly with previous studies (Kaplan 2011; Wedel, Kaplan & Jackson 2013; Wedel, Jackson & Kaplan 2013), where minimal pair counts associated with mergers were compared with those associated with non-changing contrasts, we created a comparison set of phonetically similar contrasts in each language that have *not* participated in a change. For example, we compare the functional load of the /ɔ/ ~ /ɑ/ contrast that has merged in certain varieties of American English to the functional load of phonetically similar but non-changing contrasts, such as /ɛ/ ~ /æ/, /ɑ/ ~ /æ/, etc. All phoneme pairs in this baseline dataset contrasted in just one phonological feature, e.g. *place*, *voice*, *nasality* (see Wedel A., Kaplan & Jackson, 2013; Wedel, A., Jackson & Kaplan 2013) – in other words, they are contrasts which could have plausibly merged or shifted/transphonologized in the language, but have not. Crucially, these non-changing contrasts were restricted to the same phonological context as the changing contrast they were to be compared to. For example, in our Korean dataset we identified a number of non-changing coda contrasts to compare to observed coda neutralizations, and we *also* identified a number of non-changing onset contrasts to compare to observed onset transphonologizations. The important comparison here then is 'Merged' vs. 'No Change' and 'Shifted' vs. 'No Change (we will use 'Shifted" as an abbreviation for the set of Shifted and Transphonologized contrasts).' We treat each of these contextually conditioned systems as a categorical variable and include it as a random effect in some of our statistical models.

2.2 Predictor variables

Our research question is if the functional load of a phonemic contrast predicts whether or not that contrast will be maintained when undergoing sound change, i.e. whether it will merge or shift. We operationalize our predictor variable, functional load, as the number of minimal pairs associated with the contrast, and we distinguish between within-syntactic category (e.g. noun-noun, verb-verb, adjective-adjective) versus cross-syntactic category minimal pairs (e.g., noun-verb, etc). For some words in our databases, there exist forms in multiple parts of speech, such as English *table$_{noun}$*, and *table$_{verb}$* With respect to these forms, counts were generated in two ways: (i) Each distinct part of speech for a minimal pair was included and all possible combinations of POS were counted. For example, for the English minimal pair tack$_{noun}$, tack$_{verb}$ / tap$_{noun}$, tap$_{verb}$ we counted 4 minimal pairs, one noun-noun, one verb-verb, and two noun-verb pairs. (ii) Alternatively, we just used the POS of the form with the highest lexical frequency in our corpus. For example, under this analysis, only *table$_{noun}$* would be used in counting minimal pairs as it has higher lexical frequency than *table$_{verb}$*. The results for both methods of counting minimal pairs are essentially the same. The version we show in the text here is that with all minimal pair relationships; the analysis using the highest-frequency method of counting minimal pairs is presented in Appendix B.

Minimal pairs were also counted assuming (i) each phoneme contrast is an independently changing unit, versus (ii) phoneme contrasts are grouped into featural contrasts, in which case the minimal pair count is for the group of phoneme pairs as a whole. For example, in some dialects of English in the United Kingdom, the voiceless interdental consonant /θ/ has merged with /f/, such that *think* is now pronounced as *fink*. At the same time, the voiced interdental consonant /ð/ has merged with /v/, so that *brother* is now pronounced *bruvver*. The consonants /θ/ and /ð/, and in parallel /f/ and /v/, are identical except for their voicing feature, that is, whether the vocal cords are vibrating. As a consequence, these two mergers share the same shift in articulation from use of the tongue-tip to the lower lip, with voicing remaining unchanged in each pair. Should the merger of /θ/ with /f/ and /ð/ with /v/ be therefore considered one phonological change over two pairs of sounds that are the same but for a voicing feature, in which we pool their minimal pair counts? Or two independent changes that happen to involve the same articulatory change over two pairs of sounds, in which we count their minimal pairs separately?

As a more detailed example, the table below shows minimal pair counts for two sets of parallel vocalic changes, one in Icelandic and one in Dutch. In the Icelandic vowel height merger, two pairs of front vowels merge in parallel, where the pairs differ only in lip rounding. In the Dutch example, we similarly see three parallel vowel chain shifts across different levels of tongue backness and rounding. Again, should we consider each of these changes to be independent? Or are they part of the same featural change, and therefore we should pool their counts? Crossing this distinction with the choice to count within- or cross-syntactic category minimal pairs, we have four ways of constructing the dataset. In Table 1, we illustrate these four choices with the Icelandic and Dutch examples, showing the number of minimal pairs in each case.

Table 1. Number of minimal pairs associated with sample vocalic contrasts

			Cross-Syntactic Cat.		Within-Syntactic Cat.	
Language	Change	Contrast	Phoneme	Feature	Phoneme	Freature
Icelandic	Merged	ɛː ɪː	**7**	**7**	**12**	**13**
		œː ʏː	**0**		**1**	
Dutch	Shifted	eː ɛi	**36**	**52**	**93**	**137**
		øː œy	**8**		**24**	
		oː ɔu	**8**		**20**	

2.3 Statistical Models

To compare within- vs. cross-syntactic minimal pair counts, as well as change at the phoneme vs. feature levels, we ran four Bayesian binomial models using the **brms** package in R (Bürkner, 2017). For each model, Change Type ('Merged' vs. 'Shifted') was the response variable, and Minimal Pair Count was the predictor variable. Because our count data is heavily right-skewed, we first transformed all counts using Inverse Hyperbolic Sine (HIS) transformation, as represented in the equation below.[3]

$$IHS(MPC) = \text{arsinh}(\theta * MPC)/\theta$$

For the constant θ, we use a value of 0.12, as this produces a transformation that is very similar to a natural logarithmic transformation. We use IHS transformation instead of logarithmic transformation because IHS can transform 0 counts, which logarithmic transformations cannot (see Burbige et al., 1988; Bellemare & Wichman, 2020). This is important for our data set, as some

[3] Readers are encouraged to consult the supplementary code, which includes visualizations illustrating the right skew of our count data, and compares IHS transformation with square-root transformation and several different logarithmic transformations.

contrasts are associated with no minimal pairs at all (i.e. we have meaningful zero-valued observations).

For two of the models, Minimal Pair Count corresponded to the number of within-syntactic category minimal pairs, and for the other two models, the number of cross-syntactic category minimal pairs. These were crossed with the other comparison, where two models used a data set in which each observation corresponded to a phoneme contrast (e.g., separate observations for Icelandic mergers /ɪː/ ~ /ɛː/ and /ʏː/ ~ /œː/), and the other two models used a data set in which each observation corresponded to a featural contrast (e.g., Icelandic /ɪː, ʏː/ ~ /ɛː, œː/ as one observation).

As a final step in our analysis, we compare Minimal Pair Counts for 'Merged' and 'Shifted' contrasts to those of the baseline set of phonetically similar but non-changing contrasts. We do this to make our results more directly comparable with previous research on the relationship between function load and merger (e.g. Kaplan, 2011; Wedel et al., 2013), where merged contrasts were compared to stable, non-merging contrasts. For this comparison, we use the data set (freature vs. phoneme) and predictor variable (within-category vs. cross-category minimal pair count) corresponding to the model which provided the best fit in the preceding analysis. We construct a Bayesian linear model where IHS-transformed MPC is the response variable, and Change Type (now 'Merged' vs. 'No Change' vs. 'Shifted') is a categorical predictor variable. We also include a random intercept for each system (see 2.1 above). All of our Bayesian models were run with default priors on 4 chains of 8,000 iterations each. Full model specifications are given in the supplementary data & code.

3. Results

Figure 2 below shows model predictions for the four binomial models we fit, where the response variable was change type (‘Merged’ vs. ‘Shifted’). As a reminder, the number of **within-syntactic category** minimal pairs was the explanatory variable in two of the models, while the number of **cross-syntactic category** minimal pairs was the explanatory variable in the other two. The other comparison was whether each observation in the data set corresponded to a contrast at the **phoneme** level (e.g. for the Icelandic example above, /ɪː/ vs /ɛː/ and /ʏː/ vs /œː/), or a contrast at the **featural** level (e.g. /ɪː, ʏː/ vs. /ɛː, œː/). The distribution of minimal pair counts for all four conditions are shown in the density plots above each model’s regression line. As a reminder, all figures show raw count numbers on the horizontal axis, but the axis spacing has been warped to reflect the Inverse Hyperbolic Sine (IHS) transformation we use in our statistical models.

To be able to directly compare all four models, we generated an estimate of each model’s accuracy in predicting change type. We used the *predict()* function in R to get a probability of ‘Shifted’ over ‘Merged’ for each phoneme contrast. For the models trained on the FEATURAL data set, this means that all phoneme contrasts within the same featural contrast got the same probability for ‘Shifted’ over ‘Merged,’ while for those trained on the PHONEME data set, each segmental contrast got its own probability for change type. We then used the *sample()* function weighted on these probabilities to generate a binary prediction of ‘Shifted’ vs. ‘Merged’ for each phoneme contrast. We then compared each contrast's actual change type with its predicted change type to arrive at a measure of model accuracy. This process was repeated 1,000 times to build a distribution of accuracy estimates for each model (code is included in the supplementary materials). These are shown in the smaller panel at the bottom right of each model’s main panel.

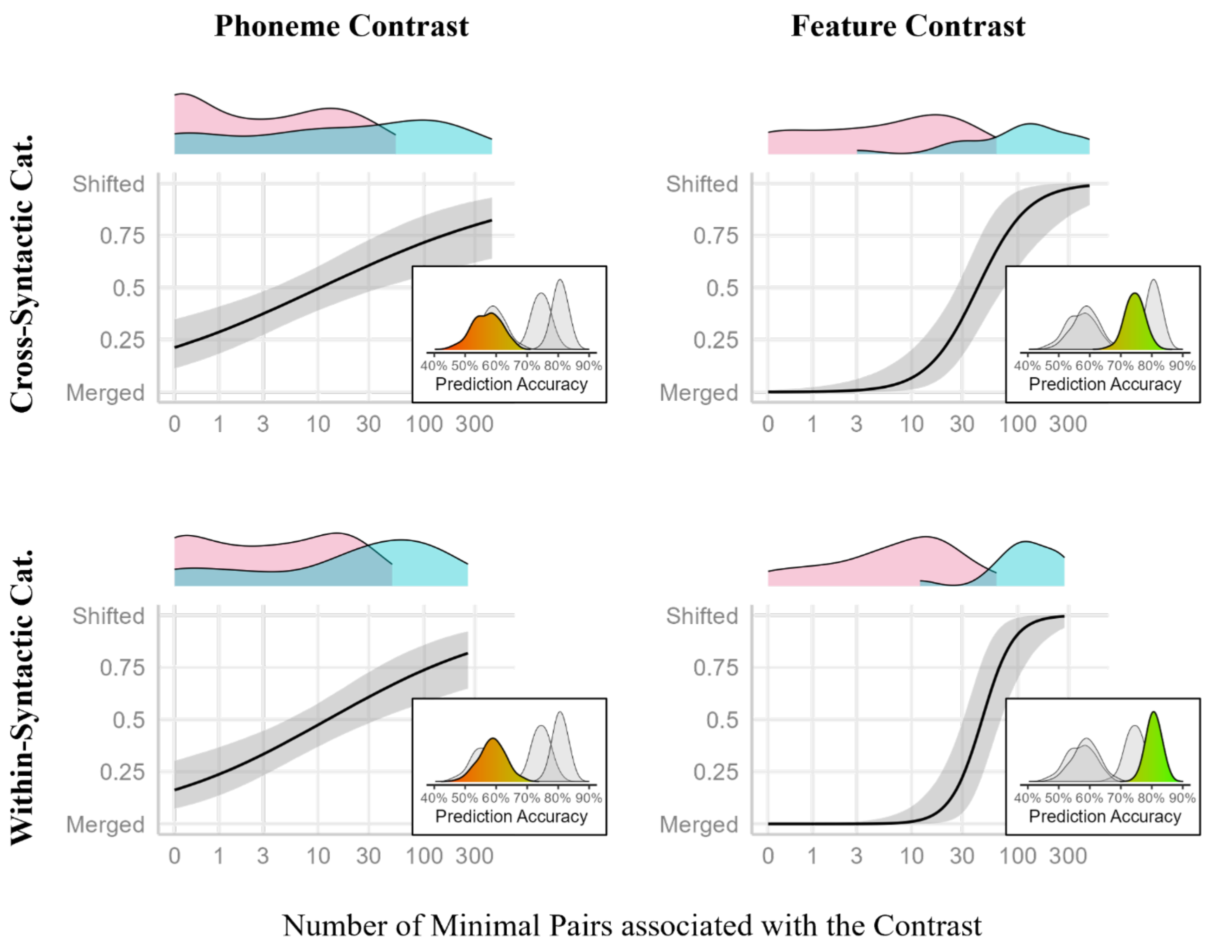

Figure 2. Model predictions and accuracy scores for four binomial models

All models show a strong effect of minimal pair count on change type in the same direction, and all have prediction accuracies significantly above chance level. This shows us that the signal for the general relationship between functional load and homophony avoidance is clear no matter what kind of MPs we count, or how we count them. In terms of prediction accuracy and model fit, WITHIN-CATEGORY models are significantly better than CROSS-CATEGORY models, consistent with the finding of Wedel, Jackson & Kaplan 2013. Further, FEATURAL models are significantly better than PHONEME models, consistent with previous experimental and theoretical work showing that phonological changes occur over featural groups rather than individually over related phonemes (see discussion).

Returning to the effects of cross-category and within-category minimal pairs, the predictive value of cross-category minimal pair counts could arise either: (i) through a direct effect of cross-category minimal pairs via a causal linkage between those minimal pairs and sound change (see discussion), or alternatively (ii), as within- and cross-category counts are correlated, it could show predictive value just by virtue of this correlation. To assess these two possibilities, Chi-square log-likelihood tests were carried out comparing a full FEATURAL model including both within- and cross category counts as predictive factors, to nested models including either of the two factors separately. The full model with both types of counts and the subsidiary model with only within-category counts were statistically equivalent ($p = .42$), while the model with just cross-category counts fit the data significantly less well than the full model with both types of counts ($p < .001$). This suggests that cross-category minimal pair counts only significantly predict merger versus shift/transphonologization on their own by virtue of their correlation with within-category counts, rather than through any independent causal effect.

Finally, we compare the minimal pair counts for 'Merged' and 'Shifted' contrasts to the comparison set of phonetically similar but non-changing contrasts ('No Change'). Including this comparison is important because one could argue that 'Shifted' contrasts are characterized by significantly more minimal pairs than 'Merged' contrasts simply because 'Shifted' contrasts behave similarly to non-changing ones with respect to functional load. The figure below shows the probability of each change type (including 'No Change') as a function of within-syntactic category minimal pair count. The data set used is the within-category, featural set.

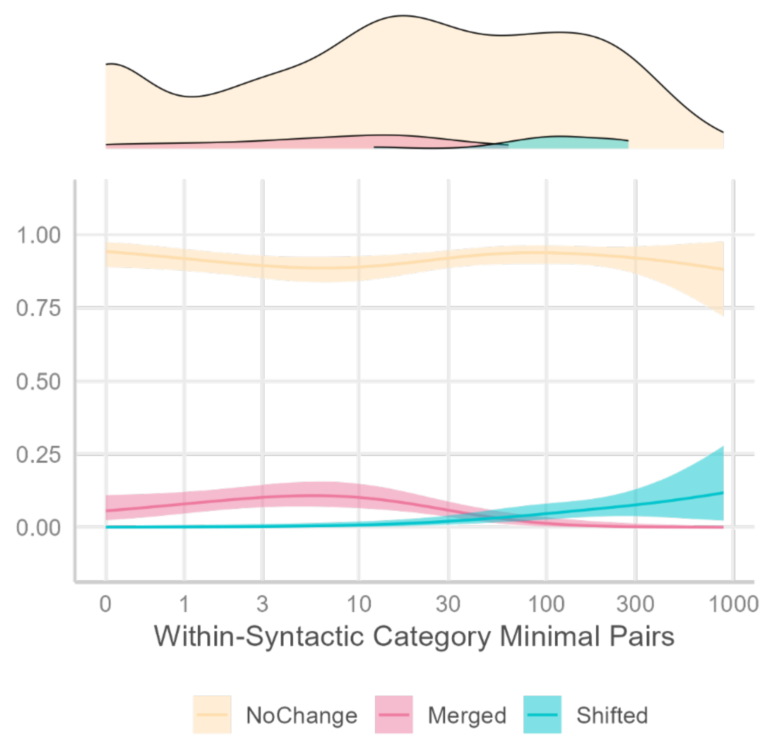

Figure 3. Model predictions including non-changing contrasts

It can be seen that relative to the distribution of minimal pair counts for 'Merged' and 'Shifted' contrasts, minimal pairs counts associated with non-changing contrasts span the entire range of observed counts. The figure below shows the results of a linear regression model in which the response variable is within-syntactic category minimal pair count (at the featural level), and the explanatory variable is change type (now ‘Merged’ vs. ‘No Change’ vs. ‘Shifted’, with ‘No Change’ as the reference level).

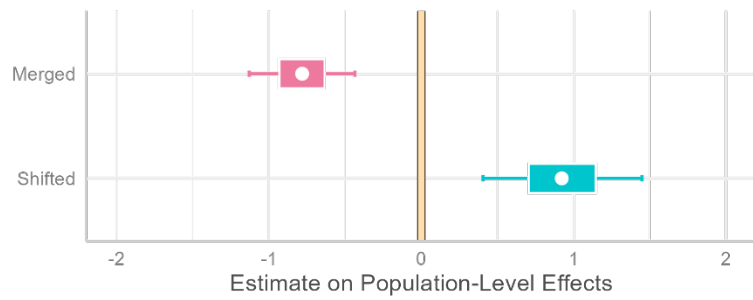

Figure 4. Effect of ‘Merged’ and ‘Shifted’ vs. ‘No Change’ on Minimal Pair Count

We show that ‘Merged’ contrasts are characterized by significantly fewer minimal pairs than non-changing contrasts, consistent with previous analyses (e.g. Wedel, Kaplan & Jackson 2013, Wedel, Jackson & Kaplan 2013, Kaplan 2011). We also show that ‘Shifted’ contrasts in our data set are characterized by significantly *more* minimal pairs than the comparison set of non-changing contrasts. This shows that it is not simply that ‘Shifted’ contrasts have similar functional load to non-changing contrasts, but are associated with *especially high* functional load.

4. Discussion

4.1 The role of functional load in sound change

Over a century of work has suggested that preservation of contrasts within a phoneme inventory is predicted by high functional load of those contrasts (e.g., Gilliéron 1918; Trubetzkoy 1939; Martinet 1952; Hockett 1967; Surendran & Niyogi 2003, Wedel, Kaplan & Jackson 2013). This is the first work to show that in addition to predicting avoidance of merger, high functional load predicts two types of changes that actively avoid loss of phonemic contrast: chain shifts and transphonologizations. These findings suggest that stabilities or changes in a language's sound system which are typologically or perceptually unexpected may find an explanation in the particularities of its lexicon. For example, the English front vowel system is very crowded, with neighboring vowels such as /ɛ/ and /ɪ/ very close in phonetic space. This should be a situation ripe for merger, yet these vowels shift instead, preserving their relatively weak perceptual contrast. This striking lack of front-vowel mergers in English dialects is explained by the very large number of minimal pairs distinguished by the front vowel system of English. On the other hand, the intervocalic /t~d/ merger to 'flap' in North American English, which creates homophony between words like *latter* and *ladder*, is unexpected because the perceptual distinction between /t/ and /d/ is particularly clear between vowels. As a consequence, this loss of contrast is surprising compared to, for example, word-final context where the distinction between /t/ and /d/ is often less clear. However, it happens that there are very few minimal pairs distinguished by /t~d/ in intervocalic context in English that could inhibit merger. In contrast, in word-final position there are many /t~d/ minimal pairs, so that even though this is a less perceptually distinct position, /t/ and /d/ in this position do not merge in English dialects, and shift instead (Purnell et al. 2005).

Earlier work has shown that beyond just a measure of minimal pairs in general, the most predictive measure of functional load is the number of minimal lexical pairs sharing a syntactic category, e.g., noun-noun, verb-verb, as opposed to noun-verb (Wedel, Jackson & Kaplan, 2013). In this expanded dataset, we confirm that within-category minimal pairs give a significantly better fit to the data for both mergers and chain shifts/transphonologizations than a measure of functional load based on between-category minimal pairs. This suggests that local predictability is a contributing factor in Further, model comparison suggests that cross-category minimal pairs may have no direct causal influence on merger, shift or transphonologization probability, and instead show an effect within this dataset solely through their correlation with within-category minimal pair counts.

A considerable body of theoretical (e.g., Trubetzkoy 1939, Jakobsen 1951, Halle & Jones, 1971) as well as experimental (Nielsen 2011; Levi 2015) work has shown that groups of similar sounds behave as single entities for many phenomena. For example, the set of English voiceless stops /p, t, k/ pattern identically with regard to the environments in which they are aspirated or unaspirated, suggesting that this is causally one pattern operating over the set of voiceless stops in

English, rather than three causally independent patterns that just happen to be identical. The work described here is consistent with the organization of sounds into sets sharing these so-called phonological features, via a novel approach. Here we find that counting minimal lexical pairs over groups of sounds that undergo the same change in parallel, instead of as individual changes, provides a significantly better fit to the data. For example, in so-called *Polder Dutch* (Stroop, 1998), the set of vowel contrasts /ɛi œy ɔu/ vs. /eː øː oː/ has shifted to [ai ay au] vs. [ɛɪ œʏ ɔʊ] (Jacobi, 2009). When the functional load of each pair of vowels is considered individually, the contrast /ɔu/ vs. /oː/ is associated with 20 minimal pairs, /œy/ vs. /øː/ with 24 minimal pairs, and /ɛi/ vs. /eː/ with 93 minimal pairs. Only the contrast /ɛi/ vs. /eː/ is predicted to shift when the data is analyzed at the phoneme level, while the contrasts /ɔu/ vs. /oː/ and /œy/ vs. /øː/ are predicted to merge. However, when the data analyzed at the level of featural contrast, the model predicts that all three phoneme contrasts shift together. As a result, treating multiple phoneme contrasts as one unitary change to a featural contrast provides a change that is associated with many minimal pairs overall. Otherwise, when counted individually, we have no explanation why some phoneme contrasts participate in a transphonologization instead of merging, since they in fact would produce few or no new homophones if they were to merge. Corresponding pathways of change in multi-level category systems have been modeled by a large body of work in agent-based multi-level exemplar category evolution (Walsh et al. 2010), in which sound change is indirectly shaped by biases for accurate transmission of the words of which they are a part (e.g. Wedel 2006, 2012; Sóskuthy 2015; Winter & Wedel 2016; Todd et al. 2019; Flego 2022; Gubian et al. 2023).

We can make a number of other predictions about minimal pair counts and the development of changes in the system of phonemic contrasts. (i) Mergers of a pair of phonemes with few minimal pairs will not broaden to include featurally similar pairs if those pairs distinguish many minimal pairs. For example, the low-back merger of /ɔ~ɑ/ in English cannot drag along the featurally similar /æ~ɛ/ pair toward merger, because the latter distinguishes many minimal pairs. When we see a group of similar sound-pairs merging in parallel, we can conclude that each of these distinguishes few minimal pairs. The converse is not true however; we cannot explain why, for example, the shift of /æ~ɛ/ does not carry along /ɔ~ɑ/ as a featurally similar pair. (ii) We can predict in part when a contextually defined, as opposed to context-free merger will occur. If a phoneme pair distinguishes many minimal pairs, but a context-defined portion of this does not, the contextually defined merger may develop. For example, in the English *pin~pen* merger, /ɪ/ and/ɛ/ merge before /n, m/ at the end of the syllable, and as expected, there are relatively few minimal pairs defined by this pair in this context. However, we do not find the context-free merger of /ɪ/ and /ɛ/ because there are many minimal pairs defined by this pair over all contexts. (iii) We can predict that a phoneme pair distinguishing few minimal pairs will only participate in a shift or transphonologization if it is part of a feature set that together has a large number of minimal pairs. More generally, we can predict whether a group of featurally similar sounds can shift/transphonologize together - a group-defined change only happens when the group as a whole has enough minimal pairs between them.

4.2 Minimal pairs and hyperarticulation

What is it about minimal pairs that might drive these patterns? A growing body of empirical evidence shows that phonetic cue(s) distinguishing a minimal pair are hyperarticulated, increasing perceptual distance between a target word and a competitor minimal pair word, whether the competitor word is in the immediate context (Baese-Berk & Goldrick 2009; Buz, Tanenhaus, & Jaeger 2016; Peramunage et al. 2011; Schertz 2013; Seyfarth, Buz, & Jaeger 2016) or not (e.g., Baese-Berk & Goldrick 2009, Sano 2018, Wedel, Nelson & Sharp 2018). For example, words beginning in a voiceless stop that have a minimal pair with a voiced stop, e.g. *cape* (*gape*), have systematically longer VOT values which increases perceptual distance to the competitor, than similar words that do not, e.g. *cake* (**gake*) (e.g., Peramunage et al., 2011, Wedel, Nelson & Sharp, 2018, but see Fox et al. 2015). This also applies in the opposite direction: words beginning with a voiced stop like *bat* (*pat*) have systematically shorter VOT which increases perceptual distance to the competitor, than similar words like *badge* (**padge*) (Nelson & Wedel, 2017; Wedel, Nelson & Sharp, 2018). The spectral distance between vowels in American English is similarly expanded in minimal pairs. For example, in minimal word pairs like *sheep* [ʃip] and *ship* [ʃɪp] which are distinguished by the /i~ɪ/ contrast, these vowels are shifted away from one another in vowel space to make the contrast more salient. Conversely, the vowels in words like *beep* [bip] and *tip* [t^hɪp] which are not minimal pairs defined by the /i~ɪ/ contrast (**bip* [bɪp], **teep* [t^hip]) are less distinct from one another (Wedel, Nelson & Sharp, 2018).

How does minimal pair driven hyperarticulation come about? There are three, non-mutually exclusive types of models that account for this within a perception/production feedback loop (see below). (i) In listener-internal accounts, comprehension filters out phonetic realizations that are not accurately understood by not leaving any, or as strong a memory trace in the intended category (e.g., Guy 1996; Lindblom 1990; Ohala 1989). Within exemplar-based models (Johnson 1997), this account assumes that a specific pronunciation will only be stored (or stored with as strong an activation) by the listener in the intended category if it is recognized with sufficient confidence (Pierrehumbert 2001; Blevins & Wedel 2009; Wedel 2006, 2012). As a consequence, ambiguous productions will contribute less to the exemplar cloud for the intended lexical category, lowering the influence of such percepts on future production of that lexical category by the listener. (ii) In listener-oriented accounts, speakers adapt their productions to minimize misunderstanding based on an implicit model of listener expectations in context (e.g, Lindblom 1990; Stent et al. 2008; Jaeger 2013; Buz & Jaeger 2016; reviewed in Jaeger & Buz 2017). In support of this model as a contributing mechanism to contrastive hyperarticulation, speakers have been shown to hyperarticulate phonetic cues to a phoneme contrast when a minimal pair competitor defined by that contrast is in the immediate context, or when the word has been previously misunderstood as its minimal pair counterpart (Buz & Jaeger 2016, Schertz 2013, Seyfarth, Buz & Jaeger 2016). (iii) Finally, in production-internal accounts, rather than an attempt to make confusable words easier to understand, contrastive hyperarticulation arises through competitive inhibition in production planning. Most recently, within Dynamical Field Theory Stern and Shaw (2023) argue that contrastive hyperarticulation can arise through inhibitory neural connections between competitor lexical representations in the process of production planning. This model proposes that the inhibition of a close lexical competitor in production planning drives dissimilation of the features distinguishing them (see also Baese-Berk & Goldrick 2009). Using the example of dissimilation of the longer VOT in a voiceless stop from the shorter VOT in a voiced stop, when producing a voiceless stop-containing target word which has a voiced-stop minimal pair competitor, inhibition of the voiced stop spills over to the neighboring voiceless consonant category of the word under production. The region of the voiceless stop VOT distribution closest to that of the voiced category distribution is more strongly inhibited, leaving the region of the voiceless distribution further away to dominate production planning. As a result, the VOT target of voiceless stop production planned during selective inhibition of the voiced category will tend to be longer, i.e. more hyperarticulated, compared to production with lower selective inhibition, as is the case without a near lexical competitor. This model is consistent with previous results suggesting that contrastive hyperarticulation of VOT distinctions in English arises through suppression of productions in the overlap region, rather than through hyperarticulation across the board (Buz & Jaeger 2016; Buz 2016).

The finding that within-syntactic category minimal pair counts better predict change type can be potentially accounted for in each of these model types, provided that local morphosyntactic structure is available at the point at which contrastive hyperarticulation takes place. In listener-internal accounts, this finding requires that lexical categories include syntactic behavior as part of their internal category representation, such that predicted syntactic category is part of the constellation of factors that contribute to lexical processing (e.g., Lester, Feldman & Moscoso del Prado Martin 2016; Berkovitch & Dehaene 2019). Likewise in listener-oriented production accounts, for this to hold syntactic category must be part of the set of predicted cues that contribute to lexical category identification, such that a speaker's prediction for the likelihood of misunderstanding is lower when target-competitor syntactic categories do not match, in turn resulting in less hyperarticulation. Turning to Stern & Shaw's speaker-internal production account, if syntactic category representations are coupled to lexical representations, minimal pair competitors that share a category will receive additional activation from that category. Minimal pair competitors sharing a category would thus exert a greater inhibitory influence on articulatory planning, causing contrastive hyperarticulation as described above. Conversely, without the overlapping category representation, contrastive hyperarticulation would be reduced or eliminated (Stern & Shaw, personal communication).

In this study we counted lemma minimal pairs rather than lexeme minimal pairs. That is, we only counted minimal pairs between citation forms for lexical entries (lemmas), e.g. the verb pair *dip* ~ *tip*, rather than counting minimal pairs between inflectionally affixed lexemes, i.e. *dip* ~ *tip*, *dips* ~ *tips*, *dipping* ~ *tipping*, *dipped* ~ *tipped*, etc. Our decision to do this was based on earlier findings for mergers (Wedel, Jackson & Kaplan 2013) that counting by lemmas provided a better fit to the data than counting by lexemes. However, this observation may have been driven by the fact that Wedel et al. (2013)'s dataset was dominated by languages with relatively regular inflectional systems, such as English. We note that the number of minimal pairs for the Icelandic consonant transphonologization is an outlier in our data set, considerably lower than other shifts and transphonologizations and within the range of mergers. In our dataset, the number of *lemma* minimal pairs in the within-category, featural model is 12 for the Icelandic transphonologization, well below the crossover point between mergers and shifts in this model at approximately 40 minimal pairs. However, if minimal pairs are counted by inflectionally affixed *lexemes* (including all word forms within inflectional paradigms), then the number of minimal pairs associated with the Icelandic transphonologization is considerably higher, well within the distribution of minimal pairs for shifts and transphonologizations.

In many languages, such as Icelandic, inflectional changes are often accompanied by stem-changes, such that minimal pairs that occur in affixed forms may not have citation forms that are minimal pairs. For example, the words *bergi* and *berki,* dative singular forms of 'rock face' and 'bark' (of tree), respectively, are clear within-syntactic (and even within-morphological) category minimal pairs in Icelandic. However, their lemma (nominative singular) forms, *berg* and *börkur*, are not, so the inflected forms *bergi* and *berki* are not counted as competitors in our analysis. Similarly, *sýnd* and *sýnt* are both nominative singular past participles of the verb *sýna*, with the only difference being that *sýnd* is feminine and *sýnt* is neuter, and these would also not be counted as competitors in our analysis. Counting by inflectionally affixed lexemes at least for some languages is supported by Blevins & Wedel (2009)'s and Flego's (2022)'s arguments that within-morphological paradigm lexeme minimal pairs are able to drive merger avoidance. The data we find here for Icelandic is consistent with this argument. This highlights the need for further research on how we go about quantifying functional load in morphosyntactic exponence, and more generally on lexical competition and access in languages with high inflectional synthesis.

4.3 How does hyperarticulation of a small number of words - those in minimal pairs - influence the broader lexicon?

The notion that variation in production of minimal pairs can influence change or stability in the wider lexicon is consistent with models that propose a causal chain linking phonetic variation at the level of individual utterances to long-term change in the abstract, sublexical category system of a speech community (e.g., Lindblom 1990, Bybee 2001, Blevins 2004, Walsh et al. 2010, Kirby 2010, Wedel 2012, Hay & Maclagan 2012, Soskuthy 2015, Flego 2022). A common prediction of these models is that a bias does not have to operate in all possible instances for it to generate change; instead it is sufficient for a bias to operate just in enough of the words, enough of the time to drive the generalization of a new pattern in the phonology of the language as a whole. In support of this prediction, exposure to variants of a sound category in specific words has been shown to result in long-perseverating shifts in listener's category boundaries for those sounds and featurally related sounds, even in words that were not used in the exposure phase (e.g., Eisner & McQueen

2005; Kraljic & Samuel 2005, Nielsen 2011; Lindsay et al. 2022, reviewed in Goldrick & Cole 2023).

To arrive at this prediction, these models draw on two related findings: (i) perceptual categories maintain some record of experienced variation rather than being fully abstract (discussed in the context of language in, among many others: Goldinger 1998; Bybee 2002; Johnson 2006; Pisoni & Levi 2007; Ernestus 2011; Foulkes & Hay 2015; Goldrick & Cole 2023); and (ii) experiencing a particular category variant influences future production and perception behavior not only for that category, but also for similar categories (reviewed Kleinschmidt & Jaeger 2015, Goldrick & Cole 2023)

This production-perception feedback loop (e.g., Pierrehumbert 2001; Wedel 2006, 2012; Soskuthy & Hay 2017, Todd et al. 2019, Flego 2022) provides a pathway by which variation in cues to linguistic category membership can spread across categories, from word to word (e.g., Bybee 2002; Maye et al. 2008) and from sound to sound (e.g., Nielsen 2011, Levi 2015, Lindsay et al. 2022) over lifetimes and generations within a speech community. As children abstract phonological grammars from the input they hear, what began as gradient, token-level shifts in adult production behavior can become reinterpreted as categorical patterns, completing the link from usage at the level of the individual, to phonology at the level of the language (e.g., Beddor 2009, see 4.3 below).

Couched within this body of models, when a word is used that has a minimal pair competitor distinguished by some cue, that cue will be hyperarticulated to enhance contrast between the word and its competitor. This enhanced sound token contributes a more contrastive exemplar to the phoneme category for that sound, thereby nudging the prototypical sound for that phoneme category away from its competing phoneme category. If two phonemes are in danger of merger because of close proximity and greater confusability, a large number of minimal pairs may suffice to preserve separation between these phonemes across the lexicon, resulting in (i) stability in phonetic space despite proximity, (ii) chain-shifting in that space, or (iii) transphonologization, that is, merger of one cue and expansion of another (Wedel 2006, 2012, Winter & Wedel 2016). Conversely, if there are few minimal pairs, their contribution to contrast maintenance may not be sufficient to prevent phonetically adjacent sounds from merging.

4.4 The roles of language usage versus acquisition in phoneme inventory evolution

Structure in phoneme systems can be thought of as a solution to the problem of how to carve up a perceptual space to map efficiently yet expressively to a categorical informational space. The domain of color naming systems presents a parallel problem. Regier and colleagues' work has shown that the Information Bottleneck Principle (Tishby, Pereira, & Bialek, 1999) - which is a formalization of the relation between efficiency versus informativeness - predicts how color systems balance how many categories exist (and where to place them), efficiency (fewer words), and informativeness (accurately distinguishing colors). Using computational modeling, they showed that generation of cross-culturally common systems depended on the interaction of two kinds of learning: that which occurs during transmission between generations, promoting efficient systems (i.e., iterated learning; Kirby, Griffiths & Smith 2014) and that occurring during usage within a generation, promoting expressive systems (Carlsson, Dubhashi & Regier 2024).

Likewise in the domain of language, Kirby et al. (2015) propose a model of cultural evolution that depends on both kinds of learning. Using both computational modeling and artificial language learning experiments, they showed that iterated-learning across generations gave rise to a pressure for simplicity (see also Fedzechkina, Jaeger & Newport 2012), while usage within a generation gave rise to pressure for expressivity. Both were required to produce language-like structure.

These conclusions parallel two observations in language pattern formation. (i) Many grammatical patterns in language mirror the kinds of errors and processes arising in adult usage within generations, rather than those associated with children in acquisition across generations (e.g., Bybee & Slobin 1982, Ohala 1989, Garret & Johnson 2013). (ii) At the same time, children have been shown to simplify variable language input into simpler, more categorical systems (reviewed in Austin et al. 2021). This supports a model in which processes in adult usage shift the mapping between phonetic cues and categories within a phoneme inventory, while children 'clean-up' the input to acquisition into an inventory that may have a more direct/simpler phonetics-phoneme category mapping. In support of the causal primacy of adult usage in explaining phoneme inventory structure, Wedel, Jackson & Kaplan (2013) found that minimal pairs with low frequency members accounted for phoneme merger probability to the same degree as minimal pairs in which both members are frequent, suggesting that it is usage by adults that create the conditions for merger, not language acquisition by children, who have small vocabularies. Similarly, Wedel, Nelson and Sharp (2018) found that the degree of minimal-pair-associated contrastive hyperarticulation of stops and vowels in casual speech was statistically unrelated to the frequency of the minimal pair competitor. The influence of low-frequency words again suggests that this hyperarticulation is not a fossilized echo of processes in language acquisition when vocabularies are small, but rather a process located in adult usage. Consistent with this, as reviewed in section 4.2 above, all extant models of contrastive hyperarticulation locate the cause of hyperarticulation in adult usage of language, whether through processes inherent to perception or production, or through audience design.

4.4 Conclusion

Languages balance pressures for efficiency and accurate information transmission. This tension has been argued to play out in the structure of phoneme inventories and how they change (e.g. Liljencrantz & Lindblom 1972). Here we show that phoneme category mergers - changes which potentially create homophony - only occur when the actual degree of resulting homophony is low. Conversely, chain shifts and transphonologizations, which preserve lexical distinctions, occur if the degree of resulting homophony under merger would be high. We confirm that only within-syntactic category minimal pairs account for this pattern, suggesting that local predictability within an utterance is a contributing factor. Phonological processes often involve parallel changes over sets of sounds that share some property, i.e. a feature. Here we show that treating parallel phoneme changes as one feature-based change gives a significantly better fit to the data, providing evidence from a novel process that phonological category change occurs over similarity-classes.

These results are consistent with evolutionary models of language change in which biases in usage toward successful communication operate at the level of meaning-bearing units such as words, not at the level of sublexical units such as phonemes (reviewed in Hall et al 2018). In other words, evolutionary pressure on phoneme inventories appears not to arise through whether a listener understands that a phoneme is /p/ or /b/, but rather through whether a listener understands that a word is *pat* or *bat*. This work contributes to phonology more broadly by providing evidence that a causally explanatory model for phoneme category structure requires an expansion of phonological theory beyond the traditional focus on sublexical units to include meaning-bearing units. In this way, patterns of change in phoneme inventories can be understood as optimizing solutions to a trade-off between successful communication of meaning and efficiency, preserving functional complexity while eliminating redundant category distinctions.

Appendix A. Database

Language	Corpus	Type/Cntxt	Type*	Phonemes	Notes	References
Engl. (RP)	CELEX	V ~ V	M	aɪ ~ ɔɪ	PRICE ~ CHOICE	Wells 1982
	Baayen et al. 1995		M	ʊə ~ ɔː	CURE ~ NORTH	
			M	ɪə ~ ɛə	NEAR ~ SQUARE	
			M	ɛː ~ ɛə	NURSE ~ SQUARE	
		C ~ C				Labov et al. 2006
			M	θ ~ f		
				ð ~ v		
		V ~ V	S	ɪ ~ ɛ	New Zealand front vowel shift	Hay 2008
				æ ~ ɛ		
				ɛ ~ i		
Engl. (Am)	CMUDict, Weide, R. L. 1995	V ~ V	M	ɑ ~ ɔ	LOT ~ THOUGHT	Labov et al. 2006
	SUBTLEX-US		M	ɔɪ ~ ɛɹ		

	Brysbaert & New 2009	V ~ Vɹ	M	ɑɹ ~ ɔɹ	START ~ NORTH	
		V ~ V[n, m].	M	ɪ ~ ɛ	PIN ~ PEN	
		V ~ V[l].	M	ɪ ~ i	HILL ~ HEEL	
			M	ʊ ~ u	PULL ~ POOL	
			M	ʊ ~ oʊ	BULL ~ BOWL	
			M	ʌ ~ ɑ	HULL ~ HALL	
			M	ʊ ~ ʌ	BULL ~ HULL	
		C ~ C	M	w ~ ʍ		
		C ~ C/V_V̌	M	t ~ d	flapping	
		C ~ Øj	M	h ~ Ø		Gordon 2008
		C ~C#	T		word-final devoicing	Purnell et al 2005
				dʒ~tʃ		
				g ~ k		
				v~ f		
				b ~ p		
				z ~ s		
				d ~ t		
		V ~ V	S	ɑ~æ	American English vowel shifts	Eckert 2008

			S	ε~ɪ		Labov et al. 2006
			S	ε ~ æ		
			S	ʌ ~ ε		
			S	e ~ ε		
			S	i ~ ɪ		
German	CELEX Baayen et al. 1995	V ~ V	M	eː ~ εː	GEBE ~ GÄBE	Wiese 2000
		C ~ C	M	ç ~ ʃ	FICHTE ~ FISCHTE	Jannedy & Weirich 2014
Dutch	CELEX	C ~ C	M	s ~ z		Kissine et al. 2003
	Baayen et al. 1995			f ~ v		
				x ~ ɣ		
		V ~ V	S	eː ~ ε	Polder Dutch shift	Jacobi 2009
				øː ~ œy		
				oː ~ ɔu		

Icelandic	Gigaword	V: ~ V:	M	ɪː ~ ɛː	*Flámæli*	Ankirskiy 2018
	Nikulásdóttir et al. 2024			ʏː ~ œː		Arnbjörns-dóttir 2006
		CC ~ CC	T	rʋ ~ r̥p	Post-Sonorant Laryngeal Trans.	Árnason 2011
				rð ~ r̥t		Hansson 2001
				rc ~ r̥c		
				rk ~ r̥k		
				lʋ ~ l̥p		
				lt ~ l̥t		
				lc ~ l̥c		
				lk ~ l̥k		
				mp ~ m̥p		
				nt ~ n̥t		
				ɲc ~ ɲ̊c		
				ŋk ~ ŋ̊k		
French	Lexique	V ~ V	M	ɛ̃ ~ œ̃	VIN ~ UN	Fagyal et al. 2006
	New et al. 2001		M	e ~ ɛ	ÉPÉE ~ ÉPAIS	
			M	œ ~ ø ~ ɔ		

Danish	Derczynski, L., Ciosici, M. R., et al. (2021)	C~C]coda	S	p~b		Basbøll 2005
			S	t~d		
			S	k~g		
Spanish	Gigaword	C ~ C	M	ʎ ~ j		Penny 2002[106]
	Mendonça et al. 2009		M	s ~ θ		Harris 1969
		C ~ C	S	p ~ b		Lewis, A. 2002
				t ~ d		
				k ~ g		
		C ~ C	M	s ~ ∅		Cotton et al. 1988
Slovak	Slovak National Corpus	C ~ C	M	ʎ ~ l		Krajčovič 1988
	Šimková (2006)	V ~ V	M	æ ~ a		
				æ ~ e		

Turkish	Sak, Güngör, & Saraçlar, 2008	C ~ Ø/V_V	M	g ~[Ø		Lewis, G. 2000
		C ~ V:Ø	T	g ~ V: Ø]	deletion with compensatory vowel lengthening	Lewis, G. 2000
Korean	Korean Academy Database	C ~ C/[-son]_	M	p ~ p͈	post-obs. tensing	Sohn 1999[173]
	Yee 2006			t ~ t͈		
				s ~ s͈		
				tɕ ~ tɕ͈		
				k ~ k͈		
		C ~ C/_]$_\sigma$	M	p ~ pʰ	coda neut.	Sohn 1999[165]
				t ~ tʰ ~ s ~ s ~ tɕ ~ tɕʰ ~ h		
				k ~ kʰ ~ k͈		
		C ~ C	T	k ~ g	VOT to F0 in the aspirated~lenis contrast	Silva 1992
				t ~ d		
				p ~ b		

Cantonese (HK)	HK Cant. Corpus	C ~ C/#_	M	n ~ l		Zee 1999
	Kwong 2004	C ~ C/_#	M	m ~ ŋ		Zee 1985
		T ~ T	M	2 ~ 5		Mok & Wong 2010
		C~ 0	M	ŋ ~ 0		To, Mcleod, & Cheung 2015
Vietnamese	Sketch Engine	V ~ V/_C#	S	iə̯ ~ i	Closed Syllable Length Contrast	Pham, 2006; Shimizu, 2021
	Kilgaroff 2004, 2014; Baroni et al. 2009			ɯə̯ ~ ɯ		
				uə̯ ~ u		
		V ~ V/_#	T	iə̯ ~ i	Open Syllable Diphthongization	Pham, 2006; 2019; Shimizu, 2021
				ɯə̯ ~ ɯ		
				uə̯ ~ u		

*M = Merger, S = Chain Shift, T = Transphonologization

Wordlists are available on request, except for Vietnamese.

Cantonese, Dutch, French, German, Korean, RP English, Slovak, Spanish, and Turkish word lists were those used in Wedel, Kaplan & Jackson 2013 and Wedel, Jackson & Kaplan 2013.

The American English lemma list is that used in Nelson & Wedel 2017.

Danish POS-tagged lemmas were obtained from The Society for Danish Language and Literature, 2020, *Lemma list of the Danish Dictionary - DDO (ELEXIS)*, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1504. Phonemic representations for these lemmas were kindly provided by Tom Brøndsted (Brøndsted, Tom: Automatic Phonemic Transcriber. http://tom.brondsted.dk/text2phoneme, visited Dec 3, 2024). A Danish frequency list was obtained from SketchEngine (Kilgarriff et al. 2004, Kilgarriff et al. 2014, Derczynski et al. 2021). Many forms in the lemma list did not have a corresponding form in the frequency list; these forms were arbitrarily assigned a frequency of 1.

Icelandic POS tagged lemmas with frequencies and pronunciations were obtained from Nikulásdóttir et al. 2024, http://hdl.handle.net/20.500.12537/331. The standard pronunciation was used.

Vietnamese POS tagged lemmas and frequencies (Baroni 2009) were obtained from SketchEngine (Kilgarriff et al. 2004, Kilgarriff et al. 2014).

Appendix B. Analysis using the highest frequency item among homophones

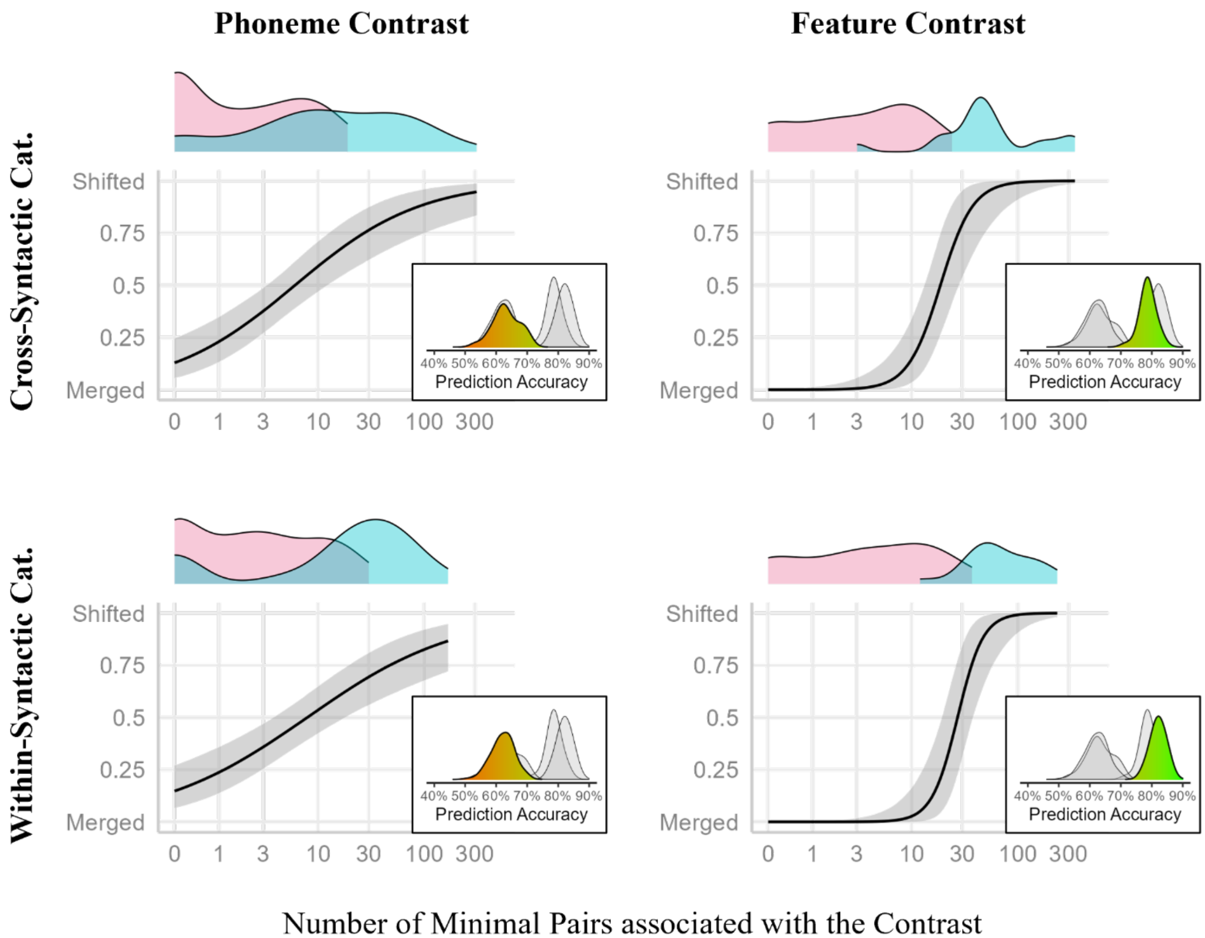

Figure B.1. Model predictions and accuracy scores for four binomial models

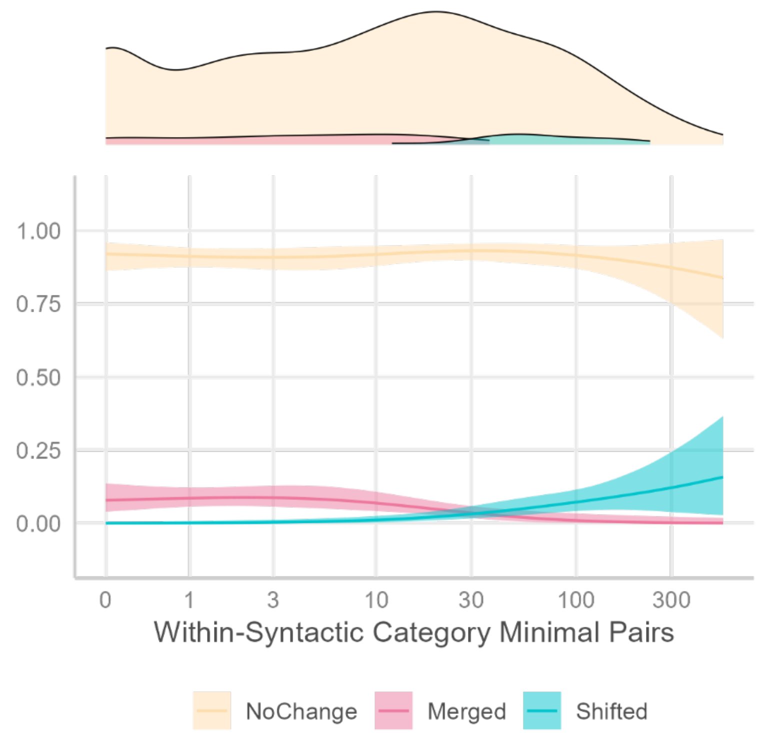

Figure B.2. Model predictions including non-changing contrasts

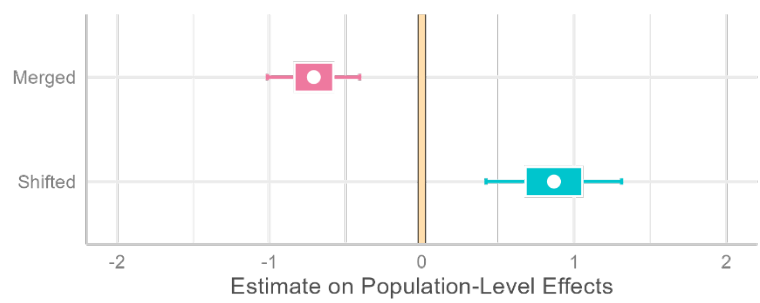

Figure B.3. Effect of 'Merged' and 'Shifted' vs. 'No Change' on Minimal Pair Count

References

Ankirskiy, A. (2018). *Investigating the potential for merger of Icelandic 'flámæli' vowel pairs through functional load* [Masters thesis, University of Georgia].

Árnason, K. (2011). *The phonology of Icelandic and Faroese*. Oxford University Press.

Arnbjörnsdóttir, B. (2006). *North American Icelandic: The life of a language*. University of Manitoba Press.

Austin, A. C., Schuler, K. D., Furlong, S., & Newport, E. L. (2021). Learning a Language from Inconsistent Input: Regularization in Child and Adult Learners. *Language Learning and Development*, 18(3), 249–277. https://doi.org/10.1080/15475441.2021.1954927

Baayen, R. H., Piepenbrock, R., & Rijn, H. van. (1995). *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium. (Release 2, CD-ROM.)

Baese-Berk, M. and Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, 24(4):527–554.

Baese-Berk, M. & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24.527-54.

Barkarson, S., Steingrímsson, S., & Hafsteinsdóttir, H. (2022). Evolving large text corpora: Four versions of the Icelandic Gigaword Corpus. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (pp. 2371-2381).

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, *43*, 209-226.

Basbøll, H. (2005). *The phonology of Danish*. OUP Oxford.

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785-821.

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50-61.

Berkovitch, L., & Dehaene, S. (2019). Subliminal syntactic priming. *Cognitive psychology*, *109*, 26-46.

Blevins, J. & Wedel, A. (2009). Inhibited Sound Change: An Evolutionary Approach to Lexical Competition. *Diachronica* 26, 143-183.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns* Cambridge: Cambridge University Press.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods, 41, 990–997.

Brøndsted, T: Automatic Phonemic Transcriber. http://tom.brondsted.dk/text2phoneme, visited Dec 3, 2024.

Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American statistical Association*, *83*(401), 123-127.

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. doi:10.18637/jss.v080.i01

Buz, E. (2016). Speaking with a (slightly) new voice: investigating speech changes, communicative goals, and learning from past communicative success. Doctoral dissertation: University of Rochester.

Buz, E. & Jaeger, T. F. (2016). The (in)dependence of articulation and lexical planning during isolated word production. *Language, Cognition, and Neuroscience* 31.404-24.

Buz, Tanenhaus, M. K. & Jaeger T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language* 89. 68-86.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14.261-90.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.

Bybee, J., & Slobin, D. I. (1982). Why small children cannot change language on their own: Suggestions from the English past tense. In A. Ahlqvist (Ed.), *Papers from the 5th International Conference on Historical Linguistics* (pp. 29–37).

Cancho, R. F. I., & Solé, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences*, 100(3), 788-791.

Carlsson, E., Dubhashi, D., & Regier, T. (2024). Cultural evolution via iterated learning and communication explains efficient color naming systems. *Journal of Language Evolution*, 9(1-2), 49-66.

Cotton, E. G. & Sharp, J. M. (1988), *Spanish in the Americas*, Washington: Georgetown University Press.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306-329.

Derczynski, L., Ciosici, M. R., et al. (2021). The Danish Gigaword Corpus. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa 2021)*.

Eckert, P. (2008). Where do ethnolects stop?. *International journal of bilingualism*, *12*(1-2), 25-42.

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, *142*, 27-41.

Eisner, F. & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics* 67. 224-38.

Fagyal, Z., Kibbee, D., & Jenkins, F. (2006). French: A linguistic introduction. Cambridge: Cambridge University Press.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, 109(44), 17897-17902.

Flego, S. (2022). *The emergence of vowel quality mutation in Germanic and Dinka-Nuer: Modeling the role of information-theoretic factors using agent-based simulation* [Doctoral dissertation, Indiana University Bloomington].

Flemming, E. (2004). Contrast and perceptual distinctiveness. In *Phonetically-Based Phonology*, Bruce Hayes, Robert Kirchner and Donca Steriade (eds.), 232–276. Cambridge: Cambridge University Press.

Foulkes, P. & Hay, J. (2015). The emergence of sociophonetic structure. *Handbook of Language Emergence*, ed. by B. MacWhinney & W. O'Grady, 292-313. Malden, MA: Wiley Blackwell.

Fox, N. P., Reilly, M., & Blumstein, S. E. (2015). Phonological neighborhood competition affects spoken word production irrespective of sentential context. *Journal of Memory and Language*, 83, 97-117.

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. *Origins of sound change: Approaches to phonologization*, 1, 51-97.

Gibson, E. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5), 389-407.

Gilliéron, J. (1918). *Généalogie des mots qui désignent l'abeille d'apres l'atlas linguistique de la france*. Paris, France: É . Champion.

Goldinger, S. D. (2000). The role of perceptual episodes in lexical processing. *Proceedings of SWAP (Spoken Word Access Processes)*, ed. by A. Cutler, J. McQueen & R. Zondervan, 155-58. Nijmegen: Max Planck Institute for Psycholinguistics.

Goldrick, M., & Cole, J. (2023). Advancement of phonetics in the 21st century: Exemplar models of speech production. *Journal of Phonetics*, *99*, 101254.

Gordon, M. J. (2002). Investigating chain shifts and mergers. In J.K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *Handbook of language variation and change*. Malden, MA: Blackwell. 244–266.

Gordon, M. J., & Schneider, E. W. (2008). New York, Philadelphia, and other northern cities: Phonology. *Varieties of English*, *2*, 67-86.

Guy, G. (1996). Form and function in linguistic variation. In G. Guy, C. Feagin, D. Schiffrin & J. Baugh (eds.), *Towards a social science of language: Papers in honor of William Labov*, 221–252. Amsterdam: John Benjamins.

Hahn, M., Jurafsky, D., & Futrell, R. (2020). Universals of word order reflect optimization of grammars for efficient communication. Proceedings of the National Academy of Sciences, 117(5), 2347-2353.

Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. (2018). The role of predictability in shaping phonological patterns. *Linguistics Vanguard*, 4(s2).

Halle, M., & Jones, L. G. (1971). *The sound pattern of Russian: A linguistic and acoustical investigation* (No. 1). Walter de Gruyter.

Hansson, G. Ó. (2001). Remains of a submerged continent. In *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics, Vancouver, 9-13 August 1999* (Vol. 215, p. 157-173). John Benjamins Publishing.

Hay, J. (2008). *New Zealand English*. Edinburgh University Press.

Hay, J. & MacLagan, M. (2012). /r/-sandhi in early 20th-century New Zealand English. *Linguistics* 50.745-63.

Hawkins, J. A., Christiansen, M. H., Collins, C., & Edelman, S. (2009). Language universals and the performance-grammar correspondence hypothesis. *Language Universals*, 54-78.

Hockett, C. (1967). The quantification of functional load. *Word*, 23, 320–339.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language learning and development*, 1(2), 151-195.

Jacobi, I. (2009). *On variation and change in diphthongs and long vowels of spoken Dutch*. Universiteit van Amsterdam [Host].

Jaeger, T. F. (2013). Production preferences cannot be understood without reference to communication. *Frontiers in Psychology* 4. 230.

Jaeger, T. F., & Buz, E. (2017). Signal reduction and linguistic encoding. *The handbook of psycholinguistics*, 38-81.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague.

Jakobson, R., Fant, C. G., & Halle, M. (1951). Preliminaries to speech analysis: The distinctive features and their correlates.

Jannedy, S., & Weirich, M. (2014). Sound change in an urban setting: Category instability of the palatal fricative in Berlin. *Laboratory Phonology*, *5*(1), 91-122.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). Academic Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34.485-99.

Kang Kwong, L. (2004). *Hong Kong Cantonese corpus*. Retrieved from http://www0.hku.hk/hkcancor/.

Kaplan, A. (2011). How much homophony is normal? *Journal of Linguistics*, 47, 631–671.

Kilgarriff, A., Baisa, V., Bušta, J, Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1: 7-36.

Kilgarriff, A., Rychlý, P., Smrž, P., Tugwel, Dl. (2004) The Sketch Engine. Proceedings of the 11th EURALEX International Congress: 105-116.

King, R. D. (1967). Functional load and sound change. *Language*, 43, 831–852.

Kirby, J. (2010). Cue selection and category restructuring in sound change: University of Chicago Doctoral dissertation.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28, 108-114.

Kirby, S., Tamariz, M., Cornish, H., and Smith, K. (2015). Compression and Communication in the Cultural Evolution of Linguistic Structure. *Cognition*, 141, 87–102.

Kissine, M., Van de Velde, H., & Van Hout, R. (2003). An acoustic study of standard Dutch/v/,/f/,/z/and/s. *Linguistics in the Netherlands*, *20*(1), 93-104.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological review*, *122*(2), 148.

Krajčovič, R. (1988). *Vy ́vin Slovenského jazyka a dialektológia*. Bratislava, Slovakia: Slovenské Pedagogické Naklatel'stvo.

Kraljic, Tanya & Arthur G. Samuel. (2005). Perceptual learning for speech: is there a return to normal? *Cognitive Psychology* 51.141-78.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Mouton de Gruyter.

Lee, Y. (2006). Sub-syllabic constituency in Korean and English. Unpublished doctoral dissertation, Northwestern University.

Lester, N. A., Feldman, L. B., & Martín, F. M. D. P. (2017). You can take a noun out of syntax...: Syntactic similarity effects in lexical priming. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 39).

Levi, S. V. (2015). Generalization of phonetic detail: Cross-segmental, within-category priming of VOT. *Language and speech*, *58*(4), 549-562.

Lewis, A. M. (2011). Contrast maintenance and intervocalic stop lenition in Spanish and Portuguese: when is it alright to lenite?. In *Current Issues in Romance Languages: Selected papers from the 29th Linguistic Symposium on Romance Languages (LSRL), Ann Arbor, 8–11 April 1999* (pp. 159-171). John Benjamins Publishing Company.

Lewis, G. (2000). *Turkish grammar*. Oxford University Press.

Liljencrants, J. & Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 839-862.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, ed. by W.J. Hardcastle & A. Marchal, 403-39. Dordrecht: Kluwer.

Lindsay, S., Clayards, M., Gennari, S., & Gaskell, M. G. (2022). Plasticity of categories in speech perception and production. *Language, Cognition and Neuroscience*, *37*(6), 707-731.

Maclagan, M. & Hay, J. (2007). Getting fed up with our feet: Contrast maintenance and the New Zealand English "short" front vowel shift. *Language variation and change*, 19, 1-25.

Martinet, A. (1952). Function, structure, and sound change. *Word*, 8, 1–32.

Maye, J., Aslin, R., & Tanenhaus, M. (2008). The Weckud Wetch of the Wast: Lexical adaptation to a novel accent. *Cognitive Science*, 32, 543-562.

Mendonça, A., Graff, D., & DiPersio, D. (2009). *Spanish gigaword second edition*. Philadelphia, PA: Linguistic Data Consortium.

Mielke, J. (2008). The emergence of distinctive features Oxford: Oxford University Press.

Nelson, N. & Wedel, A. (2017). The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics* 64, 51-70.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du Français contemporain sur internet: LEXIQUE. *L'anné psychologique*, *101*, 447–462.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39(2), 132-142.

Nikulásdóttir, A. B., Ármannsson, B. & Bergþórsdóttir, B., (2024), Icelandic Pronunciation Dictionary for Language Technology 22.01, CLARIN-IS, http://hdl.handle.net/20.500.12537/331.

Nikulásdóttir, A. B., Ármannsson, B.and Bergþórsdóttir, B. (2024), *Icelandic Pronunciation Dictionary for Language Technology 22.01*, CLARIN-IS, http://hdl.handle.net/20.500.12537/331.

Ohala, J. J. (1989). Sound change is drawn from a pool of synchronic variation. *Language change: Contributions to the study of its causes*, 173-198.

Padgett, J. (2003). Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory*, 21(1), 39-87.

Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., and Baese-Berk, M. (2011). Phonological neighborhood effects in spoken word production: An fmri study. *Journal of Cognitive Neuroscience*, 23(3):593–603.

Pham, A. H. (2006). Vietnamese rhyme. *Southwest Journal of Linguistics*, *25*(2), 107-142.

Pham, A. H. (2019). Vietnamese dialects: A case of sound change through contact. In *Interdisciplinary perspectives on Vietnamese linguistics* (pp. 31-66). John Benjamins.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526-3529.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. *Frequency effects and the emergence of lexical structure*, ed. by J.L. Bybee & P. Hopper, 137-57. Amsterdam: John Benjamins.

Pisoni, D. B. & Levi, S. V. (2007). Some observations on representations and representational specificity in speech perception and spoken word recognition. *The Oxford Handbook of Psycholinguistics*, ed. by M.G. Gaskell. Oxford: Oxford University Press.

Purnell, T., Salmons, J., Tepeli, D., & Mercer, J. (2005). Structured heterogeneity and change in laryngeal phonetics: Upper Midwestern final obstruents. *Journal of English Linguistics*, *33*(4), 307-338.

Rischel, J. (1969), Notes on the Danish vowel pattern. *Annual Report of the Institute of Phonetics* University of Copenhagen 3, 177-205.

Sak, H., Güngör, T., & Saraçlar, M. (2008). Turkish language resources: Morphological parser, morphological disambiguator and web corpus. *Proceedings of GoTAL*, *5221*, 417–427.

Sano, S. I. (2018). Durational contrast in gemination and informativity. *Linguistics Vanguard*, *4*(s2), 20170011.

Schertz, J. (2013). Exaggerating featural contrasts in clarifications of misheard speech in English. *Journal of Phonetics* 41.249-63.

Senghas, A., Kita, S., & Ozyurek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305(5691), 1779-1782.

Seyfarth, S, Buz, E. & Jaeger, T. F. (2016). Dynamic hyperarticulation of coda voicing contrasts. *Journal of the Acoustical Society of America* 139.EL31-37.

Shimizu, M. (2021). The distribution of diphthongs in Northern and Southern dialects of Vietnamese. *Studies in Geolinguistics (地理言語学研究), 1*, 106-114.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23, 287-308.

Silverman, D. (2010). Neutralization and anti-homophony in Korean. *Journal of Linguistics*, 46, 453–482.

Šimková, M., (2006). Slovak National Corpus – History and current situation. In M. Šimková (Ed.), *Insight into the Slovak and Czech corpus linguistics* (pp. 151–159). Bratislava, Slovakia: Veda.

Sóskuthy, M. "Understanding change through stability: A computational study of sound change actuation." *Lingua* 163 (2015): 40-60.

Sóskuthy, M. & Hay, J. (2017). Changing word usage predicts changing word durations in New Zealand English. *Cognition*, *166*, 298-313.

Stent, A. J., Huffman, M. K. & Brennan, S. E. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication* 50.163-78.

Stern, M. C., & Shaw, J. A. (2023). Neural inhibition during speech planning contributes to contrastive hyperarticulation. *Journal of Memory and Language*, *132*, 104443.

Surendran, D., & Niyogi, P. (2003). Measuring the functional load of phonological contrasts (Tech. Rep. No. TR-2003–12). Chicago, IL: Department of Computer Science, University of Chicago.

Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In O. N. Thomsen (Ed.) *Competing models of linguistic change: Evolution and beyond* (pp. 43–58). Amsterdam: John Benjamins.

The Society for Danish Language and Literature, 2020, *Lemma list of the Danish Dictionary - DDO (ELEXIS)*, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1504.

Tishby, N., Pereira, F. C., and Bialek, W. (1999). The Information Bottleneck Method. In Proceedings of the 37th Allerton Conference on Communication, Control and Computation, pp. 368–377

To, C. K. S., Mcleod, S. & Cheung, P. S. P. (2015). Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech

Trubetzkoy, N. (1939). *Grundzüge der phonologie*. Prague, Czech Republic: Travaux du Cercle Linguistique de Prague.

Walsh, M., Möbius, B., Wade, T. & Schütze, H. (2010) Multilevel exemplar theory. *Cognitive science* 34, no. 4. 537-582.

Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23(3):247–274.

Wedel, A. (2012). Lexical contrast maintenance and the development of sublexical contrast systems. *Language and Cognition,* 4, 319-355.

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.

Wedel, A., Kaplan, A. & Jackson, S. (2013). Lexical contrast constrains phoneme merger: a corpus study. *Cognition*, 128, 179–186.

Wedel, A., Nelson, N., & Sharp, R. (2018). The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language*, *100*, 61-88.

Wedel, Andrew, Ussishkin, Adam, King, Adam (2019). Crosslinguistic evidence for a strong statistical universal: Phonological neutralization targets word-ends over beginnings. *Language* 95(4): e428-e446.

Weide, R. L. (1995). *The CMU pronouncing dictionary*. Pittsburgh, PA: Carnegie Mellon University. Retrieved from http://www.speech.cs.cmu.edu/cgi-bin/cmudict; retrieved 2017.

Wells, J. C. (1982). *Accents of English*. Cambridge, UK: Cambridge University Press.

Wiese, R. (2000). *The phonology of German*. Oxford, UK: Oxford University Press.

Winter, B. and Wedel, A. (2016). The co-evolution of speech and the lexicon: The interaction of functional pressures, redundancy, and category variation. *Topics in Cognitive Science*, 8(2):503–513.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, *73*(1), 3-36.

Zee, E. (1985). Sound change in syllable final nasal consonants. *Journal of Chinese Linguistics*, *13*, 291–330.

Zee, E. (1999). Change and variation in the syllable-initial and syllable-final consonants in Hong Kong Cantonese. *Journal of Chinese Linguistics*, *27*, 120–167.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology.* Cambridge, MA: Addison-Wesley.