

Equilibrium Dynamics of Near-Merger: Functional Load Effects in Cantonese Tone Contrast

Tianyi Ni, Rebecca Morley

Department of Linguistics, The Ohio State University

{ni.386, morley.33}@osu.com

Abstract

We investigate the equilibrium dynamics of near-merger in the Cantonese T2–T5 contrast using an exemplar-based model that separates production and perception. Near-merger stability arises from asymmetric lexical effects: minimal-pair lexemes trigger contrast-enhancing hyperarticulation, whereas singletons allow greater variation in the absence of lexical competition. Analytical derivations and simulations show that the equilibrium production contrast depends on the proportion of minimal pairs (π_{doublet}). A strong production–perception asymmetry emerges: although production can maintain a non-zero contrast at moderate π_{doublet} , perception collapses much more rapidly as increasing numbers of ambiguous tokens pull perceptual categories toward one another under reduced functional load. When π_{doublet} reaches zero, the system converges to full merger, providing an explicit instantiation of the functional-load hypothesis. The results explain why Cantonese rising tones remain phonetically distinct yet perceptually confusable and show how limited lexical support can sustain stable near-merger states.

Index Terms: near-merger, equilibrium dynamics, functional load, exemplar models, Cantonese tones

1. Introduction

Sound change affecting phonological contrasts is a pervasive phenomenon across languages [1]. Two sounds that were once fully contrastive—allowing speakers to distinguish them categorically—may gradually lose functional distinctiveness under articulatory, perceptual, and lexical pressures [1, 2]. One trajectory of such change is **phonemic merger**, in which categories that were historically distinct lose their functional status, as in the well-known COT–CAUGHT merger in Midland English dialect [3, 4]. This collapse is rarely abrupt; contrasts typically erode gradually and may ultimately culminate in a full merger, in which the distinction is completely neutralized. Within this trajectory, near-merger represents a crucial intermediate state that occupies the asymmetric middle ground: the contrast begins to fail perceptually while a small but systematic production difference persists [5, 6]. This stage is typically accompanied by increasing acoustic overlap across speakers and generations, as well as greater within-speaker variability in production [7, 8].

Cantonese, one of the major Chinese languages, is well known for its rich tonal system and several documented tone mergers. It comprises six contrastive lexical tones (T1–T6; Figure 1) in open or nasal-final syllables, plus three checked tones restricted to stop codas [9]. Each syllable bears one lexical tone. Prior work shows that fundamental frequency (F0) is the primary cue to tonal contrast in Cantonese, with pitch height, pitch direction, and overall F0 movement distinguishing categories [10, 11, 12]. For instance, the two rising tones, T2 and T5, differ mainly in their offset pitch height while exhibiting nearly identical trajectories. Across varieties such as

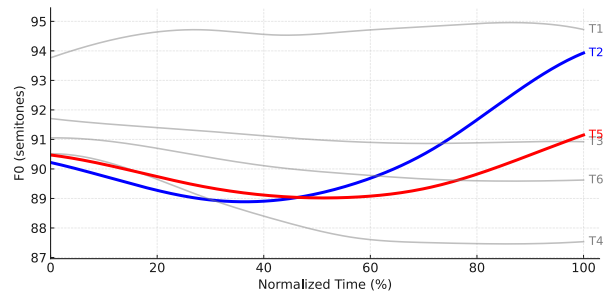


Figure 1: Mean Cantonese lexical tone contours, based on data from [13], with T2 and T5 highlighted.

Hong Kong [13], Macau [7], and Dongguan [14], the tonal space is especially crowded in the lower pitch range, where T3, T4, T5, and T6 cluster densely [13]. This compression renders several contrasts perceptually fragile and potentially unstable. Numerous studies have documented mergers or near-mergers in progress—most notably T2–T5, T3–T6, and T4–T6 [7, 13]—with some reported to be fully merged in spontaneous speech among younger speakers [15]. Within these ongoing shifts, speakers in the same community may no longer reliably perceive contrasts that remain phonetically distinguishable in production. Older speakers typically retain both perceptual sensitivity and production control, whereas younger speakers often fail to perceive the difference even when their productions still show small but systematic separation [7].

In addition, near-merger can be lexically constrained: different words may exhibit different degrees of merger depending on their lexical properties. Phonemic contrasts with higher functional load—that is, contrasts distinguishing a greater number of lexical items [16, 17]—are more resistant to loss, as minimal pairs tend to inhibit diachronic sound mergers [18, 19, 20, 21]. In the context of Cantonese, [22] demonstrated that lexical competition supports the maintenance of distinct tone categories, manifested as more categorical perceptual responses for items with lexical competitors compared to those without. This pattern aligns with the Ganong effect [23], whereby listeners’ perception of ambiguous sounds is biased toward real words, and with the Hyper- and Hypo-speech (H&H) hypothesis [24], which predicts greater articulatory distinctiveness under increased communicative or lexical contrast pressure. Together, these findings suggest that lexical competition, perceptual bias, and functional load jointly constrain the extent of phonetic variation tolerated in near-merger situations.

The current study investigates the status of near-merger through the case of the T2–T5 tone pair in Cantonese. While near-merger is often treated as a transitional stage en route to full merger, we ask under what conditions a sound sys-

tem instead stabilizes as a non-merger, a near-merger, or a full merger. We focus on how functional load, perceptual categorization, and production-based hyperarticulation jointly shape these outcomes. To address this question, we implement an exemplar-based computational model that simulates the interaction among production, perception, and lexical structure.

In standard exemplar-based models of speech representation [25, 26, 27], production and perception draw from a shared exemplar store and are treated as parallel processes. However, evidence from near-merger indicates that this assumption is too strong: speakers may maintain small but systematic acoustic distinctions in production that they no longer reliably perceive [5]. Our model therefore separates production and perception while allowing them to interact through experience-based feedback. Within this production–perception loop, perceptual biases (e.g., the Ganong effect) influence category assignment, production mechanisms (as modeled by the H&H hypothesis) modulate articulatory precision, and functional load constrains the overall stability of phonemic contrasts.

2. Data and model

2.1. Data

The dataset used in this study comes from [13, 28], originally collected to examine tone mergers in Hong Kong Cantonese. It contains monosyllabic productions from speakers exhibiting either merging or non-merging patterns. F0 contours were sampled at eight equidistant time points, converted to semitones, and the initial and final points were excluded to reduce onset and offset perturbations. For the T2–T5 pair, seven speakers identified as non-merging [13] showed clear category separation. Each target syllable was produced three times in a fixed carrier sentence. After filtering to include only non-merging productions, the dataset comprised 232 T2 and 242 T5 tokens, which served as empirical exemplars for initializing the simulation.

2.2. Exemplar space

In our model, production and perception occupy separate exemplar spaces to capture potential misalignment between what speakers produce and perceive. Each space is discretized so that acoustically different yet perceptually confusable values are represented as paired tokens across the two domains [29]. For simplicity, the exemplar space is one-dimensional, corresponding to the F0 target at tone offset (e.g., the 8th timepoint), which captures the primary contrast between Cantonese rising tones [30]. Both spaces are initialized with T2 and T5 exemplar pools obtained by random sampling from the observed data.

2.3. Lexicon

The lexicon comprises monosyllabic Cantonese words annotated for tonal category (T2, T5) and minimal-pair status. Each segment–tone combination (e.g., /dim2/ ‘point,’ from /dim/ plus its tone) forms a lexical entry: items occurring under both tones are **doublets**, and those under only one tone are **singletons**. After initialization, sampled exemplars are mapped to these entries, linking each token to its lexical identity and tone. This mapping allows the model to identify whether a token occurs in a contrastive (doublet) or non-contrastive (singleton) context, guiding the degree of hyper- or hypo-articulation during production. Each tonal category thus contains multiple subcategories, each representing an individual lexical item.

2.4. Weights

Each exemplar is initialized with a weight value of 1. During the simulation, weights decay over time at a rate r , applied equally across both categories, reflecting the gradual aging of exemplars in memory. Older exemplars therefore exert less influence on the category mean. When an exemplar re-enters production through the perception–production loop, it is stored as a new exemplar with its weight reset to 1. Memory decay is implemented as multiplicative decay:

$$w(t+1) = (1-r)w(t) \quad (1)$$

2.5. Production

Selection The model begins with LEXEME SELECTION, where the speaker selects a target word from the lexicon (e.g., /dim2/). An exemplar is randomly sampled from the corresponding subcategory (e.g., /dim2/ within tone category T2) for production.

Bias force After an exemplar is sampled from a lexeme subcategory, the model applies an UNDERSHOOT mechanism to simulate spontaneous-speech reduction due to limited duration and truncated pitch contours. The model then checks whether the lexeme is a singleton or a doublet: singletons proceed directly, whereas doublets receive additional contrast-driven HYPERARTICULATION to enhance category separation and model resistance to merger.

The undershoot mechanism follows the exponential target-approximation formulation of [31], capturing how limited duration prevents full realization of the F0 target, as shown in Equation (2):

$$x_c^{\text{undershoot}}(t) = x_c^{\text{onset}}(t) + (x_c^{\text{target}}(t) - x_c^{\text{onset}}(t))(1 - e^{-d \cdot k_c}) \quad (2)$$

where x_c^{onset} and x_c^{target} denote the onset and intended target F0 values (i.e., the actual exemplar values); d is the token’s duration, drawn from a bounded normal distribution; and k_c is a rate constant controlling the speed of approach toward the target for category c . Shorter durations or smaller k_c values yield greater undershoot, resulting in increased variability in realized contour shapes.

If the selected lexeme has a minimal-pair competitor (e.g., /ji5/ ‘chair’ vs. /ji5/ ‘ear’), the model applies contrast-driven hyperarticulation, shifting the exemplar’s target F0 away from the category boundary [30]. This enhancement is stronger when the two category means are already close: a reverse-sigmoid scaling function increases hyperarticulation as tokens approach the boundary. Doublet tokens are therefore pushed away from the midpoint between the two means—upward for T2 and downward for T5—by an amount inversely related to their distance from that midpoint. The adjustment is defined as:

$$x_c^{\text{contrast}}(t) = x_c^{\text{undershoot}}(t) + s_c \beta_{\text{max}} e^{-|x_c^{\text{undershoot}}(t) - \mu_{-c}(t)|} \quad (3)$$

$$s_c = \begin{cases} +1, & \text{if } c = 2, \\ -1, & \text{if } c = 5. \end{cases}$$

where μ_{-c} denotes the mean of the opposite category, and β_{max} represents the maximum enhancement magnitude applied when the produced exemplar lies very close to the competitor’s mean. **Imprecision force** Imprecision is modeled as Gaussian noise added during exemplar entrenchment (see below). When a stored exemplar is nudged toward the current category mean, a random perturbation drawn from $\mathcal{N}(0, \sigma^2)$ is applied to the

exemplar, introducing stochastic variation and preventing fully deterministic updates.

Entrenchment and decay Entrenchment and decay jointly stabilize category structure over time. Entrenchment pulls exemplars toward the category mean by a proportion α , while decay gradually reduces exemplar weights at rate r , limiting memory accumulation. The update for the current exemplar includes Gaussian noise, introducing stochastic variation:

$$x_c^{\text{prod}}(t) = x_c^{\text{contrast}}(t) + \alpha(\mu_c(t) - x_c^{\text{contrast}}(t)) + \epsilon(t), \quad \epsilon \sim \mathcal{N}(0, \sigma^2). \quad (4)$$

Then, the current exemplar, with its newly produced value x_c^{prod} , is then transmitted to the perception stage.

2.6. Perception

Categorization After the exemplar is passed from the production stage, the model first categorizes it. We implement this using a Gaussian kernel classifier with kernel-weighted support, shown in equation (5).

$$S_c(x^{\text{prod}}) = \sum_{i \in c} w_i \exp \left[-\frac{(x^{\text{prod}} - x_i)^2}{2h^2} \right] \quad (5)$$

and form posterior-like values:

$$P_c(x^{\text{prod}}) = \frac{S_c(x^{\text{prod}})}{S_2(x^{\text{prod}}) + S_5(x^{\text{prod}})}, \quad \hat{c} = \arg \max_{c \in \{2,5\}} P_c(x) \quad (6)$$

where x_i denotes all exemplars belonging to the category. The Gaussian kernel weights nearby exemplars more strongly than distant ones, with the bandwidth h controlling sensitivity: smaller h controlling sensitivity: smaller h yields sharper, more local decisions, whereas larger h produces smoother, more tolerant classification. Only doublets are categorized probabilistically, while singletons are assigned directly to their correct category.

Ambiguity test Since categorization above yields a hard decision, the model next tests whether the current exemplar is ambiguous, that is, lies near the boundary between the two categories. Singletons are exempt from this test, while doublets must undergo it. Ambiguity is detected when the posterior probabilities of the two categories are too close—specifically, when their difference falls below a threshold δ :

$$\Delta(x^{\text{prod}}) = |P_2(x^{\text{prod}}) - P_5(x^{\text{prod}})| \quad (7)$$

$$\text{Ambiguous}(x^{\text{prod}}) = \begin{cases} \text{ambiguous,} & \text{if } \Delta \leq \delta; \\ \text{non-ambiguous,} & \text{if } \Delta > \delta; \end{cases}$$

Crucially, exemplars flagged as ambiguous are those for which both categories are perceived as belonging to the same category. Only doublets can be ambiguous, whereas singletons cannot, due to the absence of lexical competition [22, 23].

2.7. Storage

Storage operates differently in perception and production, reflecting how Cantonese speakers integrate acoustic and non-acoustic cues during communication.

In the perception space, storage occurs immediately after ambiguity test. Unambiguous tokens are stored as exemplars of the winning category, while ambiguous tokens—those with nearly equal posterior support from both categories—are stored in both, contributing shared evidence at the boundary. Misclassified tokens are stored in the perceptually incorrect category.

This reflects how acoustic evidence near category overlap can influence multiple categories and gradually shift their boundaries.

In the production space, following perception storage, storage is guided by higher-level, non-acoustic information available to Cantonese speakers, such as contextual cues [32] that clarify intended meaning to confirm the correct lexical item and tonal category. As a result, exemplars are stored only in their correct lexical category.

2.8. Baseline model behavior

Figure 2 illustrates the temporal evolution of the weighted mean F0 values for T2 (blue) and T5 (red) in the production (left) and perception (right) grammars. At the beginning of the model’s evolution, the production and perception domains update in parallel, and the trajectories of the two categories follow a similar course. As the distributions for T2 and T5 move closer together over time, the likelihood of classifying incoming exemplars as ambiguous increases. Around epochs 600–800, this accumulation of ambiguous exemplars produces a strong bidirectional averaging effect in the perception grammar, pulling the perceptual means toward one another and visibly compressing the perceptual contrast. In contrast, the production grammar remains more robustly separated throughout the simulation, exhibiting only a gradual reduction in the mean difference due to the model’s entrenchment and replacement dynamics. This asymmetry indicates that even as speakers continue to maintain distinct production targets, their perceptual representations become increasingly confusable.

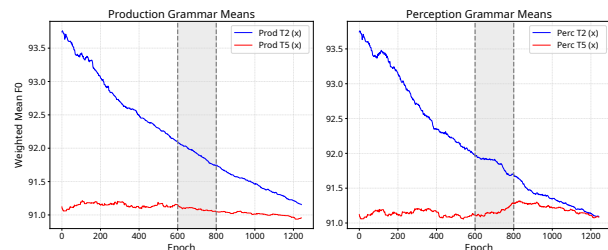


Figure 2: *Temporal evolution of category means in production and perception.*

Figure 3 plots the evolving distance between the T2 and T5 category means for production (blue) and perception (red). Consistent with Figure 2, the two domains show almost identical separation early in the simulation. The curves remain aligned until the system reaches the region around epochs 600–800, where the category distributions become sufficiently close that more exemplars are labeled as ambiguous. Once these ambiguous exemplars begin to accumulate, the perceptual distance declines more rapidly than the production distance, revealing a clear divergence between the two domains.

3. Equilibrium

We now examine the system’s equilibrium behavior to determine the conditions under which the production–perception dynamics settle into stable outcomes. Since the production grammar determines whether a residual contrast can be sustained in the long term, the equilibrium analysis focuses on production, with perception evaluated once production has stabilized.

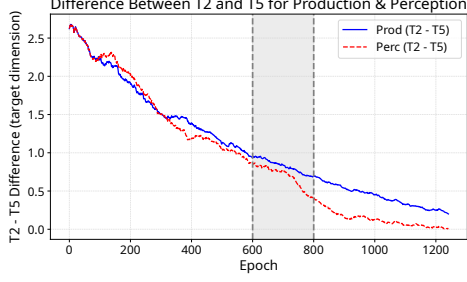


Figure 3: Temporal change of categorical distance between T2 and T5.

3.1. Equilibrium in production

At equilibrium, the category means no longer change across iterations, and the resulting contrast distance Δ^* represents the phonetic separation the system can maintain. A stable near-merger requires that production preserve a positive gap between the two categories, $\Delta^{\text{prod}}(t) = \mu_2^{\text{prod}}(t) - \mu_5^{\text{prod}}(t) > 0$, and that this gap no longer decreases across iterations. Imposing this equilibrium condition (i.e., requiring $\Delta^{\text{prod}}(t+1) = \Delta^{\text{prod}}(t)$) yields a closed-form expression that quantifies the relationship between the proportion of minimal pairs in the lexicon and the contrast distance maintained at equilibrium, as shown in Equation 8.

$$\Delta^*(t) = \pi_{\text{doublet}} \cdot \beta_{\text{max}} \left(\frac{1}{e^{d \cdot k_2}} + \frac{1}{e^{d \cdot k_5}} \right) e^{\Delta^*(t)} \quad (8)$$

where π_{doublet} is the proportion of doublets in the lexicon. This equilibrium relation is visualized in Figure 4.

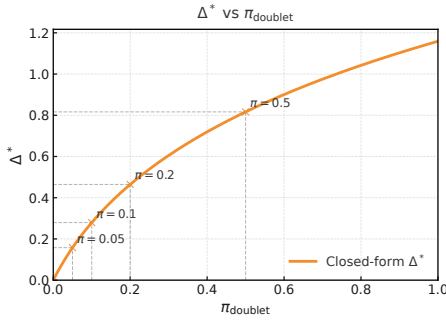


Figure 4: Closed-form equilibrium contrast distance (Δ^*) as a function of the proportion of doublets in the lexicon (π_{doublet})

In Figure 4, the curve shows a clear monotonic relationship: as the proportion of minimal pairs increases, the production system maintains a larger stable separation between the two tonal categories. When π_{doublet} is low, the equilibrium contrast becomes very small, indicating that the production system cannot sustain a phonetic gap. In the limiting case where the number of doublets falls to zero—meaning the contrast no longer serves any functional role—the equilibrium solution collapses to $\Delta^* = 0$, yielding a full merger. By contrast, when more minimal pairs remain active in the lexicon, the equilibrium distance grows, reflecting stronger functional load and greater pressure to preserve distinctiveness. The shape of the curve further sug-

gests that even a modest number of doublets can noticeably stabilize the contrast.

3.2. Perception at equilibrium

To examine perceptual behavior, we sampled 20 values of π_{doublet} (0.05, 0.10, 0.15, ...) to vary the lexical structure during model initialization. For each value, the model was run until the production system reached its equilibrium contrast Δ^* , and the perceptual category distance at that point was recorded. Figure 5 displays these perceptual contrast distances across the sampled proportions of doublets, evaluated at their corresponding production equilibria; the blue curve shows the production equilibrium contrast, and the red dotted curve shows the perceptual contrast measured at that equilibrium.

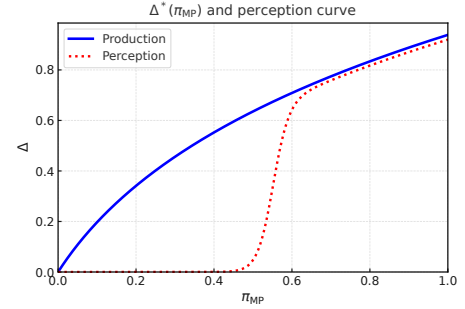


Figure 5: Perceptual contrast distance at equilibrium across different proportions of doublets (π_{doublet}).

The perceptual pattern diverges substantially from the production pattern: although production maintains a smoothly increasing contrast as π_{doublet} increases, the perceptual system collapses much more rapidly. When the proportion of doublets is low, perceptual categories converge almost completely, even in cases where production still preserves a non-zero contrast. Only once π_{doublet} becomes sufficiently large does perception begin to maintain a meaningful separation between the two categories, and even then, the perceptual contrast consistently lags behind the production contrast across most of the range. This asymmetry shows that perceptual representations are far more sensitive to reductions in functional load and are the first to collapse under diminished lexical support.

4. Discussion

The model shows that near-merger can persist only as long as some minimal pairs retain functional load; once all doublets lose their lexical function, the equilibrium contrast collapses to zero, yielding a full merger. Even in our idealized setting with equal word-selection probabilities, maintaining a small set of doublets is sufficient to preserve a positive production contrast. In more realistic lexicons, high-frequency doublets would further prolong near-merger by exerting stronger contrastive pressure and being less likely to drop out through obsolescence [18, 26, 33]. Taken together, these results provide an explicit exemplar-dynamic demonstration of the functional-load hypothesis: contrasts with greater lexical support are more resistant to merger, whereas unsupported contrasts inevitably collapse. This mechanism helps explain why Cantonese T2–T5 may remain phonetically distinct yet perceptually unstable—an equilibrium outcome sustained only by the subset of lexemes that still carry functional load.

5. References

- [1] R. Hickey, "Mergers, Near-Mergers and Phonological Interpretation," in *New Perspectives on English Historical Linguistics: Selected Papers from 12 ICEHL, Glasgow, 21–26 August 2002. Volume II: Lexis and Transmission*, ser. Varieties of English Around the World, M. Hundt, D. Schreier, and E. W. Schneider, Eds. Amsterdam: John Benjamins Publishing Company, 2008, vol. G39, pp. 125–137.
- [2] S. Todd, J. B. Pierrehumbert, and J. Hay, "Word frequency effects in sound change as a consequence of perceptual asymmetries: An exemplar-based model," *Cognition*, vol. 185, pp. 1–20, 2019.
- [3] W. Labov *et al.*, "The three dialects of English," *New ways of analyzing sound change*, vol. 5, pp. 1–44, 1991.
- [4] C. G. Clopper, D. B. Pisoni, and K. De Jong, "Acoustic characteristics of the vowel systems of six regional varieties of American English," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1661–1676, 2005.
- [5] W. Labov, M. Karen, and C. Miller, "Near-mergers and the suspension of phonemic contrast," *Language Variation and Change*, vol. 3, no. 1, pp. 33–74, 1991. [Online]. Available: <https://doi.org/10.1017/S0954394500000442>
- [6] W. Labov, S. Ash, and C. Boberg, *The Atlas of North American English: Phonetics, Phonology, and Sound Change*. New York and Berlin: Mouton de Gruyter, 2006.
- [7] J. Zhang, "Tone mergers in Cantonese: Evidence from Hong Kong, Macao, and Zhuhai," *Asia-Pacific Language Variation*, vol. 5, no. 1, pp. 28–49, 2019.
- [8] R. Y.-H. Lo and M. Babel, "How merged are Cantonese tones in spontaneous speech?" in *Proceedings of the Third International Conference on Tone and Intonation*. Herrsching, Germany: The International Conference on Tone and Intonation, May 2025, pp. 70–74.
- [9] R. S. Bauer and P. K. Benedict, *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter, 1997.
- [10] Y. Y. Fok-Chan, *A Perceptual Study of Tones in Cantonese*. Hong Kong: Hong Kong University Press, 1974.
- [11] E. Khouw and V. Ciocca, "Perceptual correlates of cantonese tones," *Journal of Phonetics*, vol. 35, pp. 104–117, 2007.
- [12] T. J. Vance, "Tonal distinctions in cantonese," *Phonetica*, vol. 34, pp. 93–107, 1977.
- [13] P. P. K. Mok, P. K. W. Wong, and A. Y. L. Leung, "Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese," *Language Variation and Change*, vol. 25, no. 3, pp. 341–370, 2013.
- [14] Y. Liang, "Merger and transfer: Tone variation and change of Dongguan Cantonese," *Lingua*, vol. 208, pp. 19–30, 2018.
- [15] N. Nagy, H. Tse, and J. N. Stanford, "Have Cantonese Tones Merged in Spontaneous Speech?" in *The Phonetics and Phonology of Heritage Languages*, R. Rao, Ed. Cambridge: Cambridge University Press, 2024, pp. 302–320.
- [16] C. F. Hockett, "The Quantification of Functional Load," *Word*, vol. 23, pp. 320–339, 1967.
- [17] D. Surendran and P. Niyogi, "Quantifying the Functional Load of Phonemic Oppositions, Distinctive Features, and Suprasegmentals," in *Competing Models of Linguistic Change: Evolution and Beyond*, O. N. Thomsen, Ed. Amsterdam: John Benjamins, 2006, pp. 43–58.
- [18] A. Wedel, A. Kaplan, and S. Jackson, "High functional load inhibits phonological contrast loss: A corpus study," *Cognition*, vol. 128, no. 2, pp. 179–186, 2013. [Online]. Available: <https://doi.org/10.1016/j.cognition.2013.03.002>
- [19] A. Wedel, S. Jackson, and A. Kaplan, "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*, vol. 56, no. 3, pp. 395–417, 2013. [Online]. Available: <https://doi.org/10.1177/0023830912443941>
- [20] A. Wedel and I. Fatkullin, "Category competition as a driver of category contrast," *Journal of Language Evolution*, vol. 2, no. 1, pp. 77–93, 2017.
- [21] A. Wedel, "Lexical contrast maintenance and the organization of sublexical contrast systems," *Language and Cognition*, vol. 4, no. 4, pp. 319–355, 2012.
- [22] R. Soo and M. Babel, "Lexical competition, sound change, and language dominance in cantonese tone category recognition," 2023.
- [23] W. F. Ganong, "Phonetic categorization in auditory word perception," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 6, no. 1, pp. 110–125, 1980. [Online]. Available: <https://doi.org/10.1037/0096-1523.6.1.110>
- [24] B. Lindblom, "Explaining phonetic variation: A sketch of the H&H theory," in *Speech Production and Speech Modelling*, W. J. Hardcastle and A. Marchal, Eds. Dordrecht: Springer Netherlands, 1990, pp. 403–439.
- [25] A. Garrett and K. Johnson, "Phonetic bias in sound change," *Origins of sound change: Approaches to phonologization*, vol. 1, pp. 51–97, 2013.
- [26] J. B. Pierrehumbert, "Exemplar dynamics: Word frequency, lenition, and contrast," in *Frequency and the Emergence of Linguistic Structure*, J. Bybee and P. Hopper, Eds. Amsterdam: John Benjamins, 2001, pp. 137–157.
- [27] R. L. Morley, *Sound Structure and Sound Change: A Modeling Approach*, second revised edition ed., ser. Conceptual Foundations of Language Science. Berlin: Language Science Press, 2023, vol. 4. [Online]. Available: <https://doi.org/10.5281/zenodo.8123702>
- [28] X. Ren and P. Mok, "The effects of contextual tonal variation on Cantonese tone merging," *Laboratory Phonology*, vol. 14, no. 1, pp. 1–27, 2023.
- [29] J. K. Kruschke, "ALCOVE: An exemplar-based connectionist model of category learning," *Psychological Review*, vol. 99, pp. 22–44, 1992.
- [30] A. Wedel, N. Nelson, and R. Sharp, "The phonetic specificity of contrastive hyperarticulation in natural speech," *Journal of Memory and Language*, vol. 100, pp. 61–88, 2018.
- [31] S. Prom-On, Y. Xu, and B. Thipakorn, "Modeling tone and intonation in Mandarin and English as a process of target approximation," *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 405–424, 2009.
- [32] C. Kung, D. J. Chwilla, and H. Schriefers, "The interaction of lexical tone, intonation and semantic context in on-line spoken word recognition: An erp study on cantonese chinese," *Neuropsychologia*, vol. 53, pp. 293–309, 2014.
- [33] J. Bybee, *Phonology and language use*. Cambridge University Press, 2003, vol. 94.