

Inventory crowding and within-context variability

Amy Li

Department of Linguistics and Philosophy, Massachusetts Institute of Technology



Case study 1: Korean dialects

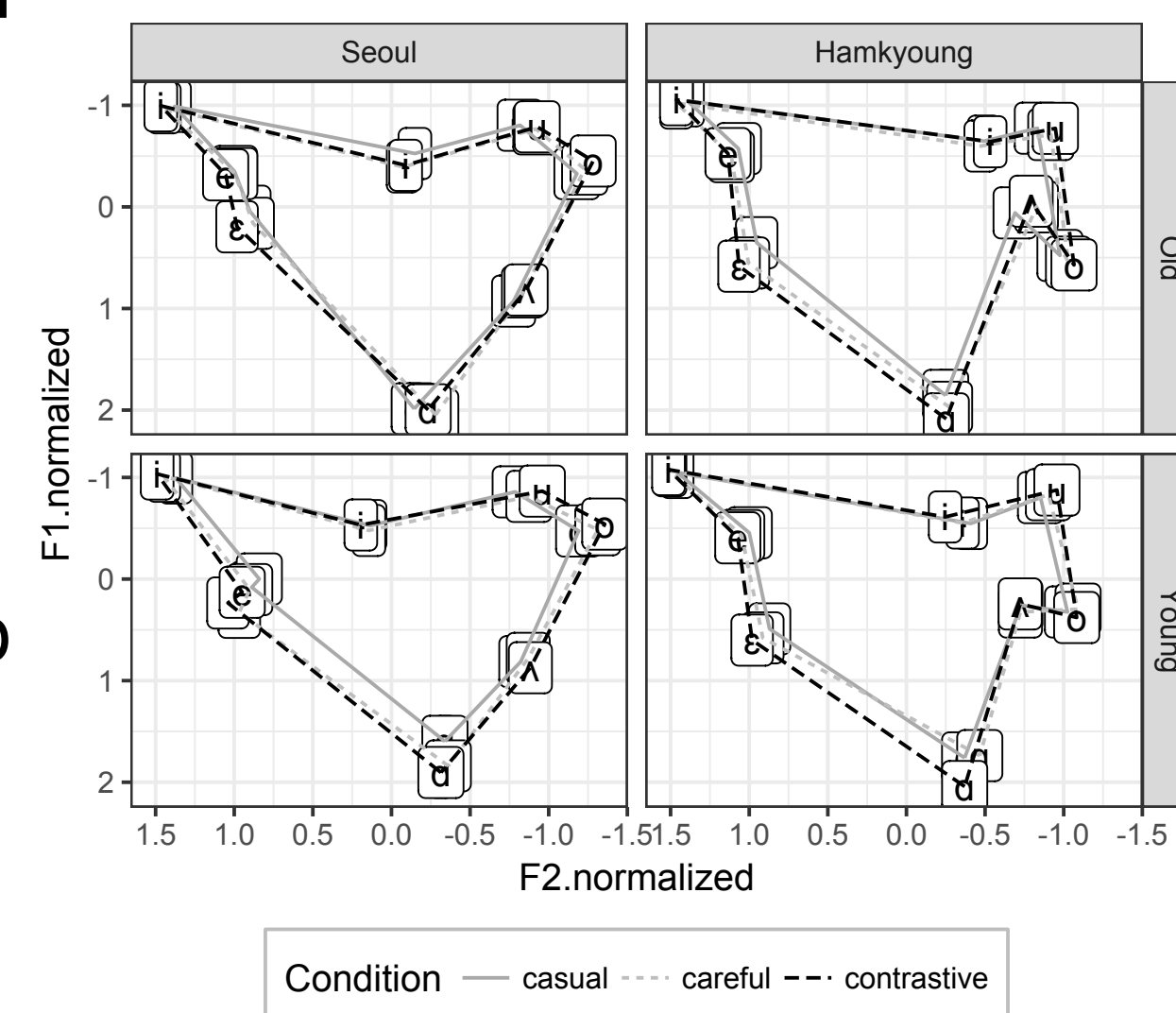
Is [i] less variable in Hamkyoung Korean than in Seoul Korean?

- ▶ Hamkyoung Korean (North Korea): retain [e]-[ɛ] contrast.
- ▶ Seoul Korean: merging [e] and [ɛ].
- ▶ Formant data from Kang et al. (2019), analyzing s.d. of [i] using two tokens in the casual condition.

Linear mixed effects model:

s.d. $\sim$ language * fn + (1 | speaker).

- ▶ Significant lower variability of f1 in Hamkyoung. ✓



Case study 2: Catalan dialects

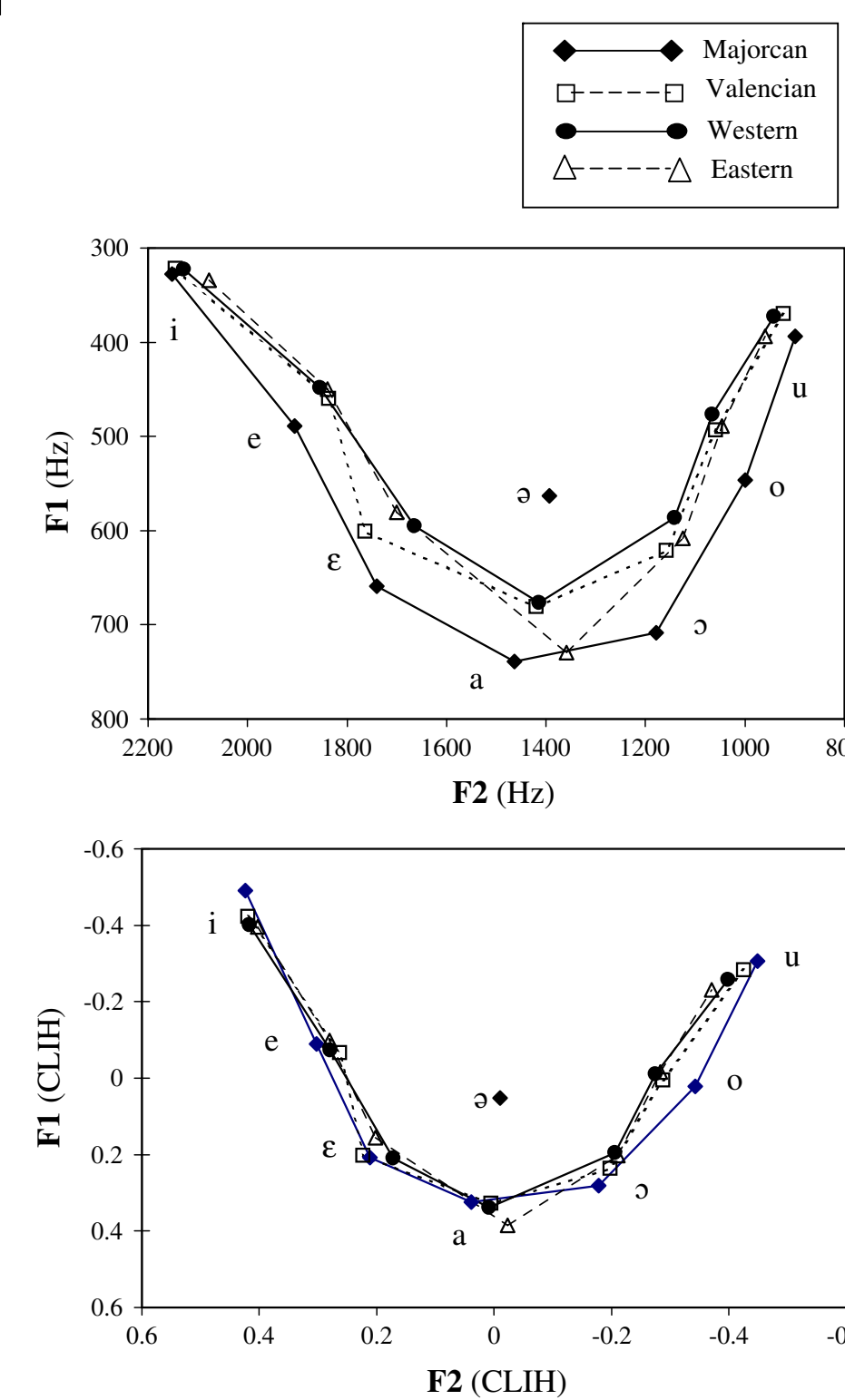
Are [a, ɛ, ɔ] less variable in Majorcan Catalan than in other dialects?

- ▶ Majorcan Catalan: contrastive schwa [ə] in stressed position for 8 total vowels.
- ▶ Valencian, Eastern, Western Catalan: no contrastive stressed [ə] for 7 total vowels.
- ▶ Formant data from target words in read meaningful sentences from Recasens and Espinosa (2006).

Linear mixed effects model:

s.d. $\sim$ language * fn * vowel + (1 | speaker).

- ▶ No significant effect for Majorcan or any interaction with [a, ɛ, ɔ]. ?



Case study 3: Italian dialects

Are [a, i, u] more variable in Italian dialects with fewer vowel height levels?

- ▶ 5-vowel Italian dialects: contrastive stressed [a, i, u, e, o].
- ▶ 7-vowel Italian dialects: contrastive stressed [a, i, u, e, o, ɛ, ɔ].
- ▶ Formant data from read meaningful sentences in CLIPS (Albano Leoni et al., 2007; Renwick, 2024).

Linear mixed effects model:

s.d. $\sim$ language * fn * vowel + (1 | speaker) + ...

- ▶ Significant greater variability of [a, i] in 5-vowel dialects. ✓

Does inventory crowdedness relate to within-context variability?

(1) Hypothesis: **more crowded** inventories have **less variable** sounds.

Limited prior empirical work:

- ▶ show null results when relating inventory size to variability (Hauser, 2022; Cortés Kandler, 2026), or
- ▶ study *contextual* variation (Lavoie, 2002; Zhang and Chodroff, 2026).

My approach: examine *within-context* formant variability of the vowel sounds *neighboring* a more/less crowded region.

What theory would account for a relationship between inventory crowdedness and variability?

No prior formal model predicts (1). Deriving (1) is not straightforward, and it only applies to languages at the universal limit of probability of confusion.

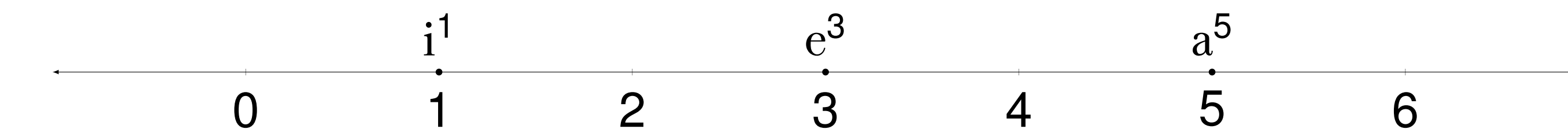
Background: Dispersion Theory of Contrast

Dispersion Theory of Contrast (Flemming, 2002, 2004) ...

- ▶ Builds on prior Dispersion Theory (Liljencrants and Lindblom, 1972; Lindblom, 1986) that contrasting sounds must be sufficiently distinct
- ▶ To make theoretical predictions about inventories of sounds.

Treat contrastive sounds as **points** along some phonetic dimension(s).

- ▶ eg. i^1 , e^3 , a^5 along the acoustic dimension *first formant*, f1.



Constraints and constraint families:

- ▶ MAXIMIZECONTRAST: n , for an inventory with n members. Higher values of n are better than lower values of n .
- ▶ MINIMUMDISTANCE[f1] = 4: Enforces distinctiveness by incurring one * per pair of sounds with distance in f1 less than 4.
- ▶ Fixed universal ranking: ... » MINDIST[f1] = 4 » MINDIST[f1] = 5 » ...

	MINDIST [f1] = 2	MINDIST [f1] = 4	MINDIST [f1] = 5	MAXIMIZE CONTRAST
(a) $i^1 e^3 a^5$		**	***	3 ()
(b) $i^1 e^2 a^5$	*	**	***	3 ()
(c) $i^1 a^5$			*	2 (*)
(d) i^1				1 (**)

Ranking MAXIMIZECONTRAST wrt. MINDIST[f1] = d selects an inventory, eg. (a) or (c), compromising on inventory size and distinctiveness.

- ▶ Poorly dispersed inventories, (b), always lose to well dispersed inventories of the same size, (a).

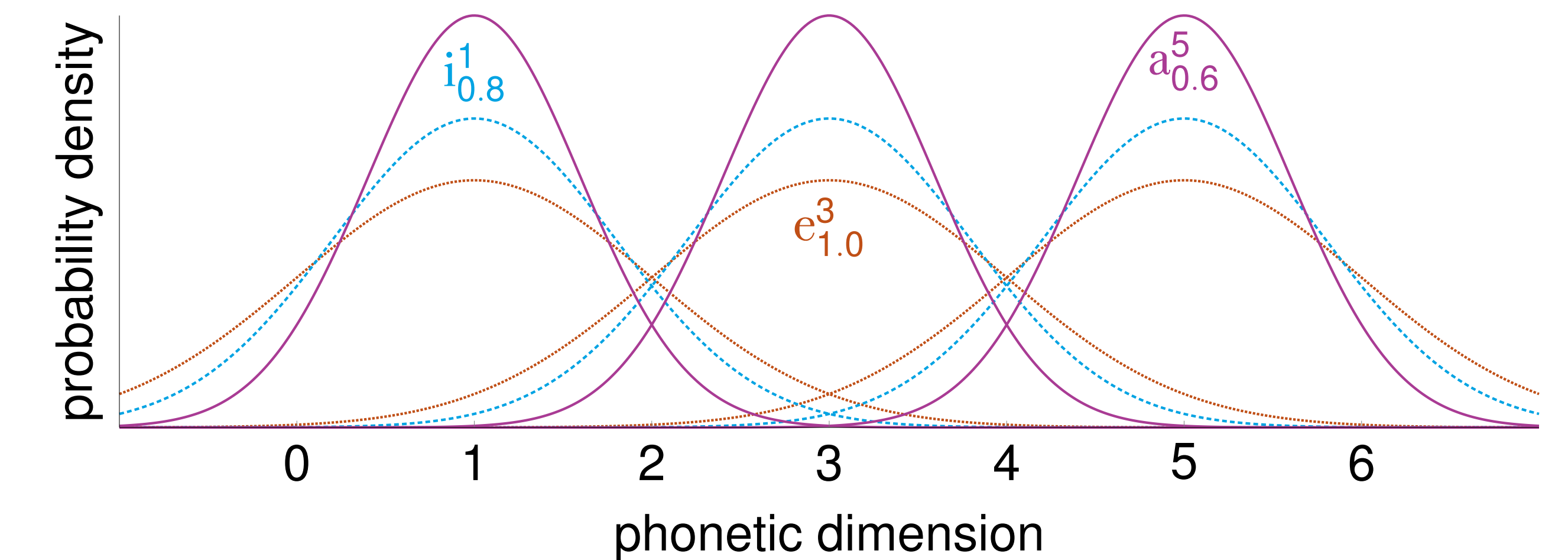
Dispersion theory with distributional representations

Articulator movement, like any muscular activation, is inaccurate.

- ▶ Let articulator inaccuracy be anticipable by the phonological system.

Treat contrastive sounds as **probability distributions**.

- ▶ For simplicity, normal distributions, such as $i_{0.8}^{1.0}$.



New constraint families, replacing MINDIST[f1] = d :

- ▶ MAXIMUMINTERSECTION[f1] = 0.2: Enforces distinctiveness by incurring one * per pair of sounds with intersection greater than 0.2.
- ▶ Fixed universal ranking: ... » MAXINT[f1] = 0.2 » MAXINT[f1] = 0.1 » ...
- ▶ $\sigma[f1] \geq 1$: One * per sound with s.d. less than 1.
- ▶ Fixed universal ranking: ... » $\sigma[f1] \geq 0.6$ » $\sigma[f1] \geq 0.8$ » $\sigma[f1] \geq 1$ » ...

	MAXINT[f1] =						$\sigma[f1] \geq$	MC
	0.22	0.1	0.05	0.02	0.01	0.8	1.0	
(a) $i^1 e^3 a^5$	**	**	**	***	***			3
(b) $i^1_{0.8} e^3_{0.8} a^5_{0.8}$		**	**	**	***		***	3
(c) $i^1_{0.6} e^3_{0.6} a^5_{0.6}$			**	**	**	***	***	3
(d) $i^1 a^5$				*	*			2
(e) $i^1_{0.8} a^5_{0.8}$					*		**	2
(f) $i^1_{0.6} a^5_{0.6}$						**	**	2

Ranking MAXIMIZECONTRAST wrt. MAXINT[f1] = p controls inventory size.

- ▶ $\sigma[f1] \geq s$ chooses between inventories of the same size.
- ▶ No prediction of (1) yet! Languages may rank MAXINT[f1] = 0.22 low.

Fix inviolable MAXINT[f1] = $p_{\max}$ at universally intolerable confusion level:

- ▶ Prediction: for **languages operating at the limit, (1) a more crowded inventory has less variable sounds**.
- ▶ The limit: except for $p_{\max}$, all MAXINT[f1] = p are ranked low.

References (abbreviated)

- Albano Leoni, F., Cutugno, F., Savy, R., Caniparoli, V., D'Anna, L., Paone, E., Giordano, R., Manfrellotti, O., Petrillo, M., and De Rosa, A. (2007). Corpora e Lessici dell'Italiano Parlato e Scritto (CLIPS).
- Cortés Kandler, M. (2026). Does vowel inventory size predict variability? LabPhon 20 presentation.
- Flemming, E. (2002). *Auditory Representations in Phonology*. Outstanding Dissertations in Linguistics.
- Flemming, E. (2004). Contrast and perceptual distinctiveness. In *Phonetically Based Phonology*.
- Hauser, I. (2022). Speech sounds in larger inventories are not (necessarily) less variable. *JASA*, 152(5).
- Kang, Y., Ryu, N.-Y., and Yun, S. (2019). Contrastive hyperarticulation of vowels in two dialects of Korean. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*, pages 43–47.
- Lavoie, L. M. (2002). Subphonemic and suballophonic consonant variation: the role of the phoneme inventory.
- Recasens, D. and Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*.
- Renwick, M. E. L. (2024). Robustness and complexity in Italian mid vowel contrasts. *Languages*, 9(4).
- Zhang, M. and Chodroff, E. (2026). A crosslinguistic analysis of vowel inventory size influence on vowel dispersion. LabPhon 20 presentation.

Inventory crowding and within-context variability

Amy Li

Department of Linguistics and Philosophy, Massachusetts Institute of Technology



Case study 1: Korean dialects

Is [i] less variable in Hamkyoung Korean than in Seoul Korean?

- ▶ Hamkyoung Korean (North Korea): retain [e]-[ɛ] contrast.
- ▶ Seoul Korean: merging [e] and [ɛ].
- ▶ Formant data from Kang et al. (2019), analyzing the casual condition.
- ▶ 1 repetition per word per speaker, 36 Hamkyoung and 21 Seoul spkr.
- ▶ Pool words containing [i] with similar contexts: *sita* [ɕita], *cito* [tɕito].
- ▶ Calculate 2-sample within-context standard deviation for each speaker.

Linear mixed effects model: $s.d. \sim \text{language} * \text{fn} + (1 | \text{speaker})$.

- ▶ Sum-coded language (Hamkyoung vs Seoul) and fn (f1 vs f2).
- ▶ (1) predicts a negative coefficient for Hamkyoung, especially in f1.
- ▶ Effect for Hamkyoung: $\beta = -38.5$, $t(55) = -1.86$, $p = 0.068$.
- ▶ Effect for H:f1 interaction: $\beta = -38.9$, $t(55) = -2.63$, $p = 0.011$.

Case study 2: Catalan dialects

Are [a, ɛ, ɔ] less variable in Majorcan Catalan than in other dialects?

- ▶ Majorcan Catalan: contrastive schwa [ə] in stressed position for 8 total vowels.
- ▶ Valencian, Eastern, Western Catalan: no contrastive stressed [ə] for 7 total vowels.
- ▶ Formant data from target words in read meaningful sentences from Recasens and Espinosa (2006).
- ▶ Calculate standard deviation of 7 repetitions per vowel in each of 4 consonantal contexts by each of 5 speakers.

Linear mixed effects model: $s.d. \sim \text{language} * \text{fn} * \text{vowel} + (1 | \text{speaker})$.

- ▶ Dummy-coded language (baseline non-Majorcan); sum-coded fn and vowel (excluding [ə]).
- ▶ (1) predicts negative coefficient for Majorcan [a, ɛ, ɔ].
- ▶ No significant effect for Majorcan or any interaction with [a, ɛ, ɔ].

Case study 3: Italian dialects

Are [a, i, u] less variable in Italian dialects with 2 levels of mid vowels?

- ▶ 7-vowel Italian dialects: contrastive stressed [a, i, u, e, o, ɛ, ɔ].
- ▶ 5-vowel Italian dialects: contrastive stressed [a, i, u, e, o].
- ▶ Formant data from read meaningful sentences in CLIPS (Albano Leoni et al., 2007; Renwick, 2024).
- ▶ Calculate standard deviation of each vowel-context-speaker for 16 speakers per dialect, 7 dialects with 7 vowels and 4 with 5 vowels.

Linear mixed effects model: $s.d. \sim \text{language} * \text{fn} * \text{vowel} + (1 | \text{speaker}) + (1 | \text{preceding place}) + (1 | \text{following place}) + (1 | \text{vowel length})$.

- ▶ Dummy-coded language (baseline 7V); sum-coded fn and vowel.
- ▶ (1) predicts positive coefficients for [a, i, u] in 5-vowel dialects, especially f1.
- ▶ Significant positive interactions: 5V:a, 5V:i.

References (complete)

Albano Leoni, F., Cutugno, F., Savy, R., Caniparoli, V., D'Anna, L., Paone, E., Giordano, R., Manfredi, O., Petrillo, M., and De Rosa, A. (2007). Corpora e Lessici dell'Italiano Parlato e Scritto (CLIPS). Available at <http://www.clips.unina.it>.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Boersma, P. and Weenink, D. (2024). Praat: doing phonetics by computer. Version 6.4.12, retrieved 12 May 2024.

Cortés Kandler, M. (2026). Does vowel inventory size predict variability? A cross-linguistic acoustic and articulatory analysis. Conference presentation at LabPhon 20, Montréal 2026.

Flemming, E. (2002). *Auditory Representations in Phonology*. Outstanding Dissertations in Linguistics. Routledge.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In Hayes, B., Kirchner, R., and Steriade, D., editors, *Phonetically Based Phonology*, pages 232–276. Cambridge University Press.

Flemming, E. (2023). Sibilant place inventories and language-specific perceptual distance. Manuscript.

Hauser, I. (2022). Speech sounds in larger inventories are not (necessarily) less variable. *The Journal of the Acoustical Society of America*, 152(5):2664–2674.

Kang, Y., Ryu, N.-Y., and Yun, S. (2019). Contrastive hyperarticulation of vowels in two dialects of Korean. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*, pages 43–47.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

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

Liljencrants, J. and Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48(4):839–862.

Lindblom, B. (1986). Phonetic universals in vowel systems.

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

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Recasens, D. and Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*, 48(6):645–666.

Renwick, M. E. L. (2024). Robustness and complexity in Italian mid vowel contrasts. *Languages*, 9(4).

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.

Zhang, M. and Chodroff, E. (2026). A crosslinguistic analysis of vowel inventory size influence on vowel dispersion. Conference presentation at LabPhon 20, Montréal 2026.

Dispersion theory with distributional representations

In addition to $p_{\max}$, other bounding limits apply to the constraint families $\text{MAXINT}[f1] = p$ and $\sigma[f1] \geq s$.

- ▶ $\text{MAXINT}[f1] = p_{\max}$ (universally inviolable): a certain amount of confusion is universally intolerable.
 - ▶ At this boundary, the tradeoff between MAXIMIZECONTRAST and $\sigma[f1] \geq s$ is that (1) more crowded inventories necessarily have narrower distributions.
- ▶ $\text{MAXINT}[f1] = p_{\min}$: reducing confusion is only helpful up to a certain amount.
 - ▶ Phonemes have no impetus to be spaced further apart than this boundary.
- ▶ $\sigma[f1] \geq s_{\min}$ (universally inviolable): only a certain amount of variability is universally achievable.
 - ▶ At this boundary, the tradeoff between MAXIMIZECONTRAST and $\text{MAXINT}[f1] = p$ reduces to the familiar tradeoff between MAXIMIZECONTRAST and $\text{MINDIST}[f1] = d$ in Dispersion Theory of Contrast (Flemming, 2002) treating phonemes as points.
- ▶ $\sigma[f1] \geq s_{\max}$: larger variability is only preferred up to a certain amount.
 - ▶ Prevents the (unlisted) inventory candidate with a single phoneme of infinite standard deviation from winning.