Dialect-specific phonological features shape perceptual generalization

Chelsea Sanker

chelsea.sanker@yale.edu

8 January 2022
Linguistic Society of America

Dialect-specific phonetics

- Different languages and different dialects can have different specific phonetic details for the realization of a sound (Lieberman 1970; Keating 1985)
- Is there sometimes also a phonological difference?
- What is the representation of a sound that is contrastive in one dialect but not in another?

Mapping across dialects

- Speakers can map vowels from other dialects onto their own vowels, even when realizations differ substantially
- Listeners will shift subsequent productions (Ross et al. 2021) and perceptual expectations (Maye, Aslin, & Tanenhaus 2008) based on exposure to other real or artificial dialects
- If two dialects have different realizations of a sound (e.g. the vowel in THOUGHT), are the phonological representations also distinct?

Generalizations in perceptual learning

Phonological organization can be reflected in feature-level shifts:

- In perceptual learning, shifted category boundaries can be extended to sounds that were not part of the exposure (e.g. Kraljic & Samuel 2006, Weatherholtz 2015)
- Prior work has demonstrated generalization across vowels (e.g. Ladefoged & Broadbent 1957; Maye, Aslin, & Tanenhaus 2008, Mitterer 2006; Chládková, Podlipskỳ & Chionidou 2017)
- What drives generalizations? Best tested with a single exposure vowel quality and several testing vowels

This study

I present results from a perceptual learning study:

- How does exposure to shifted F1 in /ɪ/ produce a shift that is generalized to other vowels, based on the vowels present in the listener's dialect (with or without the /ɑ, ɔ/ merger)?
  - Perceptual shifts generalize to vowels sharing phonological features in the domain of manipulation
  - The shift in the /ɪ-ɛ/ boundary is extended to /ʊ-ɔ/ only for speakers who have /ɔ/ rather than /ɑ/ in these items
- Interpretation?
  - Listeners' generalizations are constrained by the vowels' features in their own dialect
  - /ɑ/ has different features than /ɔ/ even when they are not contrasted

Task design

- Participants (256 English speakers) completed a two-phase perceptual learning task:
  1. **Exposure**: Consonant decisions between response options with /ɪ/ (e.g. *ship-chip*), with manipulated F1
  2. **Testing**: Vowel decisions (e.g. *pit-pet*) to test category boundaries, using vowels manipulated on a continuum.
- Two groups of participants: half from regions with distinct /ɑ, ɔ/, half with merged /ɑ, ɔ/

Exposure Phase

In the exposure phase:

- Consonant decisions between response options with /ɪ/ (e.g. *ship-chip*) – 36 items
- Across participants, two different formant manipulations; half heard these items with raised F1 and half heard them with lowered F1
- This exposure serves to shift the expected F1 of /ɪ/ (potentially also causing parallel shifts in other vowels)

Exposure Phase Stimuli

- F1 was manipulated with the Vocal Toolkit in Praat
- Stimuli in the two conditions differed only in F1, because they were made from the same recordings
  - Mean F1 in the raised F1 condition: 640 Hz
  - Mean F1 in the lowered F1 condition: 425 Hz
- Manipulations were based on 2/3 of the distance between the naturally produced F1 of /ɪ/ and /ɛ/ for the raised F1 and between F1 of /ɪ/ and /i/ for the lowered F1, calculated in Bark

Test Phase

In the test phase:

- Participants heard words with vowels manipulated along an F1 continuum and chose between response options differing in the vowel (e.g. *pit*, *pet*)
- Seven vowel contrasts differing in height or tenseness, each reflected in 4 word pairs
  - /ɪ-ɛ/ (e.g. *pit*, *pet*)
  - /ɛ-æ/ (e.g. *said*, *sad*)
  - /ei-ɛ/ (e.g. *chase*, *chess*)
  - /ɪ-ei/ (e.g. *fit*, *fate*)
  - /i-ei/ (e.g. *beak*, *bake*)
  - /i-ɪ/ (e.g. *seep*, *sip*)
  - /ʊ-ɔ/ (e.g. *foot*, *fought*)
- This phase tests listeners' new vowel boundaries

Test Phase Stimuli

In the test phase:

- All words had four F1 manipulations: F1 matching the speaker's mean F1 for each of the two vowel qualities, and two intermediate values, equally spaced in the Bark scale
- No other characteristics were altered
- Both vowels for each contrast were used to create stimuli, producing two continua (eight items): e.g. *pit, pet* items included four manipulations made from *pit* and four made from *pet*

Hypotheses

1. **Hypothesis A: Normalization**. Shifts are based on the speaker's assumed vowel space. Predicts that exposure to raised F1 in /ɪ/ will increase the F1 category boundary for *all other vowels*, with no effect of dialect.
2. **Hypothesis B: Contrast preservation.** Chain shifts occur to preserve contrasts. Predicts that exposure to raised F1 in /ɪ/ will increase the F1 category boundary for other *front vowels*, which should not be sensitive to the system of back vowels.
3. **Hypothesis C: Feature-based shifts.** Shifts are based on featural parallels. Predicts that exposure to raised F1 in /ɪ/ will increase the F1 category boundary for other *high vowels*, which is likely to be sensitive to dialect-specific vowel systems.

Statistical analysis

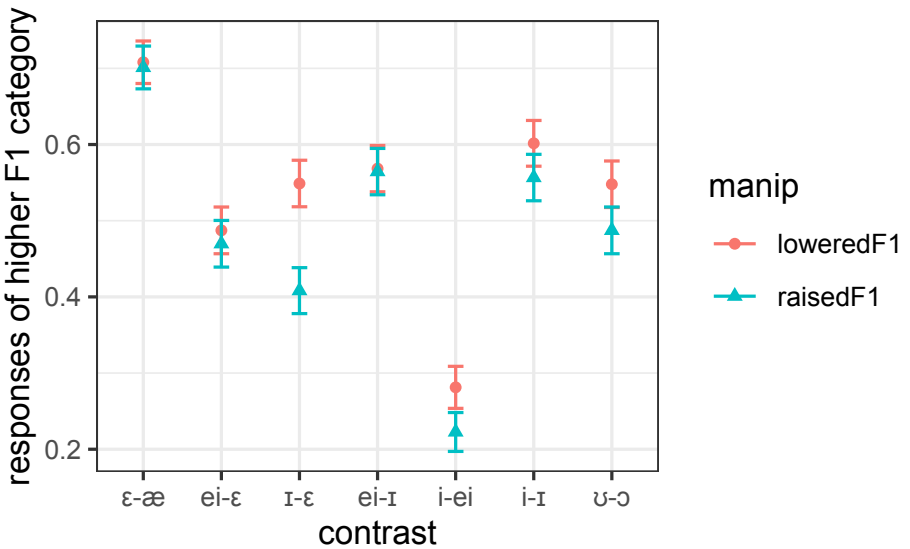
Statistical results are from a logistic mixed effects model for the odds of each response in these decisions.

To compare across vowel contrasts, responses were coded as being the category with the higher F1 or lower F1.

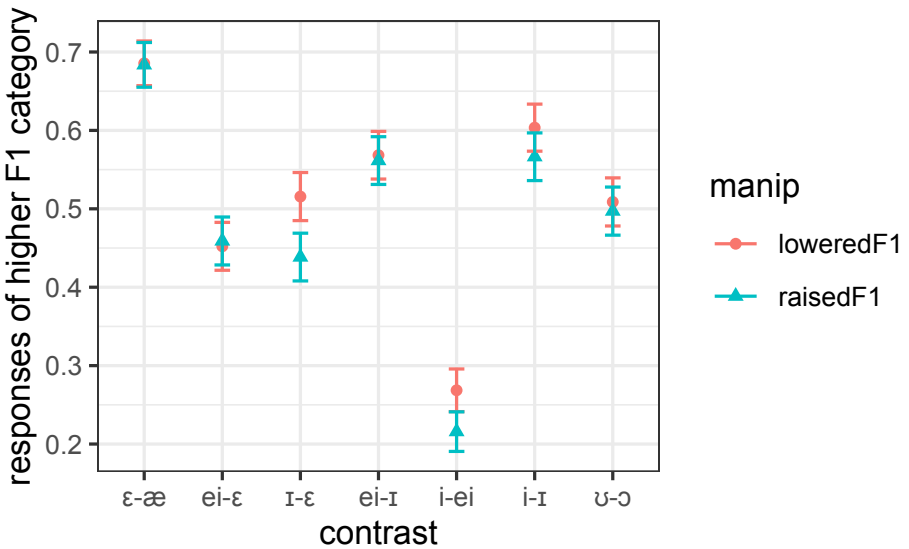
The fixed effects were vowel contrast; F1 value of the stimulus; quality of the vowel in the base recording; and the interaction between manipulation condition, vowel contrast, and dialect.

There was a random intercept for participant and for word pair.

Proportion of responses of the vowel with the higher F1, by manipulation, contrast, and dialect



Proportion of responses of the vowel with higher F1, participants with **distinct** /ɑ, ɔ/



Proportion of responses of the vowel with higher F1, participants with **merged** /ɑ, ɔ/

The different effect of manipulation on the /ʊ-ɔ/ contrast between the two dialects is significant ($\beta = 0.365$, z = 2.0, p = 0.0458)

Logistic model for responses of the vowel with the higher F1, participants with **distinct** /ɑ, ɔ/

	Estimate	Std. Error	z value	p value
(Intercept)	0.219	0.341	0.643	0.52
Contrast /ɛ-æ/	1.26	0.475	2.65	0.00813
Contrast /ei-ɛ/	-0.47	0.475	-0.99	0.322
Contrast /ɪ-ei/	0.147	0.474	0.311	0.756
Contrast /i-ei/	-2.08	0.476	-4.38	< 0.001
Contrast /i-ɪ/	0.398	0.474	0.838	0.402
Contrast /ʊ-ɔ/	-0.00954	0.474	-0.02	0.984
F1 step	0.764	0.0233	32.9	< 0.001
BaseVowel Lower F1 Category	-3.52	0.0579	-60.7	< 0.001
Manip RaisedF1 : Contrast /ɪ-ɛ/	**-1.07**	**0.133**	**-8.03**	< 0.001
Manip RaisedF1 : Contrast /ɛ-æ/	-0.0564	0.138	-0.409	0.682
Manip RaisedF1 : Contrast /ei-ɛ/	-0.131	0.132	-0.991	0.322
Manip RaisedF1 : Contrast /ɪ-ei/	-0.0303	0.132	-0.23	0.818
Manip RaisedF1 : Contrast /i-ei/	**-0.508**	**0.14**	**-3.62**	< 0.001
Manip RaisedF1 : Contrast /i-ɪ/	**-0.34**	**0.132**	**-2.58**	**0.00982**
Manip RaisedF1 : Contrast /ʊ-ɔ/	**-0.456**	**0.132**	**-3.46**	< 0.001

Intercept: Contrast /ɪ-ɛ/, Manip = lowered F1, BaseVowel = higher F1 category

Logistic model for responses of the vowel with the higher F1, participants with **merged** /ɑ, ɔ/

	Estimate	Std. Error	z value	p value
(Intercept)	-0.0945	0.324	-0.291	0.771
Contrast /ɛ-æ/	1.31	0.452	2.91	0.00363
Contrast /ei-ɛ/	-0.476	0.451	-1.06	0.291
Contrast /ɪ-ei/	0.393	0.451	0.872	0.383
Contrast /i-ei/	-1.91	0.453	-4.22	< 0.001
Contrast /i-ɪ/	0.656	0.451	1.45	0.146
Contrast /ʊ-ɔ/	-0.0516	0.451	-0.114	0.909
F1 step	0.779	0.0232	33.6	< 0.001
BaseVowel Lower F1 Category	-3.47	0.0572	-60.7	< 0.001
Manip RaisedF1 : Contrast /ɪ-ɛ/	**-0.576**	**0.129**	**-4.46**	$\mathbf{< 0.001}$
Manip RaisedF1 : Contrast /ɛ-æ/	-0.0151	0.134	-0.112	0.911
Manip RaisedF1 : Contrast /ei-ɛ/	0.0517	0.129	0.401	0.688
Manip RaisedF1 : Contrast /ɪ-ei/	-0.0512	0.129	-0.397	0.691
Manip RaisedF1 : Contrast /i-ei/	**-0.462**	**0.139**	**-3.33**	$\mathbf{< 0.001}$
Manip RaisedF1 : Contrast /i-ɪ/	**-0.278**	**0.129**	**-2.16**	**0.0312**
Manip RaisedF1 : Contrast /ʊ-ɔ/	-0.0871	0.129	-0.678	0.498

Intercept: Contrast /ɪ-ɛ/, Manip = lowered F1, BaseVowel = higher F1 category

Generalization

- Phonetic shifts are generalized based on shared phonological features – here, the realization of [+high]
- For both groups of participants, the high vs mid boundary between /ɪ/ and /ɛ/ was shifted based on exposure to raised or lowered F1 in /ɪ/, and this shift was extended to the high vs mid contrast /i-ei/

How does generalization proceed?

- The exposure vowel is shifted based on the characteristics of the exposure stimuli
- Then the shift is extended at the feature level; smaller in extent because it is not directly reinforced by the exposure stimuli, resulting in an effect of manipulation on the /i-ɪ/ contrast

Dialect-specific patterns

- Generalizations are sensitive to whether contrasts are phonologically parallel in the speaker's dialect
- For speakers with a distinct /ɔ/, it differs from /ʊ/ just in height, and the shift was extended to this high vs mid contrast too
- But for speakers with merged /ɔ/ and /ɑ/ to /ɑ/, it differs from /ʊ/ in several dimensions, and shifting /ɪ/ did not produce a parallel shift in the boundary between /ʊ/ and /ɑ/

Phonological implications

- Phonetic details are associated with sounds at the feature level
- Even though speakers can map across dialects, their representations can differ
- Different phonological organization is reflected in how shifts are generalized

References

Chládková, K., Podlipskỳ, V. J., & Chionidou, A. (2017). Perceptual adaptation of vowels generalizes across the phonology and does not require local context. *Journal of Experimental Psychology: Human Perception and Performance, 43(2)*, 414–427.

Keating, P. (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoge*d (pp. 115–131). Academic Press.

Kraljic, T., & Samuel, A. (2006). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51(2)*, 141–178.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29(1)*, 98–104.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science, 32(3)*, 543–562.

Lieberman, P. (1970). Towards a unified phonetic theory. *Linguistic Inquiry, 1*, 307–322.

Mitterer, H. (2006). Is vowel normalization independent of lexical processing? *Phonetica, 63(4)*, 209–229.

Ross, J. P., Lilley, K. D., Clopper, C. G., Pardo, J. S., & Levi, S. V. (2021). Effects of dialect-specific features and familiarity on cross-dialect phonetic convergence. *Journal of Phonetics, 86*, Article 101041.