Different Speech Styles Involve Distinct Auditory-Phonetic Planning Targets

Will Chih-Chao Chang[1], Srikantan Nagarajan[2], John Houde[2], Connor Mayer[1] & Gregory Hickok[1]

University of California, Irvine[1], University of California, San Francisco[2]

Introduction

- In speech production, speakers use auditory targets to guide motor speech planning [1-5].
- An auditory target represents an abstract sensory goal, not tied to within-category variability across utterances [6].
- However, within-category phonetic variation can be systematic. e.g., more peripheralized formant values and longer durations in clear than in casual speech [7-8].
- Systematic phonetic variation indicates that there are levels of speech planning operating within linguistic categories.

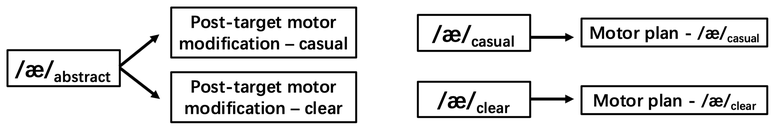
Research Questions

◆ What is the nature of auditory targets—abstract phonological categories or fine-grained, sub-categorical codes?

→ Are there distinct auditory targets for casual vs. clear speech?

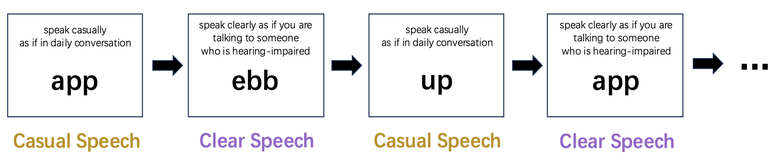
abstract, categorical target

fine-grained, sub-categorical target



Method

- Participant: 30 Native English Speakers (age 18-41, *M*=22.2)
- Task: Producing English monosyllabic vowel-initial words “app” (/æ/), “ebb” (/ɛ/), and “up” (/ʌ/) in casual vs. clear speech styles
- Design: 3 words*10 reps*3 blocks*2 speech styles = 180 trials

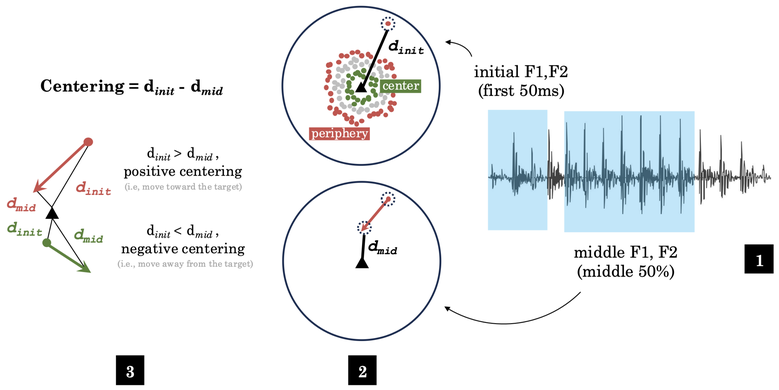


- Key examination: Centering effect [6,9]:

Initially off-target productions were adjusted toward the prototypical values of an intended sound category over time, reflecting error correction via feedback control.

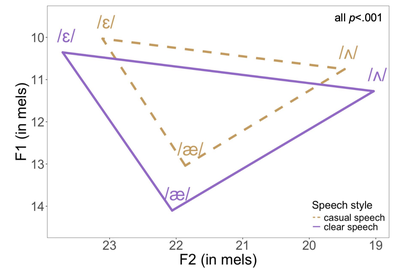
- Data analysis:

- Mean F1/F2 (mels) in initial and middle windows z-scored by speaker/vowel; medians defined as speech targets per window.
- Euclidean distances to speech targets (d_{init}, d_{mid}) computed for each production; productions classified by d_{init} as **center** (closest 1/3) or **periphery** (farthest 1/3); Centering = $d_{init} - d_{mid}$.

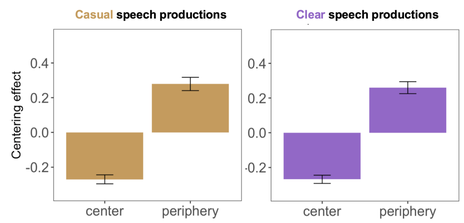


Results

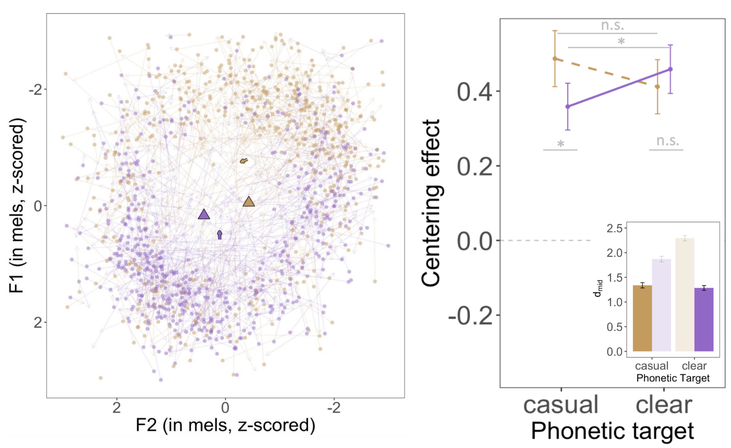
- Vowel space area: **clear > casual speech**, with significant F1/F2 differences for all vowels.



- Centering effect: **periphery** (positive) **> center** (negative) productions in both casual[6,9] and clear speech.

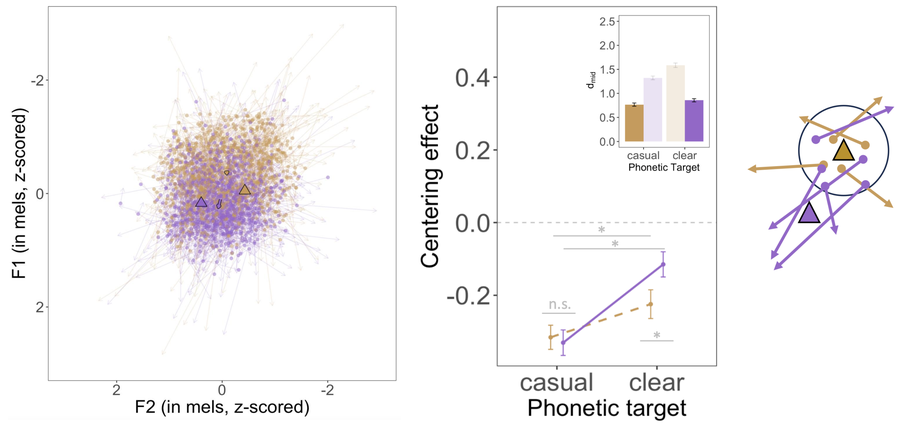


A. Periphery productions to both speech targets



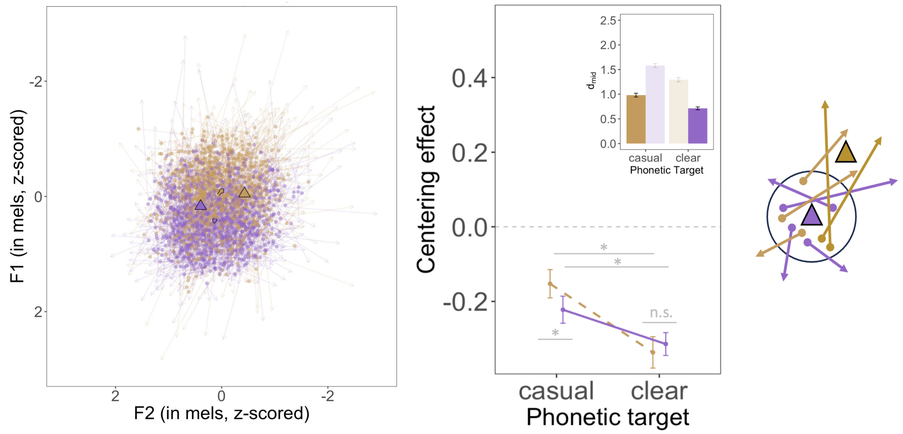
- Centering: similar to both targets in casual speech, but shifted more toward the clear target in clear speech.
- d_{mid}: ended up closer to the congruent targets for both styles.

B. Center productions to the casual speech target



- Centering: less drift from clear target for clear speech, likely because some movements consistently toward it reduce overall drift.
- d_{mid}: ended up closer to the congruent targets for both styles.

C. Center productions to the clear speech target



- Centering: less drift from casual target for casual speech, same reasoning.
- d_{mid}: ended up closer to the congruent targets for both styles.

Summary

- Auditory targets in speech planning involve sub-categorical codes that are used in real time to detect and correct speech utterances.
- Caveat: (1) centering may be driven by sub-categorical somatosensory targets[9], and (2) abstract targets may be involved in higher-level planning.

Acknowledgement

We would like to extend our gratitude to the research assistants in the UC Irvine Auditory & Language Neuroscience Lab, Lydia Cheng and Cheng-Yu Liu, for their help with data collection and annotation.

Selected Reference

[1] Houde & Nagarajan (2011). *Frontiers in human neuroscience*. **[2]** Hickok et al. (2011). *Neuron*. **[3]** Hickok (2012). *Nature reviews neuroscience*. **[4]** Tourville & Guenther (2011). *Language and cognitive processes*. **[5]** Guenther (2016). *The MIT Press*. **[6]** Niziolek et al. (2013). *Journal of Neuroscience*. **[7]** Smiljanić & Bradlow, A. R. (2005). *The Journal of the Acoustical Society of America*. **[8]** Leung et al. (2016). *The Journal of the Acoustical Society of America*.**[9]** Niziolek et al. (2015). *ICPhS*.