The relationship between production variability and perceptual acuity in explicit sensorimotor learning for speech

Hung-Shao Cheng[1], Caroline A. Niziolek[2], Adam Buchwald[1], Tara McAllister[1]
[1]*New York University, Department of Communicative Sciences & Disorders*
[2]*University of Wisconsin-Madison, Department of Communication Sciences & Disorders*

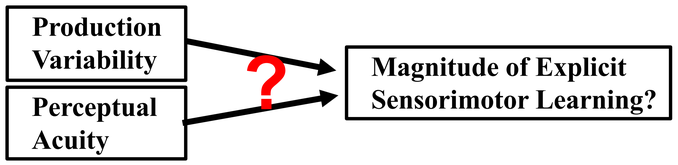
Introduction

What factors influence sensorimotor learning of speech?

- **Production variability**: ↓variability in French /o/ (L1) ⇒ ↑accuracy in Danish /ɔ/[1]
- **Perceptual acuity:** ↑acuity for /ɛ/ (AX task) ⇒ ↑ adaptation to auditory feedback perturbation[2]

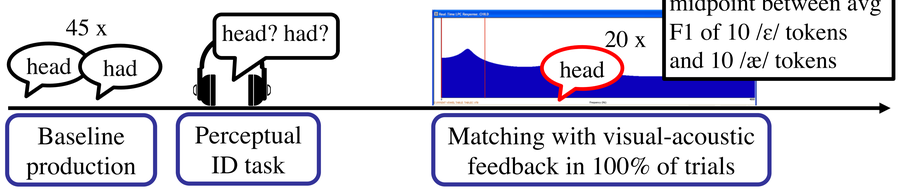
Do predictors differ in explicit versus implicit learning tasks?

- Matching a formant target in visual-acoustic biofeedback is an explicit learning task used clinically[3] and for L2 learning[1]
- Variable outcomes ⇒ need to understand individual predictors of response



Methods

- Data from Klaus et al. (2019): 37 female speakers of American English (M = 22.25 yr, SD = 3.56 yr) engaged in a task of matching a visually shifted F1 target[4]
- Data from 32 subjects were used after exclusion of outliers
- Procedure:

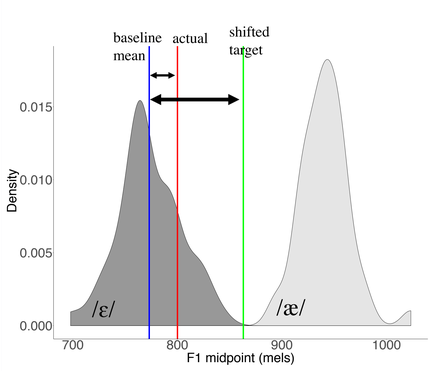


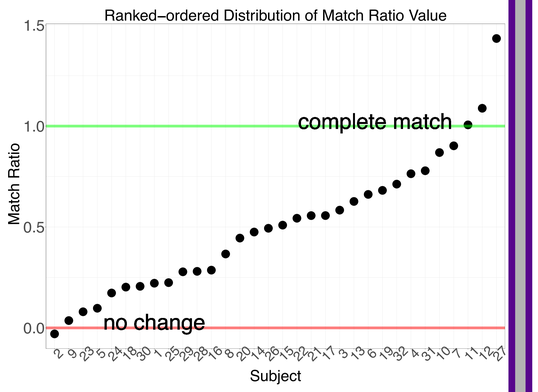
shifted F1 target = midpoint between avg F1 of 10 /ɛ/ tokens and 10 /æ/ tokens

Match Ratio as Dependent Variable

- Midpoint F1 (mel) of /ɛ/ tokens produced in the matching phase were used to calculate distance shifted as a percentage of distance from the target at baseline.

- Calculation: $\frac{F1_{actual} - F1_{baseline\ mean}}{F1_{target} - F1_{baseline\ mean}}$; higher value indicates better matching.





Analyses & Predictions

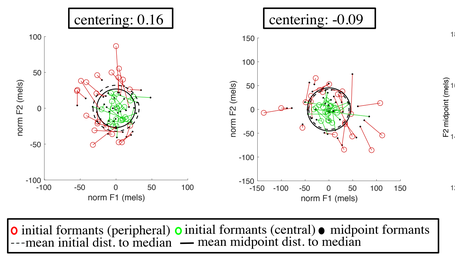
Speakers with more diffuse auditory targets → more willing to match a visually presented shifted target because those shifts are less likely to cross the phoneme boundary

Production Variability Measures

Centering Magnitude[5]:

- Correction from onset to mid vowel
- Normalized by average initial variability within each speaker
  - ➢ Larger centering (mels) suggests a smaller target in auditory space

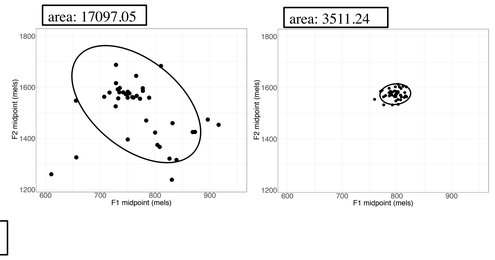
Prediction: smaller centering, higher match ratio



Area of Ellipse[6]:

- 95% CI in F1/F2 space (mels^2) for each speaker
  - ➢ Smaller area → less variable production (smaller auditory target)

Prediction: larger area, higher match ratio

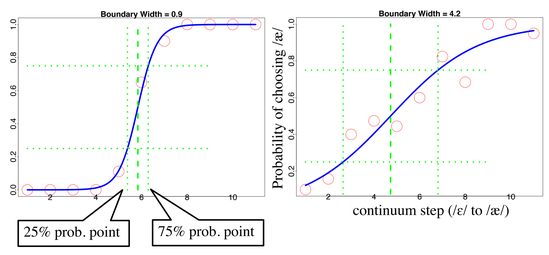


Perception Acuity Measures

Perceptual Acuity[7]:

- Identification of an 11-step continuum from /ɛ/ to /æ/
- Acuity measured as boundary width (75%-25% probability point in fitted logistic function)
  - ➢ smaller width → higher acuity

Prediction: lower acuity (i.e., larger width), higher match ratio

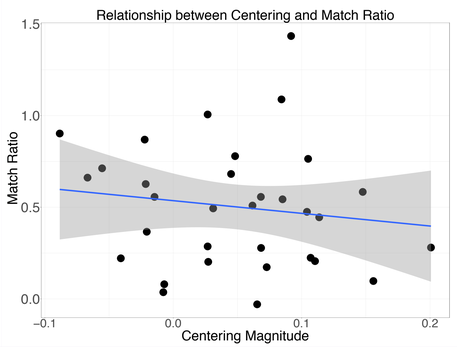


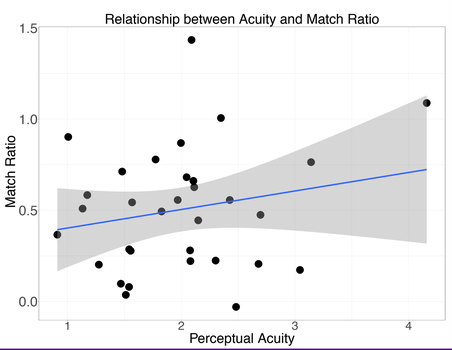
Results

Do production variability and perceptual acuity account for performance in the visual formant-matching task (match ratio)?

- **Which production variability measure better accounts for match ratio?**
  - ➢ **Two linear regression models, compared via AIC/BIC:**
  1. **Match ratio ~ centering + perceptual acuity**
  2. Match ratio ~ area of ellipse + perceptual acuity

- **Findings from Model 1, favored by AIC/BIC:**
  - ➢ Overall model fit was not significant ($F(2, 29) = 1.37$, $p = 0.27$, $R^2 = 0.09$)
  - ➢ Neither centering magnitude ($\beta = -1.06$, SE = 0.9, $p = 0.25$) nor perceptual acuity ($\beta = 0.13$, SE = 0.09, $p = 0.16$) was a significant predictor of match ratio
  - ➢ However, the sign of the coefficient of each predictor aligned with hypothesis





Discussion

- Performance in the formant-matching task (match ratio) was not significantly associated with either production variability or perceptual acuity.
  - ➢ Overall model fit was not significant
    - Suggests that other factors (e.g., somatosensory acuity, phonological awareness[6]) should be considered as predictors of matching performance.
  - ➢ However, the direction of the association of predictors with match ratio accorded with our hypotheses
- Centering was the preferred measure for production variability, based on AIC/BIC
  - ➢ Combines elements of both production variability and auditory acuity
- Future directions:
  - ➢ Provide online auditory feedback of speakers' own production during the matching task
    - Increase the reliance on using auditory feedback control mechanism
  - ➢ Continue to evaluate explicit as well as implicit sensorimotor learning of speech skills

References:
[1] Kartushina, N. et al. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics*, *57*, 21-39.
[2] Villacorta, V. M. et al. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306-2319.
[3] McAllister Byun, T., & Campbell, H. (2016). Differential effects of visual-acoustic biofeedback intervention for residual speech errors. *Frontiers in Human Neuroscience*, *10*, 567.
[4] Klaus, A. et al (2019). Can perceptual training alter the effect of visual biofeedback in speech-motor learning?. *The Journal of the Acoustical Society of America*, *145*(2), 805-817.
[5] Niziolek, C. A. et al. (2013). What does motor efference copy represent? Evidence from speech production. *Journal of Neuroscience*, *33*(41), 16110-16116.
[6] Li, J. J. et al. (2019). Individual predictors of response to biofeedback training for second-language production. *The Journal of the Acoustical Society of America*, *146*(6), 4625-4643.
[7] McAllister Byun, T., & Tiede, M. (2017). Perception-production relations in later development of American English rhotics. *PloS one*, *12*(2).

hscheng@nyu.edu