Individual differences in speech perception: Exploring cue weighting, categorization gradiency, and cognitive control

Hyoju Kim[1], Jieun Lee[2], Tzu-Hsuan Yang[2], and Phoebe Evans[2]

[1]University of Iowa (USA), [2]University of Kansas (USA)

Introduction. Listeners generally categorize speech sounds in a gradient manner [1]. However, this gradient pattern is not consistent across individuals, with some demonstrating a more categorical performance and others exhibiting a more gradient profile [2-4]. It remains uncertain whether such differences are consistent properties of individuals or vary across different types of phonological contrasts [4]. Additionally, despite recent investigations into the role of cognitive ability in shaping individual differences, its role is still inconclusive, with inconsistent results across studies and the effect size being marginal and contingent on the task type. This study further examines individual differences in speech perception, focusing on perceptual cue reliance, categorization gradiency, and cognitive control. We aim to better understand how these factors are intertwined within and across phonological contrasts. Specifically, we examined a) the dynamics of primary and secondary cue use within and across contrasts; b) the potential correlation between individuals' categorization gradiency and their reliance on secondary cues; c) the consistency of individuals' categorization gradiency and cue reliance across contrasts, and d) the associations between speech categorization and inhibitory control, exploring potential task dependencies.

Methods. Native English listeners (n=49) completed four speech perception tasks: a Visual Analog Scaling task (VAS: measurement of categorization gradiency) for stress contrasts and stop voicing contrasts, as well as a cue-weighting speech perception experiment (CW: measurement of cue reliance) for the same contrasts. The stress contrast used a stress minimal pair, *DEsert* vs. *deSSERT*, with a continuum spanning seven equidistance steps of vowel quality and pitch (step 1 being *DEsert*) [5]. The voicing contrast employed the *deer-tear*, where the continuum was manipulated into seven equidistance steps of voice onset time (VOT) and onset pitch of the following vowel (step 1 being *deer*). Each task comprised 147 trials (49 stimuli with 3 repetitions). In the VAS trials, participants heard auditory stimuli and were asked to click on a point on a horizontal line with the two endpoint words displayed at either extremity of the line. The CW task, on the other hand, was a two-alternative forced-choice task. Additionally, participants completed five inhibitory control tasks: go/no-go, flanker, spatial Stroop, color Stroop, and stop signal tasks.

Results. We adopted the rotated logistic function [2,6] to quantify listeners' categorization gradiency from the VAS response. CW responses were fitted into a logit mixed-effects model with cues as random slopes for each subject. We then quantified individuals' reliance on each cue by extracting random-slope coefficients from the model [5]. We found a significant trading relationship between primary and secondary cues for the stress contrast (Fig. 1, top-right). The relationship exhibited an opposing pattern in the stop voicing contrast (Fig. 1, bottom-right). We also found a marginal association between individuals' secondary cue reliance and their categorization gradiency for the stress contrast ($r = -0.35$, $p = .06$), with more gradient listeners showing greater reliance on secondary cues. However, this association did not hold in the voicing contrast. Notably, listeners who relied more on the pitch cue in the stress contrast relied more on the same cue in the voicing contrast ($r = 0.41$, $p < .05$) (Fig. 2, middle panel). It is also found that listeners' categorization gradiency is consistent across contrast types ($r = 0.47$, $p < .01$) (Fig. 2, right panel). None of the inhibitory control scores were related to listeners' categorization gradiency across contrast types.

Discussion. The results suggest that the dynamics of primary and secondary cue reliance may vary depending on the type of phonological contrasts. Gradient listeners tend to rely more on secondary cues, but this tendency is limited to the stress contrast. Crucially, our findings underscore the consistency in listeners' perceptual reliance on a specific cue (pitch) and their categorization gradiency across different types of phonological contrasts. Our findings also suggest that the processing of lower-level acoustic information and its integration for categorization may not be strongly associated with higher-level inhibitory control.

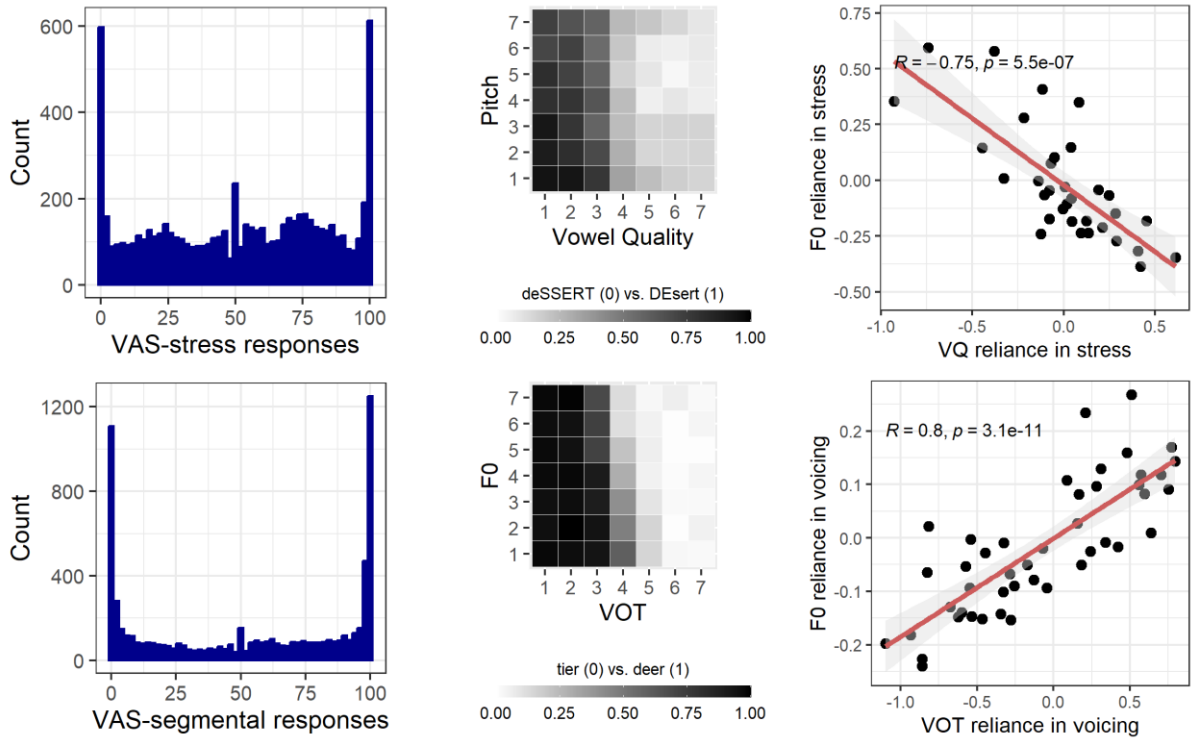

Figure 1. Listeners' performance on the tasks for the stress contrast (top panels) and the voicing contrast (bottom panels); Left-most panels show their responses to the VAS task, middle panels the cue-weighting task, and right-most panels show the relationship between the primary and secondary cues.

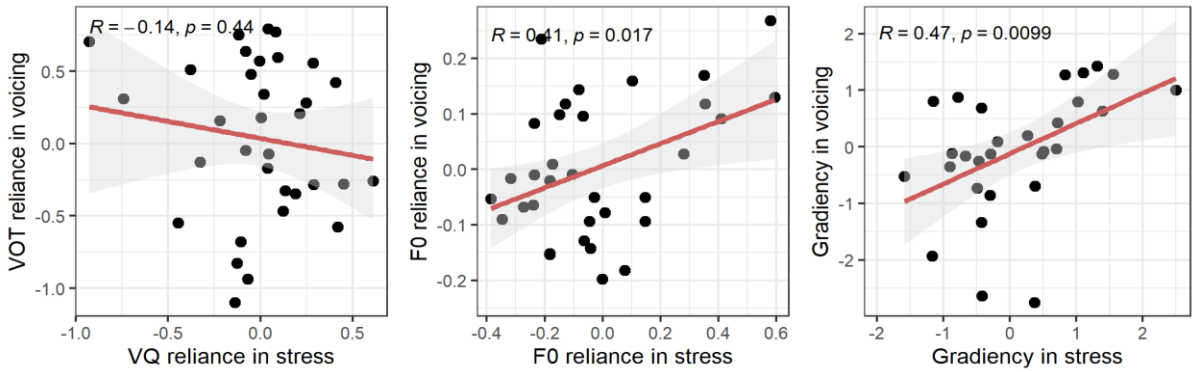

Figure 2. The relationship regarding the primary cues (left-most), the secondary cues (middle), and categorization gradiency (right-most) across contrast types.

References

[1] McMurray, B. (2022). The myth of categorical perception. *JASA*, *152*(6), 3819-3842.

[2] Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J., & McMurray, B. (2017). Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *J Exp Psychol Hum Percept Perform*, *43*(9), 1594.

[3] Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *J Phon*, *59*, 40-57.

[4] Kapnoula, E. C., Edwards, J., & McMurray, B. (2021). Gradient activation of speech categories facilitates listeners' recovery from lexical garden paths, but not perception of speech-in-noise. *J Exp Psychol Hum Percept Perform*, *47*(4), 578.

[5] Tremblay, A., Broersma, M., Zeng, Y., Kim, H., Lee, J., & Shin, S. (2021). Dutch listeners' perception of English lexical stress: A cue-weighting approach. *JASA*, *149*(6), 3703-3714.

[6] McMurray, B. (2017). *Nonlinear curvefitting for Psycholinguistics* (Version 42). Retrieved from https://osf.io/4atgv/