

# How vowel variability relates to vowel perception

Nhung Nguyen<sup>1</sup>, Jason A. Shaw<sup>1,2</sup>, Catherine T. Best<sup>1,2</sup>, and Michael D. Tyler<sup>1,3</sup>

<sup>1</sup> The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia

<sup>2</sup> School of Humanities and Communication Arts, Western Sydney University, Australia

<sup>3</sup> School of Social Sciences and Psychology, Western Sydney University, Australia  
nhung.nguyen@westernsydney.edu.au

Vowel variability is usually estimated by how spread-out the tokens of a vowel category are in F1-F2 vowel space. However, the spread of vowel tokens does not always serve as a reliable indicator of perceptual accuracy (Hillenbrand et al., 1995). Another method of estimating vowel variability is to take the magnitude of the formant mean into account (Kent, 1976; Lee et al., 1999; Nguyen & Shaw, 2014). All else being equal, the degree of variability of a formant measurement is systematically related to the magnitude of the mean of that measure. For example, vowels with high mean F1 (or F2) values are also more variable on F1 (or F2) than vowels with a low mean F1 (or F2) (Nguyen & Shaw, 2014). One way to quantify vowel variability that takes the influence of the mean measurement into account is to regress vowel means on vowel standard deviations (SDs), then consider the residuals: if a residual is positive, the vowel is variable in relation to the magnitude of the mean; if it is negative, the vowel is stable relative to its mean (Nguyen & Shaw, 2014). The relationship between this residual method of estimating vowel variability and perceptual accuracy, however, has not been evaluated. In our study, we evaluated how vowel variability relates to vowel perception in a vowel categorization task, and which index of variability (i.e., residuals versus SDs) better predicts accuracy. We hypothesized that vowel variability would correlate with categorization accuracy, but that the correlation may be different in different parts of the vowel space.

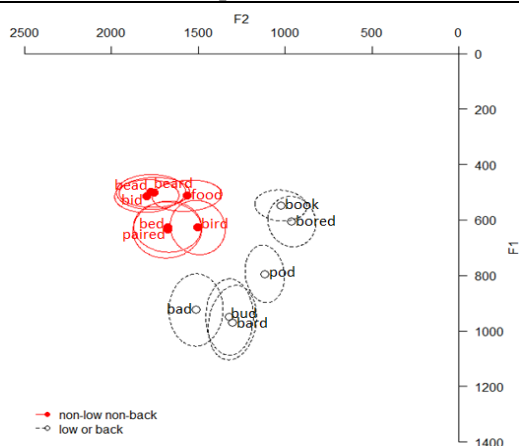
To explore how vowel variability relates to vowel perception, we considered two factors: vowel variability and location in the vowel space. Vowel variability was estimated from the means and SDs of 13 Australian English monophthongs provided in a large /hVd/ corpus of Australian English (Cox, 2006). Figure 1 shows the spread of the vowels in F1-F2 space expressed in Mel units. Note the high degree of overlap, or *crowding*, between the vowels in the non-low non-back area of the vowel space. We fitted regression lines to the means and SDs of F1 and F2 values, expressed on the Mel scale, and then added F1 and F2 residuals for each vowel to provide a single measure of variability. For the SD method, we also added F1 and F2 SDs (also in Mel) for each vowel and rescaled (centering the mean on 0) to make them directly comparable to the residuals. We then evaluated how these two indices of variability compared in predicting vowel categorization. 64 participants from the Greater Sydney community listened to 13 Australian English monophthongs (4 tokens per vowel) embedded in /hVdə/ nonce words produced by a female speaker in her 20s from Western Sydney. They were asked to choose a reference word that contained the same vowel as that in the non-word they just heard.

Binomial mixed effects models were fitted to the accuracy data (3,328 data points) in R and model comparisons were carried out. Participant and vowel token were included as random effects. Fixed effects were location of the vowel in the vowel space, which was a binary variable dividing the non-low non-back vowels (i.e., bead, beard, bid, bed, paired, food, and bird) from the others, and variability, expressed as either the sum of residuals (in one model) or of SDs (in the other). Using F1 and F2 variability as separate predictors also produced the same main results. Table 1 shows that, in an interaction with vowel location, the sum of residuals (model 2) led to a significant improvement in variance explained ( $\chi = 6.24$ ,  $p < .001$ ), compared to the sum of SDs (model 1) as the variability index. In model 2, the main effect of vowel variability was significant ( $\beta = -0.03$ ,  $p = .02$ ). Vowel location was a significant predictor in all the models, indicating that accuracy was systematically poorer in the non-low non-back area of the vowel space. Figure 2 shows that vowel variability, as quantified by the residual method, relates to categorization accuracy in two different ways: (1) When a vowel belongs to the non-low non-back area of the Australian English vowel space, listeners' categorization accuracy for a vowel decreases with increased variability. The non-low non-back area of the Australian English vowel space is a confusing area of the vowel space compared to the low or back area, as it involves a greater number of vowels undergoing change or marking accent differences (Cox, 1999; Harrington et al., 1997). In that confusing area of the perceptual space, vowel variability

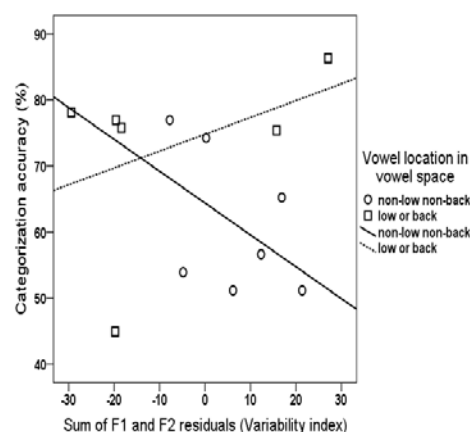
amplifies confusion. (2) Vowels in the low or back area show the opposite pattern: Variability is associated with better categorization performance in the less confusing region of vowel space. These associations are revealed most clearly when we take into account the relationship between mean formant values and their associated SDs.

**TABLE 1: Variances explained by sums of residuals versus sums of SDs.**

Model	AIC	BIC	logLik	deviance	Pr(>Chisq)
(1) with sums of SDs: Accuracy ~ (scaled) sum of SDs * vowel location + (1 Participant) + (1 Token)	3685.7	3722.4	-1836.9	3673.7	
(2) with sums of residuals: Accuracy ~ sum of residuals * vowel location + (1 Participant) + (1 Token)	3679.5	3716.1	-1833.7	3667.5	< .001



**FIGURE 1: Australian English monophthongs based on Cox (2006).**



**FIGURE 2: vowel variability by categorization accuracy for two partitions of the vowel space.**

## REFERENCES

- Cox, F. (1999). Vowel change in Australian English. *Phonetica*, 56, 1–27.
- Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics*, 26(2), 147–179.
- Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English. *Australian Journal of Linguistics*, 17(2), 155–184.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099–3111.
- Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*, 19(3), 421–447.
- Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, 105(3), 1455–1468.
- Nguyen, N., & Shaw, J. A. (2014). Why the SQUARE vowel is the most variable in Sydney. In *Proceedings of the 15th Australasian International Speech Science and Technology Conference*, 37–40.