

Phonetic convergence in Mexican Spanish: combining acoustic and perceptual assessments  
Bethany MacLeod  
Carleton University

Talkers adjust the acoustic-phonetic realization of their speech in response to being exposed to another talker, a phenomenon known as phonetic accommodation, imitation or convergence. A particular challenge in studying this phenomenon is how we assess whether convergence has taken place. One approach is to use acoustic analysis to measure particular dimensions of the two talkers' speech and determine which have changed and to what extent. However, this method does not resolve whether the observed changes are relevant in the real-world perception of convergence. Another approach is to rely on perceptual judgements from listeners to evaluate the degree of convergence. This method, however, provides no insight about which characteristics of the speech signal have changed or which listeners use to make their judgments. Understanding how listeners perceive convergence in the real world is important for developing accounts of second-dialect acquisition and sound change (Niedzielski & Giles 1996) and may also have implications for models of teaching and learning second language pronunciation (Lewandowski & Jilka 2019; Martinsen, Montgomery & Willardson 2017).

Pardo, Jordan, Mallari, Scalon & Lewandowski (2013) suggest combining both acoustic measures and a perceptual assessment of convergence to generate a more complete picture of how listeners actually make use of acoustic variation in their perception of convergence. In this combined approach, mixed-effects models determine the extent to which acoustic measures predict the perceptual assessment of convergence. A handful of studies have applied this method to analyses of convergence, but only for English (e.g. Pardo et al. 2013; Walker & Campbell-Kibler 2015). The current study applies the combined method to an exploration of phonetic convergence of Spanish vowels.

Forty-eight female native Mexican Spanish speakers participated in a speech shadowing experiment. To establish a baseline, the participants read aloud 40 bisyllabic Spanish words three times. Next, the participants shadowed one of four female native Mexican Spanish-speaking model talkers producing the same 40 words. Following the combined approach discussed above, phonetic convergence was measured via both acoustic analysis and perceptual assessment. F0, F1, F2, duration and mean intensity of the first vowel, and duration of the entire word were taken for the baseline, shadowed, and model tokens. Euclidean distance in F1-F2 space from the first vowel to model was also calculated. The difference in distance (DID: absolute value (baseline - model) - absolute value (shadowed - model)) was calculated for each of these acoustic measures.

Next, the recordings from the 48 shadower + model talker pairs were used to create 24 versions of a 4IAX perceptual discrimination experiment (Pisoni & House Lazarus 1974; Tutinetti, Whang & Escudero 2019), which a total of 87 Spanish-speaking listeners completed. In this task, listeners heard two pairs of words (XA XB), where X was always the model token and A and B were either the baseline or shadowed token from one of the production participants. The listeners' task was to decide whether the first two (XA) were more similar to each other or whether the last two (XB) were more similar to each other. The proportion of trials in which the participants chose the pair involving the shadowed token was taken to reflect the degree of convergence.

Overall, the listeners selected the pair containing the shadowed token in 53.2% of trials (significantly higher than 50%:  $p < 0.04$ ). Mixed-effects logistic regression models determined that the listeners used some, but not all, of the available acoustic measures in making their

judgements. Degree of convergence as determined by the DID measurements for Euclidean distance, F0, and word duration were significant predictors of the likelihood that the listeners would choose the pair involving the shadowed token. In contrast, the main effects of the DID measurements for F1, F2, segment duration, and intensity were not significant.

These results indicate that listeners integrate over several acoustic measures in their perception of convergence and highlight the importance of taking multiple acoustic measurements and of combining acoustic analysis with a perceptual assessment of convergence.

## References

- Lewandowski, N., & Jilka, M. (2019). Phonetic Convergence, Language Talent, Personality and Attention. *Frontiers in Communication*, 4(18), 1–19.
- Martinsen, R., Montgomery, C., & Willardson, V. (2017). The Effectiveness of Video-Based Shadowing and Tracking Pronunciation Exercises for Foreign Language Learners. *Foreign Language Annals*, 50(4), 661–680.
- Niedzielski, Nancy & Howard Giles (1996). Linguistic accommodation. In H. Goebel, P. Nelde, Z. Starý & W. Wölck (eds.), *Kontaktlinguistik – Ein internationales Handbuch zeitgenössischer Forschung*. Berlin/New York: Mouton de Gruyter. 332–342
- Pardo, J. S., Jordan, K., Mallari, R., Scalon, C., & Lewandowski, E. (2013). Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory and Language*, 69, 183–195.
- Pisoni, D. B., & House Lazarus, J. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America*, 55(2), 328–333.
- Tutinetti, A., Whang, J., & Escudero, P. (2019). Non-native vowel perception in a 4IAX task: The effects of acoustic distance. In Sasha Calhoun, Paola Escudero (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 240–244). Canberra, Australia: Australasian Speech Science and Technology Association Inc.
- Walker, A., & Campbell-Kibler, K. (2015). Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. *Frontiers in Psychology*, 6(Article 546), 1–18.