Uniformity and crosslinguistic influence in Cantonese-English bilingual stops

Khia A. Johnson & Molly Babel

While crosslinguistic influence is widespread in bilingual speech production, the nature of representation in a multilingual system remains unclear. Prior work typically examines phonetically distinct yet phonologically similar sounds, for which phonetic convergence provides evidence for crosslinguistic links (Chang, 2015). Convergence is more challenging to assess when sounds are already phonetically similar, as with English and Cantonese initial long-lag stops. Here, we leverage the articulatory uniformity framework (Chodroff & Wilson, 2017; Faytak, 2018) to assess whether bilinguals share an underlying laryngeal feature across languages. Using a Cantonese-English bilingual speech corpus (n=34; Johnson et al., 2020), we identified prevocalic word-initial /ptk/ from force-aligned transcripts refined with AutoVOT (Sonderegger & Keshet, 2012). After accounting for speech rate, there were significant correlations for mean VOT values in English (3/3 pairs: $r > 0.57$), Cantonese (2/3 pairs: $r > 0.54$), and to some extent across languages (3/9 pairs: $r > 0.55$). These moderate correlations suggest some level of uniformity but are less compelling than prior findings. Additionally, there was low adherence to the expected ordinal relationship between /ptk/ means and within-talker inconsistencies across languages. A linear mixed-effects model highlights clear VOT differences across languages but also demonstrates that talker intercepts account for substantially more VOT variation than by-talker random slopes for place of articulation or language. This analysis indicates a role for both language and individual-specific factors in accounting for the structure of VOT variation (similar to speech rate findings: Bradlow, Kim, & Blasingame, 2017) and highlights the utility of the uniformity framework.

References

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. The Journal of the Acoustical Society of America, 141(2), 886–899. doi: 10.1121/1.4976044

Chang, C. B. (2015). Determining cross-linguistic phonological similarity between segments. In E. Raimy & C. E. Cairns (Eds.), The segment in phonetics and phonology (pp. 199–217). doi: 10.1002/9781118555491.ch9

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. Journal of Phonetics, 61, 30–47. doi: 10.1016/j.wocn.2017.01.001

Faytak, M. D. (2018). Articulatory uniformity through articulatory reuse: Insights from an ultrasound study of Sūzhōu Chinese [Doctoral dissertation, University of California, Berkeley].

Johnson, K. A., Babel, M., Fong, I., & Yiu, N. (2020). SpiCE: A new open-access corpus of conversational bilingual speech in Cantonese and English. Proceedings of the 12th Language Resources and Evaluation Conference, 4089–4095.

Sonderegger, M., & Keshet, J. (2012). Automatic measurement of voice onset time using discriminative structured prediction. The Journal of the Acoustical Society of America, 132(6), 3965–3979. doi: 10.1121/1.4763995