

Language-specific /s/ acoustics for early Cantonese-English bilinguals?

Molly Babel, Victor Wong, Sabrina Luk, Kai Fong, and Ragul Loganathan

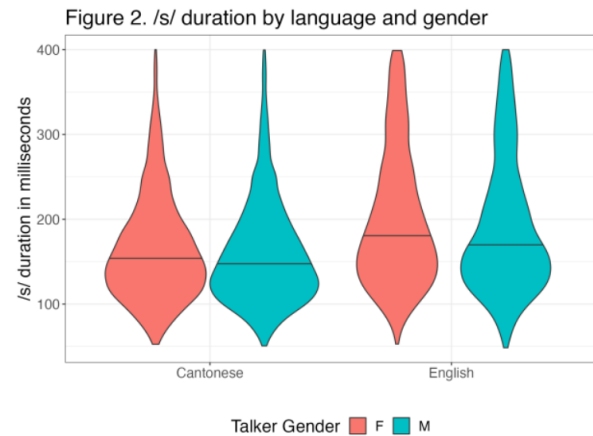
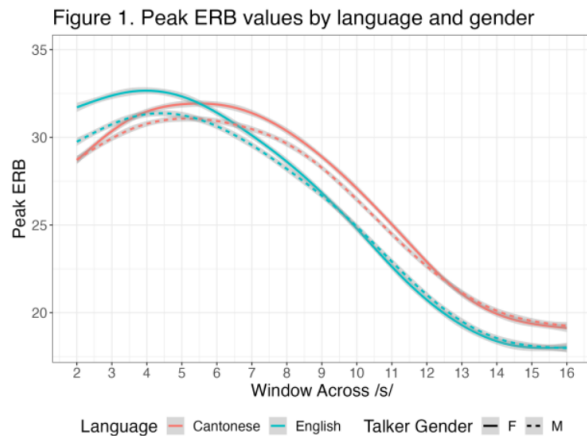
University of British Columbia

Despite sharing an IPA label, what is transcribed as an /s/ across languages can differ in its phonetic realization (e.g., Reidy, 2016). These language-specific phonetics are observed in, at least, older comparisons of /s/ in Cantonese and English. The anterior sibilant fricative in Cantonese has been variably described as [s], [ɕ], or [ʃ] (Wang, 1937; Chao, 1947). More recently, Yu (2016) characterizes spectral moments' trajectories of Cantonese /s/, observing that the values of Cantonese /s/ are [s]-like, but with considerable variation in acoustic realization conditioned by the following vowel environment. This makes the phonemic /s/ categories in English and Cantonese similar, though, when two languages have coronal sibilants that are transcribed as [s], this still, of course, does not guarantee their phonetic equivalence.

Due to the colonial history of Hong Kong, a Cantonese-speaking homeland, there is a long-standing history of Cantonese-English bilingualism. Simonet (2016) suggests that when a bilingual's languages present similar phones, it makes observable cross-linguistic influence more likely, though see Casillas (2021) for evidence that composite categories in bilinguals are rarely observed for similar, but not identical, stop inventories. The current work queries this space, building on an initial analysis by Johnson and Babel (2019). Using the interview portion of the SpiCE corpus (Johnson et al., 2021), a corpus of Cantonese and English from 34 Cantonese-English early bilinguals, we describe the acoustics of Cantonese and English /s/ to characterize the within- and across-language patterns on a group-level and for individuals.

The SpiCE corpus has been force aligned by the MFA (McAuliffe et al., 2017), but to confidently characterize trajectory differences across languages, a crucial aspect of language-specific phonetics, all word-initial pre-vocalic /s/ boundaries were manually corrected to denote the onset and offset of the fricatives (Cantonese = 4,576 tokens, English = 4,190 tokens). After being high-pass filtered at 1000 Hz, fricatives were acoustically characterized as a trajectory of peak ERB_N in 17 equally distant points based on a script developed by Reidy (2016). Peak ERB_N measures the frequencies with the greatest excitation on a multitaper psychoacoustic spectrum, and is an auditorily-informed acoustic description of fricatives (cf. Shadle, 2023).

Data analysis is ongoing, but we present a snapshot of the empirical data that we will use to answer our research question regarding the existence and potential maintenance of English-and Cantonese-specific /s/ on the group level and at the individual level. Figure 1 presents the peak ERB values across the fricative by language and talker gender. These data suggest similar peak ERB values are ultimately obtained in both languages' /s/ production, but that the timing of reaching the highest peak ERB value in Cantonese occurs later in the fricative. The duration data are presented in Figure 2, where the line represents the median value; while these values do not yet consider speech rate, they currently suggest that English /s/ may have a longer duration than Cantonese /s/. The peak ERB data will be analyzed with a functional PCA, with the identified components used to create Mahalanobis distances to quantify the by-talker distances between Cantonese and English /s/. Regardless of the outcome of the research question, these data will ultimately inform our understanding of language-specific phonetics in bilingual speech using natural spontaneous speech.



- Casillas, J. V. (2021). Interlingual interactions elicit performance mismatches not “Compromise” categories in early bilinguals: Evidence from meta-analysis and coronal stops. *Languages* 6(1), 9. <https://doi.org/10.3390/languages6010009>
- Chao, Y.-R. (1947). *Cantonese Primer* (Harvard University Press, Cambridge, MA).
- Johnson, K. A. (2021). "SpiCE: Speech in Cantonese and English", <https://doi.org/10.5683/SP2/MJOXP3>, Borealis.
- Johnson, K. A., & Babel, M. E. (2019). Bilingual sibilant acoustics in conversational Cantonese-English speech. *The Journal of the Acoustical Society of America*, 146(4), 2839-2839. <https://doi.org/10.1121/1.5136840>
- McAuliffe M, Socolof M, Mihuc S, Wagner M, Sonderegger, M (2017) Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Proceedings of the 18th Conference of the International Speech Communication Association*, 498–502. <https://doi.org/10.21437/Interspeech.2017-1386>
- Reidy, P. F. (2016). Spectral dynamics of sibilant fricatives are contrastive and language specific. *The Journal of the Acoustical Society of America*, 140(4), 2518-2529. <https://doi.org/10.1121/1.4964510>
- Simonet, M. (2016). The phonetics and phonology of bilingualism. In *Oxford Handbooks in Linguistics* (Online)(pp 1-23). Oxford, United Kingdom: Oxford University Press. doi:10.1093/oxfordhb/9780199935345.013.72
- Shadle, C. H. (2023). Alternatives to moments for characterizing fricatives: Reconsidering Forrest et al. (1988). *The Journal of the Acoustical Society of America*, 153(2), 1412-1426. <https://doi.org/10.1121/10.0017231>
- Wang, L. (1937). *Zhongguo yinyun-xue* (Chinese Phonology) (Shangwu, Shanghai), Vol. 2.