

Variable pathways in L2 sibilant category formation among Cantonese-English bilinguals in Hong Kong

Jonathan Havenhill

The University of Hong Kong

Many phonological patterns in Hong Kong English (HKE) straightforwardly derive from Hong Kong Cantonese (HKC), including *n-l* alternation, *th*-fronting, cluster simplification, and lexical tone [1,2]. As HKC exhibits a single series of contrastive sibilants, /s, ts, ts^h/, substitutions of [s] for English [ʃ] (and vice versa) are also reported [3,4]. Since the late 1970s, however, an allophonic split between anterior [ts, ts^h] and posterior [tɛ, tɛ^h] has emerged among younger HKC speakers, with the latter variants appearing before round vowels /y, œ, u, ɔ/ [5–7]. The presence of alveolo-palatal allophones for younger speakers but not older speakers raises questions for models of L2 phonological acquisition: SLM-r [8] and PAM-L2 [9] predict that acquisition begins with mapping L2 sounds to equivalent L1 sounds according to their phonetic similarity, with phones perceived as similar to an L1 phone produced using L1 targets. However, phonetic similarity may be defined by either acoustics or articulation [10], and cross-dialectal studies of L2 vowel perception suggest that L1 variation within the same language can lead to distinct L2 outcomes [15]. Through acoustic and articulatory analysis, this study examines how L2 English sibilants are produced by native HKC speakers whose L1 sibilant systems vary with respect to the presence or absence of discrete allophones differing by place of articulation. This study asks (1) To what extent do HKC speakers re-use L1 phonetic targets in producing L2 sibilants? (2) Are L1/L2 sibilants similar in acoustics, articulation, or both?

58 L1 HKC speakers participated in a two-part production experiment, including 37 women and 21 men across three age groups (18–29, 30–49, 50+). Five native English speakers participated as controls. HKC items included 72 disyllabic words with target onsets /s, ts, ts^h/ followed by /i, ɛ, ɐ, a, y, œ, u, ɔ/; English items included 96 trochaic words with onset /s, ʃ, tʃ, dʒ/ and vowel /i, ɪ, e, ɛ, u, ʊ, o, ɔ/. Midsagittal ultrasound tongue images were collected at 82 fps using an Articulate Instruments Micro system with an UltraFit headset-stabilized transducer. Front-view and side-view lip video were captured at 60 fps. Audio (48kHz) was recorded using a MixPre-10 II and an AKG C555L headset microphone. Articulatory data were automatically labeled with DeepLabCut [12,13] and dynamic multitaper spectra were extracted at 10% intervals with pymultitaper.

For younger speakers, L2 English /ʃ, tʃ, dʒ/ are acoustically similar to those of L1 English speakers (Fig. 1) and are produced with similar (but not identical) tongue positions as HKC [tɛ, tɛ^h] allophones, both of which are categorically distinct from HKC [s, ts, ts^h] (Fig. 2). For older speakers, who produce HKC sibilants at a single place of articulation in all vowel contexts, L2 English /ʃ, tʃ, dʒ/ are also acoustically native-like but are acoustically distinct from both their own L1 sibilants and from the L1 English baseline, reflecting new L2 category formation [8]. Older speakers thus achieve a similar acoustic output through a different articulatory route, as has been found for L2 vowels produced by speakers whose L1 lacks a direct phonetic equivalent [10,15]. For all HKC speakers, anticipatory lip rounding in both L1 and L2 is greater than for L1 English speakers; additionally, L2 English post-alveolars are consistently produced with secondary lip rounding regardless of vowel context, a gesture ostensibly associated with the following vowel in HKC but serving as a (native-like) secondary feature of the English post-alveolars for these speakers. These findings show that L1 variability within a single speech community, here age-dependent differences in allophonic patterning, influences the articulatory pathway to L2 category formation even when speakers produce similar acoustic outputs. Acoustic similarity between L1 and L2 categories does not entail articulatory similarity, which points to the need for additional articulatory data in studies of L2 phonology.

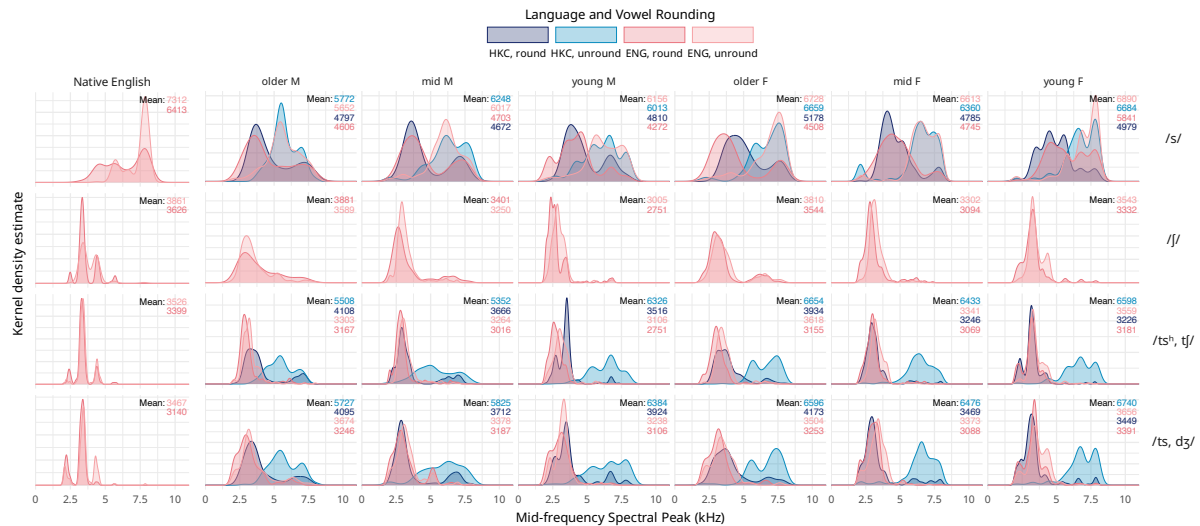


Fig. 1. Mid-frequency spectral peak (Hz) by phoneme and vowel rounding for native English and native HKC speakers

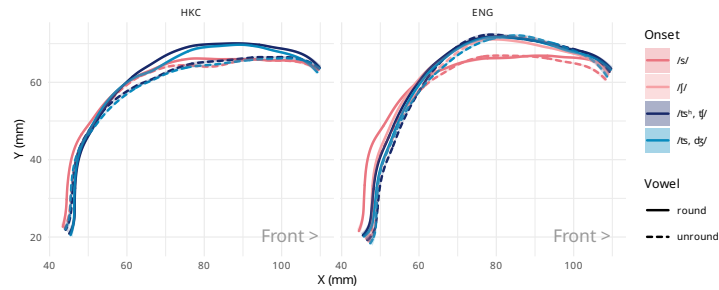


Fig. 2. Representative mid-sagittal tongue contours for young HKC speaker with rounding-conditioned affricate allophones

References

- [1] Hung, T. T. N. (2000). Towards a phonology of Hong Kong English. *World Englishes*, 19(3), 337–356.
- [2] Yeung, P. H. (2023). Contact-induced tonogenesis in Hong Kong English. *Proceedings of ICPhS 2023*.
- [3] Luke, K. K., & Richards, J. C. (1982). English in Hong Kong: Functions and status. *English World-Wide*, 3(1), 47–64.
- [4] Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language Culture and Curriculum*, 13(1), 67–85.
- [5] Cheung, K. H. (1986). *The phonology of present-day Cantonese* (Doctoral dissertation). University of London.
- [6] Chan, K.-L. (2007). *The sound change of [ts, tsʰ, s] to [tʃ, tʃʰ, ʃ] in Hong Kong Cantonese* (Master's thesis). HKUST.
- [7] Havenhill, J., Liu, M., & Li, T. W. (2023). Articulation of Cantonese and English sibilants among bilingual speakers in Hong Kong. *Proceedings of ICPhS 2023*.
- [8] Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning* (pp. 3–83). Cambridge University Press.
- [9] Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning* (pp. 13–34). John Benjamins.
- [10] Oakley, M. (2024). The articulatory and acoustic representation of second-language French vowels. *Language and Speech*, 68(4), 887–932.
- [11] McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. *Proceedings of Interspeech 2017*, 498–502.
- [12] Mathis, A., et al. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), 1281–1289.
- [13] Wrench, A., & Balch-Tomes, J. (2022). Beyond the edge: Markerless pose estimation of speech articulators from ultrasound and camera images using DeepLabCut. *Sensors*, 22, 1133.
- [14] Yu, A. C. L. (2016). Vowel-dependent variation in Cantonese /s/ from an individual-difference perspective. *Journal of the Acoustical Society of America*, 139(4), 1672–1690.
- [15] Escudero, P., & Williams, D. (2012). Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America*, 131(5).