INTERSPEECH 2021
30 August – 3 September, 2021, Brno, Czechia

Sound change in spontaneous bilingual speech: A corpus study on the Cantonese n-l merger in Cantonese-English bilinguals

Rachel Soo, Khia A. Johnson, and Molly Babel

Department of Linguistics, University of British Columbia, Canada
rachel.soo@ubc.ca, khia.johnson@ubc.ca, molly.babel@ubc.ca

Abstract

In Cantonese and several other Chinese languages, /n/ is merging with /l/. The Cantonese merger appears categorical, with /n/ becoming /l/ word-initially. This project aims to describe the status of /n/ and /l/ in bilingual Cantonese and English speech to better understand individual differences at the interface of crosslinguistic influence and sound change. We examine bilingual speech using the SpiCE corpus, composed of speech from 34 early Cantonese-English bilinguals. Acoustic measures were collected on pre-vocalic nasal and lateral onsets in both languages. If bilinguals maintain separate representations for corresponding segments across languages, smaller differences between /n/ and /l/ are predicted in Cantonese compared to English. Measures of mid-frequency spectral tilt suggest that the /n/ and /l/ contrast is robustly maintained in English, but not Cantonese. The spacing of F2-F1 suggests small differences between Cantonese /n/ and /l/, and robust differences in English. While cross-language categories appear independent, substantial individual differences exist in the data. These data contribute to the understanding of the /n/ and /l/ merger in Cantonese and other Chinese languages, in addition to providing empirical and theoretical insights into crosslinguistic influence in early bilinguals.

Index Terms: crosslinguistic influence, individual differences, sound change, Cantonese, corpus phonetics

1. Introduction

Within-community phonetic variation presents theoretical opportunities for understanding language change (e.g., identifying actuation and transmission) [1], in addition to presenting potential challenges to synchronic processing amongst listeners/speakers (e.g., do listeners map the phonetically variable pronunciations to the intended lexical meaning) [2]. Automatic speech recognition (ASR) systems and natural language processing (NLP) tools may also struggle when presented with phonetic variation—even where there is no explicit language mixing—depending on the training of the acoustic models. Cantonese is a comparatively low-resource language currently undergoing a large number of sound changes in both the tone space [3] and in the consonant series [4]. Our focus is on a sound change targeting /n/, which is produced as [l] in word-initial position [5, 6]. Specifically, we carried out a corpus study examining the acoustic productions of /n/ and /l/ in speech by Cantonese-English bilinguals in both languages.

This merging of /n/ and /l/ is observed in a number of other Chinese languages (e.g., Fuzhou, Nanjing, Chengdu, Wuhan) [7] and while speakers of many of these Chinese languages are also bilingual in Standard Mandarin—which is currently not experiencing a merger between these two consonants—the merger of /n/ and /l/ is arguably particularly interesting in the context of Cantonese. Many Cantonese diaspora communities include individuals who are highly proficient, or even dominant, in English [8]. Crucially for this study, /n/ and /l/ are separate phonemes in English. Similar categories in the L1 (Cantonese) and L2 (English) may occupy a shared phonetic space [9] and interact in different ways across individuals, with prior work demonstrating cases of convergence [10], divergence [11, 12, 13], and contrast maintenance [14]. Given the existence of a heterogeneous population of Cantonese-English bilinguals, we additionally explore individual differences in the realization of /n/ and /l/ in Cantonese and English. More broadly, individual differences are worth exploring in the context of the current study, since the data in the current study includes spontaneous speech in both languages collected from conversational interviews where bilinguals demonstrated different degrees of code-switching in each language.

Mergers-in-progress contribute to variation within a speech community, such that an individual may produce one or both variants. The /n/ and /l/ merger in Cantonese is described as a merger-by-transfer [15], whereby words that are historically produced as /n/ surface as [l] in production [4, 5, 6]. This contrasts with an acoustically gradient change wherein historical /n/ gradually loses nasality or adopts lateral airflow. Acoustic descriptions of this merger are lacking, and may elucidate the categorical or gradient nature of this change.

While the /n/-/l/ merger has been described as nearly complete in some accounts of production [4], others find that it is still variable in production and perception [16]. Speech style matters in this case, as Cantonese speakers are typically aware of the merger—and that the historical /n/-initial variant is considered a "prestige form" [16]. As a result, in formal settings, individuals may be prone to hyperarticulation and/or hypercorrection of words that are underlyingly /l/-initial. This paper presents the first analysis of Cantonese /n/ and /l/ with spontaneous speech. While prior lab-based research informs the predictions, it is important to recognize that lab and spontaneous speech behaviour can differ drastically [17]. Speech style is long-established factor in phonetic variation [18], and that interlocutors and potential observers also impact phonetic variation (e.g., Audience Design) [19]. This stylistic variation affects phonetic variation in bilinguals, presenting potentially conflicting results in the literature. For example, using single word productions, Cantonese-English bilinguals characterized as Cantonese heritage speakers were more likely to release word-final stops [13]. Conversely, a recent study of the same phenomenon in spontaneous speech found that early Cantonese-English bilinguals were *less* likely to release final stops in English than non-Cantonese-English bilinguals [20]. These conflicting outcomes simply illustrate the need to examine variation in speech across styles and registers, as this variation has maximum utility for ASR systems and the development of NLP tools for speech and language, given how little is know about how talkers interact with such systems [21].

Copyright © 2021 ISCA

http://dx.doi.org/10.21437/Interspeech.2021-1754

Below, we describe a corpus study examining acoustic characteristics of /n/ and /l/ in the spontaneous speech of Cantonese-English bilinguals in both languages. In line with the (revised) Speech Learning Model [9, 22], several studies have observed some degree of cross-language influence between the corresponding sounds in each language of a bilingual [23, 24, 25]. At the same time, the presence of similar or identical sounds in both languages does not necessarily entail a link between those two sounds, as similarity may only exist on the surface, masking independent sound categories across languages [26]. If Cantonese /n/ and /l/ are linked with English /n/ and /l/, then the merging of /n/ and /l/ in Cantonese should result in a similar narrowing of the acoustic distance between English /n/ and /l/. On the other hand, if the sound categories across languages remain independent from one another, the acoustic changes induced by the merging of /n/ and /l/ in Cantonese should not affect English /n/ and /l/. While this is perhaps overly simplistic, it highlights the framework in which we analyze sounds subject to both sound change and crosslinguistic influence. In this vein, there are two primary theoretical questions of interest for this study, and a supplementary empirical goal:

1. How are /n/ and /l/ produced in spontaneous Cantonese speech? How are /n/ and /l/ produced by the same individuals in English? This aim is largely focused on describing the variation extant in the speech communities of a bilingual's languages.
2. How are speech sounds linked with one another in the context of crosslinguistic influence and sound change? How does this vary across individuals?
3. A natural outcome of these two theoretical questions is that we provide the first acoustic description of this Cantonese merger-in-progress.

2. Method

2.1. Data

Bilingual speech was collected from the SpiCE corpus [27]. This corpus comprises speech from a heterogeneous group of 34 early Cantonese-English bilingual talkers (female = 17, male = 17). All talkers either learned Cantonese first or both languages simultaneously, and the vast majority reported comparable speaking and listening proficiency across the two languages. SpiCE talkers are united in their early and balanced bilingualism, but there is a high degree of variability in the geographic origins of the talkers (mostly Canada and/or Hong Kong), and their caregivers (mostly Hong Kong and neighboring Chinese provinces). Each bilingual talker in the SpiCE corpus completed three tasks in both languages: sentence reading, storyboard narration, and a conversational interview. Analyses here includes all three tasks. As the recording sessions were casual in nature and conducted by undergraduate student interviewers, neither the environment nor the interlocutor were likely to have induced pressure for talkers to utilize the historical "prestigious" /n/-initial forms. In other words, talkers were likely implicitly comfortable producing innovative /l/-initial forms in an environment amongst their peers. The transcripts provided with the SpiCE corpus comprise hand-corrected orthographic and force-aligned phone level annotations. The corpus is described in greater detail in [27], and in the online documentation.[1]

[1] https://spice-corpus.readthedocs.io/

2.2. Target segments and acoustic measurements

The analysis focuses on pre-vocalic, word-initial /n/ and /l/ segments in Cantonese and English.[2] Instances were only included in the initial sample if there were no other nasal or liquid segments within the word. While this criterion was included to ensure that nasality was not impacted by nasal codas (which are *not* merging with /l/ in Cantonese), it also renders the samples more comparable across languages. Prior to collecting acoustic measurements, there were a total of 13,790 instances of /n/ and /l/ matching the criteria described above.

We focus on acoustic measures shown to correlate with nasality. Following, [29], we use a mid-frequency spectral tilt measure—H4-2KHz—where nasals were expected to have a greater spectral tilt than laterals. Following [7], we also measured the F2 and F1 spacing (calculated as the difference between F2 and F1) of the initial consonant [30, 31], where laterals of the same place of articulation were expected to show greater F2-F1 spacing compared to nasals. The F2-F1 spacing also captures differences in place of articulation, which can be used to evaluate whether bilinguals use the same /n/ and the same /l/ in both languages. More velarized /l/, which we anticipate in English, will have lower F2 values [32]. All acoustic estimates were made using VoiceSauce [33]. Because of concerns about alignment accuracy with the force-aligned data, values from the middle third of the nasals and laterals are analyzed.

2.3. Exclusionary criteria

Prior to making the acoustic measurements tokens were excluded if there was another nasal or liquid segment in the word, resulting in an initial sample of 13,790.

To mitigate measurement error, tokens where the estimated average f0 for the middle third of the phone was equal to or less than 65 Hz were removed ($n = 1,187$), as were tokens where the /n/ or /l/ was exactly 30 ms ($n = 2,938$), the minimum value in forced alignment [34]. This left 9,965 tokens for analysis: Cantonese /l/ = 1,620; Cantonese /n/ = 1,519; English /l/ = 4,724, of which 3,677 are "like"; English /n/ = 1,802.[3]

3. Analysis and Results

3.1. Mid-frequency spectral tilt: H4-2KHz

All analyses were carried out using R [35]. Mid-frequency spectral tilt, measured as H4 minus the amplitude of the harmonic closest to 2000 Hz (corrected for formant frequencies [36], was used to characterize the difference between /n/ and /l/. A linear mixed effects regression model was run with H4-2KHz as the dependent variable and Consonant (with /l/ as the reference level) and Language (Cantonese as the reference level) as independent variables. Talker and word were entered as random effects with consonant and language as by-subject random intercepts without an interaction. The model intercept was significant [$\beta = 3.17, SE = 0.67, t = 4.73, p < 0.001$]. There was also a significant simple effect of Language [$\beta = -5.43, SE = 0.87, t = -6.3, p < 0.001$] and an interaction of Consonant and Language [$\beta = 10.95, SE = 1.27, t =$

[2] Cantonese words were identified in the SpiCE corpus using the pycantonese Python package [28], which implements a longest string matching algorithm.

[3] Although a large proportion of the English /l/ tokens are represented by the word "like", we ultimately decided against removing these tokens from the analysis as this would result in a loss of statistical power.

Table 1: *Tukey-adjusted estimates and p-values (estimates; p-values) from pair-wise comparisons for H4-2KHz and F2-F1 for /n/ and /l/ in Cantonese and English, organized by within-language comparisons, across-language comparisons, and cross-language/cross-phone comparisons.*

Comparison	H4-2KHz	F2-F1
Canto /l/ - Canto /n/	$-1.82; 0.29$	$-111; 0.17$
Eng /l/ - Eng /n/	$-12.77; < 0.0001$	$-619; < 0.0001$
Canto /l/ - Eng /l/	$5.43; < 0.0001$	$181; 0.001$
Canto /n/ - Eng /n/	$-5.52; < 0.0001$	$-328; < 0.0001$
Canto /l/ - Eng /n/	$-7.34; < 0.0001$	$-438; < 0.0001$
Canto /n/ - Eng /l/	$7.25; < 0.0001$	$291; < 0.0001$

$8.6, p < 0.001]$. The simple effect of Consonant was not significant [$\beta = 1.82, SE = 1.02, t = 1.78, p = 0.08$].

These results are visualized in Figure 1 by language and consonant. For the reference level /l/, English has a more negative spectral tilt, indicating higher energy in the mid-frequency range. English /n/ also has a higher spectral tilt, indicating a loss of mid-frequency energy, suggesting stronger nasalization. The Cantonese consonants are more similar to one another.

The emmeans package [37] was used for post-hoc comparisons, and these are reported in Table 1. All comparisons are differentiated by mid-frequency spectral slope except Cantonese /n/ and /l/.

These results indicate that early Cantonese-English bilinguals maintain a distinction in /n/ and /l/ in English, while this contrast is attenuated in Cantonese. Bilinguals also maintain cross-language differences in /n/ and /l/.

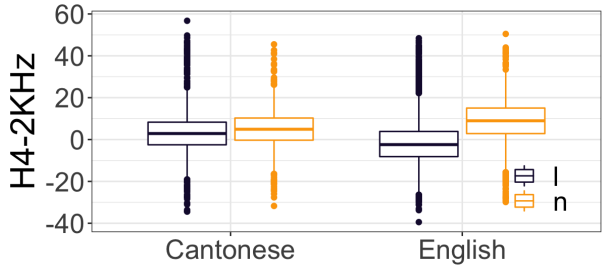

Figure 1: *Boxplots of mid-frequency spectral tilt values for Cantonese (left) and English (right) /n/ (orange) and /l/ (black).*

3.2. F2-F1 spacing

The spacing between F1 and F2 can be used to infer differences across nasals and laterals, in addition to place of articulation of /l/, as the difference between these values is related to the degree of velarization of /l/. A linear mixed effects regression model was run with F2-F1 as the dependent variable and Consonant (with /l/ as the reference level) and Language (Cantonese as the reference level) as independent variables. Talker and word were entered as random effects with Consonant and Language as by-subject random intercepts with an interaction. The model intercept was significant [$\beta = 1132.04, SE = 31.84, t = 35.56, p < 0.001$]. There were significant simple effects of Consonant [$\beta = 110.82, SE = 53.32, t = 2.08, p = 0.04$] and Language [$\beta = -180.62, SE = 46.76, t = -3.86, p < 0.001$] and an interaction of Consonant and Language [$\beta = 508.19, SE = 78.42, t = 6.48, p < 0.001$]. Given the reference levels of Cantonese and /l/, these results indicate that in Cantonese, the F2-F1 is smaller in /l/ compared to /n/, and English /l/ has smaller F2-F1 values, indicating a more velarized articulation compared to the Cantonese /l/. These results are visualized in Figure 2 by language and consonant. Within each language, /l/ has a lower F2-F1 difference than /n/, with English /l/ having the lowest values and English /n/ having the highest values.

Again, the emmeans package [37] was used for post-hoc comparisons, and these are also reported in Table 1. All comparisons are differentiated by F2-F1 except Cantonese /n/ and /l/, which notably was significantly different in the main mixed effects model. The lack of effect in the pairwise comparisons is likely due to the adjustment for multiple comparisons, and is a testament to the small size of the effect.

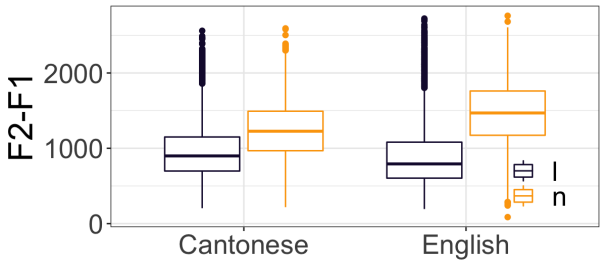

Figure 2: *Boxplots of F2-F1 values (in Hz) for Cantonese (left) and English (right) /n/ (orange) and /l/ (black).*

3.3. Individual differences

The group-level data indicates that while a spectral slope proxy for nasalization suggests no distinction between Cantonese /n/ and /l/, a formant-spacing measure suggests there may still be very small differences in Cantonese /n/ and /l/. The group-level analyses also demonstrate that cross-language differences are maintained: bilinguals' /n/ and /l/ categories are not inextricably linked across languages—English appears separate and Cantonese merged (cf. [26]). These results may, however, gloss over individual differences in the acoustic realization of these sounds, participation in or degree of merger, and cross-language links. Figure 3 presents by-subject means with standard error for /n/ and /l/ for each language and acoustic measure. This visualization makes clear the magnitude of the individual differences on all levels. While a majority have intersecting lines for H4-2KHz—such that a clear distinction is maintained in English and a merger is suggested in Cantonese—some individuals (e.g., VM21B) present what appears to be tightly coupled cross-language categories, where the contrast is similarly maintained in both languages. The same individuals do not necessarily show identical patterns in F2-F1 with respect to mergedness or cross-language linkages.

Several dimensions may underlie some of the individual differences (e.g., parent dialect, language dominance, language use patterns, code-switching etc.). For instance, with respect to code-switching, some individuals engage in code-switching at higher rates than others, and productions of /n/ and /l/ that are closer to/anticipate code-switches may show more cross-

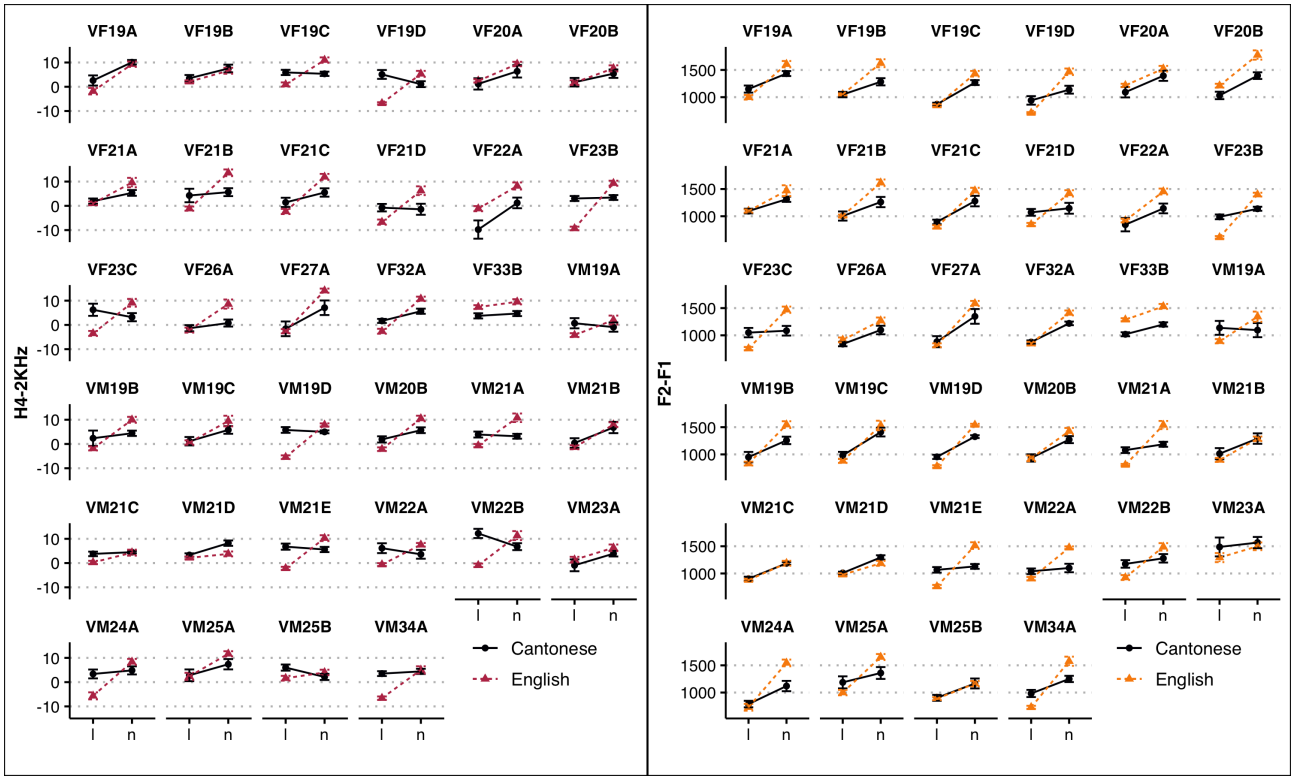

Figure 3: *By-subject plots of H4-2KHz (left) and F2-F1 (right) values (in Hz) for English (dashed lines) and Cantonese (solid lines).*

language influence than other tokens [38, 11].

4. Discussion and Conclusion

As previous studies have utilized perceptual measures to understand the merger between Cantonese /n/ and /l/ in production [4, 5, 6], the current study represents the first acoustic examination of this merger. We carried out a corpus study examining the productions of /n/ and /l/ in the spontaneous speech of Cantonese-English bilinguals in both languages. Since corresponding segments in each language occupy a shared phonetic space [9, 22] and may interact in various ways [26], we hypothesized that Cantonese /n/ and /l/ may either be linked with the same consonants in English (in which case, the merger in Cantonese would produce a similar change in English) or that /n/ and /l/ would remain independent sound categories across both languages [26]. More broadly, given the heterogeneous nature of the Cantonese-English bilingual population in the current study, we adjudicated between these hypotheses by examining the productions of /n/ and /l/ in Cantonese and English at both the group and the individual level.

Group-level comparisons revealed that Cantonese /n/ and /l/ were poorly differentiated by measures of mid-frequency spectral tilt, but somewhat separated by F2-F1 spacing, while English /n/ and /l/ were well-differentiated. As such, it appears that at the group-level, corresponding segments in Cantonese and English are not inextricably linked, rather behave as distinct sound categories. In other words, the merging of /n/ and /l/ is present in Cantonese, while the contrast between /n/ and /l/ is preserved in English. While this might suggest a categorical merging of /n/ to /l/ in Cantonese, a closer look at the data from individual speakers reveal that the surface form—evidence for the merger and cross-language links—varies considerably across individuals. This warrants a more detailed investigation. Finally, given that our English /l/-initial words included a preponderance of a single lexical item ("like"), future analyses will consider lexical frequency and its role in /l/ phonetic variation.

The acoustic measures used here are just a subset of the possible measures that may best characterize within- and cross-language patterns. Other measures that have been shown to distinguish /n/ and /l/ in English include F3-F2 spacing [31, 30] and change in F1 amplitude across the consonant and vowel [39, 30]. Analyses of vowels following the consonants of interest may use A1-P0, and A1-P1 [40]. Given that each of these measures are gathered at different time points in the syllable, a larger suite of measures will help to provide a more holistic picture of the /n/ and /l/ merger in Cantonese. English allows /n/ and /l/ in codas, and Cantonese permits just /n/ in coda position. Analysis of the consonants in coda position will enrich our understanding of language-specific phonetics, position-specific mergers, and cross-language mutual influence.

5. Acknowledgements

This work has been financially supported by a UBC Public Scholars Initiative Award to KAJ, and a SSHRC award to MB. We thank members of the Speech in Context Laboratory at UBC, particularly those who have contributed extensively to the creation of the SpiCE corpus, including Ivan Fong, Nancy Yiu, Ariana Zattera, Christina Sen, and Katherine Lee.

6. References

[1] U. Weinreich, W. Labov, and M. Herzog, *Empirical foundations for a theory of language change*. University of Texas Press

Austin, 1968, vol. 58.

[2] M. Sumner and A. G. Samuel, "The effect of experience on the perception and representation of dialect variants," *Journal of Memory and Language*, vol. 60, no. 4, pp. 487–501, 2009.

[3] P. P. Mok, D. Zuo, and P. W. Wong, "Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese," *Language Variation and Change*, vol. 25, no. 3, p. 341, 2013.

[4] C. K. To, S. McLeod, and P. S. Cheung, "Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment," *Clinical linguistics & phonetics*, vol. 29, no. 5, pp. 333–353, 2015.

[5] S. Wong, "A Chinese syllabary pronounced according to the dialect of Cantonese," *Canton: Lingnam University*, 1941.

[6] E. Zee, "Change and variation in the syllable-initial and syllable-final consonants in Hong Kong Cantonese," *Journal of Chinese Linguistics*, vol. 27, no. 1, pp. 120–167, 1999.

[7] R. Cheng and A. Jongman, "Acoustic analysis of nasal and lateral consonants: The merger in Eastern Min," *The Journal of the Acoustical Society of America*, vol. 145, no. 3, pp. 1828–1828, 2019.

[8] E. Benmamoun, S. Montrul, and M. Polinsky, "Heritage languages and their speakers: Opportunities and challenges for linguistics," *Theoretical linguistics*, 2013.

[9] J. E. Flege and O.-S. Bohn, "The revised speech learning model (SLM-r)," in *Second Language Speech Learning: Theoretical and Empirical Progress*. Cambridge, UK: Cambridge University Press, 2021, pp. 3–83.

[10] J. E. Flege, "The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification," *Journal of phonetics*, vol. 15, no. 1, pp. 47–65, 1987.

[11] B. E. Bullock and A. J. Toribio, "Trying to hit a moving target: On the sociophonetics of code-switching," in *Studies in Bilingualism*, L. Isurin, D. Winford, and K. deBot, Eds. Amsterdam: John Benjamins Publishing Company, 2009, vol. 41, pp. 189–206.

[12] J. E. Flege and W. Eefting, "Cross-language switching in stop consonant perception and production by Dutch speakers of English," *Speech Communication*, vol. 6, no. 3, pp. 185–202, 1987.

[13] M. Polinsky, *Heritage Languages and their Speakers*, ser. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press, 2018.

[14] C. B. Chang, Y. Yao, E. F. Haynes, and R. Rhodes, "Production of phonetic and phonological contrast by heritage speakers of Mandarin," *The Journal of the Acoustical Society of America*, vol. 129, no. 6, pp. 3964–3980, 2011.

[15] P. Trudgill and T. Foxcroft, "On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia," in *ociolinguistic patterns in British English*, P. Trudgill, Ed. University Park Press, Baltimore, 1978, pp. 69–79.

[16] L. S. P. Cheng, "Mergers-in-progress in Hong Kong and Vancouver Cantonese-English bilinguals," Undergraduate Honours Thesis, University of British Columbia, Vancouver, BC, 2017.

[17] S. Gahl, Y. Yao, and K. Johnson, "Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech," *Journal of Memory and Language*, vol. 66, no. 4, pp. 789–806, 2012.

[18] W. Labov, "Some principles of linguistic methodology," *Language in Society*, pp. 97–120, 1972.

[19] A. Bell, "Language style as audience design," *Language in Society*, vol. 13, no. 2, pp. 145–204, 1984.

[20] K. A. Johnson and M. Babel, "Language contact within the speaker: Phonetic variation and crosslinguistic influence," *OSF Preprints*, 2021.

[21] G. Zellou, M. Cohn, and B. Ferenc Segedin, "Age- and gender-related differences in speech alignment toward humans and voice-ai," *Frontiers in Communication*, vol. 5, p. 140, 2021.

[22] J. E. Flege, "Second language speech learning: Theory, findings, and problems," in *Speech perception and linguistic experience: Issues in cross-language research*. York Press, 1995, vol. 92, pp. 233–277.

[23] M. Antoniou, C. T. Best, M. D. Tyler, and C. Kroos, "Inter-language interference in vot production by L2-dominant bilinguals: Asymmetries in phonetic code-switching," *Journal of Phonetics*, vol. 39, no. 4, pp. 558–570, 2011.

[24] C. A. Fowler, V. Sramko, D. J. Ostry, S. A. Rowland, and P. Hallé, "Cross language phonetic influences on the speech of French-English bilinguals," *Journal of Phonetics*, vol. 36, no. 4, pp. 649–663, 2008.

[25] K.-H. Kang and S. G. Guion, "Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals," *The Journal of the Acoustical Society of America*, vol. 119, no. 3, pp. 1672–1683, 2006.

[26] J. Schertz, Y. Kang, and S. Han, "Cross-language correspondences in the face of change: Phonetic independence versus convergence in two Korean-Mandarin bilingual communities," *International Journal of Bilingualism*, vol. 23, no. 1, pp. 157–199, 2019.

[27] K. A. Johnson, M. Babel, I. Fong, and N. Yiu, "SpiCE: A new open-access corpus of conversational bilingual speech in Cantonese and English," in *Proc. 12th Language Resources and Evaluation Conference*, Marseille, France, 2020, pp. 4089–4095.

[28] J. L. Lee, "PyCantonese: Cantonese linguistic research in the age of big data," *Talk at the Childhood Bilingualism Research Centre, the Chinese University of Hong Kong*, 2015.

[29] M. Garellek, A. Ritchart, and J. Kuang, "Breathy voice during nasality: A cross-linguistic study," *Journal of Phonetics*, vol. 59, pp. 110–121, 2016.

[30] A. A. M. Prahler, "Analysis and synthesis of the american english lateral constant," Ph.D. dissertation, Massachusetts Institute of Technology, 1998.

[31] K. Johnson, *Acoustic and auditory phonetics*. John Wiley & Sons, 2011.

[32] D. Recasens, "A cross-language acoustic study of initial and final allophones of/l," *Speech Communication*, vol. 54, no. 3, pp. 368–383, 2012.

[33] Y.-L. Shue, P. Keating, C. Vicenik, and K. Yu, "VoiceSauce: A program for voice analysis," in *Proceedings of the 17th International Congress of Phonetic Sciences*, vol. 3, Hong Kong, 2011, pp. 1846–1849. [Online]. Available: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011

[34] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal forced aligner: Trainable text-speech alignment using kaldi." in *Interspeech*, vol. 2017, 2017, pp. 498–502.

[35] R. C. Team *et al.*, "R: A language and environment for statistical computing," 2013.

[36] M. Iseli, Y.-L. Shue, and A. Alwan, "Age, sex, and vowel dependencies of acoustic measures related to the voice source," *The Journal of the Acoustical Society of America*, vol. 121, no. 4, pp. 2283–2295, 2007.

[37] R. V. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means*, 2021, r package version 1.5.5-1. [Online]. Available: https://CRAN.R-project.org/package=emmeans

[38] M. Fricke, J. F. Kroll, and P. E. Dussias, "Phonetic variation in bilingual speech: A lens for studying the production-comprehension link," *Journal of Memory and Language*, vol. 89, pp. 110–137, 2016.

[39] A. S. House and K. N. Stevens, "Analog studies of the nasalization of vowels," *Journal of Speech and Hearing Disorders*, vol. 21, no. 2, pp. 218–232, 1956.

[40] M. Y. Chen, "Acoustic correlates of english and french nasalized vowels," *The Journal of the Acoustical Society of America*, vol. 102, no. 4, pp. 2360–2370, 1997.