Chapter 1

Introduction

Bilingual and monolingual linguistic experience differs drastically. This sentiment is captured by the often-repeated observation that a "bilingual is NOT the sum of two complete or incomplete monolinguals; rather, [they have] a unique and specific linguistic configuration...a different but complete linguistic entity" (Grosjean, 1989, p. 6). One of the defining characteristics of a bilingual's linguistic configuration is a shared phonetic space where both languages are produced and perceived (Flege and Bohn, 2021). Broadly, this dissertation is concerned with the implications of such a space and what aspects of sound are shared across languages. And, given that speech usually occurs in a communicative context, what is available in the bilingual speech signal to facilitate processes like talker identification and processing?

The studies presented in this dissertation approach this larger question at different levels in the phonetic space—voice quality (Chapter 3) and sound categories (Chapter 4). Yet, the studies share motivation in speech perception, as various levels of phonetic variation have been proposed to account for how listeners can identify and track a talker across languages. This introduction briefly sets up that shared motivation, tying two otherwise quite different studies together. Given the unique angle of each study, most of the literature is reviewed in the relevant chapters.

The introduction proceeds as follows. Section 1.1 gives context to the study of bilingualism in phonetics in broad strokes—that is, who is considered to be bilingual and what are some of the key characteristics that define them. The goal is to set up later chapters rather than provide a comprehensive discussion. Section 1.2 reviews some of the literature on how multilingual phonetic variation is perceived, emphasizing talker identification and how multilingual listeners process code-switching. Section 1.3 motivates the need to attend to speaking style and argues that spontaneous speech corpora greatly facilitate the study of multilingual phonetic variation. Lastly, Section 1.4 provides the specific goal or research question for each of the main content chapters—2, 3, and 4.

1.1 Bilingualism

The population of interest in this dissertation comprises early bilinguals who are comfortable speaking and comprehending their two main languages—Chapter 2 provides a detailed description. In the most general sense, a bilingual is someone with knowledge of two or more languages (Grosjean, 1989). This incredibly broad definition includes a diverse range of types of bilinguals. Different kinds of bilinguals are perhaps best described on a continuum from first language (L1) to second language (L2) dominance, the bookends of which are monolinguals and replacive bilinguals, with learners, attriters, and balanced bilinguals in the middle (Gertken et al., 2014). Using a continuum in this way reflects the heterogeneous nature of bilingualism, even if it only captures a particular facet of bilingual competence. Dominance—and other aspects like patterns of use—are affected by factors such as age of acquisition, immersion environment, frequency, social and communicative context (Gertken et al., 2014).

While a spectrum may better reflect the reality of bilingualism, much of the literature focuses on more clearly defined groups at discrete points of the spectrum, such as language learners or early (balanced) bilinguals. Typically, early bilinguals have learned both languages from their first years of life. A common cutoff is age five, or the age at which children begin regularly attending primary school

(Amengual, 2017), as this marks a qualitative change in the kind of linguistic input bilinguals receive. Regardless of when bilinguals acquire a language, they do not necessarily use their languages to the same extent across different domains. For example, a Cantonese-English bilingual in Vancouver, BC, Canada, might use English at school and Cantonese at home. Bilingual language experience varies is still other ways, including code-switching (Fricke et al., 2016a), immersion environment (Sancier and Fowler, 1997), and formal instruction (Fricke et al., 2019). Each of these factors has a demonstrated effect on speech production. Such variety leads to markedly different linguistic experiences across groups of bilinguals, and as a result, markedly different patterns in speech production.

Across different kinds of phonetics research in bilingualism, there is a common trend of comparing bilinguals to "closely matched" monolingual populations. Given the sheer heterogeneity within and across bilingual populations, there may not always be an appropriate monolingual comparison group. Further, Grosjean (1989) and many others have argued that such comparisons are often inappropriate. As a result, drawing comparisons between monolinguals and bilinguals may not always be fruitful or necessary, depending on the circumstances. This is reflected by a shift in the literature towards examining bilinguals on a within-population (e.g., Chan et al., 2020) or within-talker basis (e.g., Simonet and Amengual, 2019), or by comparing separate bilingual populations with different characteristics (e.g., Brown and Harper, 2009). This range of study designs will be apparent in the literature reviews of the following chapters. While there remains a broad range of comparisons in the literature, in all cases, there is a strong push to consider bilinguals as the complete speakers they are.

1.2 Processing bilingual talkers

Communicating in more than one language doesn't just involve the language produced by bilingual talkers; it is also impacted how listeners perceive those talkers. As noted early in this introduction, one of the major consequences of bilingualism is a shared phonetic space (Flege and Bohn, 2021), in which bilinguals presumably

(i.e., are hypothesized to) use similar voice quality to produce similar sound categories.[1] This shared phonetic space thus also impacts the perception of bilingual talkers, whether the listener is a fellow bilingual or not.

While bilingual speech perception is a large and multifaceted field (Ingvalson et al., 2014), the clearest motivation for the studies in this dissertation comes from the advantage that multilingualism offers in talker identification. Orena et al. (2019) report on a talker identification study with French-English bilingual talkers, in which bilingual listeners—particularly those with language mixing experience—were better able to generalize talker-indexical information learned in English to French and vice versa when compared to monolingual English listeners. However, all groups in the study were above chance, suggesting both linguistic and non-linguistic components to talker identification. Orena et al. offer several potential explanations for this advantage: "that there are systematic changes in indexical information...[or] systematic consistencies in linguistic information across bilingual speech" (2019, p. EL308). Bilingual listeners are highly sensitive to subtle differences in acoustic input (Ju and Luce, 2004). As a result, the presence of systematicity in both talker-indexical and linguistic information—however subtle—would be accessible to bilingual listeners, particularly those with language mixing experience. Orena et al. also suggest that the results could be explained because the bilinguals "were familiar with both languages...while the monolinguals were only familiar with one of the languages" (2019, p. EL309), though this account would be difficult to separate from the previous two.

Regardless of the particular explanation, the bilingual advantage in bilingual talker identification likely arises from their deep familiarity in listening to how talkers vary within and across languages. While these accounts emphasize the linguistic aspect of bilingual competence, there is more to the picture than that. Bullock and Toribio (2009) highlight the integral role that sociolinguistic competence plays in accounting for variability in production and that bilinguals have an

---

[1]The speech production literature discussing similarity in voice quality and sound categories will be reviewed in greater detail within Chapters 3 and 4, respectively.

expanded repertoire of forms. That is, a bilingual can produce forms canonical to either language, and they can also show divergence, convergence, interference, or hypercorrection depending on the social and cognitive circumstances. This competence is echoed by Kleinschmidt et al. (2018) in their integrated account of how social, acoustic, and linguistic variation are perceived. Learning the structure and systematicity of variation, then, is both a linguistic, talker-indexical, and social venture. This dissertation investigates the role of the former two but acknowledges the importance of attending to and considering the social component as well.

While Orena et al. (2019) point to some prior work supporting the talker-indexical and linguistic accounts of bilingual talker identification, convincing evidence remains scarce. This dissertation directly addresses these accounts from the perspective of documenting the speech signal. Chapter 3 examines voice variation—generally considered to reflect talker-indexicality. Chapter 4 focuses on the structure of phonetic category variation—a reasonably clear example of linguistic information. While using distinct methods and addressing different aspects of phonetics, each chapter represents an aspect of the signal that may facilitate crosslinguistic talker identification.

1.3   Variability in conversational speech

One of the primary goals of this dissertation is to document and investigate the structure of phonetic variation. While variability is inherent to the speech signal (cf. the lack of invariance problem Liberman et al., 1967), spontaneous speech encompasses a greater degree of variability than other speaking styles (e.g., reduction phenomena: Johnson, 2004). Spontaneous speech—in the form of conversations—is thus the focus of study in this dissertation.

Additionally, as the motivation for the studies in Chapters 3 and 4 stems from listeners' ability to identify talkers in more than one language, using conversational speech also supports the external validity of this dissertation. Conversational speech better reflects the range of forms that people use and perceive in their daily lives (cite something—ideas?). Additionally, given the potential range of variabil-

5

ity, it is also necessary to study a large enough sample such that it comprises the range of variation for the particular communicative situation. For similar reasons, Tanner et al. (2020) argue that large-scale corpus studies are uniquely valuable for understanding phonetic variation. In this vein, this dissertation leverages the study of conversational speech data from a sufficiently large speech corpus. The corpus, along with further motivation for its use, is described in Chapter 2.

1.4   Thesis goals & research questions

While each of the main content chapters in this dissertation is united by common motivation, each has a unique focus or research question. These are as follows:

Chapter 2 expands on the motivation behind studying spontaneous speech and introduces the SpiCE corpus of spontaneous bilingual **Sp**eech **i**n **C**antonese and **E**nglish (Johnson, 2021b). The development and dissemination of this corpus comprise a substantial portion of this dissertation.

Chapter 3 focuses on the structure of voice variation. In broad terms, it asks: do bilinguals have the same voice in each of their languages? More specifically, do bilinguals exhibit similar spectral properties and lower-dimensional structure in their voice across each language they speak? Chapter 3 also addresses a methodological question regarding the sample size necessary for the methods used.

Chapter 4 focuses on the structure of sound categories. In broad terms, it asks: Do bilinguals produce long-lag stops in the same way in each of their languages? More specifically, it describes the structure and sources of variation in how bilinguals produce voice-onset time in conversational speech.

Bibliography

Afouras, Triantafyllos, Chung, Joon Son, and Zisserman, Andrew. Now you're speaking my language: Visual language identification. In *Proceedings of Interspeech 2020*, pages 2402–2406, 2020. doi:10.21437/Interspeech.2020-2921.

Alderete, John, Chan, Queenie, and Yeung, H. Henny. Tone slips in Cantonese: Evidence for early phonological encoding. *Cognition*, 191:103952, 2019. doi:10.1016/j.cognition.2019.04.021.

Altenberg, Evelyn P. and Ferrand, Carole T. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice*, 20(1):89–96, 2006. doi:10.1016/j.jvoice.2005.01.005.

Amengual, Mark. Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *International Journal of Bilingualism*, 23(5):954–970, 2017. doi:10.1177/1367006917741364.

Amengual, Mark. Asymmetrical interlingual influence in the production of Spanish and English laterals as a result of competing activation in bilingual language processing. *Journal of Phonetics*, 69:12–28, 2018. doi:10.1016/j.wocn.2018.04.002.

Antoniou, Mark, Best, Catherine T., Tyler, Michael D., and Kroos, Christian. Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, 38(4):640–653, 2010. doi:10.1016/j.wocn.2010.09.005.

Antoniou, Mark, Best, Catherine T., Tyler, Michael D., and Kroos, Christian. Inter-language interference in VOT production by L2-dominant bilinguals:

Asymmetries in phonetic code-switching. *Journal of Phonetics*, 39(4):
558–570, 2011. doi:10.1016/j.wocn.2011.03.001.

Ardila, Rosana, Branson, Megan, Davis, Kelly, Kohler, Michael, Meyer, Josh,
Henretty, Michael, Morais, Reuben, Saunders, Lindsay, Tyers, Francis, and
Weber, Gregor. Common voice: A massively-multilingual speech corpus. In
Proceedings of the 12th Language Resources and Evaluation Conference,
pages 4218–4222, Marseille, France, 2020.
https://www.aclweb.org/anthology/2020.lrec-1.520.

Audacity Team. Audacity (R): Free audio editor and recorder, 2018.
https://www.audacityteam.org/.

Balukas, Colleen and Koops, Christian. Spanish-English bilingual voice onset
time in spontaneous code-switching. *International Journal of Bilingualism*, 19
(4):423–443, 2015. doi:10.1177/1367006913516035.

Barr, Dale J., Levy, Roger, Scheepers, Christoph, and Tily, Harry J. Random
effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal
of Memory and Language*, 68(3):255–278, 2013.
doi:10.1016/j.jml.2012.11.001.

Bates, Douglas, Kliegl, Reinhold, Vasishth, Shravan, and Baayen, Harald.
Parsimonious mixed models. *ArXiv Preprints*, pages 1–21, May 2018.
http://arxiv.org/abs/1506.04967.

Bauer, Robert S. and Benedict, Paul K. *Modern Cantonese Phonology*. De
Gruyter Mouton, Berlin, 1997. doi:10.1515/9783110823707.

Belin, Pascal, Fecteau, Shirley, and Bédard, Catherine. Thinking the voice:
Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3):
129–135, 2004. doi:10.1016/j.tics.2004.01.008.

Boersma, Paul and Weenink, David. Praat: Doing phonetics by computer, 2021.
http://www.praat.org/. Version 6.1.38.

Bolton, Kingsley, Bacon-Shone, John, and Lee, Siu-lun. Societal multilingualism
in Hong Kong. In *Multilingual Global Cities*, pages 160–184. Routledge,
2020. doi:10.4324/9780429463860-12.

Bradlow, Ann R, Ackerman, Lauren, Burchfield, L Ann, Hesterberg, Lisa, Luque, Jenna, and Mok, Kelsey. Language- and talker-dependent variation in global features of native and non-native speech. In *Proceedings of the 17th International Congress of Phonetic Sciences*, pages 356–359, Hong Kong, 2011. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Bradlow/Bradlow.pdf.

Bradlow, Ann R., Kim, Midam, and Blasingame, Michael. 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, 2017. doi:10.1121/1.4976044.

Brehm, Laurel and Alday, Phillip M. A decade of mixed models: It's past time to set your contrasts. *OSF Preprints*, July 2021. https://osf.io/3tgq6/.

Brown, Esther L. and Amengual, Mark. Fine-grained and probabilistic cross-linguistic influence in the pronunciation of cognates: Evidence from corpus-based spontaneous conversation and experimentally elicited data. *Studies in Hispanic and Lusophone Linguistics*, 8(1):59–83, 2015. doi:10.1515/shll-2015-0003.

Brown, Esther L. and Harper, David. Phonological evidence of interlingual exemplar connections. *Studies in Hispanic and Lusophone Linguistics*, 2(2): 257–274, 2009. doi:10.1515/shll-2009-1052.

Bullock, Barbara E. and Toribio, Almeida Jacqueline. Trying to hit a moving target: On the sociophonetics of code-switching. In Isurin, Ludmila, Winford, Donald, and deBot, Kees, editors, *Studies in Bilingualism*, volume 41, pages 189–206. John Benjamins Publishing Company, Amsterdam, 2009. doi:10.1075/sibil.41.12bul.

Burkner, Paul-Christian. brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi:10.18637/jss.v080.i01.

Burton, A. Mike, Kramer, Robin S. S., Ritchie, Kay L., and Jenkins, Rob. Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science*, 40(1):202–223, 2016. doi:10.1111/cogs.12231.

Casillas, Joseph V. Interlingual interactions elicit performance mismatches not "compromise" categories in early bilinguals: Evidence from meta-analysis and coronal stops. *Languages*, 6(1):9, 2021. doi:10.3390/languages6010009.

Ćavar, Malgorzata, Ćavar, Damir, and Cruz, Hilaria. Endangered language documentation: Bootstrapping a Chatino speech corpus, forced aligner, ASR. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pages 4004–4011, Portorož, Slovenia, 2016. https://aclanthology.org/L16-1632/.

Chan, Alice Y. W. and Li, David C. S. English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, 13(1):67–85, 2000. doi:10.1080/07908310008666590.

Chan, Leighanne, Johnson, Khia A., and Babel, Molly. Lexically-guided perceptual learning in early Cantonese-English bilinguals. *The Journal of the Acoustical Society of America*, 147(3):EL277–EL282, 2020. doi:10.1121/10.0000942.

Chang, Charles B. Determining cross-linguistic phonological similarity between segments: The primacy of abstract aspects of similarity. In Raimy, Eric and Cairns, Charles E., editors, *The Segment in Phonetics and Phonology*, pages 199–217. John Wiley & Sons, Inc., Chichester, UK, 1 edition, 2015. doi:10.1002/9781118555491.ch9.

Cheng, Andrew. Cross-linguistic F0 differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America*, 147(2): EL67–EL73, 2020. doi:10.1121/10.0000498.

Cho, Taehong and Ladefoged, Peter. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27(2):207–229, 1999. doi:10.1006/jpho.1999.0094.

Chodroff, Eleanor. *Structured variation in obstruent production and perception*. PhD dissertation, Johns Hopkins University, Baltimore, MD, 2017. https://jscholarship.library.jhu.edu/handle/1774.2/44696.

Chodroff, Eleanor and Baese-Berk, Melissa. Constraints on variability in the voice onset time of L2 English stop consonants. In Calhoun, Sasha, Escudero,

Paola, Tabain, Marija, and Warren, Paul, editors, *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 661–665, Melbourne, Austrailia, 2019. https://assta.org/proceedings/ICPhS2019/papers/ICPhS_710.pdf.

Chodroff, Eleanor and Wilson, Colin. Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, 61:30–47, 2017. doi:10.1016/j.wocn.2017.01.001.

Chodroff, Eleanor and Wilson, Colin. Predictability of stop consonant phonetics across talkers: Between-category and within-category dependencies among cues for place and voice. *Linguistics Vanguard*, 4(s2), 2018. doi:10.1515/lingvan_2017_0047.

Chodroff, Eleanor, Golden, Alessandra, and Wilson, Colin. Covariation of stop voice onset time across languages: Evidence for a universal constraint on phonetic realization. *The Journal of the Acoustical Society of America*, 145(1): EL109–EL115, 2019. doi:10.1121/1.5088035.

Clumeck, Harold, Barton, David, Macken, Marlys A., and Huntington, Dorothy A. The aspiration contrast in Cantonese word-initial stops: Data from children and adults. *Journal of Chinese Linguistics*, 9(2):210–225, 1981. https://www.jstor.org/stable/23753507.

Deuchar, Margaret, Davies, Peredur, Herring, Jon Russell, Parafita Couto, M. Carmen, and Carter, Diana. Building bilingual corpora. In Thomas, Enlli M. and Mennen, Ineke, editors, *Advances in the Study of Bilingualism*, pages 93–110. Multilingual Matters, 2014. doi:10.21832/9781783091713_008.

Ethnologue. Chinese, Yue. In Eberhard, David Mm, Simons, Gary F., and Fennig, Charles D., editors, *Ethnologue: Languages of the World*. SIL International, Dallas, TX, 24 edition, 2021. http://www.ethnologue.com. Online version.

Faytak, Matthew Donald. *Articulatory uniformity through articulatory reuse: insights from an ultrasound study of Sūzhōu Chinese*. Doctoral dissertation, University of California, Berkeley, 2018. https://escholarship.org/uc/item/0jr0010h.

Flege, James Emil and Bohn, Ocke-Schwen. The revised speech learning model (SLM-r). In Wayland, Ratree, editor, *Second Language Speech Learning: Theoretical and Empirical Progress*, pages 3–83. Cambridge University Press, 2021. doi:10.1017/9781108886901.002.

Fricke, Melinda, Baese-Berk, Melissa M., and Goldrick, Matthew. Dimensions of similarity in the mental lexicon. *Language, Cognition and Neuroscience*, 31 (5):639–645, 2016a. doi:10.1080/23273798.2015.1130234.

Fricke, Melinda, Kroll, Judith F., and Dussias, Paola E. Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language*, 89:110–137, 2016b. doi:10.1016/j.jml.2015.10.001.

Fricke, Melinda, Zirnstein, Megan, Navarro-Torres, Christian, and Kroll, Judith F. Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, 22(1):200–207, 2019. doi:10.1017/S1366728918000482.

Gahl, Susanne, Yao, Yao, and Johnson, Keith. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4):789–806, 2012. doi:10.1016/j.jml.2011.11.006.

Garellek, Marc. The phonetics of voice. In Katz, William F. and Assmann, Peter F., editors, *The Routledge Handbook of Phonetics*. Routledge, 2019. doi:10.4324/9780429056253_5.

Gelman, Andrew, Simpson, Daniel, and Betancourt, Michael. The prior can often only be understood in the context of the likelihood. *Entropy*, 19(10):555, 2017. doi:10.3390/e19100555.

Gertken, Libby M., Amengual, Mark, and Birdsong, David. Assessing language dominance with the Bilingual Language Profile. In Leclercq, Pascale, Edmonds, Amanda, and Hilton, Heather, editors, *Measuring L2 proficiency: Perspectives from SLA*, pages 208–225. Multilingual Matters, Bristol, UK, 2014. doi:10.21832/9781783092291-014.

Godfrey, J.J., Holliman, E.C., and McDaniel, J. SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the 1992 IEEE*

International Conference on Acoustics, Speech, and Signal Processing, 1992. doi:10.1109/icassp.1992.225858.

Goldrick, Matthew, Runnqvist, Elin, and Costa, Albert. Language switching makes pronunciation less nativelike. *Psychological Science*, 25(4):1031–1036, 2014. doi:10.1177/0956797613520014.

Google. Cloud speech-to-text, 2019. https://cloud.google.com/speech-to-text/.

Grieve, Jack. Observation, experimentation, and replication in linguistics. *Linguistics*, 0, 2021. doi:10.1515/ling-2021-0094.

Grosjean, François. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1):3–15, 1989. doi:10.1016/0093_934X(89)90048_5.

Grosjean, François. An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism*, 16(1):11–21, 2011. doi:10.1177/1367006911403210.

Guion, Susan G. The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, 60(2):98–128, 2003. doi:10.1159/000071449.

Haines, Nathaniel, Kvam, Peter D., Irving, Louis H., Smith, Colin, Beauchaine, Theodore P., Pitt, Mark A., Ahn, Woo-Young, and Turner, Brandon M. Theoretically informed generative models can advance the psychological and brain sciences: Lessons from the reliability paradox. *PsyArXiv Preprints*, August 2020. doi:10.31234/osf.io/xr7y3.

Hillenbrand, J, Cleveland, R A, and Erickson, R L. Acoustic correlates of breathy vocal quality. *Journal of speech and hearing research*, 37(4):769–778, 1994. doi:10.1044/jshr.3704.769.

IEEE. IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics*, 17(3):225–246, 1969. doi:10.1109/TAU.1969.1162058.

Ingvalson, Erin M., Ettlinger, Marc, and Wong, Patrick C. M. Bilingual speech perception and learning: A review of recent trends. *International Journal of Bilingualism*, 18(1):35–47, 2014. doi:10.1177/1367006912456586.

Iseli, Markus, Shue, Yen-Liang, and Alwan, Abeer. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, 121(4):2283–2295, 2007. doi:10.1121/1.2697522.

Jadoul, Yannick, Thompson, Bill, and de Boer, Bart. Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, 71:1–15, 2018. doi:10.1016/j.wocn.2018.07.001.

Järvinen, Kati, Laukkanen, Anne-Maria, and Aaltonen, Olli. Speaking a foreign language and its effect on F0. *Logopedics Phoniatrics Vocology*, 38(2):47–51, 2013. ISSN 1401-5439. doi:10.3109/14015439.2012.687764. https://doi.org/10.3109/14015439.2012.687764.

Johnson, Keith. Massive reduction in conversational American English. In Yoneyama, K. and Maekawa, K., editors, *Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium*, pages 29–54, Toyko, Japan, 2004. The National International Institute for Japanese Language. https://linguistics.berkeley.edu/~kjohnson/papers/Massive.pdf.

Johnson, Khia A. Probabilistic reduction in Spanish-English bilingual speech. In Calhoun, Sasha, Escudero, Paola, Tabain, Marija, and Warren, Paul, editors, *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 1263–1267, Melbourne, Australia, 2019.

Johnson, Khia A. Leveraging the uniformity framework to examine crosslinguistic similarity for long-lag stops in spontaneous Cantonese-English bilingual speech. In *Proceedings of Interspeech 2021*, pages 2671–2675, June 2021a. doi:10.21437/Interspeech.2021-1780.

Johnson, Khia A. SpiCE: Speech in Cantonese and English, 2021b. https://doi.org/10.5683/SP2/MJOXP3. V1.

Johnson, Khia A. and Babel, Molly. Language contact within the speaker: Phonetic variation and crosslinguistic influence. *OSF Preprints*, 2021a. doi:10.31219/osf.io/jhsfc. https://osf.io/jhsfc/.

Johnson, Khia A. and Babel, Molly. Language contact within the speaker: Phonetic variation and crosslinguistic influence. Technical report, OSF Preprints, 2021b. https://osf.io/jhsfc/.

Johnson, Khia A., Babel, Molly, Fong, Ivan, and Yiu, Nancy. SpiCE: A new open-access corpus of conversational bilingual speech in Cantonese and English. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 4089–4095, Marseille, France, 2020a. https://www.aclweb.org/anthology/2020.lrec-1.503.

Johnson, Khia A., Babel, Molly, and Fuhrman, Robert A. Bilingual Acoustic Voice Variation is Similarly Structured Across Languages. In *Proceedings of Interspeech 2020*, pages 2387–2391, 2020b. doi:10.21437/Interspeech.2020-3095. http://dx.doi.org/10.21437/Interspeech.2020-3095.

Jolliffe, I. T. *Principal Component Analysis*. Springer-Verlag, New York, 2 edition, 2002. doi:10.1007/b98835.

Ju, Min and Luce, Paul A. Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, 15(5):314–318, 2004. doi:10.1111/j.0956-7976.2004.00675.x.

Kawahara, Hideki, Agiomyrgiannakis, Yannis, and Zen, Heiga. Using instantaneous frequency and aperiodicity detection to estimate F0 for high-quality speech synthesis. In *Proceedings of the 9th ISCA Speech Synthesis Workshop*, pages 221–228, 2016. doi:10.21437/SSW.2016-36.

Keating, Patricia and Kuo, Grace. Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2):1050–1060, 2012. doi:10.1121/1.4730893.

Keating, Patricia, Kreiman, Jody, and Alwan, Abeer. A new speech database for within- and between-speaker variability. In *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 736–739, Melbourne, Australia, 2019. https://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_785.pdf.

Keshet, J., Sonderegger, M., and Knowles, T. AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction, 2014. https://github.com/mlml/autovot/. Version 0.94.

Kleinschmidt, Dave F., Weatherholtz, Kodi, and Jaeger, T. Florian. Sociolinguistic perception as inference under uncertainty. *Topics in Cognitive Science*, 10(4):818–834, 2018. doi:https://doi.org/10.1111/tops.12331.

Kreiman, Jody, Gerratt, Bruce R., Garellek, Marc, Samlan, Robin, and Zhang, Zhaoyan. Toward a unified theory of voice production and perception. *Loquens*, 1(1):e009, 2014. doi:10.3989/loquens.2014.009.

Kreiman, Jody, Lee, Yoonjeong, Garellek, Marc, Samlan, Robin, and Gerratt, Bruce R. Validating a psychoacoustic model of voice quality. *The Journal of the Acoustical Society of America*, 149(1):457–465, 2021. doi:10.1121/10.0003331.

Kruschke, John K. Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6(3):299–312, 2011. doi:10.1177/1745691611406925.

Latinus, Marianne and Belin, Pascal. Anti-voice adaptation suggests prototype-based coding of voice identity. *Frontiers in Psychology*, 2:175, 2011. doi:10.3389/fpsyg.2011.00175.

Lavner, Yizhar, Rosenhouse, Judith, and Gath, Isak. The prototype model in speaker identification by human listeners. *International Journal of Speech Technology*, 4(1):63–74, 2001. doi:10.1023/A:1009656816383.

Lee, Binna and Sidtis, Diana Van Lancker. The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, 20(3):174–185, 2017. doi:10.1080/2050571x.2016.1273572.

Lee, Jackson L. PyCantonese, 2018. https://pycantonese.org/. Version 2.2.0.

Lee, Yoonjeong and Kreiman, Jody. Within- and between-speaker acoustic variability: Spontaneous versus read speech. In *The 178th Meeting of the Acoustical Society of America*, San Diego, CA, 2019. doi:10.1121/1.5137431. Poster.

Lee, Yoonjeong and Kreiman, Jody. Language effects on acoustic voice variation within and between talkers. In *The 179th Meeting of the Acoustical Society of America*, Acoustics Virtually Everywhere, 2020. doi:10.1121/1.5146847. Poster.

Lee, Yoonjeong, Keating, Patricia, and Kreiman, Jody. Acoustic voice variation within and between speakers. *The Journal of the Acoustical Society of America*, 146(3):1568–1579, 2019. doi:10.1121/1.5125134.

Lein, Tatjana, Kupisch, Tanja, and van de Weijer, Joost. Voice onset time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism*, 20(6):732–749, 2016. doi:10.1177/1367006915589424.

Leung, Man-Tak and Law, Sam-Po. HKCAC: The Hong Kong Cantonese adult language corpus. *International Journal of Corpus Linguistics*, 6(2):305–325, 2001. doi:10.1075/ijcl.6.2.06leu.

Levi, Susannah V. Methodological considerations for interpreting the language familiarity effect in talker processing. *WIREs Cognitive Science*, 10(2):e1483, 2019. doi:10.1002/wcs.1483.

Liang, Sihua. *Language Attitudes and Identities in Multilingual China: A Linguistic Ethnography*. Springer International Publishing, 2015. doi:10.1007/978-3-319-12619_7.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 74(6):431–461, 1967. doi:10.1037/h0020279.

Liberman, Mark Y. Corpus phonetics. *Annual Review of Linguistics*, 5(1): 91–107, 2019. doi:10.1146/annurev-linguistics-011516-033830.

Lieberman, Philip and Blumstein, Sheila E. *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge Studies in Speech Science and Communication. Cambridge University Press, Cambridge, 1988. doi:10.1017/CBO9781139165952.

Lindblom, Björn and Maddieson, Ian. Phonetic universals in consonant systems. In Hyman, Larry M. and Li, Charles N., editors, *Language, Speech, and Mind: Studies in Honour of Victoria A. Fromkin*, pages 62–78. Routledge, London, 1988.

Lisker, Leigh and Abramson, Arthur S. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3):384–422, 1964. doi:10.1080/00437956.1964.11659830.

Lisker, Leigh and Abramson, Arthur S. Some effects of context on voice onset time in english stops. *Language and Speech*, 10(1):1–28, 1967. doi:10.1177/002383096701000101.

Littell, Patrick. Thank-you notes [Version 1.0: Agent focus], 2010. http://totemfieldstoryboards.org/stories/thank_you_notes/.

Llompart, Miquel and Reinisch, Eva. Acoustic cues, not phonological features, drive vowel perception: Evidence from height, position and tenseness contrasts in German vowels. *Journal of Phonetics*, 67, 2018. doi:10.1016/j.wocn.2017.12.001.

Loveday, Leo. Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech*, 24(1):71–89, 1981. doi:10.1177/002383098102400105.

Lüdecke, Daniel, Ben-Shachar, Mattan S., Patil, Indrajeet, and Makowski, Dominique. Extracting, computing and exploring the parameters of statistical models using R. *Journal of Open Source Software*, 5(53):2445, 2020. doi:10.21105/joss.02445.

Luke, Kang Kwong and Wong, May L.Y. The Hong Kong Cantonese corpus: Design and uses. *Journal of Chinese Linguistics Monograph Series*, 25: 312–333, 2015. https://www.jstor.org/stable/26455290.

Matthews, Stephen, Yip, Virginia, and Yip, Virginia. *Cantonese: A Comprehensive Grammar*. Routledge, 2013. doi:10.4324/9780203835012.

McAuliffe, Michael, Socolof, Michaela, Stengel-Eskin, Elias, Mihuc, Sarah, Wagner, Michael, and Sonderegger, Morgan. Montreal Forced Aligner, 2017. https://montrealcorpustools.github.io/Montreal-Forced-Aligner/. Version 1.0.1.

McElreath, Richard. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC, Boca Raton, 2 edition, 2020. doi:10.1201/9780429029608.

McMurray, Bob, Tanenhaus, Michael K., and Aslin, Richard N. Gradient effects of within-category phonetic variation on lexical access. *Cognition*, 86(2): B33–B42, 2002. doi:10.1016/S0010-0277(02)00157-9.

Ménard, Lucie, Schwartz, Jean-Luc, and Aubin, Jérôme. Invariance and variability in the production of the height feature in French vowels. *Speech Communication*, 50(1):14–28, 2008. doi:10.1016/j.specom.2007.06.004.

Mielke, Jeff. A phonetically based metric of sound similarity. *Lingua*, 122(2): 145–163, 2012. doi:10.1016/j.lingua.2011.04.006.

Mielke, Jeff and Nielsen, Kuniko. Voice onset time in English voiceless stops is affected by following postvocalic liquids and voiceless onsets. *The Journal of the Acoustical Society of America*, 144(4):2166–2177, 2018. doi:10.1121/1.5059493.

Munson, Benjamin and Babel, Molly. The phonetics of sex and gender. In Katz, William F. and Assmann, Peter F., editors, *The Routledge Handbook of Phonetics*. Routledge, 2019. doi:10.4324/9780429056253_19.

Munson, Benjamin, Edwards, Jan, Schellinger, Sarah K, Beckman, Mary E, and Meyer, Marie K. Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical linguistics & phonetics*, 24(4-5):245–260, 2010.

Myers-Scotton, Carol. The matrix language frame model: Developments and responses. In *Codeswitching Worldwide*, volume 126 of *Trends in Linguistics. Studies and Monographs*. De Gruyter Mouton, 2011. doi:10.1515/9783110808742.23.

Nagy, Naomi. A multilingual corpus to explore variation in language contact situations. *Rassegna Italiana di Linguistica Applicata*, 43(1-2):65–84, 2011. http://digital.casalini.it/10.1400/190440.

Navarro, Danielle. *Learning statistics with R: A tutorial for psychology students and other beginners*. University of Adelaide, Adelaide, Australia, 2015. http://ua.edu.au/ccs/teaching/lsr. Version 0.5.

Ng, Manwa L., Hsueh, Gigi, and Sam Leung, Cheung-Shing. Voice pitch characteristics of Cantonese and English produced by Cantonese-English bilingual children. *International Journal of Speech-Language Pathology*, 12 (3):230–236, 2010. doi:10.3109/17549501003721080.

Ng, Manwa L, Chen, Yang, and Chan, Ellen YK. Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers—a long-term average spectral analysis. *Journal of Voice*, 26(4):e171–e176, 2012. doi:10.1016/j.jvoice.2011.07.013.

Ng, Raymond W. M., Kwan, Alvin C.M., Lee, Tan, and Hain, Thomas. ShefCE: A Cantonese-English bilingual speech corpus for pronunciation assessment. In *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5825–5829, 2017. doi:10.1109/ICASSP.2017.7953273.

Nieuwenhuis, Rense, Manfred, te Grotenhuis, and Pelzer, Ben. Weighted effect coding for observational data with wec. *The R Journal*, 9(1):477, 2017. doi:10.32614/rj-2017-017.

Nygaard, Lynne C. and Pisoni, David B. Talker-specific learning in speech perception. *Perception & Psychophysics*, 60(3):355–376, 1998. doi:10.3758/bf03206860.

Olson, Daniel J. The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association*, 46(3): 263–285, 2016. doi:10.1017/S0025100315000468.

Ordin, Mikhail and Mennen, Ineke. Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research*, 60(6):1493–1506, 2017. doi:10.1044/2016_JSLHR-S-16-0315.

Orena, Adriel John, Polka, Linda, and Theodore, Rachel M. Identifying bilingual talkers after a language switch: Language experience matters. *The Journal of the Acoustical Society of America*, 145(4):EL303–EL309, 2019. doi:10.1121/1.5097735.

Panayotov, Vassil, Chen, Guoguo, Povey, Daniel, and Khudanpur, Sanjeev. Librispeech: An ASR corpus based on public domain audio books. In *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2015. doi:10.1109/icassp.2015.7178964.

Perrachione, Tyler K. Recognizing speakers across languages. In Frühholz, Sascha and Belin, Pascal, editors, *The Oxford Handbook of Voice Perception*, pages 514–538. Oxford University Press, 2018. doi:10.1093/oxfordhb/9780198743187.013.23.

Perrachione, Tyler K., Furbeck, Kristina T., and Thurston, Emily J. Acoustic and linguistic factors affecting perceptual dissimilarity judgments of voices. *The Journal of the Acoustical Society of America*, 146(5):3384–3399, 2019. doi:10.1121/1.5126697.

Pitt, Mark A., Johnson, Keith, Hume, Elizabeth, Kiesling, Scott, and Raymond, William. The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability. *Speech Communication*, 45(1):89–95, 2005. doi:10.1016/j.specom.2004.09.001.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. http://www.R-project.org/.

Reinisch, Eva, Weber, Andrea, and Mitterer, Holger. Listeners retune phoneme categories across languages. *Journal of Experimental Psychology: Human Perception and Performance*, 39(1):75–86, 2013. doi:10.1037/a0027979.

Revelle, William. psych: Procedures for psychological, psychometric, and personality research, 2021. https://CRAN.R-project.org/package=psych. Version 2.1.3.

Ryabov, Rashel, Malakh, Marcella, Trachtenberg, Malka, Wohl, Sherrie, and Oliveira, Gisele. Self-perceived and acoustic voice characteristics of Russian-English bilinguals. *Journal of Voice*, 30(6):772.e1–772.e8, 2016. doi:10.1016/j.jvoice.2015.11.009.

Samuel, Arthur G. Psycholinguists should resist the allure of linguistic units as perceptual units. *Journal of Memory and Language*, 111:104070, 2020. doi:10.1016/j.jml.2019.104070.

Sancier, Michele L. and Fowler, Carol A. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25(4):421–436, 1997. doi:10.1006/jpho.1997.0051.

Shue, Yen-Liang, Keating, Patricia, Vicenik, Chad, and Yu, Kristine. VoiceSauce: A program for voice analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences*, volume 3, pages 1846–1849, Hong Kong, 2011. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011.

Simonet, Miguel. The phonetics and phonology of bilingualism. In *Oxford Handbooks Online*. Oxford University Press, 2016. doi:10.1093/oxfordhb/9780199935345.013.72.

Simonet, Miquel and Amengual, Mark. Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. *International Journal of Bilingualism*, 24(2):208–221, 2019. doi:10.1177/1367006919826388.

Sjölander, Kåre. The Snack Sound Toolkit, 2004. https://www.speech.kth.se/snack/.

Sloetjes, Han and Wittenburg, Peter. Annotation by category: ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*, Marrakech, Morocco, 2008.

Soto-Faraco, Salvador, Navarra, Jordi, Weikum, Whitney M., Vouloumanos, Athena, Sebastián-Gallés, Núria, and Werker, Janet F. Discriminating languages by speech-reading. *Perception & Psychophysics*, 69(2):218–231, 2007. doi:10.3758/BF03193744.

Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual*, 2021. https://mc-stan.org.

Statistics Canada. Proportion of mother tongue responses for various regions in Canada, 2016 Census, 2017. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dv-vd/lang/index-eng.cfm.

Stewart, Douglas and Love, William. A general canonical correlation index. *Psychological Bulletin*, 70(3, pt.1):160–163, 1968. doi:10.1037/h0026143.

Stuart-Smith, Jane, Sonderegger, Morgan, Rathcke, Tamara, and Macdonald, Rachel. The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*, 6(3-4):505–549, 2015. doi:10.1515/lp-2015-0015.

Sun, Junyi. jieba, 2020. https://github.com/fxsjy/jieba. Version 0.42.1.

Sun, Xuejing. Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio. In *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages I–333–I–336, 2002. doi:10.1109/ICASSP.2002.5743722.

Sundara, Megha, Polka, Linda, and Baum, Shari. Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, 9(1): 97–114, 2006. doi:10.1017/S1366728905002403.

Tabachnick, Barbara G. and Fidell, Linda S. *Using Multivariate Statistics*. Pearson Education, Inc., 6 edition, 2013.

Tanner, James, Sonderegger, Morgan, Stuart-Smith, Jane, and Fruehwald, Josef. Toward "English" phonetics: Variability in the pre-consonantal voicing effect across English dialects and speakers. *Frontiers in Artificial Intelligence*, 3, 2020. doi:10.3389/frai.2020.00038.

Tse, Holman. *Beyond the Monolingual Core and out into the Wild: A Variationist Study of Early Bilingualism and Sound Change in Toronto Heritage Cantonese*. Doctoral dissertation, University of Pittsburgh, Pittsburgh, PA, 2019. http://d-scholarship.pitt.edu/35721/.

Tsui, Rachel Ka-Ying, Tong, Xiuli, and Chan, Chuck Siu Ki. Impact of language dominance on phonetic transfer in Cantonese–English bilingual language switching. *Applied Psycholinguistics*, 40(1):29–58, 2019. doi:10.1017/S0142716418000449.

Turk, M.A. and Pentland, A.P. Face recognition using eigenfaces. In *Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991. doi:10.1109/cvpr.1991.139758.

Vasishth, Shravan, Nicenboim, Bruno, Beckman, Mary E., Li, Fangfang, and Kong, Eun Jong. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, 71:147–161, 2018. doi:10.1016/j.wocn.2018.07.008.

Voigt, Rob, Jurafsky, Dan, and Sumner, Meghan. Between- and within-speaker effects of bilingualism on f0 variation. In *Proceedings of Interspeech 2016*, pages 1122–1126, San Francisco, CA, 2016. doi:10.21437/Interspeech.2016-1506.

Wei, Li. Translanguaging as a practical theory of language. *Applied Linguistics*, 39(1):9–30, 2018. doi:10.1093/applin/amx039.

Winterstein, Grégoire, Tang, Carmen, and Lai, Regine. CantoMap: A Hong Kong Cantonese MapTask corpus. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 2906–2913, Marseille, France, 2020. https://aclanthology.org/2020.lrec-1.355.

Wong, Wai Yi Peggy. *Syllable fusion in Hong Kong Cantonese connected speech*. Doctoral dissertation, The Ohio State University, Columbus, OH, 2006. http://rave.ohiolink.edu/etdc/view?acc_num=osu1143227948.

Xue, Steve An, Hagstrom, Fran, and Hao, JianPing. Speaking fundamental frequency characteristics of young and elderly bilingual Chinese-English speakers: A functional system approach. *Asia Pacific Journal of Speech, Language and Hearing*, 7(1):55–62, 2002. doi:10.1179/136132802805576544.

Yang, Jing. Comparison of VOTs in Mandarin–English bilingual children and corresponding monolingual children and adults. *Second Language Research*, page 0267658319851820, 2019. doi:10.1177/0267658319851820.

Yang, Yike, Chen, Si, and Chen, Xi. F0 patterns in Mandarin statements of Mandarin and Cantonese speakers. In *Proceedings of Interspeech 2020*, pages 4163–4167, 2020. doi:10.21437/Interspeech.2020-2549.

Yau, Macro. PyJyutping, 2019. https://github.com/MacroYau/PyJyutping.

Yovel, Galit and Belin, Pascal. A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences*, 17(6):263–271, 2013. doi:10.1016/j.tics.2013.04.004.

Yu, Henry. Mountains of gold: Canada, North America, and the Cantonese Pacific. In *Routledge Handbook of the Chinese Diaspora*, pages 108–121. Routledge, 2013. doi:10.4324/9780203100387.ch7.

Yuan, Jiahong, Ryant, Neville, and Liberman, Mark. Automatic phonetic segmentation in Mandarin Chinese: Boundary models, glottal features and tone. In *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2539–2543, 2014. doi:10.1109/ICASSP.2014.6854058.