Crosslinguistic (dis)similarity in Cantonese-English bilingual speech production

by

Khia Anne Johnson

B.A. Linguistics, University of Washington, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Linguistics)

The University of British Columbia
(Vancouver)

December 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Crosslinguistic (dis)similarity in Cantonese-English bilingual speech production

submitted by **Khia Anne Johnson** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Linguistics**.

Examining Committee:

Molly Babel, Linguistics, UBC
Supervisor

Kathleen Currie Hall, Linguistics, UBC
Supervisory Committee Member

Márton Sóskuthy, Linguistics, UBC
Supervisory Committee Member

TBD, Linguistics, UBC
University Examiner

TBD, Department, UBC
University Examiner

TBD, Department
External Examiner

Abstract

Bilingual speech production is highly variable. This variability arises for numerous sources, ranging from the heterogeneity of linguistic experiences to crosslinguistic influence and more. This area has historically been challenging to study, given the relative lack of high-quality bilingual speech corpora and scientific inquiry that such resources enable. This dissertation introduces the SpiCE corpus of bilingual speech in Cantonese and English and describes two corpus studies assessing crosslinguistic similarity. Chapter 2 describes how the SpiCE corpus was designed, collected, transcribed, and annotated. Broadly, it comprises recordings of 34 early Cantonese-English bilinguals conversing in both languages, hand-corrected orthographic transcripts, and force-aligned phone level annotations. Chapters 3 and 4 are motivated by a desire to understand how crosslinguistic similarity in the speech signal facilitates multilingual talker identification and discrimination. Chapter 3 addresses this question at the level of voice quality. Using 24 filter and source-based acoustic measurements over all voiced speech in the interviews, principal components and canonical redundancy analyses demonstrate that while talkers vary in the degree to which they have the same "voice" across languages, all talkers show strong similarity with themselves. To a lesser extent, talkers exhibit similarities with one another, providing further support for prototype models of voice. Chapter 4 pivots to the level of sound categories. Prior work in this area emphasizes detecting crosslinguistic influence for phonetically distinct yet phonologically similar sounds. This chapter leverages the uniformity framework to "tell sounds together"—here, the long-lag stop series in Cantonese and En-

glish. Results indicate moderate patterns of uniformity within each language and weak patterns across languages. These weak patterns were further problematized by clear crosslinguistic differences for two of the sounds, which were apparent despite their proximity in the long-lag space. Yet at the same time, more of the overall variation seems to derive from individual-specific differences. Together, Chapters 3 and 4 provide evidence for talker identification and discrimination based more on voice quality than category similarity. Altogether, this dissertation provides a novel resource and highlights the necessity of doing corpus phonetics research, both for understanding productive processes and in speculating about the bases of different mechanisms in perception.

Lay Summary

Bilingual speech is highly variable—one major source of variability arises from how bilinguals' languages influence one another. This dissertation sheds light on how languages influence each other by analyzing conversations with Cantonese-English bilinguals. In addition to contributing a new open-access data set, this dissertation examines similarity across languages. The first question deals with voice: Do bilinguals have the same voice in each language? Are voices like auditory faces? In short—yes. The second question addresses whether this same group shares P, T, and K sounds across languages—that is, do bilinguals say K the same way in English and Cantonese. The answer to this question is less clear, with variability arising from language and the person. Together, these studies clarify which aspects of speech can be used to recognize individuals speaking more than one language and give insight into how languages do and do not interact in the mind.

Preface

This dissertation is original work and I am the primary author of each chapter. Additionally, I am the sole author of chapters 1, 4, and 5. All work in this dissertation was covered by the Behavioural Research and Ethics Board at the University of British Columbia under certificate H18-02017.

Chapter 2 was a collaborative effort. I conceptualized, designed, and led all parts of the corpus development process. The corpus itself was collected by Nancy Yiu, Ivan Fong, and myself. Transcription and annotation was supported by various members of the the Speech-in-Context Lab. The writing in chapter to is based on a paper published in the proceedings of the 12th Language Resources and Evaluation (Johnson et al., 2020a), for which I did the vast majority of writing.

Chapter 3 is based on a paper published in the proceedings of Interspeech 2020 (Johnson et al., 2020b). Molly Babel contributed to the conceptualization, design, writing, and revisions. Robert A. Fuhrman contributed tko the design of the methods, and suggested the addition of cannonical correlation anayses.

Chapter 4 is based on a solo-authored paper published in the proceedings of Interspeech 2021 (Johnson, 2021b). Molly Babel provided early feedback in the design of the study, as well as feedback on an early version of the paper.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I am very fortunate to have had a deep support network throughout my degree at UBC and while writing this dissertation. First and foremost, my gratitude goes to Molly Babel—my advisor, mentor, collaborator, and friend. Molly supported me intellectually, creatively, and financially at all stages of my graduate career and dissertation writing adventure. I somehow stumbled into the perfect advising relationship, and I am so glad I did. I must also extend my gratitude to my committee members. Kathleen Currie Hall and Márton Sóskuthy both inspire me in different ways—pushing me to think deeper, pursue open science practices, develop my technical expertise, and so much more.

This dissertation included creating a speech corpus from scratch. I would not have been able to do this without a fantastic team of research assistants. Ivan Fong and Nancy Yiu were active during all stages of the project—designing, recruiting, recording, and transcribing. They did it all. I am also deeply grateful to everyone who contributed to the corpus, including Katherine Lee, Kristy Chan, Natália Oliveira Ferreira, Michelle To, Rachel Ching Fung Wong, Christina Sen, Ariana Zattera, and Rachel Soo.

Creating the corpus would not have been possible without support from the UBC Public Scholars Initiative. I was able to submit and ultimately win an award, largely thanks to Serbulent Turan's tireless support of public scholars and Molly and Kathleen's encouragement to dream big.

I would also like to thank everyone in the Speech-in-Context Lab for giving me feedback on earlier versions of my chapters, for engaging with my ideas, and for

being the first group of people to use the corpus! My dissertation also benefitted from feedback from the audiences of LREC, Interspeech, ASA meetings, and the Cantonese psycholinguistics workshop.

Writing is hard and having a community helps. Thank you to Gloria Mellesmoen for our many writing sessions, even those where we didn't write. It was the support I clearly needed, and it helped me keep a level head throughout the grad school roller coaster. While Gloria was the only person I managed regular writing time with, having an online writing community via 100 Days and Twitter was one of the more pleasant surprises of the pandemic times. Knowing that we were all in it together made it all better.

Thank you to the grad students ahead of me in the linguistics program for showing me the ropes, answering my questions, going out for beers, and being all-around awesome people: Megan Keough, Oksana Tkachman, Emily Sadlier-Brown, Alexis Black, Natalie Weber, and many others.

And lastly, to my friends and family. My parents have always supported me in following my linguistics career, and supporting me in this Ph.D. was no exception. Little did they know back when I started how this would launch my career. Thank you, in particular, to my amazing partner and husband, Brennan, for supporting me even when that meant tearing me away from a midnight writing deadline for fancy takeout. That extra mile means the world to me. And to Pepper, dog and writing companion extraordinaire—she was impeccable in reminding me to take writing breaks, ideally outside. 15/10 writing companion. [Insert dog photo here].

Chapter 1

Introduction

Bilingual and monolingual linguistic experience differs drastically and is captured in 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 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. What aspects of sound are shared across languages? And, given that speech generally occurs in a communicative context, what is available in the multilingual speech signal to facilitate processes like talker identification and processing in more than one language? The studies presented in this dissertation approach this larger question at different levels in the phonetic space but share motivation in speech perception. This introduction motivates and ties together Chapter 3 (voice quality) and Chapter 4 (sound categories). Each level of phonetic variation has been proposed to account for how listeners can track a talker across languages.

As such, this introduction will be brief, with a majority of the literature 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 a 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 the Language Familiarity Effect in 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 are necessary for 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

In the most general sense, a bilingual is someone with knowledge of two or more languages (Grosjean, 1989). Different types of bilinguals are best described on a continuum from first language (L1) to second language (L2) dominance, the bookends of which are monolinguals and replacive bilinguals, with balanced bilinguals in the middle (Gertken et al., 2014). Using a continuum effectively reflects the heterogeneous nature of bilingualism. Dominance and patterns of use are affected by factors such as age of acquisition, immersion environment, frequency, social and communicative context (Gertken et al., 2014).

Much of the bilingualism literature focuses on early bilinguals in order to draw a distinction with learners. Typically, early bilinguals have learned both languages from early childhood. A common cutoff is age five, or the age at which children begin regularly attending primary school (Amengual, 2017). Regardless of when bilinguals acquire a language, they do not necessarily use their languages in the same domains. For example, a Cantonese-English bilingual in Vancouver, Canada might use English at school and Cantonese at home. These kinds of division make for markedly different linguistic experience across groups of bilinguals, as well as in comparison to local monolingual populations. Bilingual linguistic experience differs in many other ways, including code-switching (Fricke et al., 2016), immersion environment, and formal instruction (Fricke et al., 2019). Each of these factors has a demonstrated effect on speech production. Given the sheer heterogeneity

within and across bilingual populations, there may not always be an appropriate monolingual comparison groups. 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 even necessarily appropriate, 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 bilingual populations with different characteristics (e.g., Brown and Harper, 2009).

One of the major outcomes of this experience, as noted above, is a shared phonetic space, in which bilinguals presumably (i.e., are hypothesized to) use similar voice quality to produce similar sound categories. The literature discussing similarity in voice quality and sound categories will be reviewed in greater detail within Chapters 3 and 4, respectively.

1.2 Processing bilingual talkers

Communicating in more than one language doesn't just involve the language produced by bilingual talkers; it also involves how listeners perceive those talkers. While bilingual speech perception is a large and multifaceted field (Ingvalson et al., 2014), the clearest motivation comes from the advantage that multilingualism offers in identifying talkers. 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 vise versa when compared to monolingual English listeners. Orena et al. offer 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 a boon to bilin-

gual listeners, particularly those with language mixing experience. Such bilinguals would have extensive practice at learning how individual talkers vary as they speak in more than one language and deep familiarity with how a talker varies within and across languages.

While Orena et al. (2019) point to some prior work supporting these accounts, convincing evidence remains scarce. This dissertation directly addresses these accounts of bilingual talker identification 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 clear example of linguistic information. While using different methods and addressing different aspects of phonetics, each represents an aspect of the signal that may facilitate crosslinguistic talker identification.

1.3 But make it spontaneous

1.4 Thesis goals & research questions

Chapter 2 expands on the motivation behind studying spontaneous speech and introduces the SpiCE corpus of spontaneous bilingual speech in Cantonese and English (Johnson, 2021a). The corpus comprises a substantial portion of this dissertation.

Chapter 3 focuses on the structure of voice variation. Specifically... X.

Chapter 4 focuses on the structure of sound categories. Specifically... X.

Chapter 2

Discussion & Conclusions

...

Bibliography

Mark Amengual. 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. → page 2

Esther L. Brown and David Harper. Phonological evidence of interlingual exemplar connections. *Studies in Hispanic and Lusophone Linguistics*, 2(2): 257–274, 2009. doi:10.1515/shll-2009-1052. → page 3

Leighanne Chan, Khia Johnson, and Molly Babel. 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. → page 3

James Emil Flege and Ocke-Schwen Bohn. The revised speech learning model (SLM-r). In Ratree Wayland, editor, *Second Language Speech Learning: Theoretical and Empirical Progress*, pages 3–83. Cambridge University Press, 2021. doi:10.1017/9781108886901.002. → page 1

Melinda Fricke, Melissa M. Baese-Berk, and Matthew Goldrick. Dimensions of similarity in the mental lexicon. *Language, Cognition and Neuroscience*, 31 (5):639–645, 2016. doi:10.1080/23273798.2015.1130234. → page 2

Melinda Fricke, Megan Zirnstein, Christian Navarro-Torres, and Judith F. Kroll. Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, 22(1):200–207, 2019. doi:10.1017/S1366728918000482. → page 2

Libby M. Gertken, Mark Amengual, and David Birdsong. Assessing language dominance with the Bilingual Language Profile. In Pascale Leclercq, Amanda

Edmonds, and Heather Hilton, editors, *Measuring L2 proficiency: Perspectives from SLA*, pages 208–225. Multilingual Matters, Bristol, UK, 2014. → page 2

François Grosjean. 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. → pages 1, 2, 3

Erin M. Ingvalson, Marc Ettlinger, and Patrick C. M. Wong. Bilingual speech perception and learning: A review of recent trends. *International Journal of Bilingualism*, 18(1):35–47, 2014. doi:10.1177/1367006912456586. → page 3

K. A. Johnson. SpiCE: Speech in Cantonese and English, 2021a. URL https://doi.org/10.5683/SP2/MJOXP3. V1. → page 4

Khia A. Johnson. Leveraging the uniformity framework to examine crosslinguistic similarity for long-lag stops in spontaneous Cantonese-English bilingual speech. Technical report, OSF Preprints, June 2021b. URL https://osf.io/4p23e/. → page vi

Khia A. Johnson, Molly Babel, Ivan Fong, and Nancy Yiu. 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. European Language Resources Association. URL https://www.aclweb.org/anthology/2020.lrec-1.503. → page vi

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

Min Ju and Paul A. Luce. 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. → page 3

Adriel John Orena, Linda Polka, and Rachel M. Theodore. 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. → pages 3, 4

Miquel Simonet and Mark Amengual. Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. *International Journal of Bilingualism*, 24(2):208–221, 2019. doi:10.1177/1367006919826388. → page 3