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.