Impact of language dominance on phonetic transfer in Cantonese–English bilingual language switching

RACHEL KA-YING TSUI, XIULI TONG, and CHUCK SIU KI CHAN
The University of Hong Kong

ADDRESS FOR CORRESPONDENCE
Xiuli Tong, Room 804C, Meng Wah Complex, Division of Speech and Hearing Sciences, Faculty
of Education, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: xltong@hku.hk

ABSTRACT
Bilinguals are susceptible to interaction between their two phonetic systems during speech processing.
Using a language-switching paradigm, this study investigated differences in phonetic transfer of
Cantonese–English bilingual adults with various language dominance profiles (Cantonese-dominant,
English-dominant, and balanced bilinguals). Measurements of voice onset time revealed that
unbalanced bilinguals and balanced bilinguals responded differently to language switching. Among
unbalanced bilinguals, production of the dominant language shifted toward the nondominant language,
with no effect in the opposite direction. However, balanced bilinguals' speech production was
unaffected by language switching. These results are analogous to the inhibitory control model,
suggesting an asymmetrical switch cost of language switching at the phonetic level of speech
production in unbalanced bilinguals. In contrast, the absence of switch cost in balanced bilinguals
implies differences in the mechanism underlying balanced bilinguals' and unbalanced bilinguals'
speech production.

Keywords: bilingual speech production; inhibitory control model; language dominance; phonetic
interaction

As bilingualism becomes an increasingly common reality in today's world (Lai, 2001), bilingual speakers have to cope with the differences between their native language (L1) and second language (L2) phonetic systems. A prominent phenomenon observed in bilingual speakers is language switching, which involves switching between two languages in daily conversation. During the occurrence of language switching, phonetic interaction is expected between a bilingual speaker's L1 and L2 (e.g., Antoniou, Best, Tyler, & Kroos, 2011; Flege, 1995; Goldrick, Runnqvist, & Costa, 2014; Olson, 2013, 2016; Simonet, 2014). Research has demonstrated that bilinguals can maintain a separation between their two languages in terms of phonetic production (e.g., Antoniou, Best, Tyler, & Kroos, 2010; Goldrick et al., 2014; Macleod & Stoel-Gammon, 2005; Olson, 2013, 2016). However, bilinguals differ in their patterns of language use and proficiency, which contributes to the differences found in their speech production

(Flege, Schirru, & MacKay, 2003; Guion, Flege, & Loftin, 2000; Hazan & Boulakia, 1993). Relatively few studies have examined the cross-language phonetic interaction in language switching among bilinguals whose dominance and proficiency of their L1 and L2 languages differ. Thus, the present study investigates the effect of language dominance profiles on transient phonetic interaction, with a particular focus on voice onset time (VOT) for bilabial, alveolar, and velar stops during speech production.

PHONETIC INTERACTION IN VOT

VOT, the time interval between the release of a plosive and the beginning of voicing of a following vowel (Lisker & Abramson, 1964), has been the focus of phonetic interaction studies because it embodies language-specific properties (e.g., Antoniou et al., 2011; Balukas & Koops, 2015; Goldrick et al., 2014; Olson, 2013, 2016; Piccinini & Arvaniti, 2015). Specifically, VOT is an important acoustic correlate to the distinction of voiced-voiceless phonetic contrasts across the world's languages (Lisker & Abramson, 1964; Maddieson, 1984). Three major classifications have been identified to characterize the relationship between voicing and the time interval until the onset of voicing. VOT values less than 0 ms are referred to as lead VOT; those between 0 and 30 ms are referred to as short-lag VOT; and those greater than 30 ms are referred to as long-lag VOT (Lisker & Abramson, 1964; Maddieson, 1984). The classification to which a language belongs depends on the language-specific range of VOT variations (Auzou et al., 2000; Cho & Ladefoged, 1999). For example, while English has a two-way voicing distinction (voiced vs. voiceless) contrasting between short-lag and long-lag VOTs, Cantonese maintains the voiceless unaspirated-voiceless aspirated contrast between short-lag and long-lag VOTs (Lisker & Abramson, 1964).

Because languages are different in terms of their VOT ranges, bilinguals inevitably have to accommodate to the language-specific VOT settings in each of their languages. Previous research has generally found that bilingual speakers can establish separate language-specific phonetic systems for their two languages in terms of VOT production. In that way, their production of VOTs in each of their languages corresponds to the VOT values of that language (e.g., Antoniou et al., 2010; Goldrick et al., 2014; Macleod & Stoel-Gammon, 2005; Olson, 2013; Piccinini & Arvaniti, 2015; Sundara, Polka, & Baum, 2006). For example, Antoniou et al. (2010) compared Greek–English bilinguals' production of /p, t, b, d/ stop-voicing distinctions with Greek monolinguals and English monolinguals. They found that the Greek–English bilinguals' VOT production was comparable with their monolingual peers in the word-initial position (e.g., /**p**a, **t**a, **b**a, **d**a/). Specifically, when Greek–English bilinguals were situated in a Greek monolingual mode, they produced Greek-specific VOT values, with the voiced /b, d/ stops produced with lead VOTs and the voiceless /p, t/ stops produced with short-lag VOTs. Conversely, when they were situated in an English monolingual mode, they produced English-specific VOT values (i.e., short-lag for voiced /b, d/ stops

and long-lag for voiceless /p, t/ stops). Similar results have been reported in several other studies with different groups of bilingual speakers, such as Canadian English–Canadian French bilinguals and Spanish–English bilinguals (e.g., Bullock & Toribio, 2009; Macleod & Stoel-Gammon, 2005; Piccinini & Arvaniti, 2015; Sundara et al., 2006).

However, despite maintaining a language-specific distinction in the production of VOTs, bilinguals' production does not resemble entirely those produced by their monolingual peers (e.g., Antoniou et al., 2010; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Macleod & Stoel-Gammon, 2005; Sundara et al., 2006). For instance, Antoniou et al. (2010) found that in the more complex medial postnasal context (e.g., /aˈnpa, aˈnta, aˈnba, aˈnda/), Greek–English bilinguals' VOTs deviated from their English monolingual peers: the bilinguals produced English voiced stops with lead VOTs instead of short-lag VOTs. The authors attributed this finding to the phonetic transfer from Greek to English in this more complex phonotactic context, because voiced stops in Greek are produced with lead VOTs. Thus, although separate phonological categories have been developed within bilinguals, there still exists an L1-L2 interference effect underlying the phonetic interaction between the two languages.

TWO TYPES OF PHONETIC INTERACTION

Based on previous studies showing the impact of phonetic interaction between the two phonetic systems in speech production (e.g., Antoniou et al., 2011; Caramazza et al., 1973; Flege, 1995; Goldrick et al., 2014; Olson, 2016; Piccinini & Arvaniti, 2015; Simonet, 2014), it seems plausible that interaction leads to deviations in the speech production of one language toward the phonetic properties of another language (Goldrick et al., 2014; Simonet, 2014). Such deviations could result from two types of cross-linguistic interactions.

First, phonetic interaction occurs because of long-term traces of one language influencing the other (Simonet, 2014). This is a rather static process that presumes the interference from the L1 underlies the mislearning of nonnative (L2) sound categories. As postulated in Flege's speech learning model (SLM; 1995), the L1 and L2 phonological systems exist in a common phonetic space within a bilingual speaker. As it is difficult to maintain a separation between them, the two phonological systems inevitably influence one another. Thus, SLM predicts that bilinguals tend to perceive sounds in L2 through the framework of L1 phonology. A lack of particular sound contrasts in the L1 may lead to difficulty in the perception and learning of such sound contrasts in the L2. A typical example is when native Cantonese speakers learning English have difficulty differentiating the front midvowels in English (/ɛ/ and /æ/) because that contrast does not exist in Cantonese (Chan & Li, 2000). This difficulty in establishing a phonetic representation of L2 sounds may eventually lead to phonetic deviation in the bilingual's production of L2. This long-term phonetic interaction has been established in an extensive body of research, including studies exploring whether bilingual speakers maintain and produce monolingual-like phonetic contrasts of

voicing distinction with language-specific VOT values (e.g., Antoniou et al., 2010; Caramazza et al., 1973; Macleod & Stoel-Gammon, 2005; Sundara et al., 2006).

In contrast, the second type of phonetic contrast is more dynamic and occurs during short-term operations. In this transient interface of phonetic interaction, the phonetic representations of both languages are activated and manipulated during online speech processing. It has been hypothesized that when speakers are using two languages in communication, the representations of both languages are activated simultaneously, creating competition between the two languages (Antoniou et al., 2011; Olson, 2016; Simonet, 2014). Research shows that, during online speech processing, the nontarget language imposes an impact on the target language, leading to a deviation of the target phonetic implementation toward the nontarget language (e.g., Antoniou et al., 2011; Goldrick et al., 2014; Olson, 2013, 2016; Simonet, 2014).

Transient phonetic interaction is of particular interest in the present study. While the factors influencing long-term phonetic interaction have been investigated (e.g., Antoniou et al., 2010; Caramazza et al., 1973; Flege, 1995; Macleod & Stoel-Gammon, 2005; Piske, MacKay, & Flege, 2001; Sundara et al., 2006), little evidence exists regarding the transient cross-linguistic interference of phonetic interaction (e.g., Antoniou et al., 2011; Balukas & Koops, 2015; Goldrick et al., 2014; Olson, 2013, 2016; Piccinini & Arvaniti, 2015). Moreover, transient phonetic interaction occurs more frequently when bilinguals switch between languages (Gollan & Ferreira, 2009). Thus, exploring the process of transient phonetic interaction is fundamentally important to understanding phonetic control in bilinguals.

Code switching has been widely used to test the occurrence and direction of transient phonetic interaction in bilinguals' speech production (e.g., Antoniou et al., 2011; Balukas & Koops, 2015; Olson, 2016; Piccinini & Arvaniti, 2015). Code switching is a speech style where fluent bilinguals shift between two or more languages during the same discourse (MacSwan, 2000). Most previous studies examined code switching by asking bilinguals to read or produce naturalistic speech that included code switches. However, code switching is a less-than-ideal way to examine phonetic control in bilingual speech production. One reason is that code switching occurs within a discourse context, such as a conversation where speech production may be altered by the speakers' pragmatic intent or by contextual factors, such as intelligibility and the perceived communication needs of their partners (Bell, Brenier, Gregory, Girand, & Jurafsky, 2002; Griffin & Bock, 2000; Olson, 2013). Even though code switching produces a seemingly natural speech environment that allows for simultaneous production of the two languages, it is difficult to determine which language the speakers' attention has been directed toward and the factors influencing their selection of language. Code switching thus provides a less restricted environment to systematically study bilingual speech production.

To overcome limitations of code switching, language switching was introduced as an alternative experimental paradigm where speakers, provided with specific language cues, switch between languages in a highly controlled environment, thereby restraining the influence of context and interlocutor effects (Olson, 2013).

One example involves a cued picture-naming task where certain colors represent specific languages. Participants are asked to name a series of pictures, with the response language varying unexpectedly across trials. In each trial, the response language is indicated by the color of the frame or background (Goldrick et al., 2014; Olson, 2013); as the color changes, participants are expected to change their response language. The language-switching paradigm prompts speakers about the target language such that their selection of language is experimentally manipulated and their attention is directed to execute speech production in a specific language.

The language-switching paradigm was first employed to examine lexical selection, in relation to the inhibitory mechanism in language control, during bilingual speech production. The paradigm measured switch cost, defined as the extent of disruption in the performance during unexpected language switching as measured by naming latency (Green, 1998). An asymmetrical switch cost was found to be modulated by language dominance. Specifically, greater naming latency occurred when bilingual participants switched from their nondominant to their dominant language rather than vice versa (Meuter & Allport, 1999). These findings suggest that more effort is required to suppress the dominant language in bilingual language control; therefore, during speech production, switching from the nondominant to the dominant language takes longer.

Based on such findings of an asymmetrical switch cost, Green (1998) proposed the inhibitory control model to account for language selection in bilinguals. According to this model, language schemas compete with each other to control output at the lexical level, with such control being reactive and inhibitory. Specifically, the more dominant the language, the higher the level of activation it receives. In addition, the choice of output language depends on the relative activation level of the language schemas, which is controlled by the level of inhibition acting on the nontarget language(s). For example, when switching from a nondominant to a dominant language, both language schemas are initially activated to compete for output, but the language schema of the dominant language is activated more because of reactivity. Thus, to produce the nondominant language on the first trial of a language-switching task, more active suppression on the dominant language schema is required.

Furthermore, the task set inertia hypothesis suggests that the language schema from the previous trial persists to affect subsequent trials (Allport, Styles, & Hsieh, 1994). Therefore, an active suppression of the dominant language schema should still be present at the second trial where a switch to the dominant language is required (Meuter & Allport, 1999). Performing the second trial requires a greater degree of dominant language schema reactivation and thus takes longer for speakers to switch to the dominant language. Considering language switching in the reverse direction (i.e., from dominant to nondominant language), weaker inhibition on the nondominant language during the first trial allows the speakers to switch easily and quickly to their nondominant language on the second trial. According to Meuter and Allport (1999), this discrepancy between the reaction times measured in the switch to the dominant language and to the nondominant language gives rise to an asymmetrical switch cost for unbalanced bilinguals.

To date, most studies have focused on language switching at the level of lexical selection, with little known about the mechanism used for phonetic control during bilingual speech production at the phonetic level. In principle, language switching should also be employed to address issues regarding phonetic control in bilingual speech production. The language-switching task requires bilingual speakers to make a lexical selection in speech production. Prior to their articulation of lexical words, bilingual speakers must select between their two languages and pick the one that correctly articulates the lexical choice they made (Kroll, Bobb, & Wodniecka, 2006). As phonetic sounds underlie the articulation of a given lexical word, the language-switching paradigm is expected to also provide a theoretical account of how well bilinguals can control their two phonetic systems during online speech production.

At the same time, as the inhibitory control model works as a mechanism accounting for language selection in bilinguals, it is expected that the prediction regarding bilinguals' phonetic control should be analogous to the predictions on lexical activation patterns made by the inhibitory control model. In theory, inhibition should also be responsible for controlling the activation of the competing phonemes between bilingual speakers' two languages (e.g., Lev-Ari & Peperkamp, 2013; Olson, 2013), indicating that bilingual speakers must also implement a switch at the phonetic level when switching between their two languages. In order to implement a phonetic realization in one language, the other language must be suppressed to allow for the selection of the target language. Evidence demonstrates that, in accounting for phonetic implementations in bilingual speakers, inhibitory control modulates the degree of the coactivation between the two phonetic systems (Lev-Ari & Peperkamp, 2013). Thus, the greater the bilingual speakers' inhibitory control, the better they are in maintaining the VOT norms in their dominant language and rejecting the influence from the nondominant language.

TRANSIENT PHONETIC INTERACTION IN LANGUAGE SWITCHING

A large body of literature has examined transient phonetic interaction in code switching by comparing bilingual speakers' phonetic implementations of their two languages, each of which falls within a different category along the VOT continuum (usually a short-lag vs. long-lag language contrast, such as Spanish vs. English; e.g., Balukas & Koops, 2015; Olson, 2016; Piccinini & Arvaniti, 2015). These studies demonstrated that switching between two languages at the point of code switching has a transient impact on the production of VOT at the phonetic level. In particular, a switch cost has been found at code-switching points where the contrasts in VOT values between the two languages decrease. Specifically, the VOT values of one language shift toward the norms of the VOT in the other language. For example, studies of VOT in spontaneous code-switching speech with Spanish–English bilinguals found a shortening effect of English long-lag VOT that shifted toward the duration of Spanish-like short-lag VOT values (e.g., Balukas & Koops, 2015; Piccinini & Arvaniti 2015). Yet, previous studies on

transient phonetic interaction in code switching have produced mixed results, including the conclusion of zero phonetic interaction between the two languages (Grosjean & Miller, 1994), a unidirectional transfer from short-lag to long-lag VOT (i.e., from Greek to English as in Antoniou et al., 2011, and from Spanish to English as in Bullock, Toribio, González, & Dalola, 2006), and a bidirectional transfer where the phonetic implementations of VOT in both languages were influenced by code switching (Bullock & Toribio, 2009; Olson, 2016; Piccinini & Arvaniti, 2015). Critically, the mixed results could have arisen from the differences in the methodological designs of the code-switching tasks or the code-switching corpus employed. As mentioned earlier, code switching is limited in providing clear evidence as to which language the speakers' attention has been directed toward; instead, the language-switching paradigm provides a more direct language environment for examining phonetic control in bilingual speech production.

However, surprisingly few studies have used a language-switching paradigm to examine the effect of transient phonetic interaction in bilinguals' speech production. Two available studies to date (i.e., Goldrick et al., 2014; Olson, 2013) have focused on Spanish–English bilinguals and compared their productions of stop sounds in both languages. While Spanish has a two-way voicing distinction (voiced vs. voiceless), contrasting between lead and short-lag VOT, English maintains the voiced-voiceless contrast between short-lag and long-lag VOT (Lisker & Abramson, 1964). Although both studies reported that language switching affected phonetic production, different directions of phonetic interaction were found.

Olson's (2013) study reported a unidirectional transfer that appeared to support the inhibitory control model for bilingual phonetic production. Olson (2013) tested both Spanish-dominant and English-dominant bilinguals in three experimental contexts: two monolingual (English and Spanish; 95% of the tokens were presented in the targeted language with the remaining 5% in the nontarget language) and one bilingual (with half of the tokens in English and the other half in Spanish). The results showed that the VOTs of Spanish trials increased during language switching for Spanish-dominant bilinguals (i.e., toward the direction of the VOT values of English, the nondominant language), while the VOTs of English trials decreased during language switching for English-dominant participants (i.e., toward the direction of the VOT values of Spanish, the nondominant language). This indicated a unidirectional phonetic transfer for both bilingual groups in which the dominant language was significantly influenced by the nondominant language during language switching. However, the effect was observed only in the monolingual context where the nondominant language was heavily biased. According to Olson (2013), language context was essential to "mitigate the effect of language switching" (p. 416). Furthermore, he took a gradient approach to interpret the inhibitory control model. Specifically, he argued that both languages in a bilingual context would be partially inhibited to almost the same extent because of the global reactive nature of inhibition. This balanced level of language inhibition explained why no asymmetrical phonetic transfer was found in the bilingual condition of his study. However, in a

monolingual context, the differences in global inhibition of the two phonetic systems (due to language context) together with local inhibition (due to language switching) led to asymmetrical phonetic transfer, as predicted by the inhibitory control model.

Contrary to Olson's (2013) study, where the dominant language was influenced by the nondominant language, Goldrick et al. (2014) reported transient phonetic interaction in the direction of the dominant to the nondominant language. In their study, native Spanish speakers who learned English as an L2 were exposed to a condition with balanced tokens of English and Spanish. They found that the English VOTs produced by the participants decreased during language switching (i.e., toward the VOT values of Spanish), but the opposite was not observed. This indicated that the participants' nondominant English production deviated toward the dominant Spanish phonetic system. The authors discussed the results in terms of an interactive theory of speech processing and reasoned that the nontarget representation was more active in switch trials than in stay trials. The partially activated nontarget representation would therefore influence phonetic processing and make the production deviate toward the nontarget language. However, it should be noted that these results contradicted the prediction of the inhibitory control model, which suggested that the switch cost should be higher when switching to a dominant language. That is, phonetic interference should be more apparent during a switch to the dominant language rather than in the opposite direction.

In contrast, the influence of the nontarget language on the production of the target language at switching found in Goldrick et al.'s (2014) study indicates the persistent activation of the L1 (Spanish) on the phonetic production of the L2 (English). Such a persistent influence could be partly explained by the use of cognate stimuli in their study. Cognates are translation equivalents that are phonologically similar between languages; for example, one of the English–Spanish cognate pairs used in Goldrick et al.'s (2014) study was *telephone* and *teléfono*. It has been widely hypothesized that the phonological overlap in cognate pairs triggers cross-language activation during speech production (e.g., Amengual, 2012; Costa, Caramazza, & Sebastián-Gallés, 2000; Jacobs, Fricke, & Kroll, 2016; Kroll, Michael, Tokowicz, & Dufour, 2002; Schwartz, Kroll, & Diaz, 2007). In particular, during the articulation of target words, cognates exert an effect on the activation of the nontarget language, so that increased phonetic influence from the nontarget language is transferred to the production of the target language (Amengual, 2012; Jacobs et al., 2016). Consistent with previous studies using cognate words (Amengual, 2012; Jacobs et al., 2016), Goldrick et al. (2014) demonstrated a stronger phonological influence of cognates when switching from the nontarget (L1 Spanish) to the target (L2 English) language. Considering the possible influence of cognates on bilingual speech production, it is critically important to examine whether the nontarget language influences the production of the target language at switching when no common cognates exist, as in Cantonese and English, the focus of the present study.

In addition to the different experimental conditions used in the two existing studies (Goldrick et al., 2014; Olson, 2013), the different language profiles of the

bilinguals involved may also account for their discrepant findings. Olson's (2013) study investigated bilinguals who acquired their L2 in adolescence, whereas Goldrick et al.'s (2014) study focused on those who acquired their L2 in early or middle childhood. It has been argued that the age of L2 acquisition affects the extent of the phonetic differences between production of L1 and L2 sounds (Antoniou et al., 2011). Moreover, it has been shown that individuals who learned an L2 in early childhood were more likely to establish L2 phonetic categories compared to those who learned an L2 later in life (Flege, 1991). Thus, the native Spanish speakers in Goldrick et al.'s (2014) study who acquired English as children might show a greater phonetic difference in their English and Spanish productions than those in Olson's (2013) study who acquired English as adolescents.

In line with the effect of age of L2 acquisition, it is important to note that bilingual speakers rarely possess an equal command of their languages (Grosjean, 1998). When one of the two languages within a bilingual becomes more accessible and activated in day-to-day life (Harris, Gleason, & Ayçiçegi, 2006), that language is referred to as the dominant language. Several factors have been identified as important in determining which language is the dominant one. For instance, language dominance is often associated with language proficiency (e.g., Birdsong, 2006; Tokowicz, Michael, & Kroll, 2004) and the age of L1 and L2 acquisition (e.g., Goldrick et al., 2014; Olson, 2013; Piccinini & Arvaniti, 2015). However, language abilities alone do not fully characterize language dominance. The relative strength of the two languages is perhaps critically determined by language preference and frequency of use. Consider the cases of immigrants who have learned their L1 early at home, but have been living in another country where their L2 is the dominant language used in society. As the L2 is more dominant in daily social situations, the frequency ratio between the use of L1 and L2 would skew toward the L2, such that the bilinguals would become L2 dominant (e.g., Antoniou et al., 2010, 2011; Antoniou, Tyler, & Best, 2012). In other words, variations in language dominance are highly dependent on the preferences made by individual bilinguals with regard to what language to use when and with whom (Pavlenko, 2004). Moreover, bilingual speakers vary in terms of language acquisition and language use. It has been reported that variations in bilinguals' command contributes to the differences found in bilinguals' speech production (Flege et al., 2003; Guion, Flege, & Loftin, 2000; Hazan & Boulakia, 1993). Language dominance is a critical factor when comparing the language performances of bilinguals and thus worthy of further investigation.

Critically, the mixed results in Olson's (2013) and Goldrick et al.'s (2014) studies leave unanswered the question as to whether bilinguals' language dominance profiles affect the direction of transient phonetic interactions. While Olson (2013) employed participants who acquired their L2 late, Goldrick et al. (2014) only investigated participants with Spanish-dominant backgrounds. This makes it necessary to examine bilinguals with different language-dominance profiles to verify the phonetic interaction effect in language switching. Thus, in the present study, we investigate the interaction between language dominance and transient phonetic interaction in Hong Kong Cantonese–English bilinguals who are either

dominant in Cantonese, dominant in English, or dominant in both (i.e., balanced bilinguals).

Hong Kong Cantonese–English bilinguals provide a fascinating case to further examine the transient phonetic interaction in bilingual speech production. First, most previous studies have largely focused on bilinguals who speak two typologically similar languages within the Indo-European family, such as English and Spanish (e.g., Balukas & Koops, 2015; Goldrick et al., 2014; Olson, 2013, 2016; Piccinini & Arvaniti, 2015) and English and Greek (Antoniou et al., 2011), so that they share linguistic similarities and have a large number of cognates in common. In the present study, Cantonese, which belongs to the Sino-Tibetan language family, shares very few linguistic commonalities with English in terms of phonology, lexicogrammar, and orthography (Li, 2017). Moreover, except for the set of loanwords borrowed from English, the number of common cognates is trivial (Li, 2017), minimizing the cognate effect found in Goldrick et al.'s (2014) study. Overall, a comparison of Cantonese–English bilinguals with varying language-dominant profiles provides a more direct examination of the mechanism underlying the transient phonetic interaction in bilingual speech production.

In addition, the differences in VOT between Cantonese and English phonetic inventories provide a unique window into the bilingual phonetic interaction between these two systems. English has a two-way voicing distinction (voiced vs. voiceless) that contrasts between short-lag and long-lag VOTs. The voiced stops have short-lag VOT values that are typically between 0 and 20 ms, whereas the voiceless stops have long-lag VOT values of approximately 65 to 120 ms (Auzou et al., 2000; Lisker & Abramson, 1964; Macken & Barton, 1980). Similarly, Cantonese maintains a two-way voicing distinction defined as voiceless unaspirated and voiceless aspirated (Bauer & Benedict, 1997; Lisker & Abramson, 1964). In terms of VOT, Cantonese contrast is similar to English contrast. Cantonese voiceless unaspirated stops have short-lag VOT values of less than 30 ms, while Cantonese voiceless aspirated stops have long-lag VOT values that typically exceed 75 ms (Clumeck, Barton, Macken, & Huntington, 1981; Lisker & Abramson, 1964).

However, it is important to note that the same plosive consonants are classified differently in Cantonese and English. The stop consonants /p, t, k/ are described as voiceless stops with long-lag VOT values in English, but as voiceless unaspirated stops with short-lag VOT values in Cantonese. Like the voicing distinction in Spanish and English, Cantonese has short-lag VOT voiceless (unaspirated) stops similar to Spanish, while English has long-lag VOT voiceless stops. Thus, these differences in VOT between English and Cantonese provide a key comparison with which to investigate phonetic interaction between the two languages.

THE PRESENT STUDY

The present study makes use of a clear phonetic difference between Cantonese and English (i.e., different VOT values for voiceless/unaspirated stops) in a

language-switching paradigm to investigate the phonetic control of early Cantonese–English bilingual adults. Based on SLM predictions (Flege, 1997), we hypothesize that early bilinguals across different language-dominance groups produce differentiated phonetic norms for L1 and L2 sounds. Under this assumption, this study addresses two research questions: (a) how do language-dominance profiles of unbalanced bilinguals influence the direction of phonetic transfer during language switching? and (b) do balanced bilinguals demonstrate phonetic transfer during language switching, and if so, in what way?

For the first question, we hypothesize an asymmetrical effect of language switching on the participants' dominant language, a prediction driven by the claims of the inhibitory control model. Under language-switching conditions, unbalanced participants would demonstrate a unidirectional phonetic transfer from the nondominant to the dominant language. Such an effect would be evidenced by the VOT values in the dominant language drifting toward the nondominant language's phonetic norm. For the second question, we hypothesize a lack of transient phonetic interaction between the two phonetic systems for those who are comparably dominant in both languages (i.e., balanced bilinguals).

METHOD

Participants

A total of 60 Cantonese–English bilingual adults were recruited from universities in Hong Kong. All participants were native speakers of Cantonese who began learning English before the age of 6. An additional 5 participants who were not native speakers of Cantonese were excluded. None of the participants reported a history of speech, language, hearing, or visual impairments.

Given that self-ratings have been reported to reliably reflect the linguistic performance of bilingual speakers (e.g., Dunn & Fox-Tree, 2009; Flege, Mackay, & Piske, 2002; Flege, Yeni-Komshian, & Liu, 1999), our participants were asked prior to the experiment to complete the Language Experience and Proficiency Questionnaire (LEAP-Q), which encompasses measures of language background and self-perceived language proficiency (Marian, Blumenfeld, & Kaushanskaya, 2007).

Based on the self-report questionnaire regarding their knowledge of Cantonese and English, participants were classified into three language background categories: 20 Cantonese dominant (mean age = 21.15 years, $SD = 0.49$), 20 English dominant (mean age = 22.90 years, $SD = 3.18$), and 20 balanced bilinguals (mean age = 22.20 years, $SD = 4.14$). Following previous research (e.g., Dunn & Fox-Tree, 2009; Flege et al., 1999, 2002), the language dominance of each participant in this study was determined by self-reported language proficiency in four different aspects: spoken proficiency, listening comprehension, daily exposure, and self-perceived accent. Each group's self-rated language abilities and language exposure in Cantonese and English, including spoken proficiency, listening comprehension, current daily exposure, and self-perceived accent, is reported in Table 1.

Table 1. *Means and standard deviations of self-reported Cantonese and English language backgrounds for Cantonese-dominant, English-dominant, and balanced participants*

	Cantonese		English		
	M	*SD*	*M*	*SD*	*t*
Cantonese-dominant group					
Spoken proficiency[a]	9.00	0.00	6.15	0.88	14.57***
Listening comprehension[b]	9.00	0.00	6.35	0.88	13.54***
Daily exposure (%)	71.50	15.57	19.35	11.94	8.94***
Self-perceived accent (%)	0.50	2.24	90.95	15.83	−24.27***
English-dominant group					
Spoken proficiency	5.75	1.33	9.00	0.00	−10.91***
Listening comprehension	6.25	0.97	9.00	0.00	−12.72***
Daily exposure (%)	30.00	15.39	64.00	18.04	−4.65***
Self-perceived accent (%)	66.70	35.24	5.00	20.07	6.92***
Balanced group					
Spoken proficiency	8.15	0.88	8.60	0.60	−1.69
Listening comprehension	8.30	0.66	8.60	0.60	−1.55
Daily exposure (%)	48.67	19.12	44.17	17.47	0.56
Self-perceived accent (%)	19.40	24.34	20.00	27.10	−0.08

Notes: [a]Likert scale (1 = *barely speak*, 9 = *native speaker*). [b]Likert scale (1 = *barely understand*, 9 = *native speaker*). ***$p < .001$.

The participants were asked to evaluate, on a 9-point Likert scale, their performance in spoken proficiency and listening comprehension (1 = *barely speak/barely understand*; 9 = *native speaker*). To be classified as bilingual, each participant must score greater than 2 (= *low proficiency*) on each item in both languages. No participant was classified as being functionally Cantonese or English monolingual as no items in either language were given a rating of 1 (= *barely speak/ understand*). The possible scores for each language ranged from 4, the lowest (2 in spoken proficiency +2 in listening comprehension), to 18, the highest (9 in spoken proficiency +9 in listening comprehension). All the bilinguals were then grouped according to their ratio of Cantonese to English proficiency, so the possible ratio range was from 0.22 (4 in Cantonese proficiency/18 in English proficiency) to 4.50 (18 in Cantonese proficiency/4 in English proficiency). Using a cutoff value of 7 (*good*) for both spoken proficiency and listening comprehension, participants with a ratio of 1.29 (18 in Cantonese proficiency/14 in English proficiency) to 4.50 (the highest) were categorized as Cantonese-dominant bilinguals, and those with 0.22 (the lowest) to 0.78 (14 in Cantonese proficiency/18 in English proficiency) as English-dominant bilinguals. Bilingual participants with a ratio between 0.78 and 1.29 were categorized as balanced bilinguals.

As reported in Table 1, Cantonese-dominant participants perceived themselves to be more proficient in Cantonese than in English in both spoken proficiency and listening comprehension (spoken proficiency: t (19) = 14.57, $p < .001$; listening comprehension: t (19) = 13.54, $p < .001$). Similarly, English-dominant participants perceived themselves to be more proficient in English than Cantonese in both spoken proficiency and listening comprehension (spoken proficiency: t (19) = –10.91, $p < .001$; listening comprehension: t (19) = –12.72, $p < .001$). It is important to note that the English-dominant participants studied at international schools in Hong Kong, and were therefore immersed in an English-speaking environment from a very young age. They were surrounded by friends with similar educational backgrounds and were more comfortable conversing in English. Balanced participants rated their Cantonese and English as comparable for both spoken proficiency and listening comprehension (spoken proficiency: t (19) = –1.69, $p = .107$; listening comprehension: t (19) = –1.55, $p = .137$).

In addition to spoken proficiency and listening comprehension, the classification of participants based on three language background categories was confirmed by two other factors, daily exposure and self-perceived accent, with (a) Cantonese-dominant participants demonstrating greater use and less accent in Cantonese compared to English, (b) English-dominant participants demonstrating greater use and less accent in English compared to Cantonese, and (c) balanced participants reporting comparable daily exposure and accent for both Cantonese and English (for details see Table 1).

To control for age of language acquisition, all participants were exposed to Cantonese from birth and acquired English before the age of 6. However, Cantonese-dominant, English-dominant, and balanced bilingual language groups significantly differed in their age of English acquisition, F (2, 57) = 5.13, $p < .01$, $\eta^2_p = .152$, with Cantonese-dominant participants ($M = 3.25$ years, $SD = 1.71$) acquiring English later than English-dominant participants ($M = 1.55$ years, $SD = 2.01$), $p < .05$, while English-dominant participants and balanced participants ($M = 1.80$ years, $SD = 1.70$) acquired English at a comparable age, $p = 1.000$.

Stimuli

The target stimuli consisted of 24 colored pictures of nonambiguous objects or scenes, with half used to elicit Cantonese words and the other half English words. All target words were noncognate to avoid the cognate facilitation effect on cross-language activation during speech production (e.g., Amengual, 2012; Goldrick et al., 2014; Jacobs et al., 2016). The target words, included in Appendix A, were all monosyllabic with voiceless/unaspirated initial stops, evenly distributed across three possible places of articulation. There were 6 bilabial (/p/), 6 alveolar (/t/), and 6 velar (/k/) stops in each language. All Cantonese targets were in tone 1 to avoid the effect of lexical tone on VOT (Chen, Peng, & Chao, 2009). A similar number of phonemes occurred in both the Cantonese targets ($M = 2.83$, $SD = 0.38$) and the English targets ($M = 2.94$, $SD = 0.42$), t (17) = 0.81, $p = .430$. A set of 36 additional pictures served as fillers.

Procedure

Following the procedure by Goldrick et al. (2014), in the familiarization block, participants were exposed to each picture's English and Cantonese label for 4000 ms to familiarize them with the names of the pictures. They were then instructed to name the pictures as accurately and quickly as possible. Afterward, they finished three practice sequences to become familiar with the testing procedures and the color–language pairings. These practice trials were not included in the data analysis.

In the testing block, the stimuli were presented visually using E-Prime experimental software 2.0 (Schneider, Eschman, & Zuccolotto, 2002), with each stimulus appearing one-by-one in short sequences of 5 to 13 pictures. Each trial consisted of a fixation mark (300 ms) followed by the picture to be named, which remained on the screen until the participants proceeded to the next trial by pressing the keyboard's "Enter" button. The color of the picture frame indicated the language that participants were to use on the current trial (red for Cantonese and blue for English). The response language was either the same as (stay trial) or different from (switch trial) the immediately preceding trial. The first picture in each sequence was always a filler item and was excluded from analyses. Only one target-initial stop was elicited in each sequence. The participants were offered a short break after every 6 sequences to avoid fatigue. In total, there were 216 tokens per participant (36 stimuli × 2 response types [stay or switch] × 3 repetitions).

The experiment was carried out individually for each participant in a sound-attenuated room (IAC single-wall booth) at the University of Hong Kong. All speech production was recorded at a sampling rate of 44.1 kHz, with a Shure SM58S microphone and an Edirol UA-25USB Audio Interface, which were connected to a Lenovo ThinkPad 4173DC9 laptop. A 10-cm mouth-to-microphone distance was maintained for each participant.

Acoustic analysis

Given that the focus of this study is the production of words with initial stops, VOT was the target of our acoustic analysis. We measured the VOT of tokens beginning with /p/, /t/, and /k/ using PRAAT 5.4.22 (Boersma & Weenink, 2015). Measurements were taken from the onset of the release burst to the onset of the following vowel, where voicing began. VOT was marked by hand and indicated by the time interval between the release of the plosive and the presence of the first periodic vibrations as seen in the waveform (Lisker & Abramson, 1964). The release of a plosive was perceived as a sharp spike where the waveform changed from quiescent to transient, while the periodic vocal fold vibrations were determined by repeating voicing cycles (Francis, Ciocca, & Yu, 2003). VOT measurement was taken in milliseconds (ms) and measured to 1 decimal place (e.g., Antoniou et al., 2010; Macleod & Stoel-Gammon, 2005). All measurements were conducted by the third author, and 10% of the tokens were selected randomly and recoded blindly by an independent rater. The interrater reliability was very high, $r = .995$, $p < .001$. Sample measurements are illustrated in Figures 1a and 1b.
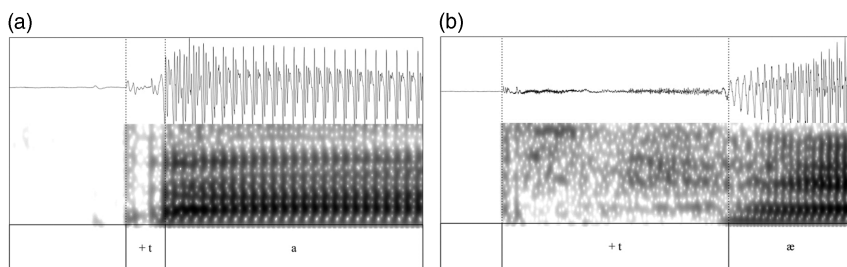
Figure 1. An example of measurement of (a) the Cantonese short-lag VOT for the alveolar stop /t/ and (b) the English long-lag VOT for the alveolar stop /t/.

Trials with production errors (including use of the wrong language, false starts, and naming errors) were eliminated from the analyses. Of the total trials, 3.86% obtained were discarded due to production errors (Cantonese-dominant group: 3.56%; English-dominant group: 5.05%; balanced bilingual groups: 2.96%). A one-way between-subjects analysis of variance demonstrated that the number of eliminated tokens per participant was similar for the Cantonese-dominant ($M = 7.70$, $SD = 8.40$), English-dominant ($M = 10.90$, $SD = 14.38$), and balanced ($M = 6.40$, $SD = 6.61$) bilingual groups, F (2, 57) = 1.00, $p = .373$.

RESULTS

Figure 2 and Table 2 show means and standard deviations of Cantonese and English VOTs for Cantonese-dominant, English-dominant, and balanced bilingual groups. To examine whether early Cantonese–English bilinguals produce differentiated phonetic norms for sounds in their two languages, a linear mixed-effects model analysis was performed using R (R Core Team, 2017) and the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). To compare the VOTs across the three places of articulation (/p/, /t/, and /k/) produced for Cantonese and English by all speakers, we entered language (Cantonese vs. English), bilingual group (Cantonese-dominant, English-dominant, or balanced), and place of articulation (/p/, /t/, and /k/) as fixed effects. In the random-effects structure, the intercept for subjects was entered as the random effect. Significance of the fixed effects in question was obtained by the analysis of variance function in the *car* package (Fox & Weisberg, 2011).

The main effect of language, χ^2 (1) = 846.20, $p < .001$, and consonant, χ^2 (2) = 48.60, $p < .001$, were found to be significant. There was also a significant interaction between language and consonant, χ^2 (2) = 22.32, $p < .001$. Post hoc tests with Tukey-adjusted comparison were conducted using the *lsmeans* package (Lenth, 2016). The post hoc analysis of the interaction between language and consonant revealed that, in general, English VOTs ($M = 79.11$ ms, $SD = 19.58$) were produced with significantly greater VOT values than Cantonese VOTs ($M = 13.12$ ms, $SD = 7.27$) across all three places of articulation, t (660) = –89.95,

$d = 4.92$, $r^2 = .92$, $p < .001$. Across both languages, it was consistently found that, overall, the velar stop /k/ ($M = 51.79$ ms, $SD = 32.88$) was produced with the greatest VOT values, while the bilabial stop /p/ ($M = 39.60$ ms, $SD = 33.68$) was produced with the least VOT values (/t/: $M = 46.96$ ms, $SD = 40.54$; /p/ versus /t/: $t (660) = -8.20$, $d = -0.20$, $r^2 = .09$, $p < .001$; /p/ versus /k/: $t (660) = -13.57$, $d = -0.37$, $r^2 = .22$, $p < .001$; /t/ versus /k/: $t (660) = -5.37$, $d = -0.13$, $r^2 = .04$, $p < .001$). As shown in Figure 2, the same pattern was found for each bilingual group, which indicated a clear categorical difference between Cantonese VOTs and English VOTs produced by all bilingual speakers. There was also a marginally significant interaction between language and bilingual group, χ^2 $(2) = 5.59$, $p = .061$. Given the inherent VOT differences in Cantonese and English initial stops, Cantonese and English VOTs were compared separately in the following analyses.

Two linear mixed-effects model analyses (one per language: Cantonese and English) were conducted to investigate whether the bilingual groups' production of VOTs varied as a function of the different response types elicited on picture-naming trials. In each model, the intercept for subjects was entered as the random effect. Bilingual group (Cantonese dominant, English dominant, or balanced), response type (stay trial vs. switch trial), and consonant (/p/, /t/, and /k/) were entered as fixed effects.

For Cantonese trials, the main effect of consonant was found significant, χ^2 $(2) = 407.61$, $p < .001$. Post hoc tests with Tukey-adjusted comparisons revealed that the VOT values of the Cantonese velar stop /k/ ($M = 21.75$ ms, $SD = 5.66$) were significantly greater than those of both the Cantonese bilabial stop /p/ ($M = 8.59$ ms, $SD = 2.67$), $t (300) = 41.43$, $d = 3.16$, $r^2 = .85$, $p < .001$,
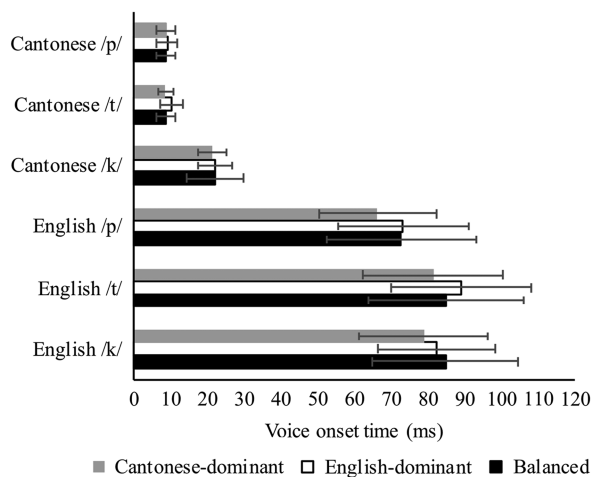

Figure 2. Means and standard deviations of Cantonese and English VOTs for the stop consonants (/p/, /t/, and /k) by Cantonese-dominant, English-dominant, and balanced bilinguals. The error bars indicate 1 SD above or below the mean VOT.

Table 2. *Range, mean, and standard deviation of Cantonese VOT and English VOT for the three language groups: Cantonese dominant, English dominant, and balanced bilingual*

	Cantonese VOT (ms)						English VOT (ms)					
	Stay trial			Switch trial			Stay trial			Switch trial		
	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*
Cantonese dominant	3.7–25.5	11.74	6.53	5.1–30.2	13.61	7.05	43.7–133.1	75.10	18.10	40.7–127.2	75.68	19.27
English dominant	4.5–39.1	13.42	7.15	4.5–30.0	13.82	6.94	42.2–137.7	83.27	18.93	41.7–126.3	79.43	18.17
Balanced	5.0–46.2	13.05	8.04	4.3–43.8	13.06	7.92	44.5–137.4	80.31	21.14	45.4–134.7	80.87	21.16

and the Cantonese alveolar stop /t/ ($M = 9.02$ ms, $SD = 2.76$), t (300) = 40.06, $d = 3.02$, $r^2 = .84$, $p < .001$. No significant difference was found between the Cantonese bilabial stop /p/ and the Cantonese alveolar stop /t/, t (300) = –1.37, $d = –0.16$, $r^2 = .01$, $p = .357$. However, the main effect of bilingual group was not significant, χ^2 (2) = 2.90, $p = .235$, nor was the main effect of response type, χ^2 (1) = 0.29, $p = .589$. No other significant interactions were found: Bilingual Group × Consonant, χ^2 (4) = 2.11, $p = .716$; Response Type × Consonant, χ^2 (2) = 0.46, $p = .796$; Bilingual Group × Response Type × Consonant, χ^2 (4) = 1.00, $p = .910$. Yet, the crucial result was a significant interaction between bilingual group and response type, χ^2 (2) = 6.58, $p < .05$, which suggests that the differences in VOT between Cantonese stay and switch trials vary across bilingual groups.

For English trials, the main effect of consonant was found to be significant, χ^2 (2) = 39.77, $p < .001$. Post hoc tests with Tukey-adjusted comparisons revealed that the VOT value of the English bilabial stop /p/ ($M = 70.61$ ms, $SD = 18.23$) was the smallest (/p/ versus /t/: t (300) = –17.16, $d = –0.75$, $r^2 = .50$, $p < .001$; /p/ versus /k/: t (300) = 13.46, $d = –0.62$, $r^2 = .38$, $p < .001$) and that of the English alveolar stop /t/ ($M = 84.90$ ms, $SD = 19.75$) was slightly longer than that of the English velar stop /k/ ($M = 81.82$ ms, $SD = 17.88$), t (300) = –3.70, $d = 0.16$, $r^2 = .04$, $p < .001$. No other significant effects were found, including bilingual group, χ^2 (2) = 1.22, $p = .542$, response type, χ^2 (1) = 1.57, $p = .211$, and interactions: Bilingual Group × Consonant, χ^2 (4) = 3.84, $p = .428$; Response Type × Consonant, χ^2 (2) = 1.47, $p = .479$; Bilingual Group × Response Type × Consonant, χ^2 (4) = 1.09, $p = .895$. Nevertheless, the crucial result was a marginally significant interaction between bilingual group and response type, χ^2 (2) = 5.17, $p = .075$, indicating that the differences in VOTs between English stay and switch trials vary among the three bilingual groups.

In both the Cantonese and English models, the interaction between bilingual group and response type was found to be (marginally) significant. Because we were interested in testing whether difference in language dominance influences the direction of transient phonetic transfer in response to language switching across the two languages within bilingual speakers, the planned contrasts with Tukey-adjusted comparisons were conducted to test a priori comparisons for the VOTs produced in stay trials versus switch trials within each language across the three bilingual groups.

Cantonese-dominant speakers produced significantly greater VOTs in Cantonese switch trials ($M = 13.61$ ms, $SD = 7.05$) than in stay trials ($M = 11.74$ ms, $SD = 6.53$), t (300) = –4.15, $d = 0.28$, $r^2 = .05$, $p < .001$. These results suggest that when Cantonese-dominant speakers switched from English to Cantonese, their VOTs for Cantonese initial stop consonants increased compared to their VOTs when staying in Cantonese from one trial to the next. However, no significant difference was found in Cantonese-dominant speakers' VOT production for English stay trials ($M = 75.10$ ms, $SD = 18.10$) and switch trials ($M = 75.68$ ms, $SD = 19.27$), t (300) = –0.49, $d = –0.03$, $r^2 = .00$, $p = .997$. These results suggest no obvious effect of language switching in English trials for Cantonese-dominant

speakers, and imply that Cantonese-dominant speakers are susceptible to the effect of language switching only in Cantonese trials.

For English-dominant speakers, no significant difference in VOTs was found between Cantonese stay trials ($M = 13.42$ ms, $SD = 7.15$) and switch trials ($M = 13.82$ ms, $SD = 6.94$), t (300) $= -0.88$, $d = -0.06$, $r^2 = .00$, $p = .950$. This indicates that English-dominant speakers were not susceptible to the effect of language switching in Cantonese trials. In contrast, English-dominant speakers produced significantly smaller VOTs in English switch trials ($M = 79.43$ ms, $SD = 18.17$) than in stay trials ($M = 83.27$ ms, $SD = 18.93$), t (300) $= 3.26$, $d = 0.21$, $r^2 = .03$, $p < .05$. These results suggest that when English-dominant speakers switched from Cantonese to English, their VOTs for English initial stops decreased compared to their VOTs when staying in English. This implies that English-dominant speakers are susceptible to the effect of language switching only in English trials.

For the group of balanced speakers, no difference in VOTs was found between Cantonese stay trials ($M = 13.05$ ms, $SD = 8.04$) and switch trials ($M = 13.06$ ms, $SD = 7.92$), t (300) $= -0.02$, $d = -0.00$, $r^2 = .00$, $p = 1.000$, nor between English stay trials ($M = 80.31$ ms, $SD = 21.14$) and switch trials ($M = 80.87$ ms, $SD = 21.16$), t (300) $= -0.48$, $d = -0.03$, $r^2 = .00$, $p = .997$. These results suggest that balanced bilingual speakers performed similarly in stay and switch trials in both Cantonese and English, and that no obvious effect of language switching occurred in either language.

In addition to the analyses of the amalgamated data, the data from each individual speaker was examined by averaging the VOT across the three places of articulation and separately for each response type (stay trial vs. switch trial) within each language (Cantonese and English).[1] Individual bilingual speakers' mean VOT values (in milliseconds) are listed in Tables 3, 4, and 5 for the Cantonese-dominant group, the English-dominant group, and the balanced-bilingual group, respectively. As revealed by the linear mixed-effects model analyses, the group of Cantonese-dominant bilingual speakers maintained longer VOT values for Cantonese switch trials than stay trials when switching from their nondominant English language to Cantonese. Individual analyses revealed that, within the Cantonese-dominant group, 15 of the 20 participants implemented a shift in Cantonese switch trials toward the direction of English-like VOT values ($M = 2.7$ ms, ranged from 0.3 to 5.3 ms). The remaining 5 Cantonese-dominant speakers did not produce longer VOT values for Cantonese switch trials when switching from English ($M = -0.5$ ms, ranged from -0.8 to -0.2 ms). As for the English-dominant group, the linear mixed-effects model analyses suggested a trend where English-dominant speakers were found to produce shorter VOT values in English switch trials, thus shifting toward Cantonese-like VOT values. It was noticed that 17 of the 20 English-dominant bilingual speakers followed this trend and produced shorter VOTs in English switch trials when switching to Cantonese ($M = -5.0$ ms, ranged from -13.5 to -0.8 ms). The remaining three English-dominant speakers did not produce shorter English VOTs in English switch trials than stay trials ($M = 2.9$ ms, ranged from 2.9 to 3.1 ms).

Table 3. *Mean VOT values for each Cantonese-dominant bilingual speaker*

Speaker	Cantonese VOT (ms)			English VOT (ms)		
	Stay	Switch	Switch–Stay	Stay	Switch	Switch–Stay
C01	12.5	14.5	2.0	117.9	122.2	4.3
C02	14.5	15.2	0.7	65.8	62.3	–3.5
C03	9.1	9.5	0.3	86.9	74.5	–12.4
C04	15.2	14.4	–0.8	90.9	88.9	–2.1
C05	11.5	11.3	–0.2	76.0	77.7	1.8
C06	11.3	10.4	–0.8	58.4	62.9	4.5
C07	11.2	10.8	–0.5	60.3	65.2	4.9
C08	10.1	11.8	1.7	71.9	75.9	4.0
C09	13.0	15.1	2.1	80.7	81.3	0.6
C10	8.7	9.2	0.4	66.1	72.0	5.9
C11	12.7	16.4	3.7	52.2	50.8	–1.3
C12	12.3	15.0	2.7	69.7	69.9	0.3
C13	9.7	13.0	3.3	70.9	71.2	0.4
C14	8.4	10.9	2.6	58.6	57.3	–1.3
C15	14.9	19.2	4.3	93.8	91.3	–2.5
C16	10.4	15.7	5.3	68.7	69.7	1.1
C17	9.3	14.1	4.8	65.9	62.1	–3.8
C18	14.3	16.6	2.3	97.9	108.7	10.8
C19	12.6	12.2	–0.4	85.2	84.9	–0.3
C20	13.2	16.9	3.7	64.4	64.8	0.3
Mean	11.74	13.61	1.9	75.10	75.68	0.6

DISCUSSION

The present study set out to investigate the effect of language dominance on transient phonetic interaction between the two phonetic systems of early Cantonese–English bilinguals in Hong Kong. We found that the early bilinguals had established clearcut phonetic categories for their L1 and L2 sounds, as indicated by the categorical difference in VOT observed when they produced the same initial stops in Cantonese versus English trials. However, Cantonese-dominant, English-dominant, and balanced bilingual participants performed differently in response to language switching. Specifically, unbalanced (i.e., Cantonese-dominant and English-dominant bilinguals) but not balanced bilinguals were prone to an effect of language switching, manifested as a shift in VOT from their dominant language toward their nondominant language. When switching from English to Cantonese, the Cantonese-dominant speakers maintained longer VOT values for Cantonese switch trials than stay trials, thus shifting toward English-like VOTs. Moreover, when English-dominant speakers switched from Cantonese to English, shorter VOT values were produced with a shift toward Cantonese-like VOTs. A short-term phonetic convergence occurred when unbalanced bilinguals switched between languages, such that the VOT values of the

Table 4. *Mean VOT values for each English-dominant bilingual speaker*

Speaker	Cantonese VOT (ms)			English VOT (ms)		
	Stay	Switch	Switch–Stay	Stay	Switch	Switch–Stay
E01	15.5	15.8	0.3	63.9	66.7	2.9
E02	14.8	15.6	0.8	106.5	99.7	–6.8
E03	11.7	11.9	0.3	69.0	64.3	–4.7
E04	10.0	10.4	0.4	87.9	85.7	–2.2
E05	10.7	9.8	–0.9	77.9	71.4	–6.5
E06	13.2	13.9	0.7	82.9	81.0	–1.9
E07	12.5	13.7	1.3	72.3	71.5	–0.8
E08	12.7	13.8	1.1	62.8	65.9	3.1
E09	15.0	15.8	0.8	82.6	85.5	2.9
E10	23.1	23.2	0.1	79.7	66.2	–13.5
E11	15.0	16.4	1.4	84.9	78.4	–6.5
E12	13.2	13.9	0.7	75.4	69.2	–6.3
E13	16.2	16.7	0.5	111.0	107.9	–3.2
E14	13.9	11.9	–2.0	78.1	74.6	–3.5
E15	17.1	16.1	–1.1	99.0	94.3	–4.7
E16	9.7	10.3	0.7	54.8	52.1	–2.7
E17	10.5	10.4	–0.1	112.7	107.0	–5.7
E18	8.7	10.4	1.7	82.0	74.9	–7.1
E19	11.5	11.1	–0.4	67.4	61.9	–5.5
E20	13.5	15.2	1.7	114.5	110.4	–4.0
Mean	13.42	13.82	0.4	83.27	79.43	–3.8

==dominant language shifted toward those of the previously used nondominant language.== Yet, this short-term phonetic convergence was absent in balanced bilinguals' speech production.

Our study clearly showed an effect of place of articulation in both Cantonese and English, with the stops produced by all three groups of speakers consistently having significantly different VOT values at different places of articulation. This finding is consistent with the general claim that VOT increases as place of articulation changes from labial to alveolar to velar (Lisker & Abramson, 1964; Volaitis & Miller, 1992). This increase in VOTs aligns with the physiological basis of speech production at different places of articulation, that is, the timing between the movement of the articulatory gesture (i.e., the tongue) and the release of air held within the vocal tract (Cho & Ladefoged, 1999).

Language switching for unbalanced bilingual participants

The present study confirmed our first hypothesis that unbalanced early bilinguals will show an asymmetrical switch cost of language switching at the phonetic

Table 5. *Mean VOT Values for each balanced bilingual speaker*

Speaker	Cantonese VOT (ms)			English VOT (ms)		
	Stay	Switch	Switch–Stay	Stay	Switch	Switch–Stay
B01	14.2	12.3	−1.8	77.5	83.1	5.6
B02	11.5	10.2	−1.3	94.4	95.0	0.6
B03	14.1	16.9	2.9	54.8	59.0	4.2
B04	24.1	26.5	2.4	98.9	104.4	5.5
B05	11.4	11.3	−0.1	88.5	92.9	4.4
B06	12.4	12.0	−0.4	70.6	69.1	−1.5
B07	11.1	10.9	−0.2	83.1	80.5	−2.6
B08	10.6	10.1	−0.5	77.9	76.9	−1.0
B09	10.5	9.5	−1.0	77.5	77.0	−0.5
B10	13.2	14.2	1.0	74.8	74.0	−0.8
B11	10.6	11.4	0.7	95.7	93.5	−2.1
B12	16.7	16.7	0.0	59.0	61.1	2.1
B13	15.5	14.7	−0.7	121.0	116.7	−4.3
B14	10.0	10.1	0.1	62.1	64.4	2.3
B15	11.3	11.6	0.3	78.9	72.6	−6.3
B16	13.2	12.8	−0.4	94.0	91.3	−2.7
B17	20.2	20.0	−0.2	121.3	126.8	5.5
B18	11.0	11.9	0.9	63.2	65.4	2.2
B19	11.2	9.9	−1.2	54.3	54.0	−0.3
B20	8.4	8.1	−0.3	58.9	59.8	0.9
Mean	13.05	13.06	0.0	80.31	80.87	0.6

level. Moreover, the switch cost arises in the direction where bilinguals switch from the nondominant to the dominant language.

In relation to the two existing studies on phonetic transfer in language switching, our results were consistent with those reported in Olson's (2013) but not with those in Goldrick et al.'s (2014) study. One possibility for the discrepant findings between our study and Goldrick et al.'s (2014) is the differences in the nature of target stimuli employed. In Goldrick et al.'s (2014) study, the use of cognate stimuli might have induced a strong coactivation effect between the target and the nontarget language in language switching. In our study, however, the use of cognate stimuli was not optimal given the lack of phonological overlap between Cantonese and English. This absence of cognate stimuli minimized the cognate effect on cross-language activation during speech production, and therefore prevented the occurrence of the same effect observed in Goldrick et al. (2014). Another possibility is the mixed bilingual status of Goldrick et al.'s (2014) participants, native Spanish speakers (all rated their Spanish proficiency as 4 on a 4-point scale) whose English proficiency covered a wide range (from 2.67 to 4 on a self-rated 4-point scale). Furthermore, half of those participants rated their English proficiency as equal to or above 3.8, making them fairly balanced in terms of language dominance. The mixed bilingual status in their sample makes it

difficult to accurately compare this present study and Goldrick et al.'s (2014). Conversely, our results are consistent with those of Olson's (2013) study, which demonstrated unidirectional phonetic transfer from nondominant to dominant language. Together, the speech production of unbalanced bilinguals' dominant language is susceptible to the influence of their nondominant language during language switching.

Our results are also analogous to the predictions of the inhibitory control model, which claims that through inhibition of the nontarget representation(s), the activated systems compete to control speech output. Therefore, the nontarget system must be inhibited for the production of speech in the target language. While previous studies have acknowledged unbalanced bilinguals' asymmetrical switch cost of language switching at the level of lexical selection (Meuter & Allport, 1999; Schwieter & Sunderman, 2008), the present study provides additional evidence for a similar mechanism at the phonetic level. Our study reveals that, in switching between two phonetic systems existing in a common phonological space (e.g., Flege, 1995), Cantonese–English bilingual speakers make a selection at the phonetic level by suppressing the activation of nontarget phonetic realization.

Specifically, our results suggest that the phonetic systems of both languages are activated in bilinguals' speech production, but the activation level is higher for the dominant language. Because the dominant language is activated and utilized more in daily life than the nondominant language, the dominant language usually needs to be inhibited more when switching between the two languages (Meuter & Allport, 1999). In other words, the higher activation level for the dominant language requires a higher level of inhibition when it is not the target language in production. To overcome this greater level of inhibition, bilingual speakers often take longer to switch to their dominant language. In contrast, switching to the nondominant language often takes less time as the level of inhibition is weaker. As a result, an asymmetrical switch cost was found where the direction is modulated by the relative dominance of the bilingual's two languages.

The dominant language's higher level of inhibition is evident in the finding of a unidirectional phonetic transfer from nondominant to dominant language. Moreover, the latent activation due to the dominant language's greater inhibition is demonstrated in the delay in resuming the implementation of the VOT norm when switching from the nondominant to the dominant language. Consider a Cantonese-dominant bilingual who switches from his/her nondominant English language to his/her dominant Cantonese language. For the initial, nondominant language trial, the Cantonese phonetic system is inhibited to allow for the production of English. This inhibition is very strong, with the amount of inhibition applied proportional to the extent to which the nontarget phonetic system is activated (Linck, Schwieter, & Sunderman, 2012). According to the task set inertial hypothesis (Allport et al., 1994), the inhibition of the Cantonese phonetic system persists until the beginning of the next, dominant language trial, which leads to phonetic transfer from the English phonetic system. Consequently, the production of the second trial deviates toward the phonetic properties of English, as seen in the increased VOT for Cantonese switch trials. However, when the

participant switches from his/her dominant Cantonese language to his/her non-dominant English language, little inhibition is initially required on the English phonetic system. The participant can therefore access easily the phonetic representation of English in the second trial, without the influence of the Cantonese phonetic system.

It should be noted that the marginally significant interaction between bilingual group and response type found in the linear mixed-effects model analysis for English trials indicates that the interaction may be subtle. This could be attributed to the individual differences in inhibitory control, which has been reported to explain variations in speech processing among monolingual speakers, bilingual speakers, and second language learners (e.g., Darcy, Mora, & Daidone, 2016; Lev-Ari & Peperkamp, 2013, 2014; Linck et al., 2012). Particularly, variations in individual inhibitory capacity modulate the degree of coactivation between bilingual speakers' two languages (e.g., Lev-Ari & Peperkamp, 2013, Linck et al., 2012). It has been reported that bilinguals with greater inhibitory skill are relatively more immune to the influence of the nondominant language on their VOT perception and production in their dominant language; on the contrary, bilinguals with lower inhibitory skill are more prone to nondominant language influence on their dominant language production (Lev-Ari & Peperkamp, 2013). Because the present study was set in Hong Kong where Cantonese is the dominant language, all participating bilinguals were consistently and inevitably exposed to Cantonese. Because inhibitory control is modulated by how much a bilingual speaker switches between the two languages, the English-dominant bilinguals in the present study may have developed greater inhibitory control than the Cantonese-dominant bilinguals. In other words, English-dominant bilingual speakers must exercise more inhibitory control in their daily lives to efficiently activate the target language and deactivate the one not in use. In contrast, Cantonese-dominant bilingual speakers switch to English only for specific occasions, such as during lectures conducted in English, which results in fewer opportunities to operate their inhibitory control. Thus, the present study has found a more robust effect of language switching in the Cantonese-dominant bilinguals' VOT production due to a relatively weaker inhibitory control. In brief, the subtle interaction found in the English trials could be the result of the English-dominant bilinguals' greater inhibitory control capacity, which allowed them to better reject the influence from the nondominant language and demonstrate fewer switch costs when switching between languages.

Language switching for balanced bilingual participants

The results obtained in the present study also indicate that, contrary to unbalanced bilinguals' production, balanced bilinguals showed no transient phonetic transfer. To the best of our knowledge, no previous research has investigated transient phonetic interactions in balanced bilinguals using a language-switching paradigm. In this study, no switch cost was found at the phonetic level in balanced bilinguals' production. The VOTs observed in stay and switch trials were similar, regardless of whether trials involved Cantonese or English. Thus, our findings

suggest that there was no effect of language switching on phonetic interaction between the two phonetic systems for balanced bilinguals.

Two plausible explanations can account for the absence of transient phonetic interaction in balanced bilinguals' production. One is derived from the inhibitory control model. According to this model, balanced bilinguals may have better inhibitory control abilities, allowing them to switch efficiently between the two languages without any apparent switch cost. Previous research has indicated that general switching abilities are improved by language-switching experiences (Prior & Gollan, 2011). Moreover, more frequent language switching has been observed among balanced bilinguals compared to unbalanced bilinguals (e.g., Gollan & Ferreira, 2009). As suggested by Gollan and Ferreira (2009), balanced bilinguals' higher proficiency in both of their languages, together with habitual language switching, may contribute to their superior inhibitory control. Balanced bilinguals may therefore deploy inhibition more rapidly, allowing for faster access to phonetic representations in another language. This may eventually minimize balanced bilinguals' switch cost, regardless of the switch direction. This hypothesis is in line with Linck et al.'s (2012) study, which posits that better inhibitory control abilities are associated with reduced switch costs in trilingual speakers.

Another plausible explanation is that balanced bilinguals may adopt a language-specific mechanism in phonetic processing, leading to the absence of transient phonetic interaction. Previous research on language switching at the level of lexical selection has shown that highly balanced bilinguals may not resort to an inhibitory control mechanism in lexical selection during speech production (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006). Instead, they may employ a language-specific selection mechanism that considers lexical representations only in the target language. Schwieter and Sunderman (2008) suggested that bilinguals' L2 lexical robustness predicted the use of different mechanisms for bilingual language selection. In particular, they found that as L2 lexical robustness increased, bilingual speakers shifted from using inhibitory control to language-specific mechanisms during speech production (Schwieter & Sunderman, 2008). Although the theoretical basis of language selection for highly balanced bilinguals is still under debate, the present study includes the language-specific mechanism of speech processing as a possible explanation for the lack of phonetic interaction seen in balanced bilinguals. Under this interpretation, phonetic representations of the nontarget phonetic system may not compete during speech processing, which is similar to the speech processing for monolingual speech production. Thus, no transient phonetic interaction was observed in balanced bilinguals.

LIMITATIONS, FUTURE DIRECTIONS, AND CONCLUSIONS

Despite the significance of our results, it should be noted that the present study employed a self-report questionnaire to evaluate participants' language-dominance profiles. Although previous research has demonstrated that self-ratings correlate reliably with language proficiency (Dunn & Fox-Tree, 2009; Flege et al., 1999,

2002), we could only obtain a relative comparison instead of an absolute quantification of language dominance across participants. Future research might consider the use of objective or standardized measures (e.g., verbal fluency) that evaluate the participants' language proficiency quantitatively (Schwieter & Sunderman, 2008).

It should also be noted that, although other aspects of language experience and frequency of use were taken into account, language dominance was based heavily on self-ratings of language proficiency, and all language information data were collected with the LEAP-Q questionnaire. Despite its popularity in evaluating language dominance (e.g., Olson, 2013, 2016; Piccinini & Arvaniti, 2015), the LEAP-Q fails to provide adequate measures that directly compare the relative strengths of bilinguals' two languages (Gertken, Amengual, & Birdsong, 2014). Future studies should allow for investigations of transient phonetic interaction along the continuum of language dominance, specifically to see whether a threshold of discrepancy exists between the proficiency of the two languages where transient phonetic interaction occurs.

In addition, language switching was the only experimental task employed in the present study. While the presence of an asymmetrical switch cost could be properly interpreted under the theoretical framework of the inhibitory control model, questions regarding the underlying mechanism of language switching for highly balanced participants remain unanswered. Thus, future studies may consider incorporating cognitive measures, such as attentional control and general inhibitory abilities, to determine whether a link between general cognitive skills and phonetic control exists during language switching.

In conclusion, our results demonstrated the role of language dominance in determining the presence and direction of transient phonetic transfer. We found that balanced bilinguals and unbalanced bilinguals responded differently to language switching in terms of phonetic control during speech production. Inhibitory control in the speech production of unbalanced bilinguals was demonstrated by a shift in VOT during trials that involved language switching from the non-dominant to the dominant language. In contrast, no switch cost was observed in balanced bilinguals, which may indicate differences in the mechanism underlying balanced and unbalanced bilinguals' speech production.

NOTE
1. We wish to acknowledge an anonymous reviewer who suggested running fine-grained analysis of individual data.

REFERENCES

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Hillsdale, NJ: Erlbaum.

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, *15*, 517–530.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, *38*, 640–653.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, *39*, 558–570.

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, *40*, 582–594.

Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: A review. *Clinical Linguistics & Phonetics*, *14*, 131–150.

Balukas, C., & Koops, C. (2015). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, *19*, 423–443.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi:10.18637/jss.v067.i01

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. Berlin: de Gruyter.

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2002). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*, 92–111.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*(Suppl.1), 9–49.

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program]. Version 5.4.22. Retrieved from http://www.fon.hum.uva.nl/praat/

Bullock, B. E., & Toribio, A. J. (2009). Trying to hit a moving target: On the sociophonetics of code-switching. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 189–206). Philadelphia: Benjamins.

Bullock, B. E., Toribio, A. J., González, V., & Dalola, A. (2006). Language dominance and performance outcomes in bilingual pronunciation. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference* (pp. 9–16). Somerville, MA: Cascadilla Proceedings Project.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, *54*, 421–428.

Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, *13*, 67–85.

Chen, L., Peng, L., & Chao, K. (2009). The effect of lexical tones on voice onset time. *Proceedings of the 11th IEEE International Symposium on Multimedia* (pp. 552–557). Piscataway, NJ: IEEE.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*, 207–229.

Clumeck, H., Barton, D., Macken, M. A., & Huntington, D. (1981). The aspiration contrast in Cantonese word-initial stops: Data from children and adults. *Journal of Chinese Linguistics*, *9*, 210–224.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology*, *26*, 1283–1296.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*, 491–511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1057–1074.

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*, *66*, 741–773.

Dunn, A. L., & Fox-Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition*, *12*, 273–289.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, *89*, 395–411.

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E. (1997). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.), *Laboratory phonology* (Vol. 9, pp. 353–382). Berlin: de Gruyter.

Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*, 567–598.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*, 467–491.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language*, *41*, 78–104.

Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage.

Francis, A. L., Ciocca, V., & Yu, J. M. C. (2003). Accuracy and variability of acoustic measures of voicing onset. *Journal of the Acoustical Society of America*, *113*, 1025–1032.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 208–225). Bristol: Multilingual Matters.

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, *25*, 1031–1036.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 640–665.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 201–258.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*, 131–149.

Grosjean, F., & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *American Psychological Society*, *5*, 201–206.

Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics*, *28*, 27–42.

Harris, C. L., Gleason, J. B., & Ayçiçegi, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Clevedon: Multilingual Matters.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, *36*, 17–38.

Jacobs, A., Fricke, M., & Kroll, J. F. (2016). Cross-language activation begins during speech planning and extends into second language speech. *Language Learning*, *66*, 324–353.

Kroll, J. F., Bobb, S., & Wodniekca, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*, 119–135.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*, 137–171.

Lai, M. L. (2001). Hong Kong students' attitudes towards Cantonese, Putonghua and English after the change of sovereignty. *Journal of Multilingual and Multicultural Development*, *22*, 112–133.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*, 1–33.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*, 320–331.

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *Journal of Phonetics*, *47*, 36–46.

Li, D. C. S. (2017). *Multilingual Hong Kong: Languages, literacies and identities*. Basel, Switzerland: Springer International.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*, 651–662.

Lisker, L., & Abramson, S. A. (1964). A cross-language study of voicing contrast in initial stops: Acoustical measurements. *Word*, *20*, 384–422.

Macken, M., & Barton, D. (1980). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language*, *7*, 433–458.

MacLeod, A. A. N., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, *3*, 118–127.

MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition*, *3*, 37–54.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multi-linguals. *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 407–420.

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, *41*, 407–420.

Olson, D. J. (2016). The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association*, *46*, 263–285.

Pavlenko, A. (2004). "Stop Doing That, Ia Komu Skazala!": Language choice and emotions in parent–child communication. *Journal of Multilingual and Multicultural Development*, *25*, 179–203.

Piccinini, P., & Arvaniti, A. (2015). Voice onset time in Spanish-English spontaneous code-switching. *Journal of Phonetics*, *52*, 121–137.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*, 191–215.

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish-English and Mandarin-English bilinguals. *Journal of the International Neuropsychological Society*, *17*, 682–691.

R Core Team (2017). R: A language and environment for statistical computing [Computer program]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime 2.0 [Computer program]*. Pittsburgh, PA: Psychological Software Tools.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, *22*, 106–129.

Schwieter, J. W., & Sunderman, G. (2008). Language switching in bilingual speech production: In search of the language-specific selection mechanism. *Mental Lexicon*, *3*, 214–238.

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, *43*, 29–37.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, *9*, 97–114.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, *7*, 255–272.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America*, *92*, 723–735.

APPENDIX A

Table A1. *Target Words and Their Phonetic Transcriptions*

Target initial stops	Cantonese		English	
	Target words	Phonetic transcriptions	Target words	Phonetic transcriptions
/p/	錶 (*watch*)	/piu1/	pig	/pɪg/
	冰 (*ice*)	/pɪŋ1/	pin	/pɪn/
	煲 (*pot*)	/pou1/	pet	/pet/
	波 (*ball*)	/pɔ1/	pen	/pen/
	爸 (*father*)	/pa1/	pot	/pɒt/
	包 (*bread*)	/pau1/	park	/pɑːk/
/t/	冬 (*winter*)	/tʊŋ1/	tea	/tiː/
	叮 (*bell*)	/tɪŋ1/	tooth	/tuːθ/
	釘 (*nail*)	/tɛŋ1/	ten	/ten/
	刀 (*knife*)	/tou1/	toe	/təʊ/
	燈 (*lamp*)	/tɐŋ1/	tie	/taɪ/
	單 (*receipt*)	/tan1/	tag	/tæg/
/k/	菊 (*chrysanthemum*)	/kʊk1/	key	/kiː/
	菇 (*mushroom*)	/ku1/	cook	/kʊk/
	乾 (*dry*)	/kɔn1/	cake	/keɪk/
	薑 (*ginger*)	/kœŋ1/	cup	/kʌp/
	雞 (*chicken*)	/kɐi1/	cat	/kæt/
	金 (*gold*)	/kɐm1/	cow	/kaʊ/

Note: English translations of Cantonese targets are shown in parentheses.