

Lexical tonal effects in code-switching: A comparative study of Cantonese, Mandarin, and Vietnamese switching with English

International Journal of Bilingualism
2024, Vol. 28(5) 799–827
© The Author(s) 2023



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13670069231181508
journals.sagepub.com/home/ijb



Katrina Kechun Li 

University of Cambridge, UK

Li Nguyen 

University of Cambridge, UK; FPT University, Vietnam; The Australian National University, Australia

Christopher Bryant 

Writer, University of Cambridge, UK

Kayeon Yoo

Amazon, UK

Abstract

Aims and objectives: Previous research has revealed much about the syntactic and social variables conditioning code-switching (i.e., the alternation between two or more languages in a discourse or utterance); however, little is known about the phonological effects. Our work explores this area by asking two main questions: (1) Does lexical tone affect code-switching between a tonal language and a non-tonal language? and (2) Is this effect (or lack thereof) observable cross-linguistically?

Methodology: We examine natural code-switching production between Cantonese and English, Mandarin and English, and Vietnamese and English. We use a semi-automatic natural-language processing method to process and extract relevant variables, including tonal categories at switch points.

Data and analysis: Data include transcribed natural speech from three bilingual corpora: the HLVC corpus (Cantonese/English, 25 speakers), the SEAME corpus (Mandarin/English, 20 speakers), and the CanVEC corpus (Vietnamese/English, 45 speakers). We use logistic mixed-effects models to examine tonal effects, taking into account other factors such as frequency and grammatical category.

Corresponding author:

Li Nguyen, The Computer Laboratory, Department of Computer Science and Technology, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, Cambridge CB3 0FD, UK.

Email: nhbn2@cam.ac.uk

Findings/conclusion: We found a robust tonal effect in Cantonese/English, a less robust effect in Mandarin/English, and no effect in Vietnamese/English. This indicates there is a tonal effect in code-switching between a tonal and a non-tonal language, but this effect is language-dependent. We also found a specific T3 ‘step-up’ pattern at Cantonese-English switch points and offered some possible phonological explanations.

Originality: This is the first study that systematically investigates tonal effects in code-switching across different language pairs, using comparable data and methods. Our finding of a Cantonese-English T3 ‘step-up’ pattern is also a novel discovery that hitherto has not been documented.

Significance/implications: Theoretically, our findings support Clyne’s ‘facilitation theory’ in code-switching at a prosodic level. Empirically, we nevertheless emphasised the complexity of different prosodic features and social variables in play, thereby rejecting the idea of ‘predicting’ code-switching solely based on linguistic factors.

Keywords

Code-switching, code-switching facilitation, tone, tonal effect, Cantonese, Mandarin, Vietnamese, English, language contact

Introduction

Code-switching (CSw) is a phenomenon whereby language users combine two or more languages in a single discourse or utterance. This linguistic behaviour has been widely documented in a variety of language-contact scenarios, from computer-mediated interaction (Laroussi, 2011; Lynn & Scannell, 2019; Papalexakis & Doğrüz, 2015; Siebenhaar, 2006; Winata et al., 2022), language teaching and learning (Ahmad & Jusoff, 2009; Carstens, 2016; Daniel et al., 2019; Nguyen et al., 2022; Wang, 2019) to natural production in multilingual communities where two or more languages coexist (Clyne, 2003; Nagy, 2011, 2018; Nguyen, 2018; Torres Cacoullous & Travis, 2018; Tse, 2019). Examples (1)–(3) provide some illustration.

1. Tweet [Irish-English]

Ceol alainn ar @johncreedon on @RTERadio1 now

music beautiful on

‘Lovely music on @johncreedon on @RTERadio1 now’

(Adapted from Lynn & Scannell, 2019)

2. Learner writing [Chinese-English]

当下 倾盆大雨时, **we will stay home**

when downfall pour basin big rain time

‘When it is pouring (basin-like big rain), we will stay at home’

(Write & Improve submission, Nguyen et al., 2022)

3. Natural speech production [Vietnamese-English]

cái concert đó có more than you expected không

CLF DET AUX

Q

‘Was that concert more than what you expected?’

(CanVEC corpus; Nguyen & Bryant, 2020)

Within linguistics, most studies to date have focussed on the syntactic and social variables conditioning code-switching; for example, Poplack (1980) et seq., Myers-Scotton, (1993) et seq., Muysken (2000) et seq., Torres Cacoullos and Travis (2018); Nguyen (2018, 2021). We still know relatively little about the phonological variables, however, despite previous evidence suggesting a link (Antoniou et al., 2011; Balukas & Koops, 2015; Fricke et al., 2016; Olson, 2013; Piccinini & Garellek, 2014). In this work, we thus explore the relationship between a particular phonological feature, namely lexical tone, and natural CSw production, in the context of three tonal languages (Cantonese, Mandarin, and Vietnamese) and a non-tonal language (English). Our research questions are therefore two-fold:

1. Does lexical tone affect code-switching between a tonal language and a non-tonal language? and
2. Is this effect (or lack thereof) observable cross-linguistically?

We attempt to answer these questions using three CSw corpora: the HLVC corpus (Cantonese/English) (Nagy, 2011), the SEAME corpus (Mandarin/English) (Lyu et al., 2010), and the CanVEC corpus (Vietnamese/English) (Nguyen & Bryant, 2020). All three datasets consist of transcribed, natural speech and are therefore highly comparable. Together, they represent a diverse range of linguistic and social scenarios where tonal and non-tonal languages come into contact.

The remainder of this paper is organised as follows. We first provide a brief description of the contact settings and tonal systems of the language varieties in question (§Background). We then move on to discuss our corpora and data processing (§Methods), before presenting our findings (§Results) and analysis (§Discussion). We finally conclude with a summary and offer recommendations for future research (§Conclusion). It is important to note that throughout this paper, we use a forward slash (/) between a pair of languages to indicate general code-switching, and a hyphen (-) to indicate a particular direction. For example, *Cantonese/English* implies code-switching between Cantonese and English in general, whereas *Cantonese-English* implies switching from Cantonese into English.

Background

The language varieties of interest in this study are Cantonese/English spoken in Hong Kong and Toronto, Canada, Mandarin/English spoken in Singapore and Malaysia, and Vietnamese/English spoken in Canberra, Australia. In what follows, we first introduce the contact settings in these areas, then describe the tonal systems of Cantonese, Mandarin, and Vietnamese, respectively. We then consider the relevant prosodic features in these varieties of English, before finally discussing how they might manifest in a code-switching scenario.

It should be noted that the contact situations in these communities are extremely complex, and it remains difficult to provide a comprehensive overview. We thus only focus on three key aspects that we deem most relevant in the contact landscape section: community formation, language use, and language attitude of the communities in question.

The contact landscape

Cantonese/English in Toronto and Hong Kong: The Cantonese-speaking community in Toronto, Canada, started growing in the 1960s, and subsequently developed into a large community that widely uses and retains Cantonese as a heritage language (Tse, 2019). Cantonese used to be the most widely spoken first language among non-official languages (4% of the population) in Toronto

according to the 2016 Census (Statistics Canada, 2017), before being surpassed by Mandarin in recent years (Statistics Canada, 2022). Recent studies have also established a generational difference: second-generation speakers tend to use fewer Cantonese words and are more likely to code-switch into English in their speech. This group is thus largely English-dominant, or considered balanced bilinguals if they also achieve good proficiency in Cantonese (Tse, 2022).

The language contact situation in Hong Kong is a little different. Cantonese has always been the predominant language in daily life, although it only became an official language along with English from 1974 onwards. While English was the sole official language before that, the English-speaking population was rather small, and it was not until the 1980s when compulsory education was institutionalised and university places were greatly expanded that the number of people who were proficient in English began to rise sharply (Joseph, 2004). This increase in educational opportunities contributed to the emergence of a distinctive English variety, known as ‘Hong Kong English’. In this variety, the stress system of British English is to some extent re-interpreted as a tonal system, which corresponds to a subset of tone patterns of Cantonese (Gussenhoven, 2012; Luke, 2000; Wee & Liang, 2016). Cantonese with English insertion also became a new norm, distinguishing itself from other Cantonese varieties spoken in mainland China (Y. S. Cheung, 1985; Gibbons, 1987; Li, 2000).

Mandarin/English in Singapore and Malaysia: The language contact situation in Singapore and Malaysia is perhaps the most complicated. This is mainly due to the array of other languages that Mandarin and English come into contact with in this area. In Singapore, English is the *de facto* national language, serving as the primary working language and inter-ethnic lingua franca. The other three official languages – Mandarin, Malay, Tamil – are designated as ethnic mother tongues of the respective ethnic groups (Kuo & Jernudd, 1993; Low, 2012). The implementation of this language policy, especially through bilingual education since the 1960s has led to effective individual bilingualism and societal multilingualism (Gupta, 1994).

Like Hong Kong English, English spoken in Singapore has been influenced by extensive, long-term language contact with indigenous languages and has taken on a distinctive form. There are generally two varieties of Singapore English: a relatively standard variety used mainly for formal purposes of communication, referred to as Standard Singapore English, and an informal colloquial variety, called Singapore Colloquial English, or Singlish. While the standard variety generally shares similar linguistic features with other established standard varieties of English such as British or American English, the colloquial variety has undergone more substantial substrate-influenced restructuring. There is still considerable inter- and intra-ethnic variation across speakers as a result of varying effects of individual bilingualism (e.g., speaking a different first languages [L1], having a different language dominance) and differences in cultural affiliation or orientation (Sim, 2019). Within the Chinese community, in particular, the shift to English is also concurrent with the shift to Mandarin from other Chinese languages. Specifically, the ‘Speak Mandarin Campaign’ launched in 1979 promoted the use of Mandarin as a lingua franca for all ethnically Chinese people, resulting in the decline of other Sinitic heritage languages (commonly referred to by locals as ‘Chinese dialects’) spoken in Singapore such as Cantonese, Teochew and Hakka.

In Malaysia, ethnic Malays make up 70% of the population. Since the 1960s, Malay has been adopted as the main (and later sole) medium of instruction and English became a second language in the education system (Ng & Cavallaro, 2019). This results in the predominance of Malay over other languages. English, however, continues to be used in the civil service and trade contexts and developed its own features (Hashim, 2020). Generally speaking, Malaysian English shares many features with Singaporean English, including the division of formal and colloquial sub-varieties. Although the substrate languages are similar, Malaysian English has more lexical borrowing from Malay, whereas Singaporean English has more from Hokkien (Low, 2012). Within the Malaysian

Chinese society, various Chinese varieties are still actively spoken today in addition to Malay and English, including Hokkien, Cantonese, and Mandarin (Khoo, 2017; Wang, 2012).

Vietnamese/English in Canberra, Australia: Compared to the aforementioned communities, the history of the Vietnamese community in Australia is more recent and/or manifests on a smaller scale. Due to the political tension following the continuous arrival of Vietnamese immigrants after the fall of Saigon in 1975, the Vietnamese community mainly built their life in Australia by setting up family businesses in the Vietnamese/Chinese-dense suburbs where English is not required on a regular basis. It is thus no surprise that Vietnamese is particularly well maintained in the community in general (Ben-Mosche & Pyke, 2012).

In Canberra, however, the situation is slightly different. Contrary to the other densely populated cities such as Sydney or Melbourne, in which Vietnamese speakers cluster in neighbourhoods and are employed in family business, the majority in Canberra work in education or the public sector, or have a partner doing so (Australian Bureau of Statistics, 2017). While official numbers are difficult to obtain, this group is well known for being ‘Canberrans’ for the most part: relatively young, well paid and well educated. This also means that they have high English proficiency in general, and code-switching within the community is more easily and naturally observed (Nguyen, 2021; Thai, 2005). Conversations within families are on average a mix of half Vietnamese, a quarter code-switching and a quarter English (Nguyen, 2021, p. 54).

In summary, the language contact situations in these three communities are clearly diverse, with different social circumstances, linguistic histories and demographic make-ups. What they all have in common, however, is an environment where language contact is ubiquitous, most of which involves a majority, (mostly) non-tonal language (i.e., English) and a heritage tonal language (i.e., Cantonese, Mandarin, Vietnamese, respectively). Together, this represents a unique collection of datasets that allow us to systematically investigate tonal effects across different contact scenarios. In what follows, we briefly introduce the tonal systems of these language varieties.

The tonal systems

Despite all being tonal languages, Cantonese, Mandarin, and Vietnamese differ noticeably in their tonal inventories. Figure 1 provides an illustration.

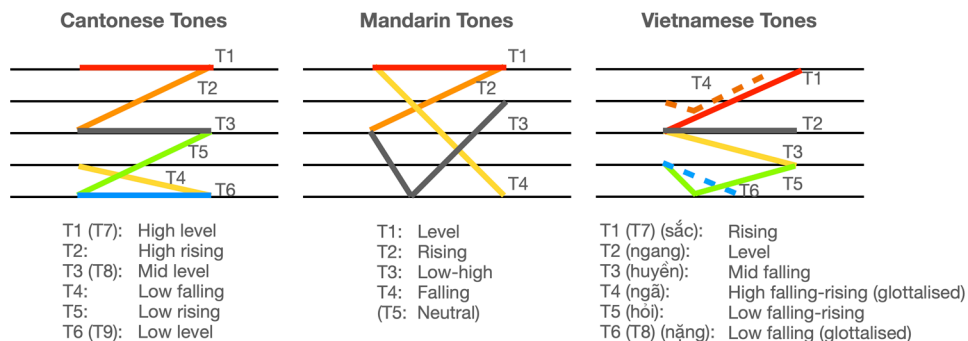


Figure 1. Tonal inventories in Cantonese, Mandarin, and Vietnamese. The shape of each tone is represented schematically using lines in colour. The dashed lines indicate that glottal constrictions are involved (a glottal stop in the middle of Vietnamese T4, and at the end of Vietnamese T6). Tones in brackets are tones with stop codas and can be merged into the same category of the previous tones.

As we can see, the tonal systems in Cantonese, Mandarin and Vietnamese differ both in terms of quantity and quality. Specifically, Cantonese has six tones, three of which are level tones (T1: high level, T3: mid level, T6: low level) and three of which are contour tones (T2: high rising, T4: low falling, T5: low rising). There are also three checked tones (high checked T7, mid checked T8, and low checked T9) that can be grouped with T1, T3, and T6, respectively, due to the matching pitch levels. The difference is that the syllables bearing checked tones end in stop consonants and are shorter in duration (compared to syllables with non-checked tones). In recent years, certain tones in Cantonese are reported to be undergoing merging (T2/T5 and T3/T6 in speech production), and this sound change is found among both native speakers in Hong Kong as well as heritage speakers in Toronto (Fung & Lee, 2019; Mok et al., 2013; Nagy et al., 2020; Zhang, 2019).

In contrast, there are five tones in Mandarin. In addition to the four tones shown in Figure 1 (T1: level, T2: rising, T3: low-high ‘dipping’, and T4: falling), there is also a neutral tone (T5), which is not shown graphically since it does not have a tonal target and its phonetic realisation is entirely determined by the surrounding tonal contexts. Mandarin spoken in Singapore and Malaysia largely follows the pronunciation of standard Mandarin spoken in the mainland China. There are a few deviations in the tonal systems; however, most of them are associated with the influence of other southern Chinese dialects (e.g., Hokkien, Cantonese) spoken in the region. First, the distribution of neutral tone is limited to fewer words compared to that in standard Mandarin (Choo, 2015). Second, T2 displays a low-level stretch in the f_0 contour before it starts to rise, which is different from standard Mandarin (L. Lee, 2010). Finally, there is a ‘fifth tone’, which is a special category reserved for a group of characters that have checked tones in Hokkien and Cantonese, and is pronounced with a shorter duration and a falling pitch (Chen, 1983). This category is less popular among young speakers, however, due to the declining presence of the dialects in Singapore and Malaysia (Choo, 2015).

In Vietnamese, tones are more commonly referred to by their ‘tone names’ instead of a numeric category as in Cantonese or Mandarin. We adopt Tuc (2003)’s assignment of numeric categories to facilitate comparison. The tonal system described here is the standard variety. There are four tones with modal phonation (T1: rising, *sắc*, T2: level, *ngang*, T3: mid falling, *huyền*, T5: low falling-rising, *hỏi*), and two tones with glottal constriction (T4: high falling-rising with mid-glottalisation, *ngã*, T6: low falling tone with ending glottalisation, *nặng*). Like Cantonese, there are two more tones, T7 and T8, which occur only in syllables with stop codas and can be grouped with T1 and T6, respectively. Unlike Chinese languages, the tonal contrast in Vietnamese not only relies on pitch and duration differences, but also tends to involve laryngeal features, such as breathiness and creakiness (Brunelle, 2009b; Pham, 2003).

Finally, English is typically classified as a non-tonal language, although some English varieties in contact with tonal languages have been known to develop tone-like features (Gussenhoven, 2012; Lim, 2009). In these contact situations, tonal categories are assigned to syllables in English words in a similar fashion as they are assigned to Chinese syllables, mostly making use of the contrast in pitch register but not pitch contours. In Hong Kong English and Singaporean/Malaysian English in particular, the exact tonal assignment also bears specific characteristics. For example, the tonal assignment in Hong Kong English seems to be reminiscent of the stress in British English in which a high tone (H) is assigned to the syllable that would bear stress in English. A mid tone (M) is assigned before the high tone and a low boundary tone (L) is assigned to the utterance-final position (W. H. Y. Cheung, 2009; Wee, 2016). This is slightly different in Singaporean/Malaysian English, where a high tone (H) is assigned to the final syllable of a content word, resulting in a series of rising melodies on the utterance level (Chong & German, 2015; Ng, 2012). This tonal development occurs in tandem with the replacement of lexical stress in the language system (Tan, 2015, 2016; Wee, 2016). Although Vietnamese-English has not been formally proposed as an established English variety, studies have suggested that Vietnamese speakers may produce and

perceive English stress as corresponding to Vietnamese tones (Nguyen, 2004; Nguyen et al., 2008). In aggregate, this seems to indicate that tonal transfer might be a strong tendency across English speakers whose L1 is a tonal language.

Tonal facilitation in code-switching

In code-switching research, one of the most relevant and influential ideas is the concept of facilitation introduced by Clyne (2003). Built on his earlier, widely known 'Triggering Hypothesis' (Clyne, 1967, 1980). Clyne stipulates that certain forms of lexical, semantic, phonetic/phonological, prosodic, tonemic, graphemic, morphological, and syntactic (or any combinations of these) similarities between the languages involved might facilitate a switch (Clyne, 2003, p.76). He specifically draws on bilingual and trilingual data from different migrant communities in Australia to show how code-switching is a matter of strong tendencies and probabilities rather than of absolute constraints as in previous works (e.g., the Equivalence Constraint, the Free Morpheme Constraint, the Government Constraint, the Conjunction Constraint, etc.). These strong tendencies are believed to be a result of one or more principles of facilitation at lexical, tonal/prosodic, or syntactic levels.

For the purpose of our research, the most relevant principle in Clyne's work is the facilitation principle at the tonal/prosodic level. This principle suggests that 'lexical items in a tonal language whose tone is identified with the pitch and stress of the non-tonal language in contact are liable to facilitate (though not necessarily cause) transversion' (Clyne, 2003; Principle 2, 'tonal facilitation', pp. 175–177). Using data from Mandarin and Vietnamese in Australia, Clyne showed how items in the same tonal range substantially increases the likelihood of a transversion (a term he prefers to code-switching). In the Vietnamese/English bilingual community, for example, he pointed out that more than 85% of switches to English occurred immediately after a Vietnamese word of a mid-to-high pitch tone (Tuc, 2003). According to Clyne, these are the tones that Vietnamese are most likely to associate with English pitch and stress (unstressed syllables with mid tones and stressed syllables with high tones). Words with these tones take speakers into an overlapping tonal zone between the two languages, thereby facilitating a switch. Similarly, most switches (96.49%, *N* not available) from Mandarin to English were also found to correlate with falling tones, which correspond to English intonation (Zheng, 1997).

While Tuc (2003) and Zheng (1997)'s studies above (i.e., the only two studies we know so far that have specifically examined tonal effects in CSw) provide some illuminating results, one essential caveat is that they did not take into account other contributing factors such as grammatical categories or frequencies in establishing the effects. There is thus a high chance that their results may be biased by the effect of certain tokens occurring more frequently at switch points and are not truly representative of tonal category effects. Furthermore, since their datasets are quite different, it is difficult to compare findings and establish whether the reported effects generalise beyond their focus. Our study thus aims to contribute in this direction by examining tonal patterns in large-scale corpora across different language pairs and applying statistical models to account for the effect of other linguistic and social variables.

Methods

Corpora

Our data comes from three corpora: the HLVC corpus for Cantonese/English, the SEAME corpus for Mandarin/English, and the CanVEC corpus for Vietnamese/English. Each of these corpora is described in more detail below and summarised in Table 1.

Table 1. Summary of the three corpora.

Language	Corpus	Clauses	CSw clauses	Speakers	Gender	Speaker background	Data type
Cantonese	HLVC	15,153	1,588 (10.5%)	25	M=13 F=12	Gen. X=8 Gen. 1=8 Gen. 2, 3=9	Sociolinguistic Interviews
Mandarin	SEAME	11,852	5,460 (46.1%)	20	M=9 F=11	Mandarin-dominant=10 English-dominant=10	Interviews Conversations
Vietnamese	CanVEC	14,047	3,313 (23.6%)	45	M=21 F=24	Gen. 1=28 Gen. 2=17	Conversations

Cantonese/English: The Cantonese/English data were drawn from the Heritage Language Variation and Change Project (HLVC; Nagy, 2011). The corpus characterises the speech of native speakers of Cantonese based in Hong Kong as a benchmark (Gen X), as well as three generations of Cantonese speakers based in Toronto, Canada. First-generation speakers (Gen 1) include those who were born in Hong Kong, moved to Toronto after the age of 18 and lived in Toronto for at least 20 years. Second-generation speakers (Gen 2) include those who were born in the Greater Toronto Area (or came from the homeland before age 6) whose parents qualify as Gen 1. Third-generation speakers (Gen 3) include those who were born in the Greater Toronto Area whose parents qualify as Gen 2. The sample was balanced for sex but varied in fluency, usage, and ethnic orientation. We specifically used the transcriptions of sociolinguistic interviews that lasted 1 hour on average for each speaker in our experiments. In total, we analysed 25 fully annotated transcripts from 25 speakers, including Gen X (4 males, 4 females), Gen 1 (4 males, 4 females), and Gen 2 and 3 speakers combined (4 males, 5 females).¹ While the transcriptions were provided both in the form of traditional Chinese characters and Jyutping romanisation, we relied on Chinese characters since part-of-speech (POS) tagging is not feasible with Jyutping romanisation.

Mandarin/English: The Mandarin data were drawn from the South East Asia Mandarin–English corpus (SEAME; Lyu et al., 2010), a Mandarin/English CSw speech corpus developed as part of a multilingual speech recognition project with a focus on CSw data. The corpus contains spontaneous speech from interviews and conversations between speakers who are residents of Singapore and Malaysia. We use the publicly available test set, which includes transcribed speech from 20 speakers.² According to the data description, half of the sample’s speech (5 males, 5 females) is dominated by English, while the other half (6 females, 4 males) is dominated by Mandarin. The transcriptions were provided in simplified Chinese characters.

Vietnamese/English: Finally, the Vietnamese/English data were from the Canberra Vietnamese–English CSw Natural Speech Corpus (CanVEC; Nguyen & Bryant, 2020). The dataset comprises 23 transcribed spontaneous conversations between 45 Vietnamese–English bilingual speakers in Canberra, Australia. First-generation speakers in this corpus include 15 males and 13 females who are adult acquirers of English and have lived in Canberra for at least 10 years. Second-generation speakers meanwhile include 6 males and 11 females who acquired both English and Vietnamese simultaneously from birth or at a very young age (before 6), and whose parents qualify as first generation. The transcriptions were provided in standard Vietnamese orthography.

Data processing

Having selected our corpora, we next applied the semi-automatic natural language processing (NLP) methodology, previously developed for the CanVEC corpus (Nguyen & Bryant, 2020), to

all three datasets. Since the transcription convention of HLVC and SEAME differed from CanVEC, we needed to do some pre-processing to standardise the data format. For example, in HLVC, we discarded any Cantonese clauses that contained redacted information or digits, since we did not know which language these items were in, and removed most punctuation, with the exception of apostrophes (e.g., don't), hyphens in compounds (e.g., T-shirt) and incomplete words (e.g., f. . .). We also removed whitespace and non-speech mark-up (e.g., laughs). Similarly, in SEAME, we removed non-speech mark-up (e.g., <v-noise>) and the whitespace between Chinese characters.

After standardisation, we developed a pipeline for processing each clause in the HLVC and SEAME corpora as described below:

1. Based on character encoding (Chinese characters or English letters), we divided the strings into the largest contiguous sequences of Chinese only or English only. There could be multiple switches in one clause.
2. Each sequence was processed by the relevant word tokeniser and POS tagging toolkit: PyCantonese (J. Lee et al., 2022) for Cantonese, Jieba and SpaCy (Honnibal & Montani, 2017; Sun, 2020) for Mandarin, and SpaCy for English.³ In addition, Cantonese and Mandarin strings were converted to Romanisation to extract tonal information (Cantonese was converted to Jyutping by PyCantonese, Mandarin to Pinyin by Pypinyin) (Huang et al., 2023).⁴
3. We assigned three language labels to tokenised data: English, Chinese, and language-neutral. Since language-neutral items refer to those not exclusive to any language (Riehl, 2005), they are not considered valid points for code-switching. The language-neutral labels were assigned to words that were POS-tagged as 'INTJ' (interjections) such as 'um' and 'okay', 'PROPN' (proper nouns) such as place names, as well as those belonging to a manually compiled list including filler words (e.g., 'eh') and sentence-final particles (e.g., 'aa3' and 'lah').⁵
4. CSw points were identified based on language labels, marked as either English-Chinese or Chinese-English. If we considered any of the top 50 most frequent words at switch points to be a language-neutral word, we added this word to the language-neutral wordlist described in the previous step and iteratively re-ran this step.
5. We extracted tonal categories at switch points based on the automatic Romanisation. For multi-syllabic tokens, TONE was extracted from the syllable adjacent to English tokens. For example, in the sentence 'send *keoi5dei6* email', T5 was extracted when counting English-Cantonese CSw, and T6 was extracted when counting Cantonese-English CSw.⁶
6. We extracted all the tokens along with the relevant information for regression analysis, including their frequency, POS-tags, speaker information, and whether they are switch points or not.

For CanVEC, only steps 4–6 were implemented as the corpus had already been tokenised, language tagged and POS tagged following steps 1–3 (Nguyen & Bryant, 2020).

Statistical analysis

We focussed on the tonal tokens (i.e., Cantonese, Mandarin, Vietnamese tokens) in the datasets to examine whether lexical tonal categories influence the probability of a token becoming a switch point, either to English or from English. To determine the tonal effect while taking into consideration other contributing factors, we used R Studio software (R Studio Team, 2020) to conduct logistic mixed-effects regressions, using the *glmer* function in the *lme4* package (version 1.1.31, Bates et al., 2015) and *lmerTest* package (version 3.1.3, Kuznetsova et al., 2017). Pairwise comparison with Tukey adjustment was conducted using the *emmeans* package (version 1.8.3,

Lenth, 2022). To validate the models, we used the *check_collinearity* function in the *performance* package (version 0.10.2, Lüdtke et al., 2021) to make sure there is no multicollinearity problem among the independent variables. We further used the *DHARMA* package (version 0.4.6, Hartig, 2022) to ensure there are no significant issues in the residual patterns.

The variables of interest are listed in Table 2, with TONE being the main fixed effect and treatment-coded. There are six tones in HLVC (T7, T8, and T9 are merged with T1, T3, and T6, respectively), five tones in SEAME (including the neutral tone T5), and six tones in CanVEC (T7 and T8 are merged with T1 and T6, respectively, see Figure 1).

Other fixed effects include frequency (FREQUENCY) and part of speech (POS) of each token, as well as GENDER and GENERATION of the speakers. POS tags were defined according to the set of Universal POS tags (de Marneffe et al., 2021). The FREQUENCY of each token was calculated as frequency per million words and then converted to Zipf scale (van Heuven et al., 2014). The scale normally ranges from 1 to 6, with the upper half (4–6) representing high-frequency words. In the Mandarin model, we had no information about the GENERATION of speakers, so we instead included a LANGUAGE variable indicating whether the speech was characterised as Mandarin-dominant or English-dominant. Results from the *check_collinearity* function in the *performance* package suggest no multicollinearity among the independent variables (all VIFs < 3) (cf. James et al., 2013). The details of model results and validations are provided in the Supplemental material.

For random effects, we included a random intercept for TOKEN and SPEAKER, and a by-speaker random slope for TONE. The TOKEN variable was generated to further control the influence of a few highly frequent tokens. Specifically, tokens with a Zipf frequency less than 4 were merged into one category, whereas those larger than 4 were treated as one category on its own. Because each token is associated with a unique tone (i.e., no variation of tone within each token), we did not include a by-token random slope for TONE.⁷

Results

We report our results in two parts. In the first part, we compare the raw frequencies of each tone appearing at switch points, similar to previous studies. In the second part, we present the results of mixed-effects logistic regression where other contributing factors are taken into consideration.

Raw frequency

Table 2 reports the raw frequency, the *distribution* (i.e., the percentage of each tone relative to the total number of switches), and the *proportion* (i.e., the percentage of each tone relative to the total occurrences of that tone in the corpus) of each tone across all switch points. In other words, the *distribution* highlights how different tones are spread across the switch points, while the *proportion* measures how often a certain tone appears at a switch point compared to a non-switch point, that is, the preference for switch points across different tones. More formally, *distribution* and *proportion* are calculated according to Equations 1 and 2, respectively.

$$\text{Distribution (\%)} = \frac{\# \text{Tone X at switch points}}{\# \text{All tones at switch points}} \times 100 \quad (1)$$

$$\text{Proportion (\%)} = \frac{\# \text{Tone X at switch points}}{\# \text{All Tone X}} \times 100 \quad (2)$$

Table 2. Regression variables.

Var. type	Var. name	Corpus		Meaning
		HLVC	SEAME CanVEC	
Dependent variable	X-ENG		Binary (0, 1)	I: The current token is not English and the next token is English; 0: Otherwise
	ENG-X		Binary (0, 1)	I: The current token is not English and the previous token is English; 0: Otherwise
Fixed effects	TONE	T1-T6 (ref: T4)	T1-T6 (ref: T4)	The total categories of non-English tokens (with reference level)
	FREQUENCY	Mean = 3.30, SD = 1.01	Mean = 3.61, SD = 0.96	Token frequency per million words converted to a Zipf-scale
Random effects	POS	11 categories	12 categories	The Universal part-of-speech tag of the token [†]
	GENDER	Male, Female		The gender of the speaker
	GENERATION	Gen X, Gen. 1, – Gen. 2,3	–	The generation of the speaker
Random effects	LANGUAGE	–	Mandarin-dominant, English-dominant	The language dominance of the speech sample
	SPEAKER	25 speakers	20 speakers	43 speakers [‡]
	TOKEN	15 categories	24 categories	15 categories Unique tokens with FREQUENCY values greater than 4 (i.e., most frequent words). All remaining tokens belong to the category 'Other'

[†]Each POS tagger uses a slightly different tagset. CanVEC includes a special CLS tag for classifiers. This does not influence the regression because we run separate models for each language.

[‡]Two speakers only spoke in English so there was no code-switching.

Cantonese/English: As we can see from Table 3, a T3 (mid-level tone) appears most often at Cantonese-English switch points, accounting for both the largest *distribution* (30.97%) and *proportion* (2.14%). A similar trend is observed in the other direction, that is, English-Cantonese, where T3 makes up 20.37% of the *distribution* and 1.16% of the *proportion* (despite T5 having a slightly larger *proportion* of 1.18%). In contrast, T4 (low-falling tone) shows the lowest *distribution* and *proportion* for both Cantonese-English switch points (*distribution*: 3.35%, *proportion*: 0.36%) and English-Cantonese switch points (*distribution*: 6.63%, *proportion*: 0.58%). Altogether, this suggests a facilitating role of T3 and a limiting role of T4 at both Cantonese-English and English-Cantonese switch points.

Mandarin/English: Results in Table 3 show that while T4 (falling tone) appears to be the most frequent tone compared to other tones at Mandarin-English switch points (44.55%), T3 (low-high ‘dipping’ tone) appears most frequently at English-Mandarin switch points (31.41%). In both directions, however, T5 (neutral tone) has the biggest *proportion* (14.74% in Mandarin-English, and 11.18% in English-Mandarin), meaning that compared to other tones, T5 has the strongest preference for switch points. T2 is identified as a limiting tone at switch points of both Mandarin-English (*distribution*: 6.76%, *proportion*: 3.76%) and English-Mandarin (*distribution*: 10.27%, *proportion*: 6.41%). The relatively large *distribution* of T4, T5, and T3 compared to T1, T2 is in line with Zheng (1997), suggesting a preference for the falling tone group at switch points in Mandarin/English.

Vietnamese/English: Results in Table 3 suggest that while T1 (rising tone) is the most frequent tone at both Vietnamese-English (35.78%) and English-Vietnamese (30.13%) switch points, T5 (low falling-rising tone) shows the strongest preference for English-Vietnamese

Table 3. Raw frequency and distribution of tones at switch points, and their proportion relative to global tone frequency. The largest numbers of *distribution* and *proportion* in each row are highlighted in orange, and the smallest numbers are in teal.

X-ENG	Statistic	T1	T2	T3	T4	T5	T6
Cantonese-English	Frequency	247	203	471	51	199	350
	Distribution	16.24	13.35	30.97	3.35	13.08	23.01
	Proportion	0.85	0.91	2.14	0.36	1.02	1.13
Mandarin-English	Frequency	646	427	1,380	2,988	1,266	–
	Distribution	9.63	6.37	20.58	44.55	18.88	–
	Proportion	5.07	3.76	6.43	9.90	14.74	–
Vietnamese-English	Frequency	1,082	627	695	101	231	288
	Distribution	35.78	20.73	22.98	3.34	7.64	9.52
	Proportion	6.73	3.84	4.90	4.67	4.95	4.64
ENG-X							
English-Cantonese	Frequency	224	249	255	83	230	211
	Distribution	17.89	19.89	20.37	6.63	18.37	16.85
	Proportion	0.77	1.11	1.16	0.58	1.18	0.68
English-Mandarin	Frequency	1,144	727	2,223	2,024	960	–
	Distribution	16.16	10.27	31.41	28.60	13.56	–
	Proportion	8.98	6.41	10.35	6.71	11.18	–
English-Vietnamese	Frequency	809	623	593	108	315	237
	Distribution	30.13	23.20	22.09	4.02	11.73	8.83
	Proportion	5.03	3.81	4.18	4.99	6.76	3.82

switch points compared to other tones (6.76%). T4 (high falling-rising tone with mid-glottalisation) is the least frequent at both Vietnamese-English (3.34%) and English-Vietnamese (4.02%) switch points, despite sharing a similar *proportion* with other tones (4.67% for Vietnamese-English and 4.99% for English-Vietnamese). This is broadly in line with the pattern previously reported by Tuc (2003, p. 98).

At this point, the general impression from raw rates seems to suggest that there exists a preference for a certain tone at switch points in each direction for each language pair. For X-Eng, these facilitating tones are identified as T3 (mid-level tone) in Cantonese, T3 (dipping tone) and T5 (neutral tone) in Mandarin, and T1 (rising tone) in Vietnamese. For Eng-X, these are T3 (mid-level tone) and T5 (low-rising tone) for Cantonese, T3 (dipping tone) and T5 (neutral tone) for Mandarin, and T1 (rising tone) and T5 (low falling-rising tone) for Vietnamese. We should emphasise, however, that the most frequently occurring tones at switch points are not necessarily the same tones that are most likely to appear at switch vs. non-switch points. For Mandarin-English, for example, while T4 is the most frequent tone at switch points compared to all other tones (accounting for 44.55% of the distribution), T5 has the strongest preference for switch points compared to other tones (14.74% of the time when it appears). This again highlights the complexity of measuring tonal effects and underscores the need to go beyond raw distribution.

Logistic regression

We next report results from our mixed-effects logistic regression models. Since we are mostly interested in the tonal effect, we focus only on the post hoc pairwise comparisons between all different combinations of tone pairs.⁸ Tests of the differences were performed on the log scale and back-transformed to the odds ratio (OR) to facilitate interpretation. In all the tables that follow, an OR larger than 1 indicates a higher probability of the first tone occurring at switch points relative to the second. SE represents the standard error of the estimate OR (the higher the number, the more uncertain the estimate), and *z-ratio* measures the deviation of the estimate OR from 1. The statistically significant results (with $ps < .05$) are highlighted in bold.

Cantonese/English: Table 4 presents the results for Cantonese/English code-switching. Looking at all the pairs involving T3 (mid-level tone), we can see that the ORs of T3 relative to other tones (T3/TX) are significantly larger than 1 (T3/T5: OR=1.85, $p = .0007$; T3/T6: OR=1.81, $p < .0001$), whereas the ORs of TX/T3 pairs are significantly smaller than 1 (T4/T3: OR=0.30, $p < .0001$; T1/T3: OR=0.43, $p < .0001$; T2/T3: OR=0.49, $p < .0001$). This indicates that T3 is indeed more likely, compared to all the other tones, to occur at Cantonese-English switch points. Furthermore, the OR for T2 (high-rising tone) to occur is significantly larger than those for T4, T5, and T6 at English-Cantonese switch points (T4/T2: OR=0.57, $p = .0062$; T2/T5: OR=2.16, $p = .0001$; T2/T6: OR=1.62, $p = .0201$). This suggests a facilitating role of T3 at Cantonese-English switch points, and a more limited facilitating role of T2 at English-Cantonese switch points, respectively. We also observe that the OR of T4 (low-falling tone) relative to T3, T5, and T6 is significantly lower than 1 at Cantonese-English switch points (T4/T3: OR=0.30, $p < .0001$; T4/T5: OR=0.56, $p = .0307$; T4/T6: OR=0.55, $p = .0113$), suggesting to some extent of a limiting effect.

Mandarin/English: Table 5 presents the post hoc pairwise comparisons for Mandarin/English code-switching. Recall from the results in Table 3 that T2 (rising tone) was consistently found to be the least frequent tone at switching points in either direction, according to both the *distribution* and *proportion* measures. However, results from the mixed-effects logistic regression (Table 5) do not support this observation. For Mandarin-English switch points, only one out of four pairs involving T2 have ORs smaller than 1 (T2/T4, OR=0.93, $p = .9260$), but even this is not statistically significant. For English-Mandarin, only the ORs of T2/T4 and T2/T5 are significantly smaller than 1 (T2/T4: OR=0.80, $p = .0067$; T2/T5: OR=0.38, $p < .0001$), indicating to some

Table 4. Post hoc comparisons in Cantonese. We begin with T4 as it is taken as the reference level in the regression model.

Pairs	Cantonese-English				English-Cantonese			
	Odds ratio	SE	z. ratio	p	Odds ratio	SE	z. ratio	p
T4/T1	0.71	0.14	-1.77	.4831	0.82	0.14	-1.22	.8252
T4/T2	0.62	0.12	-2.39	.1611	0.57	0.09	-3.50	.0062
T4/T3	0.30	0.06	-6.46	<.0001	0.78	0.17	-1.14	.8664
T4/T5	0.56	0.11	-3.02	.0307	1.22	0.22	1.09	.8859
T4/T6	0.55	0.10	-3.33	.0113	0.92	0.14	-0.57	.9932
T1/T2	0.87	0.14	-0.89	.9494	0.69	0.10	-2.53	.1146
T1/T3	0.43	0.05	-7.26	<.0001	0.96	0.16	-0.25	.9999
T1/T5	0.79	0.13	-1.48	.6774	1.49	0.25	2.39	.1585
T1/T6	0.77	0.10	-2.03	.3239	1.12	0.15	0.87	.9543
T2/T3	0.49	0.06	-5.56	<.0001	1.39	0.22	2.04	.3199
T2/T5	0.90	0.15	-0.60	.9911	2.16	0.37	4.50	.0001
T2/T6	0.89	0.11	-0.97	.9268	1.62	0.25	3.15	.0201
T3/T5	1.85	0.28	4.06	.0007	1.56	0.34	2.03	.3227
T3/T6	1.81	0.19	5.53	<.0001	1.17	0.23	0.81	.9661
T5/T6	0.98	0.14	-0.14	1.0000	0.75	0.12	-1.75	.4957

Table 5. Post hoc comparisons in Mandarin. We begin with T2 as it is taken as the reference level in the regression model.

Pairs	Mandarin-English				English-Mandarin			
	Odds ratio	SE	z. ratio	p	Odds ratio	SE	z. ratio	p
T2/T1	1.26	0.14	2.09	.2237	1.02	0.09	0.21	.9996
T2/T3	1.09	0.09	1.06	.8263	0.90	0.08	-1.23	.7336
T2/T4	0.93	0.09	-0.82	.9260	0.80	0.05	-3.37	.0067
T2/T5	1.07	0.11	0.64	.9682	0.38	0.08	-4.69	<.0001
T1/T3	0.87	0.09	-1.36	.6553	0.89	0.06	1.65	.4629
T1/T4	0.74	0.07	-3.04	.0202	0.79	0.07	-2.57	.0752
T1/T5	0.85	0.10	-1.45	.5968	0.37	0.09	-4.30	.0002
T3/T4	0.85	0.05	2.76	.0456	0.89	0.07	1.53	.5461
T3/T5	0.98	0.08	-0.27	.9988	0.42	0.10	-3.82	.0013
T4/T5	1.15	0.10	1.73	.4162	0.47	0.10	-3.65	.0024

extent a T2 limiting effect. In contrast, the facilitating effect of T5 (neutral tone) is largely confirmed in the English-Mandarin switching direction, as all the TX/T5 pairs have ORs significantly smaller than 1 (T2/T5: OR=0.38, $p < .0001$; T1/T5: OR=0.37, $p = .0002$; T3/T5: OR=0.42, $p = .0013$; T4/T5: OR=0.47, $p = .0024$). Furthermore, T4 (falling tone) also has a facilitating effect at Mandarin-English switch points, with the OR of both T1/T4 (OR=0.74, $p = .0202$) and T3/T4 (OR=0.85, $p = .0456$) significantly smaller than 1. This broadly aligns with Zheng's (1997) previous findings that Chinese falling tones facilitate the switching in and out of English by Chinese-Australian bilingual children.⁹ Although Table 3 suggests T3 also occurs frequently at switch points in both directions, it is not found to have a facilitating role after other factors are taken into consideration.

Table 6. Post hoc comparisons in Vietnamese. We begin with T4 as it is taken as the reference level in the regression model.

Pairs	Vietnamese-English				English-Vietnamese			
	Odds ratio	SE	z. ratio	p	Odds ratio	SE	z. ratio	p
T4/T1	1.42	0.21	2.31	.1919	1.03	0.17	0.20	1.0000
T4/T2	1.23	0.18	1.41	.7186	1.22	0.17	1.41	.7198
T4/T3	1.17	0.18	1.02	.9099	1.17	0.19	0.98	.9260
T4/T5	1.04	0.16	0.28	.9998	1.01	0.20	0.04	1.0000
T4/T6	1.13	0.20	0.70	.9825	1.18	0.20	0.99	.9228
T1/T2	0.87	0.08	-1.42	.7162	1.18	0.10	2.03	.3242
T1/T3	0.83	0.08	-2.06	.3109	1.13	0.10	1.45	.6985
T1/T5	0.74	0.08	-2.83	.0522	0.98	0.12	-0.20	1.0000
T1/T6	0.80	0.10	-1.80	.4649	1.14	0.12	1.24	.8182
T2/T3	0.95	0.10	-0.50	.9961	0.96	0.09	-0.48	.9970
T2/T5	0.85	0.10	-1.39	.7315	0.83	0.10	-1.54	.6396
T2/T6	0.92	0.11	-0.72	.9801	0.96	0.10	-0.35	.9993
T3/T5	0.89	0.09	-1.11	.8767	0.86	0.12	-1.09	.8868
T3/T6	0.97	0.11	-0.29	.9997	1.01	0.12	0.07	1.0000
T5/T6	1.09	0.14	0.62	.9900	1.17	0.17	1.07	.8939

Table 7. Summary of results. The parenthesis indicates that the effect does not hold for all the pairwise comparisons.

	Facilitating	Limiting
Cantonese-English	T3	(T4)
English-Cantonese	(T2)	-
Mandarin-English	(T4)	-
English-Mandarin	T5	(T2)
Vietnamese-English	-	-
English-Vietnamese	-	-

Vietnamese/English: Table 6 reports the pairwise comparisons for Vietnamese/English code-switching. As we can see, the models do not support any of the previously observed patterns. In other words, there is no clear preference for any particular tone at switch points in Vietnamese/English. This contrasts with what we initially saw in the distribution of raw frequencies reported in Table 3 and also with Tuc's (2003) previous work. This suggests that the previously claimed tonal effect in Vietnamese/English CSw (Tuc, 2003) is likely a side effect of other factors, thereby reinforcing the importance of going beyond raw rates and considering other interacting factors in our evaluation.

Table 7 summarises the logistics regression results across the three contact settings in this study. As we can see, results show some positive evidence for tonal effects in Cantonese/English and Mandarin/English but not for Vietnamese/English. This indicates that there exists a tonal effect in code-switching between a tonal and a non-tonal language, but this effect is language-dependent. In what follows, we focus on the effects found in Cantonese/English and Mandarin/English and discuss their implications.

Discussion

Tonal effects in CSw: evaluating evidence for prosodic facilitation

The most robust tonal effect found in this study is that of Cantonese-English code-switching, where T3 (mid-level tone) was unanimously identified as exhibiting clear facilitating effects, as reflected in both sections ‘Raw frequency’ and ‘Logistic regression’. Situating this in the context of the Cantonese communities in question, we suspect that this might be due to the tonal assignment of Hong Kong English,¹⁰ in which either a high (H) or a mid (M) tone (i.e., level tones) is assigned to English syllables (§The tonal systems). This means that compared to contour tones, a Cantonese level tone at a switch point would bear more resemblance to the prosody of subsequent English syllables and therefore be more conducive to a switch to English. This is further supported by the fact that there appears to be a limiting effect of T4 – a contour tone – on the other end of the spectrum. Ultimately, the pattern accords with Clyne’s tonal facilitation hypothesis (§Tonal facilitation in code-switching), thereby reaffirming the role of overlapping prosodic conditions in code-switching. A natural question which may emerge at this point is that if level tones in Cantonese (T1, T3, T6) are indeed preferred over contour tones (T2, T4, T5) at switch points, why does only T3 demonstrate a facilitating effect? We will return to this in the later discussion. For now, the key point to note is that T3, a mid-level tone which shares some prosodic features with Hong Kong English, displayed robust facilitating effects in Cantonese-English code-switching.

It is reasonable to suspect that the effect of T3 is not directly related to tone, but rather to certain POS or specific words that happen to carry that tone. Indeed, as Table 8 illustrates, there are only 3 unique particles (PART) with T3 that appeared at Cantonese-English switch points, but they have very high frequency ($N=88$). While this might suggest some kind of overlapping effect between tone and POS/specific token, it is clear from the table that T3 is also well spread across a range of unique words and many other POS. This means that even though we cannot completely tease the effect of tone and POS apart, we can be reasonably confident that T3 is indeed a strong contributing factor to a switch from Cantonese to English.¹¹

Table 8. Joint frequency table of Part of Speech (POS) and Tone at Cantonese-English switch points.

CAN-ENG POS	T1		T2		T3		T4		T5		T6	
	Unique	Total	Unique	Total	Unique	Total	Unique	Total	Unique	Total	Unique	Total
ADJ	5	7	4	7	4	4	0	0	3	3	8	10
ADP	0	0	2	8	2	2	2	6	0	0	2	2
ADV	13	38	5	75	11	45	10	18	1	3	19	60
AUX	2	2	1	1	2	14	1	2	2	24	0	0
CCONJ	0	0	3	8	1	2	1	3	1	4	5	56
NOUN	17	74	19	21	10	161	9	11	7	15	18	19
NUM	5	18	2	2	3	3	0	0	0	0	1	1
PART	2	8	3	21	3	88	2	3	1	3	1	1
PRON	12	74	5	11	10	72	1	1	4	73	6	20
VERB	20	26	25	48	25	80	6	6	10	74	20	181
X	0	0	1	1	0	0	1	1	0	0	0	0

For Mandarin/English, we also identified some evidence for the facilitating role of T5 (neutral tone) in English-Mandarin CSw and T4 (falling tone) in Mandarin-English CSw, but the interpretation is less conclusive. Specifically, the occurrence of T5 at English-Mandarin CSw is limited to six words, with two particles taking up a large portion (*de*: 77%, $N=735/960$; *le*: 17%, $N=165/960$). This is in significant contrast with T3 at Cantonese-English switch points, where the occurrences are distributed among 65 different words. It is thus less clear whether the facilitating role of T5 at English-Mandarin switch points can be solidly attributed to the tonal effect. Similarly, while we found a facilitating effect of T4 for Mandarin-English code-switching, this effect is not entirely robust. Zheng (1997) previously attributed the facilitating effect of T4 in Mandarin/English code-switching to the similarities between Mandarin falling tones and the English falling intonation contour. Although this explanation fits with Clyne's (2003) hypothesis as well as what we proposed for Cantonese/English, it may not satisfy the Mandarin/English data we have at hand. In particular, the intonation contour in Singaporean/Malaysian English mainly consists of a series of rising melodies across an utterance, and a falling contour can only appear on the last syllable of the content word in an intonation phrase (Chong & German, 2015; German & Chong, 2018). This unique intonation pattern leads us to believe that the tonal effects in Mandarin/English code-switching observed here are less likely to be attributed to the similarity between Mandarin falling tones and (Australian) English intonation as suggested in Zheng (1997).

Furthermore, we should take note that the speakers of the SEAME corpus come from a more linguistically diverse environment than the bilingual children in Australia as in Zheng (1997). The SEAME speakers' speech is thus likely, to a larger extent, subject to the influence of other languages in contact. For example, Singaporean/Malaysian Mandarin contains traces of other southern Chinese languages such as Hokkien and Cantonese (Chin & Cavallaro, 2021), while Singaporean/Malaysian English is often influenced by Malay (Ng, 2012). The extent of these influences further differs depending on the presence/dominance of various dialects in different sub-regions.

Finally, it is worth mentioning that we found no facilitating or limiting tonal effect for Vietnamese/English, contrary to what we have seen in the other language pairs. Given that Vietnamese T2 is similar to Cantonese T3 (Figure 1), this finding may seem to run counter expectation. It should be noted, however, that the presence of a mid-level tone is not the sole determining factor in facilitating a switch to English. There remain some non-trivial distinctions in the Cantonese and Vietnamese tonal systems that should be taken into account. For example, the Vietnamese tonal system involves more glottalisation features, while the Cantonese tonal system relies more on pitch. This means that a level tone in Cantonese may have different statuses in the system compared to a level tone in Vietnamese. Furthermore, the prosody in Vietnamese-English may not resemble that in Hong Kong English, thereby contributing to different phonological interactions in a code-switching context.

In all cases, further examination of the acoustic data and speakers' language background is needed to arrive at a more robust conclusion.

Tonal effects in Cantonese-English: T3 step-up pattern

Having discussed the tonal effects in all the language pairs, we now return to a question previously raised in relation to the results for Cantonese/English. Namely, if level tones in Cantonese (T1, T3, T6) are indeed generally preferred over contour tones (T2, T4, T5) at switch points, why does only T3 demonstrate a facilitating effect? Thanks to the availability of the HLVC audio data, we were able to cast some light in this direction.

First, we observe that there exists a specific 'step-up' pattern at Cantonese-English switch points; that is, the pitch of the English word starts at a higher level than the previous Cantonese

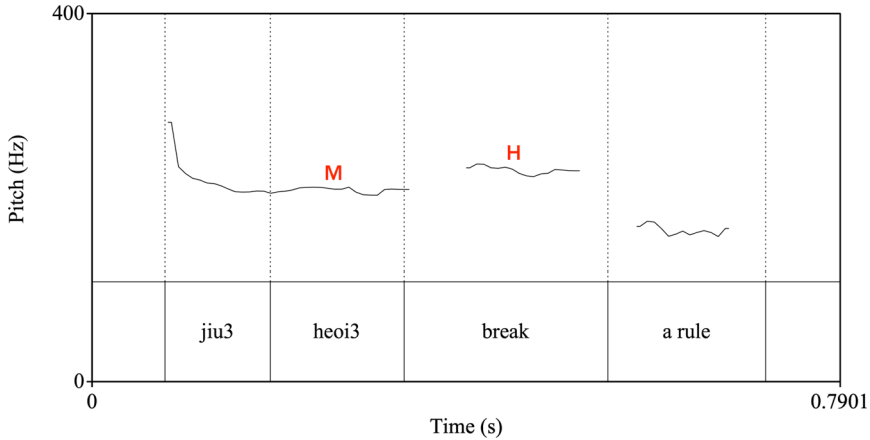


Figure 2. F0 contour of a Cantonese-English CSw at Cantonese T3. Translation of the sentence: '(...) is going to break a rule'.

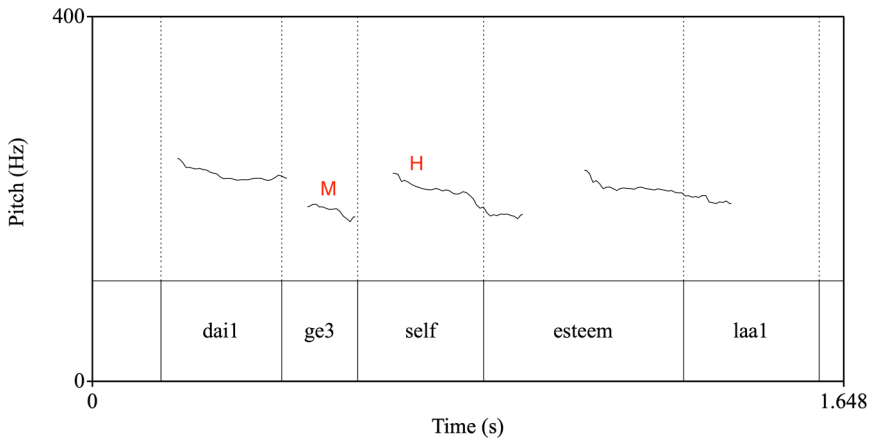


Figure 3. F0 contour of a Cantonese-English CSw at Cantonese T3. Translation of the sentence: '(It is) low self-esteem *la* (sentence-final particle)'.

syllable, as illustrated in Figures 2 and 3. These graphs represent some typical examples of the f0 pattern at Cantonese-English switch points where T3 is present on the Cantonese syllable. As we can see, the f0 contours show a 'step up' from the Cantonese mid-level T3 to the English monosyllabic words 'break' and 'self' (which are associated with an H tone). This 'step-up' pattern can be consistently observed for T3 throughout the corpus. Although the f0 difference identified in the 'step up' is not huge compared to the overall tonal space of Cantonese (the difference is about 20 Hz in both Figures 2 and 3), this level of pitch raise is perceptually clear to native speakers.

Previous research has suggested that in Hong Kong English, tones are specified underlyingly for syllables that usually bear stress in British English, and M tones are associated with any word-initial syllables that do not have underlying H tones (Wee, 2016). In other words, an M tone is the unmarked default choice in tonal assignment, unless a H tone is specified in the lexicon. When an English word after a switch point has an initial H tone, which is a common scenario, a preceding

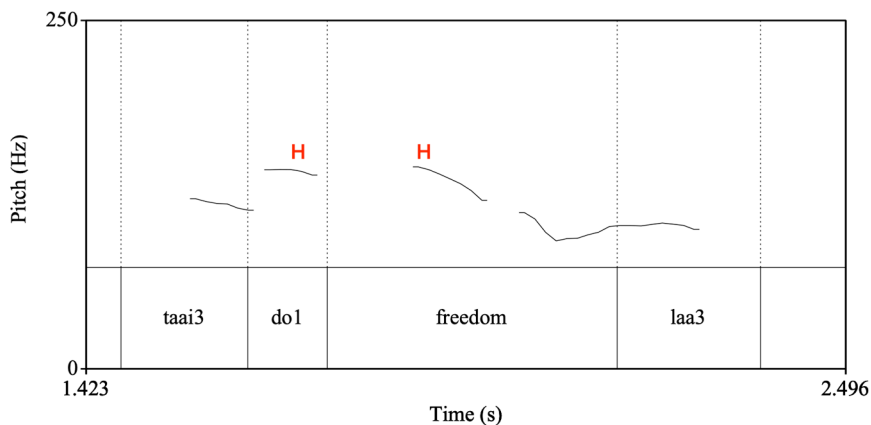


Figure 4. F0 contour of a Cantonese-English CSw at Cantonese T1. Translation of the sentence: ‘Too much freedom’.

M tone would best fit the tonal pattern of English words. In Cantonese, T3 has in fact been found to be associated with an M tone in terms of the pitch pattern (Wee & Liang, 2016). T1, on the other hand, is associated with H tones, and T6 is associated with neither H nor M tones in Hong Kong English. We illustrate in Figure 4, for example, that the ‘step-up’ pattern observed at T3 switch points is not borne out at the less frequently occurring T1 switch points. This exclusive resemblance between T3 and M tone possibly explains why only T3, but not other level tones, has a facilitating effect in Cantonese-English CSw.

Conclusion

We began this study by asking whether lexical tone has an impact on code-switching between a tonal language and a non-tonal language, and whether this effect (or lack thereof) is observable cross-linguistically. Our research has found some positive evidence that lexical tone indeed has an effect on Cantonese/English and Mandarin/English but not on Vietnamese/English. This suggests that although such an effect exists, this effect is language-dependent. For Cantonese/English specifically, we observed the most robust finding, where T3 (mid-level tone) plays a clear leading role in facilitating a switch from Cantonese to English. We explained this in reference to Clyne’s concept of prosodic facilitation (Clyne, 2003), and furthermore observed a specific ‘step-up’ pattern occurring at this particular switch point. We ultimately highlighted the prosodic similarity between Cantonese T3 and Hong Kong English M tone in facilitating their code-switching behaviour.

As for Mandarin/English and Vietnamese/English, although the previously reported tonal effects were not (fully) confirmed in our study, we should keep in mind the complex contact background of these communities when drawing conclusions. This lack of effect could be a result of other factors being captured in our study (e.g., grammatical categories, token frequency, etc.), but could also equally be a result of factors NOT being captured (e.g., regional variation, individual linguistic background, etc.). In any case, what is ostensibly clear is that even if there might be tonal effects in CSw for these language pairs, they were not robust enough to rise above interactions with other variables in play and thus stand as significant.

Looking forward, we rely on future research to address some of the limitations that we were not able to address in this study. Specifically, we hope the specific acoustic features of these varieties

can be more accurately captured using audio data. This will allow a more detailed and comprehensive phonological analysis of tonal effects at both social and individual levels. The more limited distribution of neutral tone (T5) in Singaporean/Malaysian Mandarin (Choo, 2015), the ongoing tone merger in Cantonese (Cheng, 2017; Nagy et al., 2020), or the regional tonal variation in Vietnamese (Brunelle, 2009a), for example, can be more effectively reflected using the f0 data from the audio recordings. Furthermore, the extent to which the ‘step-up’ pattern is prevalent in Cantonese/English CSW needs to be more systematically examined. Future investigations could also look at tonal effects between two tonal languages, or between a tonal language and another non-tonal (i.e., non-English) language to offer a more comprehensive understanding of this matter.

Finally, we would like to emphasise that although the lack of consistent effects across all language pairs in this work hints towards an absence of coarse generalisation, it has opened up possibilities for more fine-grained categorisation; for example, which TYPE of tonal systems might be more conducive to a switch to and from English (or other languages in future work). This line of enquiry should be of interest to researchers not only in tonal languages, but also in language typology more broadly. Ultimately, we hope to have shed some light on the intriguing question of tonal effects in language contact, thereby encouraging further interest in this area.

Acknowledgements

The authors are grateful to Naomi Nagy for granting access to the HLVC corpus.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was funded by Language Sciences Incubator Fund and Issac Newton Trust (PI: L.N.).

ORCID iDs

Katrina Kechun Li  <https://orcid.org/0000-0003-1611-0450>

Li Nguyen  <https://orcid.org/0000-0001-8632-7909>

Christopher Bryant  <https://orcid.org/0000-0002-1015-9467>

Supplemental material

Supplemental material for this article is available online: https://github.com/kc-li/Tonal_Aspects_CSW.

Notes

1. Gen 2 and Gen 3 Cantonese speakers are combined due to the low number of Gen 3 available ($N=1$).
2. <https://github.com/zengzp0912/SEAME-dev-set>
3. We manually evaluated the performance of the POS taggers on a subset of 100 randomly selected sentences. We report an accuracy rate of 82.5% for the HLVC corpus and 85.7% for the SEAME corpus.
4. We compared the Jyutping transcriptions by the HLVC human transcriber against the Jyutping romanisation of characters by PyCantonese on a subset of 100 randomly selected sentences and found an agreement rate of 88.4%. We could not do the same evaluation for Pinyin due to the lack of audio recordings. However, we did adapt the dictionaries to better fit the limited distribution of T5 (neutral tone) in Singaporean/Malaysian Mandarin, as previously described in Choo (2015).

5. We discounted sentence-final particles first because there are multiple tonal possibilities associated with one character transcription, see Sybesma and Li (2007). Furthermore, it is not clear whether these particles straightforwardly belong to only Cantonese or Mandarin in these communities, as some have been documented as part of the target English varieties in contact (e.g., ‘lah’ in Singaporean English).
6. We also calculated the three-way switches, for example, Chinese-Neutral-English, or English-Neutral-Chinese, but found their frequencies to be negligible.
7. We used the *drop1()* function to conduct model selection. Since the AIC values in general did not deviate much when dropping individual variables, the full model was kept for all three datasets to facilitate comparison.
8. For full results including other variables, please refer to Appendix 1. Among other things, the consistently significant effect of FREQUENCY (Zipfreq) appears to be one of the most noticeable. Specifically, with the exception of English-Mandarin, the regression models strongly suggest that the more frequent a word is, the more likely it is to facilitate a switch. Various POS were also identified as significant predictors; however, this set of results is not directly comparable to existing findings. This is because we used treatment coding in which ADJ was set as the baseline in the statistical models, while the majority of other existing findings on POS were based on sum coding with no baseline. In other words, the former shows whether a particular POS has a significant effect compared to the baseline, while the latter shows whether POS overall has an effect. For the frequency and proportion of each POS at switch points, see Appendix 2.
9. Zheng (1997)’s category of falling tone includes not only T4, but also T5 (neutral tone), and T3 that precedes T1, T2 and T4 (due to tone sandhi).
10. Although it remains unclear whether contact speakers in Canada speak Hong Kong English in monolingual contexts, it is not uncommon for these speakers to use Hong Kong English prosody in code-switching contexts, for example, to pronounce English words within Cantonese strings. Despite very little work documenting this tendency, we were able to confirm this observation through both the HLVC recordings and verification from various Cantonese heritage speakers.
11. The frequency and proportion of each POS at switch points across all datasets can be found in Appendix 2.

References

- Ahmad, B., & Jusoff, K. (2009). Teachers’ code-switching in classroom instructions for low English proficient learners. *English Language Teaching*, 2(2), 49–55.
- 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(4), 558–570. <https://doi.org/10.1016/j.wocn.2011.03>
- Australian Bureau of Statistics. (2017). *Census 2016: Australian Capital Territory*. Technical Report, ABS, Canberra, ACT, Australia.
- Balukas, C., & Koops, C. (2015). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, 19(4), 423–443. <https://doi.org/10.1177/1367006913516035>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Ben-Mosche, D., & Pyke, J. (2012). *The Vietnamese diaspora in Australia: Current and potential links with the homeland*. Report of an Australian Research Council Linkage Project. Technical Report, The Australian Research Council, Canberra, ACT, Australia.
- Brunelle, M. (2009a). Northern and Southern Vietnamese tone coarticulation: A comparative case study. *Journal of the Southeast Asian Linguistics Society*, 1, 49–62.
- Brunelle, M. (2009b). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics*, 37(1), 79–96. <https://doi.org/10.1016/j.wocn.2008.09.003>
- Carstens, A. (2016). Translanguaging as a vehicle for L2 acquisition and L1 development: Students’ perceptions. *Language Matters*, 47(2), 203–222. <https://doi.org/10.1080/10228195.2016.1153135>
- Chen, C. Y. (1983). A fifth tone in the Mandarin Spoken in Singapore. *Journal of Chinese Linguistics*, 11(1), 92–119.

- Cheng, K. S. K. (2017). Beginning or on-going? 2b-3a tone change in Hong Kong Cantonese revisited. *Journal of Chinese Linguistics*, 45(2), 313–343. <https://doi.org/10.1353/jcl.2017.0015>
- Cheung, W. H. Y. (2009). Span of high tones in Hong Kong English. *Annual Meeting of the Berkeley Linguistics Society*, 35(1), 72. <https://doi.org/10.3765/bls.v35i1.3599>
- Cheung, Y. S. (1985). Power, solidarity, and luxury in Hong Kong: A sociolinguistic study. *Anthropological Linguistics*, 27(2), 190–203.
- Chin, N. B., & Cavallaro, F. (2021). The curious case of mandarin Chinese in Singapore. In R. Jain (Ed.), *Multilingual Singapore* (pp. 159–178). Routledge.
- Chong, A. J., & German, J. S. (2015). Prosodic phrasing and F0 in Singapore English. In *ICPHS 2015*. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1010.pdf>
- Choo, Y. (2015). *Variant and invariant: A preliminary approach to the exploration of the emergence of Singapore Mandarin* [PhD thesis]. Nanyang Technological University.
- Clyne, M. (2003). Dynamics of transversion. In M. Clyne (Ed.), *Dynamics of language contact: English and immigrant languages, Cambridge approaches to language contact* (pp. 159–192). Cambridge University Press.
- Clyne, M. G. (1967). *Transference and triggering. Observations on the language assimilation of postwar German-speaking migrants in Australia*. Martinus Nijhoff.
- Clyne, M. G. (1980). Triggering and language processing. *Canadian Journal of Psychology/revue Canadienne de Psychologie*, 34, 400–406.
- Daniel, S. M., Jiménez, R. T., Pray, L., & Pacheco, M. B. (2019). Scaffolding to make translanguaging a classroom norm. *TESOL Journal*, 10(1), e00361.
- de Marneffe, M. C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal dependencies. *Computational Linguistics*, 47(2), 255–308. https://doi.org/10.1162/coli_a_00402
- Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of Memory and Language*, 89, 110–137.
- Fung, R. S. Y., & Lee, C. K. C. (2019). Tone Mergers in Hong Kong Cantonese: An asymmetry of production and perception. *The Journal of the Acoustical Society of America*, 146(5), EL424–EL430. <https://doi.org/10.1121/1.5133661>
- German, J., & Chong, A. (2018). Stress, tonal alignment, and phrasal position in Singapore English. In *Sixth international symposium on tonal aspects of languages* (pp. 150–154). https://www.adamjchong.com/uploads/8/9/5/9/89596735/tal_germanchong_2018_final.pdf
- Gibbons, J. (1987). *Code-mixing and code choice: A Hong Kong Case study*. Multilingual Matters.
- Gupta, A. F. (1994). *The step-tongue: Children's English in Singapore*. Multilingual Matters.
- Gussenhoven, C. (2012). Tone and intonation in Cantonese English. In *The third international symposium on tonal aspects of languages* (p. 4). https://isca-speech.org/archive_v0/tal_2012/presentations/tl12_O3-04_p.pdf
- Hartig, F. (2022). *DHARMa: Residual diagnostics for hierarchical (multi-level / mixed) regression models* (R package version 0.4.6). <https://CRAN.R-project.org/package=DHARMa>
- Hashim, A. (2020). Malaysian English. In K. Bolton, W. Botha, & A. Kirkpatrick (Eds.), *The handbook of Asian English* (pp. 373–397). John Wiley. <https://doi.org/10.1002/9781118791882.ch16>
- Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. <https://sentometrics-research.com/publication/72/>
- Huang, H., Kong, M. S. K. I., bors-homu Chen, W., Yang, Y., Wang, D., Bot, S., Gates, T., Lee, A., Bekkers, D., Timperi, J., Wang, M., Badger, T. G., Yang, W., Su, Y., & Guan, Z. (2023) *Pypinyin*. Zenodo. <https://doi.org/10.5281/zenodo.7538192>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer.
- Joseph, J. E. (2004). Case study 1: The new quasi-nation of Hong Kong. In J. E. Joseph (Ed.), *Language and identity: National, ethnic, religious* (pp. 132–161). Palgrave Macmillan.
- Khoo, K. U. (2017). Malaysian mandarin variation with regard to mandarin globalization trend: Issues on language standardization. *International Journal of the Sociology of Language*, 2017(244), 65–86. <https://doi.org/10.1515/ijsl-2016-0057>

- Kuo, E. C. Y., & Jernudd, B. H. (1993). Balancing macro-and micro-sociolinguistic perspectives in language management: The case of Singapore. *Language Problems and Language Planning*, 17(1), 1–21. <https://doi.org/10.1075/lplp.17.1.01kuo>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- Laroussi, F. (2011). *Code-switching, languages in contact and electronic writings*. Peter Lang Verlag.
- Lee, J., Chen, L., Lam, C., Lau, C. M., & Tsui, T. H. (2022). PyCantonese: Cantonese linguistics and NLP in Python. In *Proceedings of the thirteenth language resources and evaluation conference* (pp. 6607–6611). European Language Resources Association.
- Lee, L. (2010). The tonal system of Singapore Mandarin. In *Proceedings of 22nd North American conference on Chinese Linguistics NACCL-22 & the 18th international conference on Chinese linguistics IACL-18 (IACL-18) (Vol. 1, pp. 345–362)*. <https://naccl.osu.edu/sites/naccl.osu.edu/files/25%20lee.pdf>
- Lenth, R. V. (2022). *Emmeans: Estimated marginal means aka least-squares means*. <https://cran.r-project.org/web/packages/emmeans/emmeans.pdf>
- Li, D. C. S. (2000). Cantonese-English code-switching research in Hong Kong: A Y2K review. *World Englishes*, 19(3), 305–322. <https://doi.org/10.1111/1467-971X.00181>
- Lim, L. (2009). Revisiting English prosody: (Some) new Englishes as tone languages? *English World-Wide*, 30(2), 218–239. <https://doi.org/10.1075/eww.30.2.06lim>
- Low, E. L. (2012). English in Singapore and Malaysia. In A. Kirkpatrick (Ed.), *Routledge handbooks in applied linguistics: The Routledge handbook of World Englishes*. Routledge.
- Luke, K. (2000). Phonological re-interpretation: The assignment of Cantonese tones to English words. *International Conference on Chinese Linguistics*. <https://hub.hku.hk/handle/10722/109030>
- Lynn, T., & Scannell, K. (2019). Code-switching in Irish tweets: A preliminary analysis. In *Proceedings of the Celtic language technology workshop* (pp. 32–40). European Association for Machine Translation.
- Lyu, D. C., Tan, T. P., Chng, E. S., & Li, H. (2010). SEAME: A mandarin-English code-switching speech corpus in South-East Asia. In *Interspeech 2010* (pp. 1986–1989). ISCA.
- Mok, P. P. K., Zuo, D., & Wong, P. W. Y. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, 25(3), 341–370. <https://doi.org/10.1017/S0954394513000161>
- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.
- Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching*. Clarendon.
- Nagy, N. (2011). A multilingual corpus to explore variation in language contact situations. *Rassegna Italiana di Linguistica Applicata*, 43(1–2), 65–84.
- Nagy, N. (2018). Linguistic attitudes and contact effects in Toronto’s heritage languages: A variationist sociolinguistic investigation. *International Journal of Bilingualism*, 22(4), 429–446. <https://doi.org/10.1177/1367006918762160>
- Nagy, N., Stanford, J., & Tse, H. (2020). Tone mergers in spontaneous speech and gaps in the Tone Inventory. https://sophia.stkate.edu/english_fac/47/
- Ng, B. C., & Cavallaro, F. (2019). Multilingualism in Southeast Asia: The post-colonial language stories of Hong Kong, Malaysia and Singapore. In *Multidisciplinary perspectives on multilingualism* (pp. 27–50). De Gruyter Mouton.
- Ng, E. C. (2012). Chinese meets Malay meets English: Origins of Singaporean English word-final high tone. *International Journal of Bilingualism*, 16(1), 83–100. <https://doi.org/10.1177/1367006911403216>
- Nguyen, L. (2018). Borrowing or code-switching? Traces of community norms in Vietnamese-English speech. *The Australian Journal of Linguistics*, 38(4), 443–466. <https://doi.org/10.1080/07268602.2018.1510727>
- Nguyen, L. (2021). *Cross-generational linguistic variation in the Canberra Vietnamese heritage language community: A corpus-centred investigation* [Thesis, University of Cambridge].

- Nguyen, L., & Bryant, C. (2020). CanVEC -the Canberra Vietnamese-English Code-switching natural speech corpus. In *Proceedings of the 12th language resources and evaluation conference* (pp. 4121–4129). European Language Resources Association.
- Nguyen, L., Yuan, Z., & Seed, G. (2022). Building educational technologies for code-switching: Current practices, difficulties and future directions. *Languages*, 7(3), 220.
- Nguyen, T. A. T. (2004). *Prosodic transfer: The tonal constraints of Vietnamese Acquisition of English stress & rhythm* [PhD thesis]. The University of Queensland.
- Nguyen, T. A. T., Ingram, C. L. J., & Pensalfini, J. R. (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics*, 36(1), 158–190. <https://doi.org/10.1016/j.wocn.2007.09.001>
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41(6), 407–420. <https://doi.org/10.1016/j.wocn.2013.07.005>
- Papalexakis, E., & Doğruöz, A. S. (2015). Understanding multilingual social networks in online immigrant communities. In *Proceedings of the 24th international conference on World Wide Web Companion (WWW'15)* (pp. 865–870). Association for Computing Machinery. <https://doi.org/10.1145/2740908.2743004>
- Pham, A. H. (2003). *Vietnamese tone: A new analysis*. Routledge.
- Piccinini, P., & Garellek, M. (2014). Prosodic cues to monolingual versus code-switching sentences in English and Spanish. In *Speech prosody 2014* (pp. 885–889). ISCA. <https://pages.ucsd.edu/~ppiccinini/publications/PiccininiGarellek2014.pdf>
- Poplack, S. (1980). Sometimes i'll start a sentence in Spanish y termino en espanol: Toward a typology of codeswitching. *Linguistics*, 18(7–8), 581–618. <https://doi.org/10.1515/ling.1980.18.7-8.581>
- Riehl, C. M. (2005). Code-switching in bilinguals: Impacts of mental processes and language awareness. In J. Cohen, J. MacSwan, K. Rolstad, & K. T. McAlister (Eds.), *ISB4: Proceedings of the fourth international symposium on bilingualism* (pp. 1945–1959). Cascadilla Press.
- RStudio Team. (2020). *RStudio: Integrated development environment for R*.
- Siebenhaar, B. (2006). Code choice and code-switching in Swiss-German Internet Relay Chat rooms. *Journal of Sociolinguistics*, 10(4), 481–506. <https://doi.org/10.1111/j.1467-9841.2006.00289.x>
- Sim, J. H. (2019). 'But you don't sound Malay!': Language dominance and variation in the accents of English-Malay bilinguals in Singapore. *English World-wide*, 40(1), 82–112. <https://doi.org/10.1075/eww.00023.sim>
- Statistics Canada. (2017). *Toronto census metropolitan area, Ontario and Ontario Province*. Census Profile, Statistics Canada.
- Statistics Canada. (2022). *Census of population*. Census Profile, Statistics Canada.
- Sun, J. (2020). Jieba: Chinese words segmentation utilities. <https://pypi.org/project/jieba/>
- Sybesma, R., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua*, 117(10), 1739–1783. <https://doi.org/10.1016/j.lingua.2006.10.003>
- Tan, R. S. K. (2016). How do we stress? Lexical stress in Malaysian and British English. In T. Yamaguchi, & D. Deterding (Eds.), *English in Malaysia, Brill's studies in language, cognition and culture* (Vol. 14, pp. 65–80). Brill.
- Tan, Y. Y. (2015). 'Native' and 'non-native' perception of stress in Singapore English. *World Englishes*, 34(3), 355–369. <https://doi.org/10.1111/weng.12151>
- Thai, B. D. (2005). Code choice and code convergent borrowing in Canberra Vietnamese. In T. Le (Ed.), *Proceedings of the international conference on critical discourse analysis: Theory into research*. University of Tasmania.
- Torres Cacoullos, R., & Travis, C. E. (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge University Press.
- Tse, H. (2019). *Beyond the monolingual core and out into the wild: A variationist study of early bilingualism and sound change in Toronto heritage Cantonese* [PhD dissertation]. University of Pittsburgh.
- Tse, H. (2022). What can Cantonese heritage speakers tell us about age of acquisition, linguistic dominance, and sociophonetic variation? In R. Bayley, D. R. Preston, & X. Li (Eds.), *Variation in second and heritage languages: Crosslinguistic perspectives* (pp. 97–126). John Benjamins.
- Tuc, H. D. (2003). *Vietnamese-English bilingualism: Patterns of code-switching*. Routledge Curzon.

- van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67(6), 1176–1190. <https://doi.org/10.1080/17470218.2013.850521>
- Wang, D. (2019). *Multilingualism and Translanguaging in Chinese language classrooms*. Palgrave Macmillan.
- Wang, X. (2012). *Mandarin spread in Malaysia*. The University of Malaya Press.
- Wee, L. H. (2016). Tone assignment in Hong Kong English. *Language*, 92(2), e67–e87. <https://doi.org/10.1353/lan.2016.0039>
- Wee, L. H., & Liang, Y. (2016). The nature of Hong Kong English tones as seen through a comparison of F0 profiles with Hong Kong Cantonese. *Yuyanxue Luncong*, 54, 257–276.
- Winata, G. I., Aji, A. F., Yong, Z. X., & Solorio, T. (2022). The decades progress on code-switching research in NLP: A systematic survey on trends and challenges. http://aixpaper.com/view/the_decades_progress_on_codeswitching_research_in_nlp_a_systematic_survey_on_trends_and_challenges
- Zhang, J. (2019). Tone mergers in Cantonese: Evidence from Hong Kong, Macao, and Zhuhai. *Asia-Pacific Language Variation*, 5(1), 28–49. <https://doi.org/10.1075/aplv.18007.zha>
- Zheng, L. (1997). Tonal aspects of code-switching. *Monash University Linguistics Papers*, 1(1), 53–63. <https://doi.org/10.4225/03/5930c7d962034>

Author biographies

Katrina Kechun Li is currently a PhD candidate in Theoretical and Applied Linguistics at the University of Cambridge. Her research interest is prosody and its role in conveying communicative functions, with a particular focus on the interaction between tone and intonation.

Li Nguyen is a postdoctoral research fellow at the University of Cambridge and the Australian National University. She is also a visiting lecturer at FPT University, Vietnam, currently leading their research network on Linguistics and Language Technology. Her work focuses on multilingualism, language contact, heritage language speakers, language variation and change and computational tools for low-resource language varieties.

Christopher Bryant is an Applied AI Research Scientist at Writer, where he researches and develops various text-based NLP services, and a Visiting Researcher at the University of Cambridge. His main research interests include grammatical error detection and correction, computational approaches to code-switching, and computer-aided language learning.

Kayeon Yoo is a Language Engineer at Amazon, developing Text-To-Speech voices based on Deep Learning Models. Prior to that, she completed a PhD in Linguistics with a specialism in Phonetics and Phonology. Her research interests include natural language processing, speech prosody, language variation and change, and dialects of Korean.

Appendix I

Full regression model of Cantonese/English

Predictors	Cantonese-English			English-Cantonese		
	Odds ratios	CI	<i>p</i>	Odds ratios	CI	<i>p</i>
(Intercept)	0.00	0.00–0.00	<.001	0.00	0.00–0.00	<.001
T1	1.40	0.97–2.03	.076	1.22	0.89–1.70	.221
T2	1.61	1.09–2.39	.017	1.77	1.29–2.44	<.001
T3	3.29	2.29–4.73	<.001	1.28	0.84–1.95	.256
T5	1.78	1.22–2.60	.003	0.82	0.58–1.17	.276
T6	1.82	1.28–2.59	.001	1.09	0.80–1.48	.571
Zipfreq	1.43	1.29–1.59	<.001	1.19	1.06–1.32	.002
POS [ADP]	1.92	1.05–3.52	.035	1.08	0.54–2.16	.834
POS [ADV]	1.82	1.23–2.70	.003	2.10	1.38–3.21	.001
POS [AUX]	1.86	1.14–3.04	.013	1.27	0.72–2.24	.410
POS [CCONJ]	1.11	0.68–1.82	.678	1.93	1.20–3.12	.007
POS [NOUN]	1.61	1.08–2.42	.020	1.05	0.67–1.62	.839
POS [NUM]	1.34	0.77–2.30	.298	1.57	0.92–2.68	.100
POS [PART]	1.33	0.81–2.19	.253	0.91	0.52–1.58	.739
POS [PRON]	3.28	2.19–4.89	<.001	3.63	2.37–5.56	<.001
POS [VERB]	2.25	1.53–3.29	<.001	1.65	1.10–2.48	.016
POS [X]	0.78	0.19–3.31	.739	4.88	2.43–9.81	<.001
GENDER [F]	0.61	0.39–0.93	.023	0.87	0.54–1.41	.566
GENERATION [1]	3.16	1.78–5.60	<.001	4.40	2.30–8.41	<.001
GENERATION [23]	5.15	2.91–9.09	<.001	9.59	5.06–18.17	<.001
Random effects						
σ^2	3.29			3.29		
τ_{00}	0.72	SPEAKER		0.47	SPEAKER	
	1.05	TOKEN		0.38	TOKEN	
τ_{11}	0.07	SPEAKER.T1		0.10	SPEAKER.T1	
	0.16	SPEAKER.T2		0.11	SPEAKER.T2	
	0.08	SPEAKER.T3		0.34	SPEAKER.T3	
	0.01	SPEAKER.T5		0.07	SPEAKER.T5	
	0.05	SPEAKER.T6		0.03	SPEAKER.T6	
	ρ_{01}	-0.10	SPEAKER.T1		0.12	SPEAKER.T1
	-0.95	SPEAKER.T2		-0.62	SPEAKER.T2	
	-0.78	SPEAKER.T3		-0.26	SPEAKER.T3	
	-0.81	SPEAKER.T5		0.91	SPEAKER.T5	
	-0.97	SPEAKER.T6		0.40	SPEAKER.T6	
<i>N</i>	25	SPEAKER		25	SPEAKER	
	15	TOKEN		15	TOKEN	
Observations	98,645			98,645		

Full regression model of Mandarin/English

Predictors	Mandarin-English			English-Mandarin		
	Odds ratios	CI	<i>p</i>	Odds ratios	CI	<i>P</i>
(Intercept)	0.03	0.02–0.06	<.001	0.06	0.03–0.10	<.001
T1	0.79	0.64–0.99	.036	0.98	0.83–1.16	.835
T3	0.91	0.77–1.08	.289	1.11	0.94–1.31	.219
T4	1.08	0.90–1.30	.415	1.25	1.10–1.42	.001
T5	0.94	0.76–1.15	.521	2.63	1.76–3.95	<.001
Zipfreq	1.18	1.12–1.24	<.001	0.97	0.92–1.03	.297
POS [ADP]	0.81	0.52–1.26	.344	2.40	1.62–3.55	<.001
POS [ADV]	1.07	0.72–1.57	.743	1.93	1.36–2.74	<.001
POS [CCONJ]	1.06	0.53–2.10	.874	2.89	1.65–5.06	<.001
POS [DET]	2.16	1.44–3.25	<.001	1.77	1.21–2.61	.003
POS [NOUN]	1.13	0.76–1.66	.552	1.10	0.76–1.57	.619
POS [NUM]	2.41	1.60–3.62	<.001	1.44	0.97–2.15	.074
POS [PART]	1.29	0.82–2.02	.265	1.40	0.88–2.24	.156
POS [PRON]	1.52	1.02–2.27	.041	3.36	2.33–4.82	<.001
POS [SCONJ]	0.45	0.23–0.88	.020	3.90	2.49–6.11	<.001
POS [VERB]	1.45	0.99–2.13	.056	1.00	0.70–1.42	.982
POS [X]	3.14	1.55–6.37	.001	1.08	0.41–2.87	.879
GENDER [F]	1.50	1.27–1.78	<.001	1.26	1.01–1.57	.043
LANGUAGE [sge]	1.47	1.24–1.74	<.001	1.20	0.96–1.50	.109
Random effects						
σ^2	3.29			3.29		
τ_{00}	0.55	TOKEN		0.19	TOKEN	
	0.19	SPEAKER		0.14	SPEAKER	
τ_{11}	0.11	SPEAKER.T1		0.06	SPEAKER.T1	
	0.04	SPEAKER.T3		0.06	SPEAKER.T3	
	0.08	SPEAKER.T4		0.01	SPEAKER.T4	
	0.04	SPEAKER.T5		0.07	SPEAKER.T5	
	0.19	SPEAKER.T1		0.76	SPEAKER.T1	
ρ_{01}	-0.89	SPEAKER.T3		0.39	SPEAKER.T3	
	-0.81	SPEAKER.T4		-0.48	SPEAKER.T4	
	-0.80	SPEAKER.T5		-0.86	SPEAKER.T5	
<i>N</i>	20	SPEAKER		20	SPEAKER	
	24	TOKEN		24	TOKEN	
Observations	57,594			57,594		

Full regression model of Vietnamese/English

Predictors	Vietnamese-English			English-Vietnamese		
	Odds ratios	CI	<i>p</i>	Odds ratios	CI	<i>p</i>
(Intercept)	0.03	0.01–0.05	<.001	0.02	0.01–0.04	<.001
T1	0.71	0.53–0.95	.021	0.97	0.71–1.33	.839
T2	0.81	0.61–1.08	.157	0.82	0.62–1.08	.158
T3	0.85	0.63–1.15	.305	0.86	0.63–1.17	.330
T5	0.96	0.71–1.30	.783	0.99	0.68–1.46	.967
T6	0.88	0.62–1.25	.487	0.85	0.61–1.18	.324
Zipfreq	1.11	1.03–1.21	.010	1.23	1.13–1.35	<.001
POS [ADV]	1.42	1.04–1.93	.027	1.08	0.86–1.37	.513
POS [CLS]	2.41	1.69–3.43	<.001	0.58	0.40–0.84	.004
POS [CONJ]	0.96	0.66–1.39	.815	0.88	0.67–1.17	.391
POS [DET]	0.79	0.51–1.22	.293	1.45	1.11–1.89	.006
POS [NOUN]	1.59	1.22–2.08	.001	0.56	0.45–0.70	<.001
POS [NUM]	3.26	2.45–4.35	<.001	0.99	0.77–1.28	.965
POS [PART]	0.50	0.22–1.13	.095	0.90	0.58–1.40	.652
POS [PREP]	1.98	1.47–2.67	<.001	2.07	1.66–2.59	<.001
POS [PRON]	1.44	1.08–1.94	.014	0.48	0.37–0.62	<.001
POS [VERB]	1.72	1.32–2.23	<.001	0.56	0.46–0.69	<.001
POS [X]	1.56	0.74–3.32	.246	0.46	0.18–1.14	.094
GENDER [F]	0.96	0.60–1.52	.850	0.91	0.56–1.47	.697
GENERATION [2]	1.53	0.94–2.48	.088	2.12	1.29–3.49	.003
Random effects						
σ^2	3.29			3.29		
τ_{00}	0.71	SPEAKER		0.96	SPEAKER	
	0.39	TOKEN		0.16	TOKEN	
τ_{11}	0.08	SPEAKER.T1		0.25	SPEAKER.T1	
	0.08	SPEAKER.T2		0.09	SPEAKER.T2	
	0.09	SPEAKER.T3		0.20	SPEAKER.T3	
	0.07	SPEAKER.T5		0.53	SPEAKER.T5	
	0.29	SPEAKER.T6		0.23	SPEAKER.T6	
ρ_{01}	-0.52	SPEAKER.T1		-0.41	SPEAKER.T1	
	-0.44	SPEAKER.T2		-0.42	SPEAKER.T2	
	-0.23	SPEAKER.T3		-0.40	SPEAKER.T3	
	-0.45	SPEAKER.T5		-0.06	SPEAKER.T5	
	-0.33	SPEAKER.T6		0.27	SPEAKER.T6	
<i>N</i>	43	SPEAKER		43	SPEAKER	
	15	TOKEN		15	TOKEN	
Observations	55,940			55,940		

Appendix 2

The frequency of each POS and its proportion at switch points across all datasets

POS	CAN-ENG		ENG-CAN		MAN-ENG		ENG-MAN		VIE-ENG		ENG-VIE	
	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
ADJ	31	2.04	27	2.16	30	0.45	38	0.54	69	2.28	154	5.74
ADP	18	1.18	12	0.96	161	2.40	233	3.29	–	–	–	–
ADV	239	15.71	245	19.57	1274	19.00	1412	19.95	180	5.95	313	11.66
AUX	43	2.83	24	1.92	–	–	–	–	–	–	–	–
CCONJ	73	4.80	78	6.23	13	0.19	24	0.34	–	–	–	–
CLS	–	–	–	–	–	–	–	–	522	17.26	153	5.70
CONJ	–	–	–	–	–	–	–	–	131	4.33	195	7.26
DET	–	–	–	–	396	5.90	193	2.73	46	1.52	188	7.00
NOUN	301	19.79	115	9.19	470	7.01	393	5.55	369	12.20	279	10.39
NUM	24	1.58	28	2.24	263	3.92	113	1.60	198	6.55	140	5.21
PART	124	8.15	112	8.95	969	14.45	986	13.93	35	1.16	66	2.46
PREP	–	–	–	–	–	–	–	–	170	5.62	369	13.74
PRON	251	16.50	300	23.96	912	13.60	2316	32.72	323	10.68	222	8.27
SCONJ	13	0.19	65	0.92	–	–	–	–	–	–	–	–
VERB	415	27.28	299	23.88	2192	32.68	1300	18.37	973	32.18	601	22.38
X	2	0.13	12	0.96	14	0.21	5	0.07	8	0.26	5	0.19