

The Influence of Cantonese on the Phonetic Transfer of English Consonants: Taking the Error of Fricatives as an Example

Ruoyu Li^{1*†}, Zhuomin Peng^{2†}

¹*BeiJing Normal University Zhuhai Campus, Zhuhai, China*

²*Guangdong University of Foreign Studies, Guangzhou, China*

**Corresponding Author. Email: lry20030503@163.com*

†These authors contributed equally to this work and should be considered as co-first author.

Abstract. Language transfer is a crucial factor in second language acquisition (SLA). While the influence of Mandarin Chinese on English phonology has been studied, the specific role of regional dialects, particularly Cantonese, requires further investigation. This study focuses on the phonetic transfer from first language (L1) Cantonese to English fricative consonants (/f/, /v/, /s/, /z/, /ʃ/, /ʒ/, /θ/, /ð/) among Chinese English L2 learners. Sixty participants will be recruited and divided into three groups based on their language backgrounds: (1) L1 Mandarin L2 English, (2) L1 Mandarin & Cantonese (bilingual) L2 English, and (3) L1 Cantonese L2 Mandarin L3 English. Participants will complete a language background questionnaire and a recorded production task involving minimal pairs and sentence contexts containing target fricatives and fillers. Acoustic analysis using Praat will measure the spectral properties of the produced fricatives. We hypothesize that (1) L1 Cantonese speakers will exhibit greater deviation from native-like English fricative production compared to L1 Mandarin speakers, (2) the bilingual group (Mandarin & Cantonese) will show an intermediate pattern, and (3) Specific fricatives absent or differing in Cantonese phonology will be most susceptible to transfer errors. The findings aim to contribute to understanding the role of specific L1 phonological systems and language status in SLA, with implications for tailored pronunciation instruction for learners with diverse Chinese dialect backgrounds.

Keywords: phonetic transfer, Cantonese, English consonants, fricatives, second language acquisition

1. Introduction

Multilingualism is no longer an exception but a norm in the 21st-century globalized classroom. In the context of China's greater degree of openness to the outside world, English has become a language that most students must master. In particular, with the rapid development of foreign trade and other sectors, the Guangdong–Hong Kong–Macao Greater Bay Area (GBA) has positioned English as a strategic regional language for commerce, media, and daily communication. At the

same time, in recent years, the number of students and professionals from the Chinese mainland (especially from Guangdong) studying or working in Hong Kong has increased. Therefore, the demand for learning English in the Cantonese-speaking region has become even greater than before.

Cantonese, a major Chinese dialect, is prevalent in Guangdong Province, the southwestern part of Guangxi, Hong Kong, and Macau. Its phonological system exhibits significant differences from both Mandarin and English, which may result in unique phonetic transfer patterns among Cantonese-speaking English learners [1]. Therefore, it is extremely necessary to conduct research on the impact of learning Cantonese as a native language on the acquisition of a second or third language.

Fricatives, particularly, play a crucial role in speech clarity. Both Mandarin and English phonetic systems contain fricatives in terms of articulation [2]. In comparison, the fricative system in English is more complex than that of Mandarin and Cantonese. English fricatives include /f/, /v/, /θ/, /ð/, /s/, /z/, /ʃ/, /ʒ/, and /h/, among which several sounds are either absent or differ phonemically in Cantonese and Mandarin. For example, the dental fricatives /θ/ and /ð/ do not exist in the phonemic inventory of either Mandarin or Cantonese. Moreover, English fricatives can be categorized into voiceless and voiced fricatives, but the voiceless-voiced distinction is not phonemic in either Mandarin or Cantonese fricatives, meaning that native speakers do not perceive the differences sensitively, leading to difficulties in producing voiced fricatives in English such as /v/, /z/, and /ʒ/.

It is worth noting that due to different environments and educational backgrounds, learners show differences in the order of acquiring dialects and Mandarin. Some learners may study Cantonese and Mandarin as common mother tongues, while others may first acquire Cantonese as their mother tongue. Existing studies have rarely discussed the impact of this difference on L3 acquisition.

2. Literature review

There are relatively few historical studies in the area of linguistics about English learning in the Cantonese-speaking area. Only in the past few decades have studies focused about some phonetic phenomena of English learners in these areas, such as Guangdong and Hong Kong. Previously, some scholars have conducted studies on the perception of English phonetics by Cantonese learners as a second language [3], but research on English phonetic production remains scarce.

2.1. Language transfer

First is the theory that our research is based on, which is called "language transfer". "Language transfer" refers to the psycholinguistic process whereby learners subconsciously rely on phonological, lexical, syntactic or pragmatic features of their mother tongue (or any previously acquired language) when producing or comprehending a second/foreign language. When L1 and L2 structures coincide, positive transfer facilitates acquisition; when they diverge, negative transfer arises and fossilized errors may occur [4]. Kong shows that Cantonese learners of English experience a double transfer effect: L1 gaps such as the absence of /r/, /v/ and /θ/ trigger substitutions like /l/, /w/ and /s/, while L2 Mandarin interference, such as retroflex versus dental distinctions, glide deletion, further distort English segments, creating a Cantonese-Mandarin-English cascade of interference [4]. Huang adds that the lack of final consonants and consonant clusters in Cantonese leads to systematic final-consonant deletion and epenthesis, so *guests* becomes /ges/ and *please* becomes /pə'li:z/ [5]. These findings reveal that transfer is not a unidirectional L1-to-L2 phenomenon; it can operate across multiple already-learned languages and may be horizontally reinforced by an intermediate second language. Pedagogically, teachers must employ contrastive

analysis, high-variability perceptual training and explicit corrective feedback to block negative transfer and help learners establish independent L2 phonetic categories.

2.2. Language status and language distance

The concepts of language status and psychotypology (perceived language distance) are crucial for understanding transfer patterns. Learners are more likely to transfer structures they perceive as similar between their L1 and the target language. Regarding status, an L2 may influence L3 acquisition, a phenomenon known as L2 transfer [6,7]. For bilinguals, the phonological systems of both languages can interact and influence L3 learning [8]. Defining bilingualism for research purposes is complex. For this study, a bilingual (Mandarin & Cantonese) participant is operationally defined as an individual who (1) Acquired both Mandarin and Cantonese naturally from childhood (sequential or simultaneous bilingualism) (2) Uses both languages regularly in daily life (at least 30% usage for each based on self-report questionnaire adapted from language background protocols like those discussed on LHQ [9]), and (3) Self-reports a high proficiency (minimum 4 out of 5 on a Likert scale) in listening and speaking for both languages. This definition emphasizes early acquisition, regular use, and high self-assessed proficiency, ensuring participants have robust phonological representations in both Chinese varieties.

2.3. Cantonese influence on English phonology

Existing research on Cantonese ESL learners has identified specific challenges. Common errors include the substitution of /θ/ and /ð/ [4,10], devoicing of English voiced fricatives like /v/ and /z/ [4,10,11], and difficulties with the /s/-/ʃ/ distinction due to differences in the Cantonese inventory [4]. However, many studies focus on perception or general error analysis. Fewer studies employ detailed acoustic analysis to objectively quantify production differences, and even fewer directly compare Cantonese-dominant learners with Mandarin-dominant learners and Mandarin-Cantonese bilinguals within the same experimental framework. This gap highlights the need for the present study, which aims to provide a more nuanced understanding of how specific L1/L2 backgrounds modulate phonetic transfer, using acoustically rigorous methods.

3. Research questions and hypothesis

Based on the literature review, this study will manage to answer the following research questions:

1) Does L1 background (Mandarin vs. Cantonese vs. Mandarin-Cantonese bilingual) significantly affect the accuracy of English fricative production by Chinese L2/L3 learners?

2) Compared to the others, do simultaneous bilingual groups (group 2) exhibit unique production patterns, resulting in a certain hierarchical order in the degree of deviation in pronunciation output among the three groups?

3) Are certain English fricatives more susceptible to transfer errors from Cantonese than others?

In the native language phoneme hypothesis, also known as the native language priority hypothesis [12,13], in trilingual acquisition, the native language has a greater influence on L3 than L2 and is more dominant. Previous research on trilingual learners has also found that in Chinese-English-Japanese trilingual learning, the percentage of native language transfer is higher than that of second language transfer [14]. Therefore, this study hypothesizes that, regardless of the previous bilingual acquisition pattern, the native language is likely to exert the strongest influence on L3. That is, for learners in Group 3 whose L1 is Cantonese, their English pronunciation is more affected

by Cantonese, which is reflected in the data as a greater deviation in the articulation of some fricatives.

Hypotheses were made accordingly: H1: The L1 Cantonese group (Group 3) will exhibit significantly greater deviation from native-like spectral features in English fricative production compared to the L1 Mandarin group (Group 1). H2: The Mandarin-Cantonese simultaneous bilingual group (Group 2) will exhibit output patterns that fall between those of Group 1 and Group 3. The degree of deviation in English fricative production among the three groups of participants is ranked from lowest to highest as Group 1, Group 2 and Group 3. H3: Fricatives that are absent or markedly different in Cantonese will show the highest degree of production inaccuracy across all groups, with the most significant deviations observed in Group 3.

4. Methods

4.1. Participants

In this study, a total of 60 university students (aged 18-24) will be recruited. They will be divided into three groups based on their language background:

Group 1 (L1 Mandarin, n=20): The sequential bilingualism group. Native speakers of Mandarin Chinese with limited/no proficiency in Cantonese. L2 is English.

Group 2 (L1 Mandarin & Cantonese Bilingual, n=20): The simultaneous bilingualism group. Early bilinguals in Mandarin and Cantonese, meeting the operational definition outlined in section 1.2.2. L2 is English.

Group 3 (L1 Cantonese, n=20): The sequential bilingualism group. Native speakers of Cantonese, who acquired Mandarin later as an L2. L3 is English.

All participants will be required to have comparable English proficiency levels to control for the effect of general L2 proficiency on pronunciation accuracy. This will be ensured by recruiting participants with similar IELTS scores (target band: 5.5-6.5, with minimal variation between groups). Participants must have normal hearing and no reported speech disorders. Informed consent will be obtained prior to the experiment.

Each group of participants is required to have a balanced number of females and males, while minimizing the interference of factors such as participants' years of education and prior testing experience. Furthermore, it is necessary to collect data on how many individuals have received speech training before the experiment begins, as well as whether their accents are closer to the official English RP (Received Pronunciation) or exhibit other accentual tendencies.

4.2. Materials

4.2.1. Questionnaire

LHQ3.3, the Latest updated version of the Language History Questionnaire, will be chosen as a tool for investigating the language background of the subjects. According to the official website, this series of questionnaires has been cited 1,333 times by studies according to Google Scholar and has been widely used and recognized around the world [15]. Except for assessing the linguistic background and language proficiency of multilingual or second language learners, LHQ3 has an automatic scoring system, which allows researchers to assign different weights to different modules when calculating aggregated scores. This addresses the issue of varying focus areas among different

researchers regarding the language use and background of multilingual users [9,16,17]. It has been widely utilized and cited in authoritative literature.

Language history questionnaire (LHQ). Go to <https://blclab.org/lhq3/> to use the online version and for reference

(1) Participant ID number		(2) Age	
(3) Gender	☐ Male ☐ Female	☐ Non-binary	☐ Non-relevant
(4) Education	☐ Graduate school (Doctor) ☐ Graduate school (Master) ☐ College (Bachelor)		
	☐ High school ☐ Middle school ☐ Elementary school ☐ Other		
(5) Parents' Education	Father	☐ Graduate school (Doctor) ☐ Graduate school (Master) ☐ College (Bachelor)	
		☐ High school ☐ Middle school ☐ Elementary school ☐ Other	
	Mother	☐ Graduate school (Doctor) ☐ Graduate school (Master) ☐ College (Bachelor)	
		☐ High school ☐ Middle school ☐ Elementary school ☐ Other	
(6) Handedness		☐ Right-handed ☐ Left-handed	☐ Ambidextrous

(7) Indicate your native language(s) and any other languages you have studied or learned, the age at which you started using each language in terms of listening, speaking, reading, and writing, and the total number of years you have spent using each language.
*Notes For "Years of use", you may have learned a language, stopped using it, and then started using it again. Please give the total number of years.

Language	Listening	Speaking	Reading	Writing	Years of use*

(8) Country of origin

(9) Country of residence

(10) If you have lived or traveled in countries other than your country of residence for three months or more, then indicate the name of the country, your length of stay (in Months), the language you used, and the frequency of your use of the language for each country.
* You may have been to the country on multiple occasions, each for a different length of time. Add all the trips together

Country	Length of stay (in Months)*	Language	Frequency of use						
			1	2	3	4	5	6	7
			☐ 1, ☐ 2, ☐ 3, ☐ 4, ☐ 5, ☐ 6, ☐ 7.						
			☐ 1, ☐ 2, ☐ 3, ☐ 4, ☐ 5, ☐ 6, ☐ 7.						
			☐ 1, ☐ 2, ☐ 3, ☐ 4, ☐ 5, ☐ 6, ☐ 7.						
			☐ 1, ☐ 2, ☐ 3, ☐ 4, ☐ 5, ☐ 6, ☐ 7.						

(11) Indicate the way you learned or acquired your non-native language(s). Check one or more boxes that apply.
* e.g., immigrating to another country where the dominant language is different from your native language so you learn this language through immersion in the language environment.

Non-native Language	Immersion*	Classroom instruction	Self-learning
	☐	☐	☐
	☐	☐	☐
	☐	☐	☐
	☐	☐	☐

Figure 1. LHQ3 questionnaire

4.2.2. Production task stimuli

The recording task will consist of two parts: The first part is Sentence Contexts. In this part, the presented sentences will embed the target words in carrier sentences to elicit production in a more natural and contextual environment. There will be a total of 36 target word items. Based on the literature review, the focus will be on two contrast pairs: labiodental fricative /f/ vs. dental fricative /θ/, and alveolar fricative /s/ vs. post-alveolar fricative /ʃ/, with 4 words per pair, each embedded in 3 types of contexts. Additionally, 12 filler items containing non-target sounds will be included, making a total of 36 stimuli. The second part is Minimal Pairs word production, contrasting the target English fricatives. There will be 6 specific contrast word pairs for the two contrasts, each pair appearing twice. In addition, 12 filler items containing non-target sounds will be included, making a total of 36 stimuli.

Sentence contexts will be presented to the participants first, followed by the Minimal Pairs, in order to avoid task monotony and reduce participants' attention to specific phonological targets. The presence of filler items also helps achieve this goal. All target words and sentences will be presented

on the screen in a random order in the form of a PowerPoint to ensure participants produce the expected targets.

Table 1. Fricatives minimal pairs

Minimal Pair	Word Pair	Minimal Pair	Word Pair
/f/ vs. /θ/	fin thin	/s/ vs. /ʃ/	sip ship
	free three		ass ash
	first thirst		sort short
	Sentence Context		Sentence Context
	She dropped the ship into the sea. I can see why you sip the tea slowly.		The tickets are free until three o'clock. A thin fin broke the water's surface.

4.2.3. Tools and equipment

TASCAM DR44WL linear PCM recorder (with 44.1 kHz sampling rate, 16-bit quantization level) and AKG C544L headset microphones will be used [4]. Praat 6.0 will be used to create the phonetic stimuli for the identification task [11]. Each stimulus was preceded by a 400-ms period of silence to space out the stimuli. A synthetic sound with a fundamental frequency of 500 Hz and a duration of 400 ms was added as a start signal to alert the participants. Finally, a 1000-ms period of silence was added to the end of each stimulus so that the participants would have enough time to understand the provided options on the computer screen [10].

The experiment will be conducted in a quiet room. Participants will wear closed-back headphones to clearly hear the start signal and minimize background noise interference. The microphone will be kept at a constant distance from the participant's mouth, approximately 3 centimeters.

4.3. Procedure

The experiment is divided into three procedures:

Distribute and Collect the LHQ3 Questionnaire: Select 60 experimental participants who meet the language background requirements through the questionnaire.

Informed Consent: Participants will read and sign the consent form, then complete the language background questionnaire.

Familiarization: Before each task, Participants will undergo a short practice session to familiarize themselves with the task. Then, the participants were asked to introduce themselves in the target language (saying, for example, their name, university) in a way of semi-structured interview. The purpose of the self-introduction was to ease them into their target language mode.

Production Task: Participants will be presented with the visual prompts for each item. They will be instructed to repeat the word or sentence clearly and naturally. Each target item will be produced three times to ensure sufficient data for analysis. The entire session will be recorded.

4.4. Analysis

4.4.1. Acoustic analysis

Recordings will be analyzed using Praat (version 6.0). The pronunciation model will be selected from Received Pronunciation English (RP) pronunciation from a reputable source in the Longman Pronunciation Dictionary sound bank.

The spectral characteristics of fricatives mainly depend on the place of articulation and the shape of the resonating cavity [18]. Previous research has shown [19] that spectral peak location, spectral moments, and both normalized and relative amplitude can serve to distinguish all four places of fricative articulation. In this study, the accuracy and variance of the discrimination task will be used as computational measures. Prior to analysis, data corresponding to 12 interfering words will be excluded.

4.4.2. Statistical analysis

First, the recorded audio files will be imported into the Windows edition of the Praat system, and the target fricative segments will be annotated. The spectral peak location, spectral moments, and both normalized and relative amplitude of the two sets of fricatives will be analyzed. In the /f/ vs. /θ/ group, special attention will be paid to the spectral centroid and spectral slope, while in the /s/ vs. /ʃ/ contrast, special attention will be given to the F3 formant transition. These analyses will be conducted through computation. Next, the perceptual accuracy for each of the 12 sets in the three groups will be calculated to summarize differences in consonant perception among speakers from different dialect regions. Third, G-power will be used to perform a one-way ANOVA to verify whether the L1 background of speakers from Cantonese-speaking regions has a significant effect on consonant production, and the variance of the three groups will be calculated and ranked to address the research questions.

5. Discussion and implication

5.1. Discussion

The expected research results will be discussed in conjunction with the hypotheses and existing literature. If the results support Hypothesis H1, it will indicate that the transfer effect from L1 Cantonese is stronger than that from L1 Mandarin, which can be attributed to specific phonetic differences between Cantonese and English. If Hypothesis H2 is supported, it will suggest a complex interaction within the bilingual phonetic system, with gradient ranking evidencing the transfer effect of a Cantonese L1 or L2 background on the pronunciation of fricatives. If results support Hypothesis H3, they will highlight areas where Cantonese learners are prone to errors in learning fricatives. The discussion will also address the theoretical implications of language distance and native language status in transfer. The use of acoustic measurements provides objective evidence beyond perceptual error coding and underscores the practical significance of this study for English learners in Cantonese-speaking regions.

5.2. Implication

This proposed study aims to investigate the specific effects of L1 Cantonese on the production of English fricatives, comparing them to effects from L1 Mandarin and Mandarin-Cantonese

bilingualism. By employing a controlled group design, detailed language background assessment, and acoustic phonetic analysis, it seeks to acquire. The findings are expected to contribute valuable insights to theories of L2 phonetic transfer and have practical significance for English language teaching in Cantonese-speaking regions and for learners with Cantonese heritage worldwide.

5.3. Limitation

Firstly, the findings may have limited generalizability due to the specific phonological focus on fricatives and the particular language backgrounds of the participants and the operational definition of bilingualism. Secondly, the questionnaire, relying heavily on self-reported proficiency, may cause problems in the process of selecting participants.

Acknowledgement

Li Ruoyu and Peng Zhuomin contributed equally to this work and should be considered co-first authors.

References

- [1] Chen Peimiao. (2021). A Study on the Negative Transfer of Cantonese on English Phonetic Learning of English Majors—Taking Plosives and Fricatives in English Consonants as Examples. In Proceedings of the 14th Graduate Academic Seminar and the 4th Forum on Foreign Language and Literature Doctorate held by Xiamen University (pp. 56-67). School of Foreign Languages, Xiamen University
- [2] Wu Jianxiao. (2008). A Comparative Analysis of Chinese and English Fricatives. *Journal of Mudanjiang Normal University (Philosophy and Social Sciences Edition)*, (05), 69-70. [https://doi.org/10.13815/j.cnki.jmtc\(pss\).2008.05.023](https://doi.org/10.13815/j.cnki.jmtc(pss).2008.05.023).
- [3] CHAN, A. Y. W. (2012). Cantonese English as a Second Language Learners' Perceived Relations Between "Similar" L1 and L2 Speech Sounds: A Test of the Speech Learning Model. *The Modern Language Journal*, 96(1), 1–19. <http://www.jstor.org/stable/41478802>
- [4] Kong Guiying. (2016). A contrastive Analysis on Mandarin and English Pronouncing Errors in Cantonese Region. *Overseas English*, (04), 9-11.
- [5] Huang Yifan. (2009). How to Reduce the Negative Transfer of Cantonese on English Pronunciation Learning. *Science and Technology in Western China*, 8(17), 83-84.
- [6] Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Mouton de Gruyter.
- [7] Fouser, R., Problems and Prospects in Third Language Acquisition Research [J]. *Language Research*, 1995, (387-415).
- [8] Cenoz, J., The Effect of Linguistic Distance, L2 Status and Age on Cross-Linguistic Influence [M]. 2001, (171-203).
- [9] Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, 23(5), 938-944.
- [10] Tong Chaihong. (2017). Analysis of English Pronunciation Errors in English Major Students Whose Native Language is Cantonese (Master's Thesis, Guangxi Normal University).
- [11] Tian Shaoping. (2009). Exploration of English Pronunciation Errors Among Students in Cantonese Regions. *Modern Education Forum*, (02), 60-63.
- [12] Cai Jinting. Research on the Transfer in Trilingual Grammar Acquisition: Theory, Evidence, and Reflection [J]. *Foreign Language Research*, 2020, 37(6): 1-8, 42
- [13] Fan Linzhou. Evolutionary Causes and Transformation Strategies of Native Language Transfer [J]. *Chinese University Teaching*, 2022(10): 36-41.
- [14] Tang Jianmin, Zhan Di, Zhu Lixia. Native and Second Language Transfer in Japanese Word Usage by Chinese-English-Japanese Trilinguals [J]. *Foreign Language Journal*, 2022(1): 84-91.
- [15] LCLab. What's New [Internet]. LHQ-BLCLab. 2025 [cited 2025 Oct 31]. Available from: <https://lhq-blclab.org/faq/>

- [16] Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(3), 673-680.
- [17] Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior research methods*, 38(2), 202-210.
- [18] LI Shanpeng, GU Wentao. Acoustic characteristics of Mandarin affricates. *Journal of Tsinghua University (Science and Technology)*, 2016, 56(11): 1202-1208.
- [19] Jongman A, Wayland R, Wong S. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 2000, 125(6): 3962–3973.