Studies in English Language Teaching
ISSN 2372-9740 (Print) ISSN 2329-311X (Online)
Vol. 12, No. 3, 2024
www.scholink.org/ojs/index.php/selt

Original Paper

Pronunciation Difficulties in Cantonese English Learners: A Focus on Segmental Errors

Xinting Zhang[1], Wenhui Zhu[1] & Yu Chu[1]

[1] Southern University of Science and Technology, China

Received: August 3, 2024 Accepted: September 20, 2024 Online Published: September 30, 2024

doi:10.22158/selt.v12n3p111 URL: http://dx.doi.org/10.22158/selt.v12n3p111

Abstract

This study investigates the pronunciation difficulties faced by Cantonese-speaking English as a Foreign Language (EFL) learners, focusing on segmental errors in both vowels and consonants. A production experiment was conducted with 13 Cantonese university students, examining their performance on carefully selected minimal pairs embedded in short sentences. The results revealed common pronunciation errors, including the substitution of /æ/ with /ɛ/, final consonant devoicing, and /v/ substituted with /w/. These errors were analyzed in relation to the functional load (FL) principle, which emphasizes the importance of phonological contrasts that contribute most to intelligibility. The findings highlight that errors with high FL values, such as /æ/-/ɛ/ and /p/-/b/, should be prioritized in pronunciation instruction to enhance learners' overall intelligibility. The study concludes with pedagogical implications for teaching pronunciation more effectively in Cantonese EFL contexts.

Keywords

Pronunciation difficulties, Pronunciation teaching, Teaching priorities, Functional Load Principle

1. Introduction

In the teaching of English as a Foreign Language (EFL), pronunciation plays a vital role in enhancing learners' communicative abilities. Clear and comprehensible pronunciation not only helps learners express their ideas more effectively but also helps them engage in academic, social, and professional interactions. However, many university English courses do not offer specialized pronunciation classes, and learners face a wide range of pronunciation issues that teachers cannot address comprehensively within the limited class time. Therefore, it is crucial to prioritize key areas of instruction to maximize teaching effectiveness.

In recent years, the focus of pronunciation teaching has shifted from reducing accents to improving

intelligibility (Levis, 2018; Munro & Derwing, 1995). While accents may persist, research shows that accented speech can still be understood. Hence, the goal of pronunciation teaching has evolved from achieving native-like pronunciation to ensuring that learners can be clearly understood. This study investigates the segmental pronunciation errors of Cantonese EFL learners, using a production experiment to provide detailed insights into their difficulties. The findings will inform pedagogical recommendations for improving pronunciation instruction and guiding learners toward greater intelligibility.

2. Literature Review

2.1 Functional Load Principle

Functional Load Principle, or FL, has been used to rank segmental contrasts according to their importance in English pronunciation. The basic definition of FL refers to the "amount of functional work" a phonemic contrast carries within a language system. The simplest and most intuitive method of quantifying FL was proposed by Martinet (1952), known as the "minimalist" measurement, which evaluates FL based on the number of minimal pairs. Catford (1987) built a functional load list using this method. Brown (1988), however, adopted a more detailed analysis strategy that not only counted minimal pairs but also considered their frequency of occurrence and the number of word pairs (such as "live" and "leave") that are distinguished solely by phonemic contrast within the same word class. Despite some differences in methodology, both approaches reveal a high degree of consistency when identifying contrasts with high FL, such as /l/ vs. /n/ and /i/ vs. /ɪ/.

In terms of pronunciation teaching, research has shown that errors in high-FL contrasts (e.g., substituting /n/ for /l/) have a greater impact on listeners' perception of foreign accents and intelligibility, while low-FL errors have less impact (Munro & Derwing, 2006). For example, phonetic errors with /θ/ and /ð/ are common among Chinese learners, but these phonemes have relatively low FL. Mispronouncing them usually has little effect on communication, as /θ/ may be pronounced as /s/ or /f/, and /ð/ as /z/ or /d/, without obstructing meaning. In contrast, confusion between /i/ and /ɪ/ carries a higher FL. For instance, errors between /i/ and /ɪ/ in pairs like "slip" and "sleep" can lead to meaning confusion, thus affecting communication.

Therefore, in pronunciation teaching, educators should prioritize high-FL phonemes based on the functional load principle, helping students master these critical phonemic contrasts. This FL-based teaching strategy ensures efficient use of resources, enabling students to maximize their improvement in pronunciation clarity and communicative ability within a short period.

2.2 Review of the Literature on the Pronunciation Difficulties of Cantonese EFL Learners

Numerous studies have explored the pronunciation difficulties of Cantonese speakers learning English. While much of the research describes these errors, fewer studies employ experimental designs. Additionally, many of these studies focus on Hong Kong English, a variety influenced by Cantonese, whereas other Cantonese-speaking regions may present different challenges. Two experimental studies

have been particularly influential in designing the test stimuli for the current investigation.

2.3 Vowels

Zhang (2012) investigated the intelligibility of Cantonese-accented English among four Cantonese college students using a transcription task. The study found that regarding vowels, university students from Cantonese-speaking regions did not seem to experience significant difficulties. The primary issue identified was confusion between the long vowel /iː/ and the short vowel /ɪ/, resulting in nine errors. In contrast, Chan (2010) conducted a systematic study of 40 Hong Kong ESL learners, who participated in three reading tasks involving isolated words, minimal pairs, and short passages. The findings on Hong Kong English revealed that /æ/ was the most inaccurately produced vowel, with an accuracy rate of only 16.4%. On the other hand, /ɪ/ and /e/ had the highest accuracy rates, both above 95%. Other problematic vowels included /iː/ (57.5% accurate), /ɔː/ (56% accurate), and /uː/ (64.8% accurate).

2.4 Consonants

According to Zhang (2012), for consonants produced by Cantonese college students in Guangdong, the primary issues observed were sound addition, confusion, and elision. Common confusions included mixing /n/ and /m/ in word-final positions, and confusion between voiceless and voiced pairs, such as /t/ and /d/ or /p/ and /b/. Additionally, there was confusion between /r/ and /l/, /r/ and /w/, /n/ and /l/, and /n/ and /ŋ/, where these sounds were often substituted for one another. Consonant elision typically involved the omission of final consonants, as in pronouncing “mine” as “my.”

Chan (2010) also found that the most problematic consonants for Hong Kong learners were voiced obstruents, which had low accuracy rates. For example, /ð/ had an accuracy of only 3.7%, /z/ 7.8%, /ʤ/ 11.1%, /v/ 15.3%, and /d/ 70%. The sound /ð/ was often substituted by /d/, /θ/, or /f/, while /z/ was frequently devoiced as /s/, and /ʤ/ was substituted by [tʃ]. Another challenging sound was the dark /l/ in the final position, as in “bill,” which was often pronounced with some form of vocalization.

The present study builds on this body of research, targeting problematic sounds identified through previous experimental studies to further analyze the pronunciation difficulties of Cantonese EFL learners at the segmental level.

3. Method

3.1 Participants

Thirteen participants were recruited via campus social media, all of whom were students from a university in Shenzhen, China. The sample included seven undergraduate and six graduate students (M = 22.58, SD =3.18). All participants were native speakers of Cantonese coming from different regions in Guangdong such as Maoming, Zhaoqing, and Foshan, and they predominantly communicate in Cantonese with their family members. They all had similar English learning experiences, having started their English education in school at the age of 9 or 10, with English mainly used in academic settings. Although there was considerable variability in their pronunciation proficiency, the researcher’s evaluation indicated that all participants exhibited characteristic features of Cantonese-accented English,

including the devoicing of final consonants, syllable-timed rhythm, and certain vowel confusions. Consequently, they were included in the study and categorized into two groups: high proficiency and low proficiency.

3.2 Materials

The focus of this study is to identify segmental pronunciation difficulties for Cantonese speakers of English. Materials were selected through a contrastive analysis of problematic sounds and common substitutions identified in previous research to keep the experiment within a practical time limit. The sound pairs chosen reflect phonological patterns typical of Cantonese-accented English, as discussed in the literature. Specifically, these sound pairs and their corresponding minimal pairs were selected as test items.

Table 1. Sound Pairs and Test Words for Vowels

Problematic sound pairs	Test words
/i/, /ɪ/	sheep, ship
/u/, /ʊ/	pool, pull
/ɔ/, /ɑʊ/	shots, shouts
/æ/, /ɛ/	pan, pen
/æ/, /aɪ/	back, bike
/ɛ/, /eɪ/	pen, pain

Table 2. Sound Pairs and Test Words for Consonants

Problematic sound pairs	Test words
/b/, /p/	bear, pear
/d/,/t/	dry, try
/ð/,/d/	they, day
/θ/, /f/	three, free
/θ/, /s/	think, sink
/z/, /s/	zip, sip
/ʤ/, /tʃ/	junk, chunk
/v/, /f/	van, fan
/v/, /w/	vine, wine
/r/, /l/	right, light
/l/, /n/	low, no
final /p/, final /b/	cap, cab
final /t/, final /d/	cart, card

final /k/, final /g/	back, bag
final /s/, final /z/	price, prize
final /ʃ/, final /tʃ/	wash, watch
final /n/, final /m/	gun; gum
final /n/, final /ŋ/	win, wing
final /l/, dropping /l/	cold, code

Minimal pairs were selected based on two criteria: first, they consist of common everyday words, and second, they can be embedded in the same meaningful sentence frame, which is necessary for the subsequent intelligibility test. Although this was not a strict requirement for the production experiment, it was important to ensure consistency across both experiments. As a result, the words were mostly selected from the same part of speech. However, in a few cases—such as with 'pull' and 'pool'— words from different parts of speech had to be used due to the lack of suitable minimal pairs in the same word category.

Previous studies often used simple word-reading tasks or sentence frames such as "Say ___ again," to compare minimal pairs, which lack naturalness. Drawing inspiration from a sound discrimination task in *Pronunciation Pairs*, this study created two semantically meaningful short sentences (5-7 words) for each minimal pair. Each word appeared in a different sentence to avoid repetition. For example:

- The fan/van is broken.
- She bought a new van/fan.

In both cases, the words "fan" and "van" fit naturally into their sentences. However, if the same sentence were used for both words, it would need to be repeated, potentially causing the speaker to focus too much on distinguishing them. To avoid this, two different sentences were used for each minimal pair, with one for each word. This design helped participants engage with coherent sentences while reducing the use of contextual cues that could make target words too predictable. It should be noted, however, that controlling for the frequency of target words within sentences was not always feasible. For instance, in the sentence "I got a cold/code," while both "cold" and "code" are grammatically acceptable, the phrase "got a cold" is more commonly used than "got a code." Despite this limitation, this method still provides an improvement over context-free, unnatural sentences. By using more natural sentence structures, the design reduces listeners' reliance on context while preserving sentence comprehensibility, enhancing the validity of the test stimuli.

3.3 Recording Procedure

The sentences were recorded in individual sessions (20–40 minutes long) in a sound-proofed room. Prior to the recording, each speaker was given enough time to familiarize themselves with the speech materials and opportunities to ask questions. The 50 sentences were quasi-randomized to ensure that no minimal pairs appeared consecutively. Sentences were presented to the participants via PowerPoint slides displayed on an iPad. The order of the sentences remained consistent for all speakers.

The recordings were made digitally using a microphone and Praat software at a sampling rate of 44.1 kHz. Each sentence was recorded twice, with the second recording used for data analysis. Additionally, a passage reading and spontaneous speech sample were recorded, but the results of those recordings are not discussed in this study.

3.4 Stimulus Preparation

The initial step involved processing each audio file to extract the individual recorded sentences. A 500-millisecond period of silence was added to both the beginning and end of each sentence file to ensure consistency in timing. In total, 700 sound files were created (50 sentences*14 speakers (Note 1)). To minimize variations in perceived loudness across the recordings, all files were normalized for peak intensity using a Praat script.

3.5 Segmental Analysis

The recorded speech was analyzed by systematically comparing the participants' perceived pronunciation with the expected forms produced by a native speaker of American English. Acoustic analysis was not employed, as the focus of the study was on the perception and comprehension of non-native speech, which can often differ from its acoustic characteristics (Levis, 1999). The analysis was conducted by the primary researcher, an experienced non-native English teacher with phonetic training. Each target word in the recorded sentences was orthographically transcribed by the researcher, with consultation from a native speaker of American English.

4. Results

4.1 Vowel Errors

Table 3 presents the results of vowel production by Cantonese speakers of English, organized by the frequency of errors. The results highlight several vowel pronunciation errors made by Cantonese speakers of English. The most frequent error involved the substitution of /æ/ with /ɛ/ (76.9%), followed by errors such as /ɛ/ realized as /eɪ/ or /æ/ (61.5%) and /aʊ/ substituted with /ɔ/ (46.2%). These findings indicate challenges with certain vowel distinctions, particularly those that do not exist or are not prominent in Cantonese, such as the distinction between /æ/ and /ɛ/.

In relation to the FL principle, which emphasizes focusing on phonological contrasts that carry a higher load in terms of distinguishing meaning in communication, the high FL value of 10 (highest on the Brown (1988) ranking) with /æ/-/ɛ/, 9 with /ɛ/-/eɪ/, and 8 with /ɪ/-/i/, suggest that these distinctions should be prioritized in pronunciation instruction. These contrasts are likely to have a significant impact on intelligibility, as they often distinguish key lexical items in English. On the other hand, errors such as /u/ realized as /ʊ/, though having an error rate of 38.5%, have a lower FL value of 3, implying that while still important, these errors could be deprioritized compared to the more frequent and impactful vowel substitutions. This approach ensures that teaching efforts are directed toward the most critical areas for improving learner intelligibility.

Table 3. Vowel Errors

Vowels	Realization	Error Rate
/æ/	/ɛ/	76.9%
/ɛ/	/eɪ/, /æ/	61.5%
/aʊ/	/ɔ/,	46.2%
/ʊ/	/u/	38.5%
/ɔ/	/aʊ/	30.8%
/ɪ/	/i/	26.9%

4.1 Consonant Errors

Table 4. Top 10 Consonant Errors

Consonants	Realization	Error Rate
Final /g/	Final /k/	92.3%
Final z	Final s	84.6%
Final /ŋ/	Final /n/	84.6%
/v/	/w/	84.6%
Final /m/	Final /n/	69.2%
/tʃ/	/tr/	69.2%
Final /b/	Final /p/	69.2%
/θ/	/f/	61.5%
Dark /l/	vocalization	61.5%
Final /t/	dropping	46.2%

Consonant errors were more frequent, with final voiced-voiceless distinctions being particularly problematic. The highest error rate involved final /g/ realized as /k/ (92.3%), followed by final /z/ realized as /s/ (84.6%) and final /ŋ/ realized as /n/ (84.6%). These errors suggest a difficulty in maintaining voiced/voiceless distinctions in coda positions, a well-documented issue linked to the absence of voiced obstruents in final positions in Cantonese. The substitution of /v/ with /w/ (84.6%) and final /m/ with /n/ (69.2%) were also frequent.

From a functional load perspective, contrasts such as /p/-/b/, /m/-/n/, /g/-/k/, and /z/-/s/ carry high FL values, indicating that these errors could significantly reduce intelligibility and should be prioritized in pronunciation instruction. On the other hand, errors with lower functional load, such as /θ/-/f/ (FL value of 1), although frequent, may have a lesser impact on overall intelligibility and could be deprioritized in teaching.

5. Conclusion

The results of this study highlight key segmental pronunciation challenges faced by Cantonese EFL learners. The most problematic vowel and consonant contrasts, such as /æ/-/ɛ/ and /g/-/k/, should be prioritized in pronunciation instruction due to their high functional load and significant impact on intelligibility. By focusing on these critical areas, teachers can help learners improve their communicative effectiveness in English without aiming for accent elimination. These findings highlight the value of adopting an intelligibility-focused approach to pronunciation, in line with modern trends in English language teaching.

The findings of this study have significant implications for pronunciation teaching in EFL contexts, particularly for Cantonese-speaking learners. The most frequent and impactful errors, especially those involving high-FL contrasts, should be prioritized in classroom instruction. By addressing distinctions such as /æ/-/ɛ/ and final voiced/voiceless consonants, teachers can help learners improve their intelligibility more efficiently.

In future research, further exploration into the role of prosody and suprasegmental features could complement the segmental analysis presented here, providing a more comprehensive understanding of Cantonese speakers' pronunciation challenges.

Funding

This work was supported by the Guangdong Provincial Co-construction Project for Philosophy and Social Sciences: "A Study on Factors Affecting the Intelligibility of English Pronunciation by Native Cantonese Speakers" under Grant GD20XWY23.

References

Baker, A., & Goldstein, S. (2008). *Pronunciation pairs: an introduction to the sounds of English*. New York: Cambridge University Press.

Boersma, Paul & Weenink, David. (2024). *Praat: doing phonetics by computer* [Computer program]. Version 6.4.20.

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly*, *22*(4), 593-606.

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. In J. Morley (Ed.), *Current Perspectives on Pronunciation: Practices Anchored in Theory* (pp. 87-100). Alexandria, VA: TESOL.

Chan, A. Y. (2010). Advanced Cantonese ESL learners' production of English speech sounds: Problems and strategies. *System*, *38*(2), 316-328.

Levis, J. M. (1999). Intonation in theory and practice revisited. *TESOL Quarterly*, *33*(1), 37-63.

Levis, J. M. (2018). Teaching for Intelligibility: Guidelines for Setting Priorities. In *Intelligibility, Oral Communication, and the Teaching of Pronunciation* (pp. 185-216). chapter. Cambridge:

Cambridge University Press.

Martinet, A. (1952). Function, Structure, and Sound Change. *Word*, *8*, 1-32.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97.

Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed, & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 377-96). NJ: Wiley Blackwell.

Munro, M. J., & Derwing, T. M. (2006). The Functional Load Principle in ESL Pronunciation Instruction: An Exploratory Study. *System*, *34* (4), 520-531.

Zhang, L. (2012). *International Intelligibility of Chinese-accented English*. [Unpublished doctoral dissertation]. Central China Normal University.

Note 1. A native speaker of American English also recorded the test sentences, primarily for reference purposes, but these recordings were excluded from the analysis.