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

Identifying Pronunciation Challenges for Cantonese EFL Learners: A Focus on Intelligibility

Xinting Zhang[1] & Yuan Li[1]

[1] Southern University of Science and Technology, China

Received: August 8, 2024 Accepted: September 22, 2024 Online Published: September 30, 2024

doi:10.22158/selt.v12n3p120 URL: http://dx.doi.org/10.22158/selt.v12n3p120

Abstract

This study investigates the intelligibility of Cantonese-accented English by examining segmental pronunciation errors and their impact on native English listeners' ability to accurately transcribe spoken sentences. Building on a previous production experiment, this research focuses on identifying which vowel and consonant mispronunciations most hinder intelligibility. Ten native English listeners transcribed 1,000 sentences produced by Cantonese speakers, and transcription accuracy was calculated for each target sound. The results reveal that dark /l/ (e.g., "cold"), final consonants like /g/ (e.g., "bag"), and the vowel /æ/ (e.g., "pan") had the lowest accuracy rates, while sounds like /n/ (e.g., "no") and /ʃ/ (e.g., "sheep") were consistently transcribed correctly. These findings highlight the importance of targeting specific pronunciation difficulties in Cantonese EFL learners to improve intelligibility. The study supports the prioritization of intelligibility over accent reduction in pronunciation instruction, providing practical implications for teaching pronunciation in Cantonese-speaking contexts.

Keywords

Intelligibility, Pronunciation teaching, Teaching priorities

1. Introduction

1.1 Background and Rationale

Intelligibility, defined as the degree to which a listener can understand the speech of a non-native speaker, is a central goal in second language (L2) pronunciation teaching (Levis, 2018). It has long been recognized that pronunciation errors can significantly impact a speaker's intelligibility, affecting not only comprehension but also communication effectiveness in real-world contexts. In contrast to the earlier emphasis on achieving native-like pronunciation, contemporary approaches to English as a Foreign Language (EFL) teaching prioritize intelligibility over accent reduction, especially in settings

where English functions as a global lingua franca. This shift emphasizes the need to identify specific phonological features that contribute to communication breakdowns for non-native speakers.

For learners whose first language (L1) differs significantly from English, such as Cantonese, certain segmental features—individual vowel and consonant sounds—may pose particular challenges. Previous research has highlighted common pronunciation difficulties for Cantonese speakers, such as devoicing of final consonants, vowel confusion, and the replacement of certain English sounds with Cantonese approximations. While these studies provide valuable insights into the general pronunciation tendencies of Cantonese speakers, fewer studies have systematically examined how these segmental errors impact intelligibility from the listener's perspective. Identifying which sound errors hinder comprehension is crucial for developing targeted pronunciation instruction that improves learners' communicative competence.

1.2 Research Focus

The current study aims to address this gap by focusing on the intelligibility of Cantonese-accented English at the segmental level. Specifically, the purpose of this research is to identify which segmental errors made by Cantonese speakers of English most affect intelligibility for native listeners. Building on a previous production experiment, which analyzed segmental errors in the speech of Cantonese EFL learners, this study extends the analysis to include listener perceptions. In the production experiment, participants were recorded reading sentences containing minimal pairs of sounds commonly mispronounced by Cantonese speakers. These same sentences are to be evaluated by native English listeners, who will transcribe what they hear. By comparing the listeners' transcriptions to the intended target words, the study will determine which sound confusions lead to the most frequent intelligibility breakdowns. This will provide a comprehensive view of the pronunciation challenges that most affect intelligibility, offering important pedagogical implications for pronunciation teaching in Cantonese-speaking contexts.

2. Literature Review

2.1 Pronunciation Teaching Based on Intelligibility

In second language (L2) pronunciation research, intelligibility, comprehensibility, and accentedness are key concepts used to analyze and understand the perception of L2 speech. While these terms represent distinct aspects of speech, they are interconnected and provide valuable insights into how listeners interpret and evaluate non-native speech. According to Munro and Derwing (1995), intelligibility refers to the extent to which a listener can accurately understand the speaker's intended message. It is often assessed by having listeners transcribe or interpret speech, with a focus on how effectively the message is conveyed. Comprehensibility, in contrast, relates to the ease or difficulty of understanding the speaker. It reflects the effort required by the listener. Even if the listener correctly understands the content, if the process is laborious, the rating of comprehensibility may be lower. Lastly, accentedness describes the degree of deviation from native-like pronunciation, involving segmental (e.g., vowel and consonant

mispronunciations) or suprasegmental features (e.g., tone, length, and rhythm). While accentedness is a common feature of L2 speech, it does not always negatively impact intelligibility or comprehensibility. Listeners may notice a speaker's accent yet still fully understand the message, especially if they are familiar with that particular non-native accent.

Munro and Derwing's research highlighted that in L2 pronunciation teaching, the pursuit of native-like pronunciation is not only difficult but also unnecessary. Achieving a native-like accent is often unrealistic and may not enhance real-world communication. The focus, instead, should be on improving intelligibility—ensuring that the speaker can be understood effectively. This approach reflects a shift in pronunciation pedagogy toward more practical goals, emphasizing the impact of speech on communication rather than striving for perfection in accent. As long as the speaker's pronunciation does not cause significant misunderstandings or communication breakdowns, some degree of accent is acceptable and even natural in L2 speech.

This shift has important implications for pronunciation teaching. Teachers should prioritize the phonological features that truly affect communication rather than overemphasizing accent reduction. Many studies support the view that intelligibility should be the primary goal of pronunciation instruction. Munro and Derwing (2015) and Jenkins (2000) have demonstrated that accented speech can be understood in many contexts, indicating that reducing an accent should not be the central focus. Instead, ensuring that speech is intelligible is key. Jenkins further argues that in English as a Lingua Franca (ELF) settings, many pronunciation errors do not impede communication, as long as the speech remains clear and intelligible.

2.2 Experimental Design

Intelligibility tasks are designed to assess a listener's ability to accurately transcribe or repeat speech, measuring how closely their responses match the intended utterances. These tasks offer an objective assessment of performance by providing a binary evaluation—whether the listener correctly identified the target or not. As a result, these tasks measure the listener's capacity to achieve the ultimate goal of speech perception: accurately recognizing the linguistic elements, such as phonemes, words, and sentences, produced by the speaker.

The most common method for evaluating intelligibility in L2 speakers involves the transcription of their speech. Various test items can be used, including isolated words, sentences, read passages, and spontaneous speech, each with distinct advantages and drawbacks. Isolated words may highlight specific pronunciation errors, while sentences and passages offer a more contextualized view of how these errors affect intelligibility.

However, listeners often rely on contextual clues to infer meaning, even when pronunciation errors occur, which can make it challenging to assess intelligibility accurately. To address this, semantically anomalous sentences (e.g., "A dark nail zaps a ready reason") (Kennedy & Trofimovich, 2008), which are syntactically correct but lack semantic meaning, are used to minimize contextual support. This forces listeners to rely solely on syntactic relationships between words rather than meaning. While this method

reduces the influence of context, it creates artificial listening conditions that may not fully reflect real-world speech perception.

Research has shown that semantically anomalous sentences are indeed more difficult for listeners to identify. Kennedy and Trofimovich (2008) found that transcription accuracy for these sentences was the lowest compared to both true-false sentences and semantically meaningful sentences.

3. Method

3.1 Participants

3.1.1 Speakers

Thirteen university students from various regions of Guangdong, China, participated in the study. All were native Cantonese speakers with similar English learning backgrounds, having begun formal English instruction at the age of 9–10.

3.1.2 Listeners

Ten native American English speakers, aged 28-45, with no prior exposure to China, were recruited through personal connections. The listeners represented various professional fields and reported no hearing impairments. They were compensated for their participation.

3.2 Materials

This study focuses on identifying specific sounds in the speech of Cantonese speakers of English that affect intelligibility. To achieve this, we carefully selected 25 commonly confusing sound pairs observed among Cantonese speakers and created minimal pairs for these sounds, embedding the minimal pair words into short, semantically coherent sentences. Unlike traditional word-reading tasks, this sentence-based format is more natural. The method also represented an improvement over the use of semantically anomalous or context-free, unnatural sentences. By using more natural sentence structures that minimize the listeners' reliance on context, such an approach enhances the validity of the test stimuli.

The test materials used in this study were the same as those from the production experiment. For a full list of the 25 test sound pairs and their corresponding minimal pairs, refer to the article on the production experiment in the same issue.

3.3 Stimuli

Due to the limited availability of listeners, recordings from 10 speakers, selected from the original 13, were used for the transcription task. Each speaker's utterances were repeated twice during the task, resulting in a total of 1,000 recorded sentences to be transcribed. These sentences were divided into 10 sets, one for each listener, with each set containing 100 sentences—10 distinct sentences from each speaker. This ensures that each listener heard all 50 sentences twice spoken by all ten speakers, with both iterations of the same sentence spoken by different speakers.

The 100 sentences in each set were embedded into a PowerPoint presentation, with five audio files per slide. The presentations were then sent to the ten listeners for transcription. Listeners completed the

task remotely, in their own spaces. They were instructed to transcribe what they heard as accurately as possible, and to make an educated guess if they were unsure of any words by typing their response into a Word document.

4. Results

The transcription results from each listener were compiled into an Excel file. A Python script was used to compare the transcribed sentences with the original sentences, highlighting any discrepancies. Words that did not match the original were extracted and paired with the corresponding correct words. These pairs were manually reviewed by the researcher before proceeding with the analysis.

Intelligibility was assessed by calculating the accuracy rate for each target sound, rather than using the conventional overall intelligibility score. This approach allowed for direct comparison between commonly confused sounds. Spelling errors and added grammatical endings, such as "-ed" and "-s," were disregarded, but omitted endings were counted as errors, as they could indicate the omission of a crucial sound.

4.1 Intelligibility of Vowels

Table 1 presents the transcription accuracy rates for English vowel sounds produced by Cantonese speakers of English. The vowels are listed along with words that contain the sounds.

Table 1. Correct Transcription of English Vowels by Cantonese Speakers of English

Vowels	Transcription	Error Rate
/æ/ (pan)	/ɛ/ pen)	5%
/i/ (sheep)	/ɪ/ (ship, strip)	45%
/ɔ/ (shots)	/aʊ/ (shouts), /əʊ/ (shows)	50%
/ɛ/ (pen)	/eɪ/ (pain), /æ/ (pan)	50%
/aʊ/ (shouts)	/ɔ/ (shots), /ɚ/ (shirts), /ɛ/ (shells)	70%
/eɪ/ (pain)	/ɛ/ (pen)	95%
/aɪ/ (bike)	/æ/ (back)	95%

The results show a wide range of accuracy rates, with some vowels being pronounced correctly at very high rates, while others present significant difficulties.

The vowel sounds with the highest accuracy are /eɪ/ (as in pain) and /aɪ/ (as in bike), both with an accuracy rate of 95%. Mid-range accuracy is observed for /ɔ/ (as in shots) and /ɛ/ (as in pen), both achieving 50%.

The vowel /æ/ (as in pan), which often presents challenges for non-native speakers, particularly those from language backgrounds without a similar low-front vowel, has the lowest accuracy rate at just 5%. This indicates a significant issue with the production of this vowel.

Overall, the data suggest that the pronunciation of English vowels varies widely among the speakers, with diphthongs generally being more accurately produced than monophthongs, and some sounds—such as /æ/—posing a particular challenge. These findings point to potential areas of focus for pronunciation instruction, especially for vowels with low accuracy rates like /æ/ and /ɛ/.

4.1 Intelligibility of Consonants

Table 2 presents the transcription accuracy rates for English consonant sounds produced by Cantonese speakers of English.

Table 2. Correct Transcription of English Consonants by Cantonese Speakers of English

Consonants	Transcription	Error Rate
Dark /l/ (cold)	code, coat, call	5%
Final /g/ (bag)	back, bed	5%
Final /ŋ/ (wing)	win	10%
Final /b/ (cab)	cap, car	15%
Final /z/ (prize)	price	15%
/tʃ/ (chunk)	trunk,	30%
Final /m/ (gum)	gun	35%
Final /d/ (card)	car	45%
/v/ (vine)	wine	52.5%
/dʒ/ (junk)	drunk, trunk	60%
Final /sh/(wash)	watch	65%
/z/ (zip)	say, see, sit	65%
/d/ (dry)	try	70%
Final /t/ (cart)	car, cap	70%
Clusters (box)	boss, boat	80%
Final d dropping (waited, removed)	wait, remove	80%
Final /n/ (gun)	gum	85%
/n/ (no)	low	95%
/l/ (loud, light)	now, nice	95%
/ʃ/ (sheep)	seat	95%

The transcription accuracy rates for consonants in this intelligibility experiment reveal varying degrees of difficulty for native speakers when interpreting L2 English speech. Dark /l/ (as in cold) and final /g/ (as in bag) had the lowest accuracy rates at just 5%, indicating significant challenges in accurately perceiving these sounds. Other problematic areas included final /ŋ/ (as in wing) and final /b/ (as in cab),

both with low accuracy rates of 10-15%. In contrast, consonants like /n/ (as in no) and /ʃ/ (as in sheep) showed much higher accuracy rates, with both reaching 95%, demonstrating that these sounds were much easier for native speakers to transcribe. Notably, consonant clusters (as in box) and final sound dropping (as in waited) were moderately challenging, achieving accuracy rates between 65% and 80%. Overall, the results highlight particular consonants and phonological features, such as dark /l/ and devoicing of final consonants, as key areas of difficulty for Cantonese speakers of English.

4.3 Discussion

Participants generally exhibited a low level of comprehension when listening to Cantonese-accented English, particularly in sentences involving final consonants and vowel substitutions. Listeners encountered significant difficulties in understanding the pronunciation of native Cantonese speakers. This result shows a higher error rate compared to previous experiments, which may be attributed to the design of the speech materials. Since the study directly examined minimal pairs, participants were required to accurately differentiate between contrasting sounds. When the context provided identical cues for both contrasting sounds, listeners had to rely solely on phonetic distinctions. This result underscores the importance of contextual cues in listening comprehension. However, when context fails to clearly favor one option and creates competition between two, pronunciation accuracy becomes critical. In such instances, even subtle phonetic differences can significantly affect understanding, with clear and precise pronunciation being crucial for listeners to distinguish between sounds. This finding highlights that in real-world communication, pronunciation accuracy and contextual cues work together to ensure the effectiveness of information transmission.

5. Conclusion

The findings of this study provide valuable insights into the specific segmental pronunciation errors made by Cantonese speakers of English that significantly impact intelligibility for native English listeners. Both vowels and consonants presented varying degrees of difficulty, with the lowest accuracy rates observed for sounds such as dark /l/ (e.g., “cold”), final /g/ (e.g., “bag”), and the vowel /æ/ (e.g., “pan”). These errors likely stem from phonological differences between Cantonese and English, particularly in the production of final consonants and low-front vowels, which have no direct equivalent in Cantonese. On the other hand, some sounds, such as /n/ (e.g., “no”) and /ʃ/ (e.g., “sheep”), were consistently transcribed with high accuracy, suggesting that certain English sounds are more easily acquired by Cantonese speakers.

These findings have important pedagogical implications for pronunciation teaching. First, pronunciation instruction should prioritize sounds that most frequently lead to intelligibility breakdowns, such as dark /l/, final consonants, and vowels like /æ/. Teachers can use targeted exercises and phonetic training to address these specific difficulties, helping students reduce errors that hinder communication. Additionally, the results suggest that while some degree of accentedness is acceptable in real-world communication, emphasis should be placed on ensuring that learners’ speech is intelligible rather than

aiming for native-like pronunciation.

Moreover, the findings support the growing trend in pronunciation teaching that prioritizes intelligibility over accent reduction. By focusing on sounds that directly impact understanding, teachers can enhance learners' communicative competence in diverse English-speaking environments, where clear communication is the primary goal. These insights can be integrated into pronunciation curricula for Cantonese-speaking learners, encouraging more effective teaching strategies that promote intelligibility and improve overall communication outcomes.

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.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, *64*(3), 459-489.

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.

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

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.