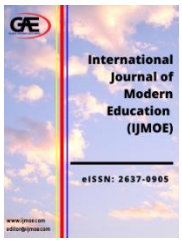
International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046



INTERNATIONAL JOURNAL OF MODERN EDUCATION (IJMOE)
www.ijmoe.com

GLOBAL ACADEMIC EXCELLENCE

L1 PHONOLOGICAL TRANSFER IN L2 ENGLISH PERCEPTION: A MIXED STUDY OF CANTONESE-SPEAKINGUNDERGRADUATES

Ye Yizhou[1*], Latisha Asmaak Shafie[2], Razlina Razali[3]

[1] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email: 2023516063@uitm.edu.my
[2] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email: ciklatisha@uitm.edu.my
[3] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email: razlinarazali@uitm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 18.06.2025
Revised date: 09.07.2025
Accepted date: 13.08.2025
Published date: 10.09.2025

To cite this document:

Ye, Y., Shafie, L. A., & Razali, R. (2025). L1 Phonological Transfer in L2 English Perception: A Mixed Study of Cantonese-Speaking Undergraduates. *International Journal of Modern Education, 7* (26), 689-708.

DOI: 10.35631/IJMOE.726046

Abstract:

This study examines the extent to which the phonological system of Cantonese influences the perception of English segmental and suprasegmental features among second language (L2) learners. The research is motivated by persistent perceptual challenges faced by Cantonese-speaking learners, particularly concerning phonemes absent in their L1 and prosodic mismatches between tonal and stress-based systems. Grounded in two theoretical frameworks—the Revised Speech Learning Model (SLM-r) and the Perceptual Assimilation Model for L2 learners (PAM-L2)—this study adopts a mixed-methods approach, combining a phoneme perception questionnaire with follow-up interviews. Data were collected from 60 Cantonese-speaking undergraduates in Guangdong, China. Results indicate systematic segmental confusion (e.g., /θ/ → /s/, /v/ → /w/, /æ/ → /ɛ/) and suprasegmental misperception stemming from tonal transfer, such as the reliance on pitch and syllable duration to identify English lexical stress. Lexical familiarity was found to modulate perception accuracy, with unfamiliar words prompting increased L1-based inference. These findings confirm that L2 speech perception is heavily filtered through L1 categories and align with assimilation types predicted by SLM-r and PAM-L2. The study underscores the need for perception-based instruction tailored to L1-specific constraints and contributes empirical evidence to the growing body of research on cross-linguistic phonological transfer in tonal-language learners.

Keywords:

Cantonese Learners, L1 Transfer, L2 Perception Models Phonological Acquisition, Speech Perception, Suprasegmentals

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Introduction

The Significance of Perception in L2 Phonological Development

Accurate speech perception plays a foundational role in second language (L2) phonological development, particularly for learners whose first language (L1) differs markedly from the target language in phonemic inventory, syllable structure, and prosodic systems. As pointed out by Darcy (2024), perception is not a passive precursor to production but often a persistent bottleneck, especially when it comes to building robust phonolexical representations. Learners with underdifferentiated or "fuzzy" phonolexical encoding frequently experience lexical competition failures, leading to inefficiencies in real-time speech processing (Kapnoula, 2021). These difficulties are magnified in adverse listening conditions, such as fast speech or background noise, where accurate perception is crucial for comprehension.

Recent research emphasizes that even highly proficient L2 learners may continue to experience perceptual challenges long after production appears fluent (Simonchyk & Darcy, 2023). This perception–production misalignment suggests that perceptual encoding is more resistant to change and may require explicit training or high-quality input to improve. Moreover, the degree of perceptual fuzziness directly affects lexical access and inhibition processes, as shown in studies on non-native word recognition (Rocca & Darcy, 2022). These findings underscore the necessity of focusing on perception—not just production—in L2 phonological research and pedagogy.

Phonological Divergences Between Cantonese and English

Cantonese-speaking learners of English face specific perceptual challenges due to systemic phonological mismatches. Cantonese has a relatively small vowel inventory with only six contrastive monophthongs, which limits its capacity to distinguish English tense–lax vowel pairs such as /ɪ/–/iː/ and /æ/–/ɑː/ (Escudero, 2007). These contrasts are often assimilated into a single L1 category, resulting in perceptual overlap and high confusability. In consonantal terms, the absence of interdental fricatives /θ/ and /ð/ in Cantonese leads to substitutions with alveolar or labiodental sounds, such as /s/ or /f/, as reported by Salheen and Yap (2023). These substitutions are not only articulatory but also perceptual in nature, indicating a deeper phonological transfer.

The syllable structure of Cantonese is also highly restrictive compared to English. It generally follows a (C)V(C) pattern, with limited use of consonant clusters. Learners often employ vowel epenthesis or coda deletion as repair strategies when processing or producing English syllables (Zhao et al., 2024). However, not all transfer is negative. He et al. (2025) found that Cantonese rusheng syllables, which feature unreleased final stops, may facilitate the perception and production of English coda stops in certain contexts. Nonetheless, such positive transfer effects are highly context-dependent and do not generalize across all phonological domains.

Additionally, orthographic influence must be considered. The Chinese logographic system lacks systematic grapheme–phoneme correspondence, which hinders learners' ability to develop accurate sound–symbol mappings in English (Jiang et al., 2024). This can delay the development of phonological awareness and reduce the effectiveness of phonics-based instruction.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Suprasegmental Transfer and Lexical Stress Difficulties

Beyond segmental mismatches, suprasegmental features also pose significant learning barriers. Cantonese is a tonal language in which pitch contour is used to distinguish lexical meaning, whereas English is a stress-timed language that uses stress for prominence and information structure. This typological contrast often leads Cantonese learners to misinterpret English lexical stress as tonal variation (Cheng, 2023), resulting in unnatural prosody, misplaced stress, and reduced intelligibility.

Wang (2025) reported that many Cantonese-speaking L2 learners struggle to identify stressed syllables in multisyllabic English words and tend to rely on pitch height and syllable duration as perceptual cues—strategies that are maladaptive in English. Xie (2024) further found that Cantonese-English bilingual children exhibit delayed acquisition of English lexical stress and prosodic rhythm, especially in continuous speech. These difficulties are not simply due to lack of input but reflect deep-rooted L1 transfer effects that affect the perception and processing of suprasegmental cues. Such evidence underscores the need to incorporate suprasegmental training in L2 curricula, especially for learners from tonal language backgrounds.

Perception–Production Asymmetry and Transfer Mechanisms

The mismatch between perception and production is a well-documented phenomenon in L2 phonology. Learners may be able to produce target-like sounds without consistently perceiving them accurately—a pattern known as perception–production asymmetry. Melnik-Leroy, Turnbull, and Peperkamp (2022) highlighted that even when learners achieve native-like production of difficult L2 contrasts, perceptual discrimination may remain poor. This supports the theoretical predictions of the Speech Learning Model—Revised (SLM-r), which posits that successful L2 acquisition requires the perceptual formation of new phonetic categories prior to accurate motor execution (Flege & Bohn, 2021).

Similarly, the Perceptual Assimilation Model for L2 learners (PAM-L2) argues that non-native sounds will be assimilated into the closest L1 categories unless new categories are formed (Best & Tyler, 2007). For sounds that fall between L1 categories or lack a clear counterpart—such as English /v/ or /θ/—learners may experience persistent confusability and substitution. These models provide a coherent explanation for why certain L2 contrasts are more difficult than others and why perception-focused interventions are necessary.

Sociophonetic factors also modulate perceptual transfer. As Tse (2022) notes, heritage speakers often exhibit perceptual patterns shaped by their dominant language—even when their heritage language was acquired early. This suggests that age of acquisition alone cannot account for transfer effects and that linguistic dominance plays a crucial role in shaping phonological perception.

Research Gaps and Study Objectives

Despite these theoretical insights, empirical research focusing specifically on the perceptual difficulties of Cantonese-speaking EFL learners remains limited. Much of the existing literature emphasizes production-oriented outcomes or general ESL populations, with insufficient attention to L1-specific perceptual pathways and transfer patterns (Xie, 2024). To address this gap, the present study employs a mixed-methods design to investigate both segmental and suprasegmental perception of English among Cantonese-speaking undergraduates.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Drawing on the frameworks of SLM-r and PAM-L2, this study seeks to identify systematic transfer effects in phoneme perception, elucidate the strategies learners adopt when processing unfamiliar L2 sounds, and contribute to a more nuanced understanding of how L1 phonology shapes L2 speech processing in typologically distinct language pairs.

Specifically, the objectives of this study are:

1. To identify which English phonemes and prosodic features are most susceptible to misperception among Cantonese-speaking undergraduates.
2. To examine how specific aspects of Cantonese phonology—such as syllable structure, tone, and limited vowel inventory—contribute to these perceptual challenges.
3. To evaluate whether the observed patterns of perception errors align with theoretical predictions made by the Revised Speech Learning Model (SLM-r) and the Perceptual Assimilation Model for L2 (PAM-L2).

Accordingly, the study addresses the following research questions:

RQ1: Which English segmental (e.g., /θ/, /æ/, /ɪ/) and suprasegmental features (e.g., lexical stress, linking) are most frequently misperceived by Cantonese-speaking learners?

RQ2: Which phonological features of Cantonese contribute most significantly to these perception difficulties?

RQ3: Do these perceptual errors exhibit systematic transfer patterns that can be explained by SLM-r and PAM-L2?

Based on these questions, the following hypotheses are proposed:

H1: Learners will show lower accuracy in perceiving English phonemes that lack clear Cantonese equivalents, such as interdental fricatives and tense–lax vowel pairs.

H2: Suprasegmental perception difficulties—particularly in stress and rhythm—will reflect interference from Cantonese tonal processing.

H3: The observed perceptual patterns will exhibit systematic transfer effects consistent with the predictions of SLM-r and PAM-L2, including single-category and category assimilation types.

Literature Review

Phonological Characteristics of Cantonese and Their Implications for L2 Perception

Cantonese presents several phonological features that influence how its speakers perceive English. Notably, it has a smaller vowel inventory (six monophthongs) and a restrictive (C)V(C) syllable structure (Zhao et al., 2024). These characteristics limit the learner's ability to distinguish English vowel contrasts such as /æ/–/ɑː/ or /ɪ/–/iː/, often leading to assimilation into a single L1 category (Escudero & Boersma, 2004). The absence of certain consonants—such as /θ/, /ð/, and /v/—further contributes to misperception and substitution with acoustically similar L1 sounds (Salheen & Yap, 2023).

Moreover, Cantonese's tonal system assigns lexical meaning to pitch contour, unlike English, which uses stress for prosodic prominence. This mismatch leads to difficulty in identifying lexical stress and rhythm in English, as learners tend to over-rely on pitch height and duration instead of intensity or vowel reduction (Kang, 2024; Cheng, 2023). He et al. (2025) noted that the presence of unreleased stops in Cantonese rusheng syllables may positively transfer to English coda perception, though such effects are context-dependent.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Segmental Perception Challenges among Cantonese Learners

Numerous studies have documented systematic segmental perception errors among Cantonese-speaking ESL learners. For instance, /θ/ is often perceived as /s/, and /v/ as /w/—patterns that reflect L1-perceptual assimilation rather than random confusion (Archibald, 2021; Salheen & Yap, 2023). Archibald (2021) highlights that such segmental difficulties are not arbitrary but follow predictable patterns rooted in phonemic inventory mismatches and perceptual assimilation hierarchies.

The compact vowel system of Cantonese also results in frequent merging of English vowel pairs, such as /æ/ and /ɛ/, or /ʌ/ and /ɑː/, leading to perceptual ambiguity (Jiang et al., 2024; Escudero & Boersma, 2004). Jiang et al. (2024) further demonstrate that this ambiguity is exacerbated by weak grapheme–phoneme correspondence awareness among Chinese learners, suggesting that orthographic interference plays a role in phonological processing and L2 decoding fluency.

ERP studies confirm that L2 phonemes resembling L1 patterns elicit stronger neural responses, suggesting that perceptual filtering occurs at early cognitive stages (Gosselke et al., 2021). Importantly, even when learners can produce these sounds correctly, their perceptual discrimination remains limited—highlighting the need for perception-based instruction (Melnik-Leroy et al., 2022).

Suprasegmental Transfer and Tonal Interference

Suprasegmental perception poses additional challenges for Cantonese learners. Studies show that learners often misidentify stressed syllables by relying on pitch contours, a strategy influenced by their native tonal system (Cheng, 2023; Wang, 2025). Xie (2024) found that even bilingual Cantonese–English children struggle with English prosody, particularly in continuous speech contexts.

Moreover, sociophonetic variation within Cantonese may also influence suprasegmental transfer. Soo and Babel (2024) observed that regular exposure to multiple pronunciation variants (e.g., tone mergers, variant finals) facilitates dual-mapping in bilingual learners. This suggests that Cantonese's internal variation may either buffer or intensify transfer effects depending on learners' sociolinguistic input.

Such suprasegmental transfer results in unnatural speech rhythm, misplaced stress, and reduced intelligibility. Learners tend to use duration and pitch height as cues for stress, despite these being secondary in English stress perception. These findings align with PAM-L2 predictions, which suggest that unfamiliar prosodic categories are often assimilated to the closest L1 perceptual equivalents (Best & Tyler, 2007).

Theoretical Models of L2 Phonological Perception

Two major frameworks explain how L1 phonological structures influence L2 speech perception: the Revised Speech Learning Model (SLM-r) and the Perceptual Assimilation Model for L2 learners (PAM-L2). SLM-r posits that L2 category formation depends on the learner's ability to perceive phonetic distinctions as sufficiently different from existing L1 categories (Flege & Bohn, 2021).

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

PAM-L2 classifies assimilation patterns (e.g., single-category, category-goodness, two-category) and predicts discrimination difficulty accordingly (Best & Tyler, 2007). For example, /θ/ and /s/ often fall under single-category assimilation for Cantonese learners, resulting in low discrimination accuracy. McLaughlin et al. (2024) demonstrated that high-variability phonetic training (HVPT), grounded in SLM principles, improves perceptual accuracy for such problematic contrasts.

Escudero and Boersma (2004) offer a complementary perspective by modeling L2 perception as a dynamic adjustment of perceptual grammars based on L1–L2 similarity. Their work bridges L2 phonological theory and practical acquisition, showing how L2 learners recalibrate perception based on statistical learning from input.

Tse (2022) emphasizes that sociophonetic dominance—not just age of acquisition—affects perceptual transfer. Heritage speakers may show dominant-language interference even when their early exposure favors the heritage language. This suggests that L2 perception is shaped by both phonological compatibility and sociolinguistic context.

Summary of Recent Empirical Studies (2020–2025)

To consolidate the reviewed literature, Table 1 summarizes recent empirical studies (2020–2025) that directly inform the present research. These studies cover segmental, suprasegmental, orthographic, and sociophonetic aspects of L1 transfer in L2 English perception among Cantonese speakers.

Table 1: Summary of Recent Empirical Studies on Cantonese–English Phonological Transfer (2020–2025)

Author(s)	Year	Focus	Participants	Key Findings	Relevance
He et al.	2025	Positive transfer	Cantonese learners	Rusheng syllables aid coda stop perception	Context-dependent facilitation
Wang	2025	Accent perception	Cantonese EFL learners	Difficulty in perceiving stress under accent	Confirms suprasegmental interference
Soo & Babel	2024	Dialect variation & input	Cantonese-English bilinguals	Variant exposure enhances dual-mapping	Input-driven modulation
Xie	2024	Suprasegmental acquisition	Cantonese-English children	Delayed stress acquisition	Tonal interference
Zhao et al.	2024	Syllable structure	Cantonese EFL learners	Epenthesis and coda deletion	Structural simplification transfer
Salheen & Yap	2023	Segmental transfer	Cantonese EFL learners	/θ/ → /s/, /v/ → /w/ substitutions	Confirms PAM-L2 assimilation
Cheng	2023	Stress misperception	Cantonese EFL learners	Stress perceived as pitch	Confirms tonal transfer

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Research Gaps and Rationale

The preceding review reveals several underexplored areas in the study of L1 phonological transfer among Cantonese learners of English. First, while segmental transfer has been widely studied, suprasegmental features such as stress perception remain less understood, especially in continuous speech settings. Second, the role of dialectal variation within Cantonese (e.g., tone mergers, variant finals) in shaping L2 perception has received limited empirical attention. Third, although theoretical models like SLM-r and PAM-L2 are frequently cited, few studies have systematically applied them to Cantonese-English contexts using perception-based tasks. Finally, orthographic interference and sociophonetic dominance remain emerging but understudied factors in L2 phonemic perception. These gaps indicate the need for integrated frameworks that account for both linguistic and sociolinguistic sources of transfer.

Most prior studies have addressed production errors or segmental-level transfer, with relatively little focus on real-time perception or dialectal variation. In particular, few have applied SLM-r or PAM-L2 to explain how different Cantonese dialects may shape L2 perception differently. Moreover, perceptual training studies remain limited in scope and are often not tailored to tonal-language learners. This study seeks to fill these gaps by examining how Cantonese phonology affects learners' perception of English phonemes, using SLM-r and PAM-L2 as interpretive frameworks.

This study adopts two complementary theoretical models to explain the perceptual challenges faced by Cantonese speakers in acquiring English segmental contrasts: the Revised Speech Learning Model (SLM-r) and the Perceptual Assimilation Model for L2 learners (PAM-L2). Both models provide a principled account of how first language (L1) phonological structures influence second language (L2) speech perception, particularly in the early stages of category formation and phonemic discrimination.

The SLM-r, developed by Flege (2018), posits that successful acquisition of L2 phonetic categories depends on the learner' s ability to reorganize their existing phonological space. When L1 and L2 sounds are perceived as highly similar, learners may fail to form a distinct new category, instead assimilating the L2 sound into an existing L1 category. This often results in persistent perceptual confusion and production errors. In the context of Cantonese speakers, this model helps explain the frequent merging of English vowel pairs such as /ɪ/–/iː/ and /æ/–/ɑː/, which are absent in the more compact Cantonese vowel inventory.

The PAM-L2 (Best & Tyler, 2007) builds on the original PAM framework by specifying how L2 sounds are perceptually assimilated into L1 categories. It classifies L2 contrasts into different assimilation types — such as single-category, two-category, and uncategorized assimilation — each predicting varying degrees of discrimination difficulty. For example, English interdental fricatives / θ / and /ð/ are often assimilated to Cantonese alveolar or labiodental sounds due to the lack of equivalent phonemes in the L1 inventory, leading to misperception and substitution.

While both models have been widely applied in L2 phonological acquisition research, their use in explaining Cantonese-specific perceptual transfer remains relatively limited. In particular, few studies have systematically applied these models to account for the perceptual assimilation patterns among speakers of different Cantonese dialects. This study seeks to address this gap by using SLM-r and PAM-L2 as interpretive tools to examine how Cantonese learners perceive English phonemes, and how these perceptions lead to systematic transfer effects in segmental

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

processing.

Figure 1 illustrates the theoretical structure of the study, showing how L1 phonological characteristics interact with L2 perceptual mechanisms to produce observable transfer patterns. This framework forms the basis for the research questions, experimental design, and interpretation of findings in subsequent chapters.

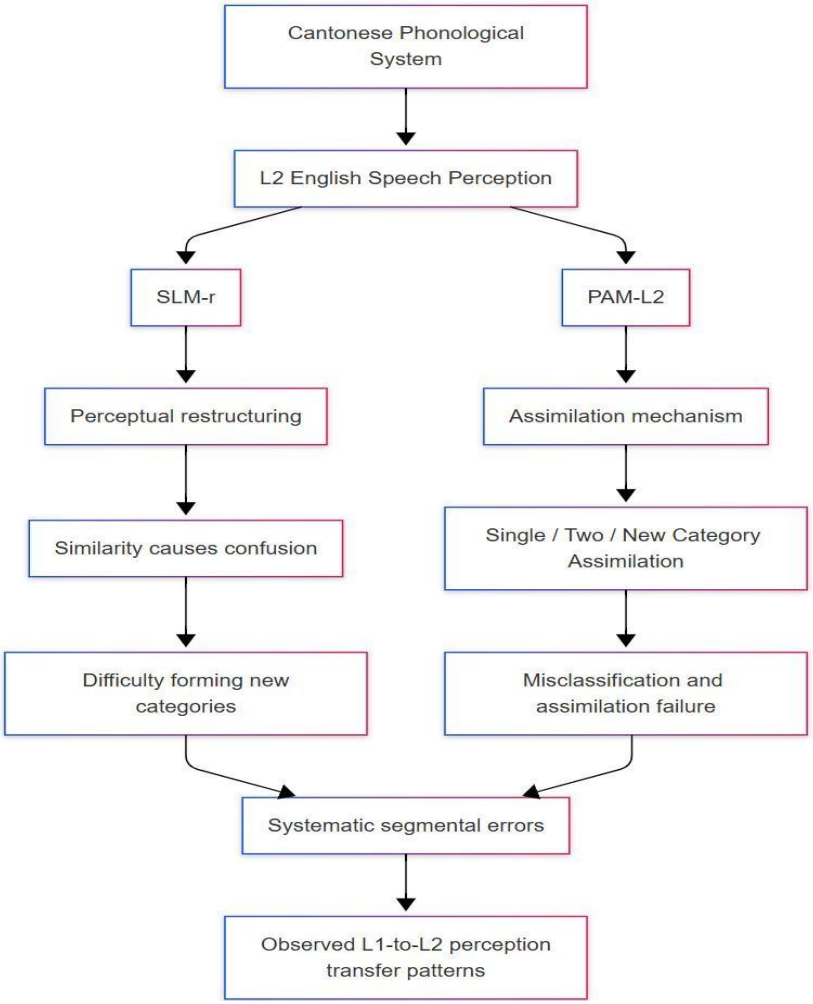


Figure 1: Theoretical Framework

Methodology

This study adopted a mixed-methods research design to explore the perceptual transfer of English phonological features among Cantonese-speaking learners (See Figure 2). The quantitative component consisted of a phoneme perception questionnaire to identify accuracy patterns anderror types. The qualitative component, comprising semi-structured interviews, providedinsights into learners' subjective awareness, interpretive strategies, and perceived L1 influences. All participants were informed of the study's purpose and provided written informed consent. Participation was voluntary, and all responses were anonymized. The study was conducted in accordance with the ethical guidelines of the researchers' institution.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

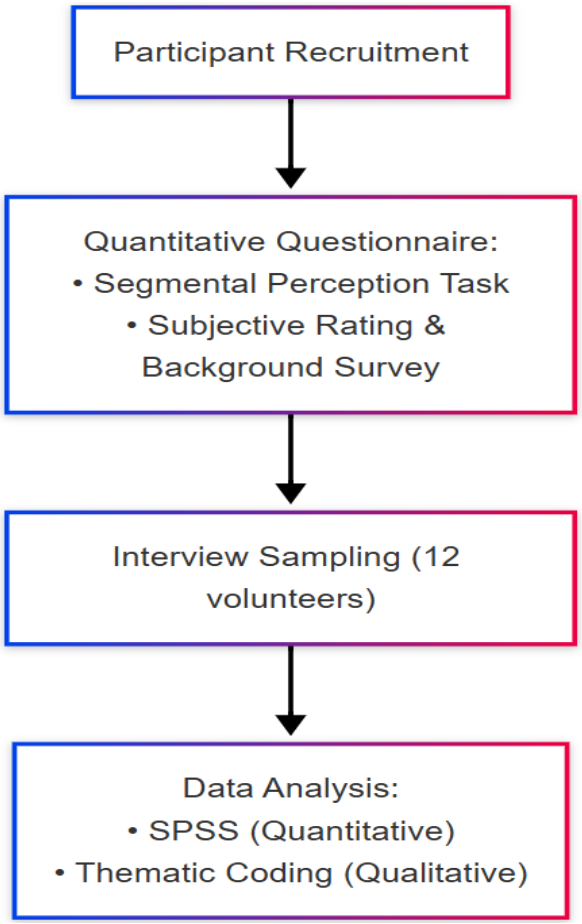


Figure 2: Research Procedure Flowchart

A mixed-methods approach was chosen to capture both observable behavioral patterns and underlying cognitive and perceptual reasoning, offering a more holistic understanding of the transfer phenomena.

The study targeted 60 Cantonese-speaking undergraduate students from a university in Guangdong Province. Participants were required to have intermediate or higher English proficiency. All data were collected remotely. Quantitative results were analyzed using SPSS, focusing on frequency counts, error patterns, and contrast-level differences. Interview data underwent thematic analysis to identify recurring themes related to L1-based transfer effects. The selection of English segmental and suprasegmental features for the questionnaire was guided by prior literature and the conceptual framework detailed below.

Conceptual Framework

This study investigates the influence of the Cantonese phonological system on the perception of English segmental and suprasegmental features. The conceptual framework (see Figure 3) outlines the key factors involved in L2 speech perception by Cantonese learners, including consonant and vowel system differences, syllable structure constraints, and prosodic mismatches. These L1-specific features are hypothesized to lead to systematic perception errors in English, particularly in identifying phonemic contrasts and stress patterns.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

The framework categorizes perception errors into two main types: segmental (e.g., misperception of / θ /, /æ/, /ɪ/) and suprasegmental (e.g., lexical stress, connected speech). These are examined through both quantitative and qualitative methods. The conceptual modelalso incorporates learners' dialectal background as a potential moderating variable. This framework serves as the foundation for task design, variable selection, and data analysis in this study.

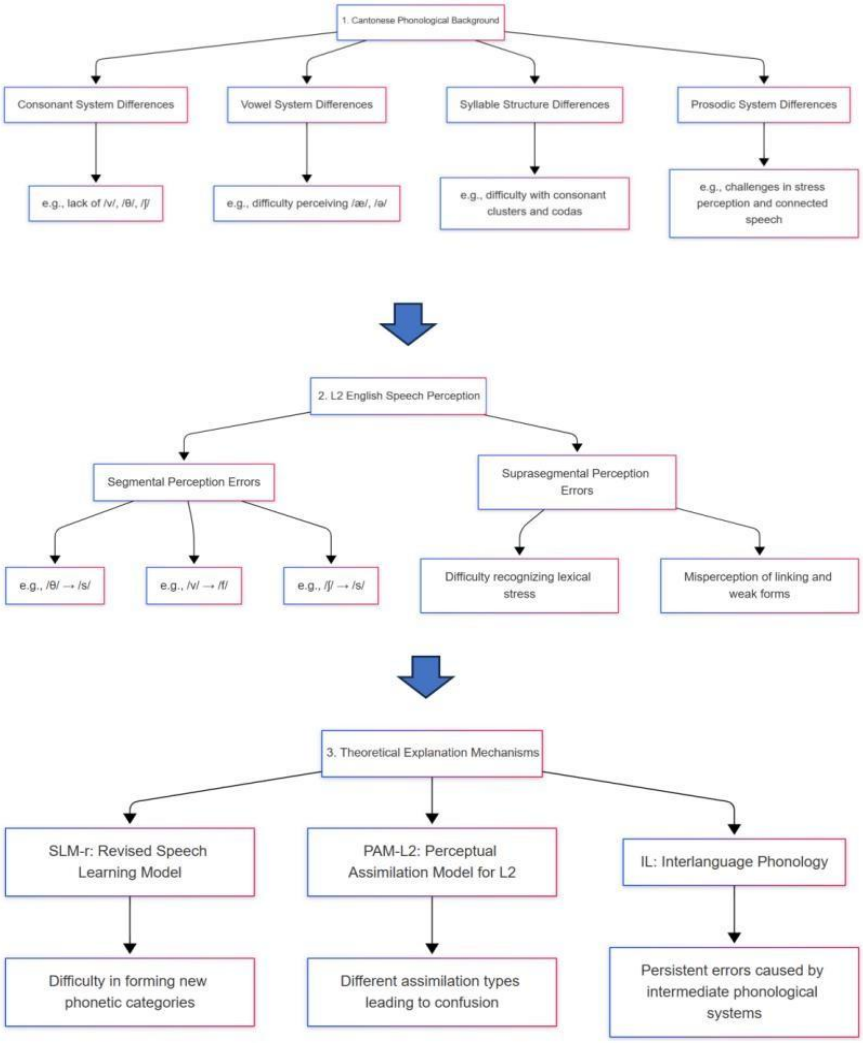


Figure 3: Conceptual Framework

Participants

A total of 60 undergraduate students were recruited from a university in Guangdong Province, China. All participants were native speakers of Cantonese, aged between 18 and 24, with intermediate to upper-intermediate proficiency in English (based on self-reported CET-4 or IELTS equivalence). Participants had no known hearing impairments and no formal training in phonetics or linguistics.

To ensure a degree of dialectal representation, participants were drawn from both Guangzhou Cantonese and other regional varieties (e.g., Taishanese, Zhongshan dialect), although no subgroup analysis was conducted due to sample size limitations. Participants completed a language background questionnaire detailing their exposure to English, age of acquisition, and use of Mandarin or other dialects.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Instruments

Questionnaire

A questionnaire was developed to examine Cantonese-speaking learners' perception of selected English segmental contrasts. The design of the questionnaire was informed by previous literature on L1-L2 phoneme perception and targeted contrasts commonly reported as challenging for Cantonese learners. The questionnaire was distributed via Google Forms and consisted of three main sections.

Section 1 called Segmental Perception Tasks.This section included multiple-choice listening items that tested participants' ability to identify target words from English minimal pairs. Each item played a short audio clip containing a single word, and participants were asked to choose the word they heard from four written options. The contrasts focused on phonemes that are often difficult for Cantonese speakers, such as /θ/–/s/, /æ/–/ɛ/, /v/–/w/, and /l/–/n/. All audio stimuli were recorded by a native speaker of General American English to ensure clarity and consistency. For example:/θ/ vs. /s/: thin vs. sin； /v/ vs. /w/: vet vs. wet.

Section 2 is Subjective Perception Ratings. In this section, participants were asked to rate how difficult they found different pairs of English phonemes to distinguish. Using a 5-point Likert scale (1 = very easy, 5 = very difficult), they evaluated contrasts such as /θ/–/s/, /æ/–/ɛ/, /v/–/w/, and /l/–/n/. The internal consistency of the scale was tested using Cronbach's alpha, which yielded a value of 0.712. This indicates acceptable reliability for exploratory research. Construct validity was supported by expert review and pilot testing, with an overall validity coefficient of 0.70, suggesting that the scale was sufficiently aligned with the intended phonological constructs.

Participants were also asked which phoneme they were more likely to confuse when uncertain (e.g., whether they tend to hear /θ/ as /s/), and which cues they relied on to make perceptual judgments.

Section 3 named Language Background and Experience.This section collected demographic and language-related background information, including participants' native language(s), age of English learning onset, prior phonetic or pronunciation training, self-assessed English proficiency, and frequency of English use in daily life.

Additionally, participants were asked whether they believed their Cantonese background influenced their ability to perceive English sounds, and whether they subconsciously used Cantonese phonological categories when they were uncertain about what they heard. An open-ended question was set at last to invit participants to describe experiences ofmishearing an English word and to reflect on possible reasons for the misperception (e.g., speed, accent, similar-sounding words, or prior mislearning).

Structured Interview Protocol

To gain deeper insights into learners' perceptual experiences, follow-up written interviews were conducted with a voluntary subset of 12 participants who had previously completed the questionnaire. The interview protocol comprised five open-ended questions, prompting participants to elaborate on their listening difficulties, perceived L1 influence, and coping strategies. Sample questions included "Which English sounds do you find most difficult to

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

distinguish, and why?" "Do you think your Cantonese pronunciation habits affect how you perceive English sounds?" "What strategies have helped you improve your English sound perception?" All responses were analyzed qualitatively using thematic coding to identify recurring patterns and learner beliefs related to segmental and suprasegmental perception.

Data Collection and Analysis

Data collection was conducted over a two-week period through remote and asynchronous means. A total of 60 Cantonese-speaking university students participated in the study by completing a questionnaire designed to assess their perception of selected English segmental contrasts. The questionnaire integrated a perception task, subjective difficulty ratings, and a background survey.

Following the questionnaire, 12 participants were invited to participate in follow-up written interviews. These interviews aimed to elicit deeper insights into learners' awareness of L1 transfer, perceptual strategies, and experiences of phoneme misperception.

Quantitative data from the questionnaire were analyzed using SPSS software. Descriptive and inferential statistics were calculated to examine perception accuracy and its relationship to subjective ratings. Reliability and construct validity of the rating scale had been confirmed during instrument development. Interview responses were examined using inductive thematic analysis. Recurring themes were identified and cross-compared with the quantitative findings to support data triangulation. This mixed-methods approach allowed for a more comprehensive understanding of how L1 phonological transfer affects L2 segmental perception.

Findings

This section presents and interprets the results of a mixed-methods analysis examining the influence of Cantonese phonological structure on English speech perception. Quantitative findings from a phoneme perception task are integrated with qualitative insights from semi-structured interviews, and the analysis is framed by two theoretical models: the Revised Speech Learning Model (SLM-r) and the Perceptual Assimilation Model for L2 learners (PAM-L2). The discussion addresses the three research questions and confirms the predicted transfer patterns at both segmental and suprasegmental levels.

Segmental and Suprasegmental Performance

The perception task revealed consistent difficulty with several English phoneme contrasts, particularly those absent in the Cantonese inventory. The /v/ versus /w/ contrast yielded the lowest accuracy at 25.0%, with over 60% of participants misidentifying the nonword "vill" as the familiar word "will." This pair also received the highest mishearing tendency and a subjective difficulty rating above 2.9, indicating strong perceptual confusion. Similarly, the /θ / versus /s/ contrast showed only 60.0% accuracy, with many learners unable to distinguish "thin" from "sin" or "thing." In contrast, phoneme pairs that exist in the Cantonese inventory or closely align with existing categories— such as /l/ versus /n/— achieved higher accuracy (66.7%), suggesting that phonological compatibility facilitates perceptual accuracy. The vowel contrast /æ/ versus /ɛ/ showed moderate difficulty (68.3% accuracy), reflecting partial perceptual distinction.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Correlation Patterns

As shown in Figure 4, the subjective difficulty ratings negatively correlate with actual performance for most contrasts, especially /θ/ and /v/.

	Q1 accuracy	Q2 accuracy	Q3 accuracy	Q4 accuracy	Q5 accuracy	Q6 accuracy	Q7/ θ / vs /s/	Q7/æ/ vs /ɛ/	Q7/v/ vs /w/	Q7/l/ vs /n/
Q1 accuracy thin	1	0.176	0.289	−0.236	0.314	0.091	−0.182	−0.057	−0.191	0.031
Q2 accuracy bad	0.176	1	0.355	−0.103	0.512	−0.016	−0.04	0.057	0.036	0.134
Q3 accuracy lap	0.289	0.355	1	−0.163	0.229	0.158	−0.234	0.039	−0.044	0.074
Q4 accuracy vill	−0.236	−0.103	−0.163	1	−0.257	−0.052	0.092	−0.04	0.045	−0.156
Q5 accuracysleen	0.314	0.512	0.229	−0.257	1	0.199	−0.244	0.081	−0.054	−0.067
Q6 accuracy wet	0.091	−0.016	0.158	−0.052	0.199	1	0.006	−0.068	−0.132	−0.174
Q7 / θ / vs /s/	−0.182	−0.04	−0.234	0.092	−0.244	0.006	1	0.199	0.451	0.399
Q7 /æ/ vs /ɛ/	−0.057	0.057	0.039	−0.04	0.081	−0.068	0.199	1	0.278	0.447
Q7 /v/ vs /w/	−0.191	0.036	−0.044	0.045	−0.054	−0.132	0.451	0.278	1	0.549
Q7 /l/ vs /n/	0.031	0.134	0.074	−0.156	−0.067	−0.174	0.399	0.447	0.549	1

Figure 4: Correlation Heatmap of Perception Accuracy (Q1–Q6) And Perceived DifficultyRate (Q7)

Subjective ratings of difficulty correlated negatively with objective accuracy, particularly for /v/ and /θ/, indicating that learners were aware of the contrasts they found most challenging. These trends confirm that segmental misperception is systematic and rooted in L1 phonological structures.

Although suprasegmental features were not the primary target, performance on nonwords and unfamiliar prosodic items revealed clear evidence of tonal transfer (see Table 2). In the case of the nonword "sleen" (Q5), which contained atypical stress and no lexical support, accuracy dropped to 78.3% — considerably lower than for familiar words — and subjective difficulty ratings peaked at 3.45. Conversely, when the phoneme contrast /v/ versus /w/ was embedded in a familiar item like "wet" (Q6), accuracy rose to 83.3%, underscoring the role of lexical familiarity in facilitating phoneme identification. Interview data confirmed that participants often relied on pitch height, syllable duration, and even syllable loudness to identify stress, a strategy consistent with tonal processing in Cantonese but mismatched with English stress-timed rhythm. Several learners explicitly stated that they guessed stress placement based on how "long" or "strong" a syllable sounded, signaling a transfer of suprasegmental attention from pitch-based to stress-based systems.

Table 2: Accuracy Rates and Difficulty Ratings for Q1-Q6

Item	Correct Answer	Accuracy (%)	Mean
Q1 – thin	thin	60.00	1.900
Q2 – bad	bad	68.33	1.850
Q3 – lap	lap	66.67	1.650
Q4 – vill	vill	25.00	1.683
Q5 – sleen	sleen	78.33	3.450
Q6 – wet	wet	83.33	1.167
/θ/ vs /s/ – Difficulty rating (Q7)	—	—	2.617
/v/ vs /w/ – Difficulty rating (Q7)	—	—	2.583
/l/ vs /n/ – Difficulty rating (Q7)	—	—	2.167
/æ/ vs /e/ – Difficulty rating (Q7)	—	—	2.417
/s/ vs /θ/ – Mishear tendency (Q8)	—	—	3.000
/v/ vs /w/ – Mishear tendency (Q9)	—	—	2.917

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Lexical Familiarity and Error Types

Lexical familiarity emerged as a significant modulating factor (see Table 3). The sharp contrast between the perception of the novel nonword "vill" (25.0% accuracy) and the known word "wet" (83.3%)—both testing the /v/-/w/ contrast—demonstrates that word recognition can scaffold phoneme perception. This finding aligns with previous research suggesting that top-down lexical access interacts with bottom-up auditory processing, allowing learners to supplement phonemic uncertainty with contextual knowledge (Cutler et al., 2004; Melnik et al., 2022). Interviewees reported greater confidence in identifying sounds embedded in familiar words and often resorted to semantic inference when uncertain. One learner described recognizing "think" only because it appeared in the phrase "in water," implying thatlexical cues can override segmental ambiguity. Such reliance on context was particularly prevalent in low-frequency or unfamiliar items, where phonemic boundaries were unclear.

Table 3: Summary Of Substitution Patterns and L1 Equivalents

English Phoneme	Common Misheard Form	L1 Equivalent or Strategy	Example from Interview
/v/	/w/, /f/	No /v/ in Cantonese	"very" → "fei-rei", "wery"
/θ/	/s/, /f/	No interdental fricatives	"think" → "sink"
/æ/	/ɛ /	Vowel overlap	"bad" → "bed"
/l/	/n/	Partial overlap	"light" → "night" (less frequent)

Phonological Transfer and Theoretical Models

Overall, error types were not random but reflected L1-based substitution strategies (See Table 4). The most frequent substitutions included /θ/ to /s/, /v/ to /w/, and /æ/ to /ɛ/, all of which correspond to the closest perceptual equivalents in the Cantonese phoneme inventory. These patterns mirror the assimilation types described in PAM-L2 (Escudero, 2007), where /θ/ and /s/ fall under single category assimilation—both being mapped onto one Cantonese category, yielding poor discrimination. The /v/ and /w/ pair represents category goodness assimilation: the unfamiliar /v/ is perceived as a variant of /w/, preventing the formation of a new category.

The /æ/ and /ɛ/ contrast reflects category overlap, where learners can partially distinguish the two but still demonstrate boundary confusion. In contrast, /l/ and /n/, which exist as distinct phonemes in Cantonese, show high discrimination accuracy, consistent with PAM-L2's two-category assimilation. These patterns also align with SLM-r ' s emphasis on phonetic similarity and category formation: when L2 sounds are perceived as too similar to L1 categories, learners fail to establish new phonetic representations (Flege, 2018; Salheen & Yap, 2023).

Table 4: Mapping of Error Types to PAM-L2/SLM-r Categories

Contrast	Error Pattern	PAM-L2 Category	SLM-r Interpretation
/θ/ vs /s/	Confused as same sound	Single Category	New Category not formed
/v/ vs /w/	Overlapping	Category Goodness	Category assimilation
/æ/ vs /ɛ/	Partial confusion	Category Overlap	Incomplete phonetic boundary
/l/ vs /n/	Mostly accurate	Two Category	Existing L1 categories support

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Thematic Analysis of Learner Strategies

Thematic analysis of learner strategies revealed three dominant trends: tonal-based stress inference, structural simplification, and semantic guessing. Many learners described using pitch movement or syllable duration to locate English stress, reflecting transfer from Cantonese's tonal system. Others reported frequent omission of final consonants or insertion of epenthetic vowels to conform to Cantonese's (C)V(C) syllable structure. For example, "street" was often perceived as "seut" and "next" as "nek," suggesting a restructuring of English syllables to fit L1 phonotactics. This finding aligns with Optimality Theory predictions on markedness constraints (Zhao et al., 2024) and confirms that structural simplification is a perceptual—not just productive—strategy. Interviewees also reported heavy reliance on context when unable to identify a phoneme. Notably, 71.7% of learners acknowledged guessing based on context, while only 5% explicitly recognized L1 influence, suggesting that phonological transfer often operates below the level of conscious awareness (see Table 5).

Table 5: Learners' Reported Perce ption Strategies (Q10–Q17)

Strategy Reported	% of Learners
Guessed based on context	71.7%
Used Cantonese to "guess" when unsure	68.3%
Paid attention to pitch/duration for stress	62.5%
Explicitly compared with Cantonese phonemes	5.0%

These themes are further illustrated in the keyword cloud generated from interview data, where strategies such as guessing, tone-based stress inference, and phoneme confusion emerge as salient (see Figure 5).

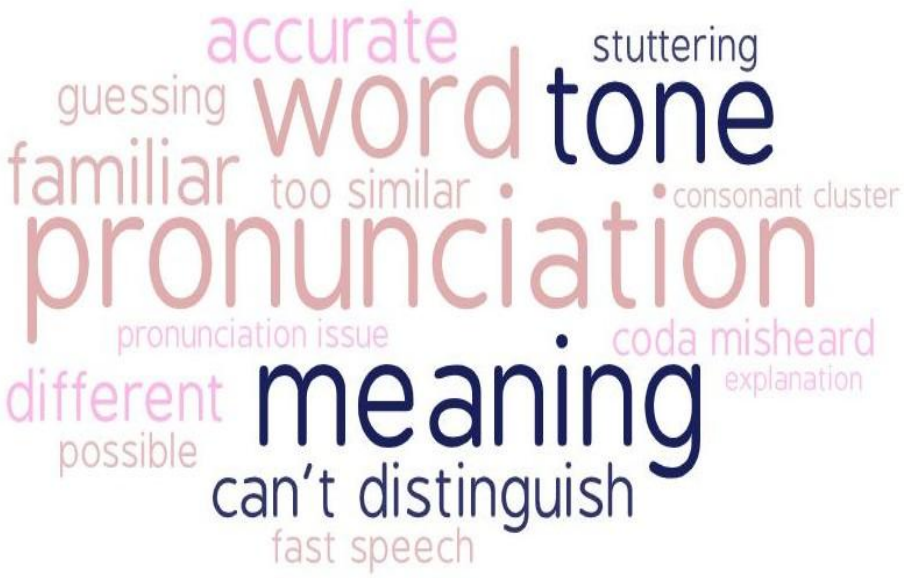


Figure 5. Keyword Cloud of Learner-Reported Strategies and Perceptual Difficulties During English Listening Tasks

Thematic analysis of interview responses revealed several recurring strategies used by learners when processing unfamiliar English phonemes. Many participants reported relying on pitch movement and syllable duration to identify lexical stress, a strategy that reflects their native language's use of tone. Others mentioned modifying the structure of unfamiliar words by

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

omitting final consonants or inserting vowels, aligning with the (C)V(C) structure typical of Cantonese syllables. These modifications were particularly common in words containing consonant clusters or coda consonants. In addition, a large number of learners described using semantic or contextual information to guess the identity of unfamiliar sounds, especially when they encountered low-frequency or nonword stimuli. While some learners were aware of these strategies, most did not consciously attribute their perceptual decisions to their L1 background.

These findings confirm all three hypotheses proposed in the study (see Table 6). Hypothesis 1 (H1), which predicted lower accuracy for phonemes absent in Cantonese, is supported by the poor performance on /θ/, /v/, and /æ/. Hypothesis 2 (H2), which anticipated suprasegmental interference from tonal transfer, is confirmed through learners' reported use of pitch and duration instead of intensity and vowel reduction. Hypothesis 3 (H3), which proposed systematic alignment with SLM-r and PAM-L2, is substantiated by the error patterns and substitution types that match theoretical predictions.

Table 6: Summary of Hypothesis Testing

Hypothesis	Statement	Result	Supporting Evidence
H1	Learners will show lower accuracy for English phonemes not present in Cantonese.	Supported	/v/, /θ/, /æ/ had lowest accuracy and highest difficulty ratings.
H2	Suprasegmental perception is influenced by tonal transfer.	Supported	Learners used pitch/duration cues for stress; tonal strategies reported in interviews.
H3	Perceptual errors align with SLM-r and PAM-L2 predictions.	Supported	/θ/→/s/(single category),/v/↔/w/ (assimilation),/æ/vs/ɛ /(overlap) matched model types.

More broadly, the results demonstrate that L1 phonological structure is a powerful and predictive force in shaping L2 perception. Segmental misperception stems from category assimilation and failed category formation, while suprasegmental misperception arises from the misapplication of tonal strategies to stress-based prosody. These effects are further compounded by structural simplification and orthographic interference. The logographic nature of Chinese writing, which lacks phoneme-grapheme correspondence, may delay phonological awareness and contribute to perceptual ambiguity in English (Jiang et al., 2024).As such, L1 transfer is not a peripheral issue but a central cognitive mechanism in second language acquisition—a position supported by both behavioral and neurocognitive evidence (Gosselke et al., 2021).

Discussion and Limitations

Based on the observed patterns of segmental confusion and suprasegmental misperception, several instructional considerations emerge. Training that focuses on auditory discrimination may help learners form more accurate phoneme boundaries, particularly for contrasts that are absent in the Cantonese inventory (Bradlow et al., 1999; McLaughlin et al., 2024). Exercises involving minimal pairs and varied speaker input have been shown to support the development of robust phonetic categories (Lively et al., 1993), and may reduce perceptual confusability in L2 learners.

For suprasegmental features, instructional materials that emphasize intensity, vowel reduction, and rhythmic timing could help learners shift attention away from pitch-based cues, which are often transferred from tonal L1s such as Cantonese (Cutler, 2012; Kang, 2024; Cheng, 2023).

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Visual tools such as pitch contour diagrams and rhythm tapping may recalibrate learners' attention to English stress patterns.

In addition, incorporating high-frequency lexical items into listening and pronunciation tasks can provide top-down support for phoneme identification (Field, 2009; Tse, 2022), while also leveraging the familiarity effect to enhance intelligibility. Finally, raising learners' metalinguistic awareness of cross-linguistic transfer may encourage more conscious strategy use during listening (Vandergrift & Goh, 2012; Tse, 2022), particularly in tasks that require segmental discrimination and stress perception under real-time processing demands.

Despite these contributions, the study has limitations. The sample was drawn from a single university group, limiting generalizability. Future research should explore dialectal variation within Cantonese speakers, as subgroups like Taishanese or Zhongshanese may exhibit distinct transfer patterns (Tse, 2022). Moreover, the task relied on isolated words, which lack ecological validity; sentence-level or discourse-based stimuli could better capture real-world perception (Wang, 2025). Finally, while Likert ratings and interview data offer insights into cognitive strategies, real-time measures such as eye-tracking or ERP would provide a more dynamic view of L2 perceptual processing (Gosselke et al., 2021).

The present findings offer theoretical implications for models of L2 speech perception. Specifically, the observed segmental confusions and stress misalignments align with predictions from SLM-r (Flege & Bohn, 2021), which emphasizes the role of L1 phonetic categories in shaping L2 perception. The data also support PAM-L2's (Best & Tyler, 2007) account of assimilation patterns, particularly in cases of category overlap or single-category assimilation. Notably, the findings extend these models by illustrating how tonal L1 experience may interfere not only with segmental perception but also with suprasegmental stress patterns—an area that has received relatively less attention in both models. As such, the study contributes to a more nuanced understanding of how L1 prosodic structures shape L2 phonological acquisition.

In sum, the results provide compelling evidence that L1 transfer systematically constrains L2 phonological perception. These constraints are not arbitrary but follow predictable pathways rooted in phonological compatibility and perceptual mapping. The findings validate the explanatory power of SLM-r and PAM-L2 and offer practical guidance for teaching English pronunciation more effectively in tonal-language contexts.

Conclusion

This study set out to examine how L1 Cantonese phonological structures influence the perception of English segmental and suprasegmental features among university learners. The findings successfully addressed the research objectives by identifying systematic patterns of segmental confusion and prosodic misperception, which can largely be attributed to cross-linguistic transfer.

Theoretically, the study provides empirical support for the Speech Learning Model-revised (SLM-r) and the Perceptual Assimilation Model for L2 (PAM-L2), extending their explanatory power to a tonal-language context. Practically, the results offer pedagogical insights into how instruction can target specific perceptual gaps, especially through varied input, metalinguistic training, and prosody-focused activities.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Several challenges were encountered. The task format, based on isolated word stimuli, limited ecological validity, and the reliance on post-task ratings constrained the analysis of real-time processing. To address these obstacles, future studies could incorporate sentence- or discourse-level stimuli and adopt online measures such as eye-tracking or ERP to capture dynamic perceptual responses.

Future research should also consider intra-dialectal variation among Cantonese speakers (e.g., Taishanese, Zhongshanese) and explore how prosodic interference interacts with lexical tone systems in L1. Longitudinal studies may further clarify how perceptual patterns evolve with increased exposure and training, contributing to a more comprehensive understanding of L2 phonological development in tonal-language learners.

In total, this study not only charts the perceptual footprint of L1 transfer in Cantonese-English contexts but also offers a roadmap for more effective and theoretically grounded approaches to second language phonological development. To teach pronunciation effectively, we must first learn to listen—not only to the sounds, but to the systems learners bring with them.

Acknowledgements
The authors would like to thank the faculty members and student participants from the undergraduate program in Guangdong Province for their time and participation, which significantly contributed to the study.

References

Alfadhil Attaher Salheen, D., & Yap, N. T. (2023). Acquiring L2 phonology: Foreign accent and L1 influence. International Journal of Academic Research in Business and Social Sciences, 13(4), 1451–1459.

Archibald, J. (2021). Ease and difficulty in L2 phonology: A mini-review. Frontiers in Communication, 6(1), 1–7.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. Perception & Psychophysics, 61(5), 977–985.

Chen, W. (2021). A phonetic and phonological analysis of the production and perception of English post-alveolar fricatives by Mandarin and Mandarin/Shaoxing Wu speakers [Master's thesis, Shanghai International Studies University].

Cheng, J. (2023). Acquisition of English liaison among Chinese EFL learners from the perspective of language transfer. Heliyon, 9(1), e20418. https://doi.org/10.1016/j.heliyon.2023.e20418

Cutler, A. (2012). Native Listening: Language Experience and the Recognition of Spoken Words. The MIT Press.

Darcy, I. (2024). Second language speech processing: A guide to conducting experimental research. Taylor & Francis.

Escudero, P. (2007). Second language phonology: The role of perception. In Phonology in context (pp. 109–134).

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. Studies in Second Language Acquisition, 26(4), 551–585.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046
Field, J. (2009). Listening in the Language Classroom. Cambridge: Cambridge University Press.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), Second language speech learning: Theoretical and empirical progress (pp. 3–83). Cambridge University Press.

Gosselke Berthelsen, S., Horne, M., & Shtyrov, Y. (2021). Phonological transfer effects in novice learners: A learner's brain detects grammar errors only if the language sounds familiar. Bilingualism: Language and Cognition, 24(4), 656–669.

He, X., Tao, J., & Chen, Z. (2025). An experimental study of the positive transfer of Cantonese Rusheng syllable structures on unreleased stops in English. Advances in Education, Humanities and Social Science Research, 2(2), 79–83.

Jiang, Y., Gai, X., & Wang, Z. (2024). The influence of segmental and suprasegmental phonological awareness on word and pseudoword reading: A comparison between native English speakers and native Chinese speakers of English. Frontiers in Psychology, 15(1), 1–12.

Kang, Y. (2024). A study on suprasegmental acquisition of Chinese EFL learners from the perspective of L1 transfer. Taylor & Francis Online, 1(1), 1–15.

Kapnoula, E. C. (2021). On the locus of L2 lexical fuzziness: Insights from L1 spoken word recognition and novel word learning. Frontiers in Psychology, 12, 689052. https://doi.org/10.3389/fpsyg.2021.689052

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. The Journal of the Acoustical Society of America, 94(3 Pt 1), 1242–1255.

McLaughlin, D. J., Baese-Berk, M. M., & Van Engen, K. J. (2024). Exploring effects of brief daily exposure to unfamiliar accent on listening performance and cognitive load. Frontiers in Language Sciences, 2(1).

Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French /u/-/y/ contrast. Second Language Research, 38(4), 581–605. https://doi.org/10.1177/02676583211049360

Rocca, B., & Darcy, I. (2022). Comprehensibility improvements in integrated pronunciation instruction: A comparison of instructional methods and task effects. Journal of Second Language Pronunciation, 8(3), 328–362.

Simonchyk, A., & Darcy, I. (2023). Development of production skills in the absence of precise phonolexical representations. Second Language Research, 39(3), 623–648.

Soo, R., & Babel, M. (2024). Regular exposure facilitates dual-mapping of Cantonese pronunciation variants. Canadian Journal of Linguistics / Revue canadienne de linguistique, 69(1), 99–117.

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 (Vol. 28, pp. 97–126). John Benjamins.

Vandergrift, L., & Goh, C. C. M. (2012). Teaching and learning second language listening: Metacognition in action. Routledge.

Wang, Z. (2025). The effect of accent on listening comprehension: Chinese L2 learners' perceptions and attitudes. THAITESOL Journal, 38(2), 47–69.

Xie, Y. (2024). Phonological transfer from Cantonese to English for Cantonese-English bilingual children: A scoping review. English Language and Literature Studies, 14(4), 1–17. https://doi.org/10.5539/ells.v14n4p1

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 26 (September 2025) PP. 689-708
DOI: 10.35631/IJMOE.726046

Zhao, R., Gong, Q., & Chen, Y. (2024). Examining negative transfer in Cantonese learners of English as a foreign language: An optimality theory approach to English syllable acquisition. IRAL – International Review of Applied Linguistics in Language Teaching, 62(1).