

**ROMANIZATION-INDUCED PHONETIC DEVIATION IN MANDARIN
AFFRICATES OF INDONESIAN MIGRANT WORKERS' INTERLANGUAGE
IN TAIWAN**

¹Arjuna Satria Dewa Bagaskara*

¹Universitas Brawijaya, Indonesia

¹dewaagaskara@student.ub.ac.id,

²Ressi Maulidina Delijar

²Universitas Brawijaya, Indonesia

²ressimd@ub.ac.id

*Corresponding author

Received : April 15, 2026

Revised : June 17, 2026

Accepted : June 25, 2026

Published : June 30, 2026

How to Cite (in APA Style):

Bagaskara, A. S. D., & Delijar, R. M. (2026). Romanization-induced phonetic deviation in Mandarin affricates of Indonesian migrant workers' interlanguage in Taiwan. *Pioneer: Journal of Language and Literature*, 18(1), 180-199.

<https://doi.org/10.36841/pioneer.v18i1.8140>

© 2026 The Authors. Licensed by CC BY-SA 4.0

Abstract: This study investigates how a non-standard romanization system used in a private Indonesian training institution shapes the production of Mandarin affricates by Indonesian Migrant Workers (IMW) in Taiwan. It asks how L1-based spellings interact with Indonesian phonology to produce systematic deviations in Mandarin affricate interlanguage. A descriptive case study was conducted with five female workers, a purposively bounded sample reflecting the female-dominated composition of the caregiving sector in which they are employed, combining acoustic analysis of /z/, /c/, /j/, /q/, /zh/, /ch/ measuring Voice Onset Time (VOT), Frication Duration (FD), and Center of Gravity (CoG) with semi-structured interviews and analysis of the training module. The results show pervasive over-aspiration of non-aspirated affricates, variable aspiration of aspirated segments, and consistent dentalization of retroflex affricates, indicating a collapse of aspiration contrasts and a shift toward alveolar/dental articulations. Module inspection reveals that modified romanization visually maps Mandarin affricates onto Indonesian graphemes for /tʃ/ and /dʒ/, while interviews indicate a strong reliance on these spellings and hyper articulation strategies. The study concludes that local romanization, L1 phonology, and communicative pressures jointly foster a fossilized interlanguage affricate system, and argues for revising romanization practices and integrating explicit articulatory instruction and feedback in migrant worker Mandarin training.

Keywords: *acoustic phonetics, affricate consonants, Indonesian migrant workers, interlanguage, Romanization*

INTRODUCTION

Standard Mandarin functions as a major global lingua franca and serves as a key medium of communication in transnational labour markets (Campbell & Grondona, 2008). Taiwan has become one of the principal destinations for Indonesian Migrant Workers (IMW), and employers in the caregiving, service, and industrial sectors increasingly expect basic Mandarin proficiency (BP2MI, 2025; RTI, 2024). In such high-stakes workplace contexts, accurate realization of segmental contrasts is critical, because even subtle phonemic misproductions can impede intelligibility and disrupt interaction, which makes the phonetics and phonology of Mandarin particularly relevant for the linguistic preparation of IMW (Saville-Troike & Barto, 2016).

Within Mandarin segmental phonology, affricate consonants represent a locus of difficulty for many L2 learners (Reetz & Jongman, 2020). These segments form contrastive series across place of articulation, e.g., alveolar, retroflex, and alveolo-palatal, and state of aspiration, with voiceless aspirated and unaspirated pairs carrying the phonemic load (Lee-Kim, 2014; Li et al., 2019). Aspiration is largely encoded in temporal parameters such as Voice Onset Time (VOT) and Frication Duration (FD), while place contrasts are primarily cued by spectral measures such as Center of Gravity (CoG) (Lee-Kim, 2014; Li et al., 2014). For learners whose L1 lacks aspirated obstruents and retroflex consonants, including Indonesian speakers, acquiring these fine-grained temporal and spectral distinctions is therefore demanding and provides a backdrop for understanding how IMW approximate Mandarin affricates (Li et al., 2019; Reranta & Huntley, 2022).

Indonesian employs a voiceless-voiced plosive system without a phonemic aspiration contrast and does not include retroflex consonants in its inventory (Alwi et al., 2000; Li et al., 2019). Empirical work on Indonesian plosives has shown that speakers typically produce relatively short VOT values and implement stop contrasts mainly through voicing, rather than through differences in aspiration (Li et al., 2019). When such speakers acquire Mandarin, aspirated and unaspirated affricates are often assimilated to existing L1 categories, which weakens their intended phonemic distinctiveness. In practical settings, this L1-L2 phonological gap is further shaped by the instructional practices of private job-training centers (*Lembaga Pelatihan Kerja Swasta*, LPKS) that prepare prospective IMW.

One salient instructional factor is the choice of romanization system. *Hànyǔ Pīnyīn* has functioned as the dominant international standard for transcribing Standard Mandarin and is widely used in formal education and research (Afrina & Cleveresty, 2021; Zhou, 2019). Beyond its role as a transcription tool, *Pīnyīn* provides an orthographic representation of Mandarin's segmental and tonal structure, and work on orthographic effects has shown that learners internalize grapheme-phoneme mappings and may develop non-target pronunciations when spelling patterns diverge from the phonological reality of the L2. Even standard *Pīnyīn* can induce systematic mispronunciations when its spellings underrepresent certain phonological segments (Basseti, 2007, 2008; Escudero et al., 2008). These findings suggest that any romanization system used in instruction has the potential to shape, and in some cases distort, learners' phonetic categories.

At LPKS Eka Management Cipta Sejahtera in Ponorogo, East Java, Mandarin teaching materials employ a modified romanization system drawing on elements of the International Phonetic Alphabet (IPA) and Indonesian orthography rather than adopting *Hànyǔ Pīnyīn* (LPKS, n.d.). Standard *Pīnyīn* forms such as *měicì* and *zìjǐ* are printed as “Mei Jhe” and “Ce Ci”, prompting learners to associate Mandarin affricates with Indonesian grapheme-phoneme correspondences (Bagaskara, 2026). In practice, this design encourages learners to interpret Mandarin /z/ as analogous to Indonesian /c/ [tʃ] and “zh” as comparable to Indonesian “j” [dʒ], thereby reinforcing L1-based equivalence classifications and raising the question of how such romanization choices affect the phonetic realization of Mandarin affricates in migrant speech.

Rather than developing full theoretical groundwork here, this study draws lightly on the Speech Learning Model and on research into orthographic effects, both of which treat L2 segments as being learned in relation to existing L1 categories and as sensitive to written input (Basseti, 2008; Flege, 1995; Flege & Bohn, 2021). Prior research on Mandarin affricates among Indonesian learners offers useful but partial insight, and none of it has examined how a non-standard, institution-specific romanization shapes affricate production among migrant workers specifically (Fitriyah & Subandi, 2020; Li et al., 2019; Li & Gu, 2015; Wang et al., 2013). To close part of that gap, five IMW trained at the same institution were examined through a descriptive case study, an approach chosen deliberately because it allows sustained acoustic and interview-based attention to a small, tightly bounded group rather than shallow coverage of a large one (Creswell & Poth,

2018). All five happened to be women, which mirrors rather than distorts the demographic reality of the caregiving sector in which most of them work (BP2MI, 2025). This article accordingly asks how a modified romanization system interacts with L1 phonology and learner strategy to shape the acoustic realization of Mandarin affricates in IMW interlanguage, and it identifies the pedagogical and phonological factors behind the deviations found.

REVIEW OF LITERATURE

Acoustic Phonetics and Mandarin Affricates

Acoustic phonetics examines the physical properties of speech sounds, including the temporal and spectral parameters that encode phonological contrasts (Reetz & Jongman, 2020). For obstruents, measures such as VOT, FD, and CoG are central to distinguishing the place and manner of articulation as well as the state of aspiration (Irawan, 2017; Lee-Kim, 2014). In Mandarin, affricates form aspirated-unaspirated pairs within the obstruent system, and their acoustic realization combines a stop-like closure with a fricative-like release measurable in both time and frequency domains.

Large-scale acoustic studies have shown that normalized FD and amplitude are primary cues to aspiration, whereas spectral energy distribution contributes mainly to the place of articulation among Mandarin affricates (Li & Gu, 2015). Spectral analyses of Mandarin sibilants similarly indicate that retroflex segments tend to have lower CoG values than their alveolar counterparts, reflecting a longer back cavity due to tongue retraction (Lee-Kim, 2014). Research on Taiwanese and Singapore Mandarin further reports reduced or shifted alveolar-retroflex contrasts under certain prosodic or sociolinguistic conditions, visible as systematic differences in the CoG and related spectral parameters (Chang, 2012; Goh et al., 2024). These findings establish acoustic benchmarks against which L2 productions, including those of Indonesian learners, can be evaluated.

Acquisition studies have investigated how L2 learners from different L1 backgrounds produce Mandarin affricates. Thai learners, for instance, tend to narrow the aspirated-unaspirated contrast, leading to foreign-accented realizations that nonetheless preserve some temporal differentiation (Wu & Sang, 2026). Other studies have documented negative transfer from Malay and Japanese, where learners assimilate

Mandarin affricates to familiar L1 categories, a pattern observable through deviations in VOT and spectral properties (Li et al., 2019; Wang et al., 2013). However, these studies generally involve classroom learners with access to standard *Pīnyīn*, and there is limited acoustic research focusing specifically on Indonesian adult learners or migrant workers and their affricate production in real-world communicative settings.

Romanization, Orthography, and L2 Phonology

Romanization systems provide a bridge between spoken Mandarin and alphabetic scripts, with *Hànyǔ Pīnyīn* functioning as the internationally accepted standard for transcribing Standard Mandarin (Afrina & Cleverest, 2021; Zhou, 2019). *Pīnyīn* is widely regarded as offering relatively transparent grapheme-phoneme correspondences. However, its conventions sometimes underrepresent phonological segments, especially in certain rimes, which can obscure aspects of the underlying phonological structure (Bassetti, 2007). Experimental work has shown that learners of Chinese often delete vowels not represented in *Pīnyīn*, indicating that orthographic input shapes mental phonological representations and can induce systematic segmental omissions (Bassetti, 2007, 2008).

Orthographic effects in L2 phonology have been widely documented in studies across languages. Orthographic input interacts with acoustic input and can trigger segmental additions, deletions, and substitutions that do not occur in native speech (Bassetti, 2008, 2024). Such orthography-induced pronunciations arise because literate learners rely on grapheme-phoneme mappings established in their L1 when interpreting L2 spellings (Escudero et al., 2008). In English as a second language, opaque orthography can lead to persistent mispronunciations, and similar mechanisms have been observed in L2 Chinese when *Pīnyīn* or other romanization systems diverge from phonetic transparency (Bassetti, 2007, 2008). These findings imply that deviations from phonologically faithful romanization can systematically redirect learners' phonological development.

The risk of misguidance increases when institutions develop local, non-standard romanization systems that approximate L2 sounds with L1-like spellings. In vocational training contexts such as LPKS for IMW, modified romanization that aligns Mandarin segments with Indonesian orthographic conventions may be intended to facilitate decoding, but can in fact anchor learners' phonological categories more strongly in L1

(Bagaskara, 2026; LPKS, n.d.). Despite this potential impact, empirical studies explicitly linking such modified romanization systems to measurable acoustic deviations in learners' speech remain scarce, particularly in migrant-worker training settings. Earlier work on Indonesian Mandarin learners has shown that pronunciation errors cluster around segments that are absent or differently patterned in Indonesian, such as retroflex initials, aspirated obstruents, and the palatal series *j*, *q*, *x*, with error sources traced to mother-tongue transfer and insufficient understanding of pronunciation rules (Liu et al., 2021).

Second Language Speech Learning and Interlanguage

The Speech Learning Model (SLM) proposes that L2 sounds are acquired against the backdrop of an existing L1 phonetic system, with L2 categories established or merged depending on their perceived similarity to L1 sounds (Flege, 1995; Flege & Bohn, 2021). When an L2 sound is perceived as equivalent to an L1 sound, learners may fail to form a new category and instead map the L2 sound onto the L1 category, resulting in equivalence classification and a persistent foreign accent (Flege, 1995). When an L2 sound is perceived as sufficiently different, learners are more likely to create a new category, which can approach native-like realizations with adequate input and use (Flege & Bohn, 2021). This framework provides a useful lens for understanding how L1-based romanization may bias category formation for Mandarin affricates.

From this perspective, interlanguage phonology is viewed as a dynamic system shaped by L1 transfer, the quantity and quality of L2 input, and internal restructuring, but it can become fossilized when certain patterns are repeatedly reinforced and not corrected (Saville-Troike & Barto, 2016; Sinurat and Herman, 2019). For adult learners, especially those in constrained instructional environments, the fossilization of segmental deviations is common when L2 contrasts lack support in the L1 inventory or when feedback on pronunciation is limited (Han, 2004). Psycholinguistic strategies such as hyperarticulation or over-correction may be adopted to increase perceived clarity. However, without accurate articulatory targets these strategies can further distort the acoustic realization of L2 segments rather than bringing them closer to native norms (Chang, 2012; Reranta & Huntley, 2022). Therefore, the role of orthography must be considered within this broader interplay of transfer, input, and strategy.

Orthographic input interacts with these learning mechanisms in a systematic manner. When romanization systems visually assimilate L2 sounds to L1 graphemes, they may bias learners' perception of similarity and promote equivalence classification in contexts where a new L2 category is required (Bassetti, 2008). For Indonesian learners of Mandarin, a modified romanization that maps Mandarin affricates onto Indonesian “c” or “j”, for instance, is likely to strengthen L1-based categories and hinder the formation of distinct aspirated and retroflex categories in L2 (Bagaskara, 2026; Li et al., 2019). This suggests that the combined effects of L1 phonology and romanization practices must be examined empirically to understand how IMW interlanguage affricate systems emerge and become stable.

Indonesian Migrant Workers (IMW) and Mandarin Training

IMW represent a large and growing population in Taiwan, employed mainly in the domestic care, manufacturing, and service sectors (BP2MI, 2025; RTI, 2024). Language training for prospective IMW is commonly conducted through LPKS in Indonesia, which provides short, intensive courses focused on functional workplace communication rather than comprehensive linguistic competence (Purba et al., 2020; Widodo & Belgradoputra, 2019). In similar private course settings, speaking-oriented programs tend to prioritize practice-based techniques over explicit phonetic instruction, and materials are often adapted to perceived learner needs, including locally developed romanization systems (Afrina & Cleveresty, 2021; Farizi & Herwiana, 2022).

Research on Indonesian learners of Mandarin has largely concentrated on university students with access to Hànyǔ Pīnyīn and professional instructors, documenting segmental and tonal errors and frequent mispronunciations in segments absent from Indonesian, such as ü, apical-i, aspirated plosives, and the retroflex and palatal series (Fitriyah & Subandi, 2020; Li et al., 2019; Liu et al., 2021). Studies specifically targeting IMW in Taiwan have tended to focus on sociolinguistic aspects or general communicative competence rather than detailed acoustic analysis, and only a small number of works have attempted to link acoustic data with training materials to show how institutional practices, such as non-standard romanization, shape IMW interlanguage phonetics (Widodo & Belgradoputra, 2019).

METHOD

This study employed a descriptive case study design that combines acoustic phonetic analysis with interviews and document analysis (Creswell & Poth, 2018). The case was bounded by one private training institution, LPKS Eka Management Cipta Sejahtera in Ponorogo, whose alumni share a relatively homogeneous training background and exposure to a modified romanization system (Bagaskara, 2026; LPKS, n.d.). Five IMW (all female, aged 37–49) who had completed training at this institution and were working in Taiwan were selected through homogeneous purposive sampling to reduce physiological variation and align instructional histories. A small, purposively selected sample of this kind is standard practice in case study and acoustic phonetic research, where depth of analysis across many tokens per speaker takes priority over breadth across many speakers, and where the goal is analytic insight into a bounded case rather than statistical generalization to a wider population (Etikan et al., 2016). The exclusively female composition of the sample was not an artificial restriction imposed by the researcher but a reflection of the workforce itself, since national placement data show that the caregiving and domestic sectors, where the great majority of this institution's graduates are employed in Taiwan, are staffed predominantly by women (BP2MI, 2025). Restricting the sample to female participants therefore also served to control gender as a potential source of physiological variation in the acoustic measures under study.

Data comprised three sources: recorded productions of target words containing Mandarin affricates /z/, /c/, /j/, /q/, /zh/, /ch/ in citation form, the LPKS Mandarin training module employing modified romanization, and semi-structured interviews on linguistic backgrounds, learning experiences, and pronunciation difficulties (Bagaskara, 2026). Speech was recorded in WAV format and analyzed in Praat to obtain VOT, FD, and CoG, summarized as mean values and ranges per consonant category and compared with native reference values from previous acoustic studies (Boersma & Weenink, 2024; Lee-Kim, 2014; Li & Gu, 2015; Narhan et al., 2023). Interview transcripts were coded thematically to identify reliance on romanization, awareness of articulatory targets, and strategies such as over-correction, while document analysis traced mappings between Mandarin affricates and their modified spellings. Triangulation across these sources was then used to relate observed phonetic deviations to pedagogical, L1-phonological, and psycholinguistic factors (Bassetti, 2008; Flege, 1995; Saville-Troike & Barto, 2016).

FINDINGS AND DISCUSSION

Acoustic Characteristics of IMW Affricate Production

Before turning to qualitative interpretation, this section outlines the main acoustic patterns observed across speakers and affricate categories. Table 1 summarizes the ranges of VOT, FD, and CoG for each Mandarin affricate in comparison with native reference values from previous studies, providing the quantitative basis for the subsequent discussion of contrast erosion and place shifts in IMW interlanguage.

Table 1. Summary of Acoustic Parameters for Mandarin Affricates

Affricate	Type	Parameter	Native Range (literature)	IMW Range (R1–R5)	Main Deviation Pattern
/z/	alveolar, non-asp	VOT	20–40 ms (Lee-Kim, 2014)	77–197 ms	Strong VOT over-lengthening
		FD	0–20 ms	42–167 ms	Frication overly long
		CoG	4000–8000 Hz	4039–6649 Hz	Still alveolar region
/c/	alveolar, asp	VOT	80–120 ms	88–190 ms	Overlap with /z/
		FD	60–100 ms	63–155 ms	Often longer than native
		CoG	4000–8000 Hz	3961–6346 Hz	Alveolar
/j/	palatal, non-asp	VOT	20–40 ms	98–145 ms	Over-aspiration
		FD	0–20 ms	48–123 ms	Excessive frication
		CoG	3000–5000 Hz	4417–6124 Hz	Higher spectral energy
/q/	palatal, asp	VOT	80–120 ms	50–185 ms	Under/over-aspiration (variable)
		FD	60–100 ms	20–106 ms	Sometimes shorter than /j/
		CoG	3000–5000 Hz	839–6102 Hz	Highly variable
/zh/	retroflex, non-asp	VOT	20–40 ms	91–155 ms	Over-aspiration
		FD	0–20 ms	44–126 ms	Long frication
		CoG	2000–4000 Hz	4092–4591 Hz	CoG too high → dentalization
/ch/	retroflex, asp	VOT	80–120 ms	87–205 ms	Large variation, overlap with /zh/
		FD	60–100 ms	22–157 ms	Unstable duration
		CoG	2000–4000 Hz	2674–6144 Hz	Very high CoG for some tokens

As Table 1 indicates, non-aspirated affricates /z/ and /j/ are consistently produced with VOT values well above native norms, frequently overlapping with or even exceeding those of aspirated counterparts /c/ and /q/. This over-lengthening of VOT erodes the temporal distinction between aspirated and unaspirated affricates that is phonemic in Mandarin. In the spectral domain, retroflex affricates /zh/ and /ch/ exhibit CoG values that are systematically higher than the 2000–4000 Hz range reported for native speakers, often reaching 4000–6000 Hz, which is typical of dental or alveolar sibilant constrictions (Lee-Kim, 2014; Li & Gu, 2015). These patterns point to a weakening of both the aspiration contrast and the retroflex–alveolar place contrast in IMW interlanguage, a

configuration that has not been reported in studies of classroom learners using standard *Pīnyīn* (Wang et al., 2013; Wu & Sang, 2026).

To illustrate how these patterns manifest in individual tokens, the following figures present representative Praat displays of affricate productions by selected participants. Figure 1 shows waveform and spectrogram snapshots of the alveolar affricates /z/ and /c/ produced by participant R2. These tokens exemplify the temporal convergence of non-aspirated and aspirated alveolar affricates that was consistently observed across the data.

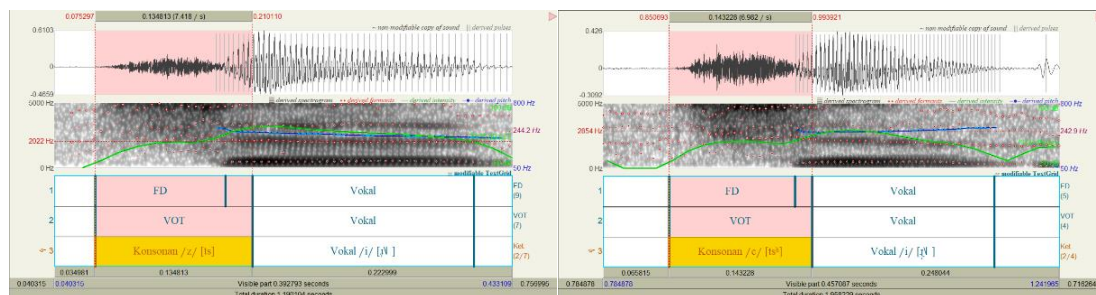


Figure 1. Praat /z/ and /c/ by R2

Figure 1 illustrates the productions of /z/ and /c/ by participant R2. In both tokens, the waveform shows a clear stop release followed by a long stretch of aperiodic frication noise, and the spectrogram reveals densely packed high-frequency energy extending over more than 100 ms. The measured VOT values are 134 ms for /z/ and 143 ms for /c/, with frication durations that are similarly extended, producing almost identical temporal envelopes. This near-perfect overlap indicates that, for this speaker, the aspirated–unaspirated contrast that should differentiate /z/ [ts] and /c/ [ts^h] in native Mandarin has effectively collapsed in the alveolar series, so that both segments are realized as heavily aspirated affricates rather than as a minimal pair with distinct timing cues.

While the alveolar series demonstrates a collapse of the aspiration contrast through generalized over-aspiration, the palatal series reveals a different configuration of temporal deviation. Figure 2 presents Praat displays of /j/ and /q/ as produced by participant R4, highlighting how non-aspirated /j/ can become over-aspirated and aspirated /q/ under-aspirated. This figure serves to illustrate the instability of aspiration timing within the palatal affricate pair.

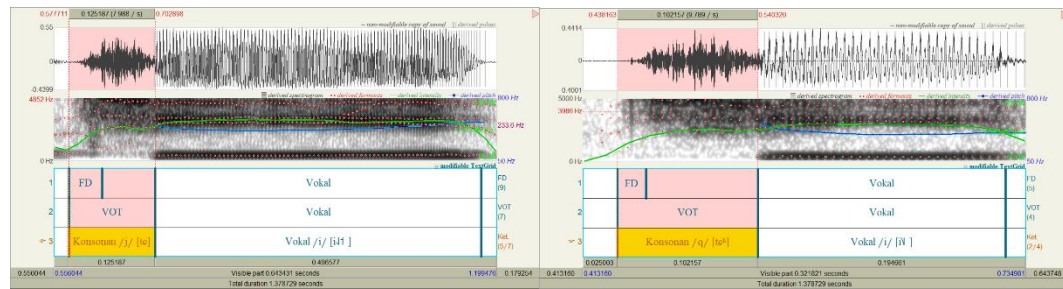


Figure 2. Praat /j/ and /q/ by R4

Figure 2 presents the productions of /j/ and /q/ by participant R4, highlighting a different but related pattern in the palatal series. The /j/ token, which should be a non-aspirated [tɕ], shows an extended frication interval and a VOT of approximately 125 ms, accompanied by strong high-frequency energy in the spectrogram. In contrast, the /q/ token, expected to be an aspirated [tɕ^h], displays a shorter VOT of around 102 ms and a less salient aspirated burst. In other words, /j/ is over-aspirated and /q/ is under-aspirated, such that their temporal profiles converge instead of forming a robust oppositional pair. This pattern suggests that IMW do not consistently map palatal affricates onto the native-like aspiration categories, but rather use variable lengthening as a general strategy without clear phonological targets.

The spectral properties of the retroflex series provide particularly clear evidence of place-of-articulation shifts. Figure 3 displays waveform, spectrogram, and spectral slice information for /zh/ and /ch/ produced by participant R3. These tokens are representative of a broader tendency in the corpus for retroflex affricates to exhibit high-frequency energy concentrations, consistent with dental or alveolar rather than retroflex constrictions, thereby visualizing the dentalization inferred from the CoG measurements.

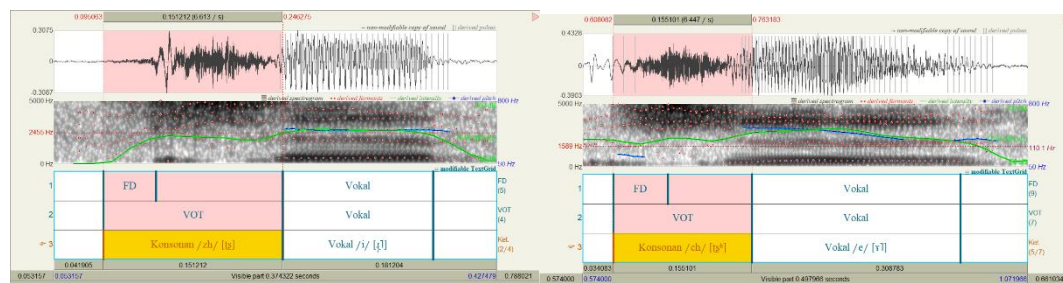


Figure 3. Praat /zh/ and /ch/ by R3

Figure 3 shows /zh/ and /ch/ produced by participant R3 and focuses on spectral characteristics. In both tokens, spectral slices taken over the fricative portion exhibit

centers of gravity reaching up to 6144 Hz, far above the 2000–4000 Hz range typically associated with Mandarin retroflex affricates in native speakers. The spectrograms also show concentrated energy in the higher frequency bands rather than the relatively lower-frequency emphasis expected for retroflex constrictions with a long back cavity. These acoustic properties are more consistent with dental or alveolar affricates, such as [tʃ] or [dʒ], than with [tʂ] or [tʂʰ], providing strong evidence of systematic dentalization in the retroflex series.

Taken together, Figures 1-3 demonstrate that IMW do not merely show attenuated aspiration contrasts or slightly shifted places of articulation, as reported in previous work on L2 Mandarin affricates (Wang et al., 2013; Wu & Sang, 2026), but display a more radical restructuring of the affricate system. Temporal cues that should support aspirated-unaspirated oppositions are neutralized through pervasive over- and under-aspiration, while spectral cues associated with retroflexion are replaced by patterns typical of alveolar/dental articulations. This constellation of patterns is best understood as an interlanguage system reshaped by L1 categories and orthography, rather than as a simple “approximation” of the native Mandarin affricate inventory.

Romanization-Induced Pedagogical Interference

Document analysis of the LPKS Mandarin module reveals a consistent strategy of replacing standard *Pīnyīn* spellings with locally designed romanization that closely resembles Indonesian orthography. For instance, *měici* is printed as “Mei Jhe”, *zìjǐ* as “Ce Cì”, *qìchē* as “Jì Jhe”, and *zhīdào* as “Ce Tao”, with the consonant clusters ‘c’, ‘j/jh’, and “che” functioning as primary cues to learners’ pronunciation. These spellings do not simply transcribe Mandarin sounds, but actively recast them in terms of familiar Indonesian grapheme–phoneme correspondences. Table 2 presents selected examples of these mappings, juxtaposing standard *Pīnyīn* forms, the LPKS romanization, and the likely L1 categories activated by these spellings, in order to clarify how the orthographic system may orient learners’ mental phonetic representations.

Table 2. Examples of *Pīnyīn*–LPKS Romanization–L1 Category Mappings

Mandarin Word	<i>Pīnyīn</i>	Target Affricate(s)	LPKS Romanization	Likely L1 association	Potential phonetic effect
每次	měici	/c/ [tʂʰ]	Mei Jhe	“jh” ≈ /dʒ/	Retroflex/aspiration pulled to /dʒ/
自己	zìjǐ	/z/ [ts]	Ce Cì	“c” ≈ /tʃ/	/z/ assimilated to /tʃ/
汽车	qìchē	/q/, /ch/ [tʂʰ, tʂʰ]	Jhi Jhe	“jhe” ≈ /dʒ/	Palatal/retroflex merged to /tʃ/
知道	zhīdào	/zh/ [tʂ]	Ce Tao	“ce” ≈ /tʃ/	Retroflex replaced by /dʒ/

As shown in Table 2, the LPKS romanization systematically uses Indonesian-like spellings (“c”, “j/jh”, “che”) to represent Mandarin affricates, effectively inviting learners to categorize L2 segments via L1 grapheme–phoneme mappings. This design goes beyond the *Pīnyīn*-induced segmental misrepresentations documented in classroom settings by directly anchoring affricates in established Indonesian categories /tʃ/ and /dʒ/. In SLM terms, such orthographic choices bias learners toward equivalence classification by visually encoding L2–L1 similarity where meaningful phonological distinctions exist (Flege, 1995; Flege & Bohn, 2021).

The acoustic patterns described in Table 1 align closely with these orthographic mappings. Learners who see *zǐjǐ* represented as “Ce Ci” produce /z/ with long VOT and high CoG more typical of Indonesian /tʃ/, while those taught to read “Jhi” for /zh/ realize retroflex affricates with CoG values exceeding 6000 Hz, consistent with dental or alveolar constrictions. Interview comments such as “I just read it like the Indonesian ‘c’” or “zh is like our j” make explicit that the module’s romanization is treated as the primary guide to segmental production, rather than standard *Pīnyīn* or articulatory descriptions.

Interview excerpts support this interpretation. One participant stated, “For ‘zh’ we were taught to read it as ‘jh’, so I remember it as the Indonesian ‘j’ sound. I never knew the tongue had to be curled back.” Another remarked that because the module writes *zǐjǐ* as “Ce Ci,” she “just pronounces it like the Indonesian ‘c’ but with a bit more hiss.” These comments show that the module’s romanization serves as a mental phonetic manual: learners rely on it not only to decode words, but also to infer how to produce them. The elevated CoG values for /zh/ and /ch/ in IMW speech, especially the extreme case of 6144 Hz observed for R3, align with this reliance on dental/alveolar L1 constrictions instead of retroflex articulations (Lee-Kim, 2014; Li & Gu, 2015).

Compared with prior studies where learners used standard *Pīnyīn* and often retained at least partial affricate contrasts (Wang et al., 2013; Wu & Sang, 2026), the present data suggest that a non-standard romanization can fundamentally reshape the emergent L2 phonological system. Rather than merely under-realizing aspiration or slightly shifting place of articulation, IMW trained with the LPKS system appear to have re-categorized Mandarin affricates as variants of Indonesian /tʃ/ and /dʒ/, with aspiration and retroflexion failing to form robust categories.

L1 Phonology, Interlanguage Fossilization, and Psycholinguistic Strategies

The observed deviations are also consistent with the influence of Indonesian phonology. Given that Indonesian lacks an aspiration contrast and retroflex series, learners have no L1 template for Mandarin aspirated affricates or retroflex place distinctions (Alwi et al., 2000; Li et al., 2019). In this context, the over-lengthening of VOT for non-aspirated segments can be seen as an attempt to increase articulatory “strength” without a clear understanding of how aspiration functions as a phonological feature in Mandarin. Similarly, the dentalization of retroflex affricates reflects an approximation of unfamiliar [tʂ] or [tʂʰ] sounds by familiar /tʃ/ or /dʒ/ categories, which are easier to produce and already entrenched in L1.

Over time, the combination of L1 transfer and romanization-induced bias appears to have led to fossilization. Participants had been living and working in Taiwan for extended periods, yet their affricate productions remained systematically non-native-like, suggesting that their interlanguage had stabilized (Bagaskara, 2026). This aligns with theoretical accounts of fossilization in adult L2 speech, particularly when learners receive limited corrective feedback on segmental errors (Han, 2004; Saville-Troike & Barto, 2016). Participants reported that Taiwanese employers rarely corrected their pronunciation as long as the communicative intent was clear from context, indicating that the communicative ecology does not challenge entrenched phonetic patterns.

Psycholinguistic strategies further interact with these structural factors. Several participants described “making the sound longer” or “exaggerating” consonants when they felt misunderstood. This behavior resembles hyperarticulation, a strategy intended to increase clarity by enhancing acoustic cues (Chang, 2012; Reranta & Huntley, 2022). Previous research on speaking courses in Indonesia has also highlighted that sustained practice in a supportive environment, coupled with encouragement to speak without fear of making mistakes, is seen by tutors as a key condition for improving oral skills (Farizi & Herwiana, 2022). However, in the absence of accurate articulatory targets, hyperarticulation in this context often amplifies deviations: VOT and FD increase even further, while CoG remains high due to continued use of dental/alveolar constrictions. Unlike classroom learners in previous studies, who sometimes use hyperarticulation to move toward native norms (Wu & Sang, 2026), IMW in this study appear to

hyperarticulate along an L1-shaped trajectory, reinforcing rather than correcting their interlanguage patterns.

Theoretical and Pedagogical Implications

The present study extends existing research in three ways. First, it documents a configuration of Mandarin affricate production which characterized by collapsed aspiration contrasts and systematic dentalization of retroflexes, in a population of adult migrant workers trained with a non-standard romanization system, a context rarely examined in prior work. While earlier studies reported partial contrast reduction or moderate place shifts in learners using standard *Pīnyīn* (Wang et al., 2013; Wu & Sang, 2026), the current findings suggest a more profound re-categorization driven by orthographic design.

Second, the study empirically links modified romanization to measurable acoustic deviations, thereby concretizing arguments about orthographic effects in L2 phonology beyond tone and vowel phenomena (Bassetti, 2007; Li et al., 2014). By combining acoustic measurements, module analysis, and learner interviews, the article shows how specific grapheme choices (“c”, “j/jh”, “che”) guide learners toward particular L1 categories and how these choices surface as elevated CoG and over-lengthened VOT in actual speech. This triangulated evidence refines theoretical models by highlighting the role of locally designed orthographies in shaping equivalence classification and fossilization within SLM.

Third, the study foregrounds the ecological conditions of migrant worker language learning, where short-term vocational training, functional communicative goals, and limited feedback jointly shape interlanguage development. Unlike university learners in previous Indonesian-Mandarin studies, IMW in Taiwan operate in an environment where intelligibility suffices and phonetic accuracy is rarely monitored (Widodo & Belgradoputra, 2019). Situating orthographic effects within this ecology underscores that correcting romanization alone is insufficient. Effective intervention must also include explicit articulatory instruction and mechanisms for ongoing feedback.

Pedagogically, the findings suggest three priorities for Mandarin training programs targeting migrant workers. First, romanization systems should align with *Hànyǔ* *Pīnyīn* or, at minimum, avoid directly mapping Mandarin affricates onto Indonesian graphemes that encode different phonological categories. Second, pronunciation

instruction should incorporate articulatory descriptions and visual tools e.g., tongue diagrams, video models, Praat-based feedback to help learners differentiate aspirated vs. unaspirated and retroflex vs. alveolar/palatal articulations. Third, training providers and employers should establish simple feedback routines so that segmental errors that undermine contractiveness can be noticed and addressed rather than allowed to fossilize in interlanguage.

CONCLUSION AND SUGGESTIONS

Conclusion

This study has shown that IMW trained with a non-standard, L1-based romanization system produce Mandarin affricates with substantial deviations from native norms in both temporal and spectral dimensions. Non-aspirated affricates exhibit over-lengthened VOT that overlaps with aspirated counterparts, while retroflex affricates are systematically dentalized, as indicated by elevated CoG values, effectively collapsing phonemic contrasts that are crucial to Mandarin intelligibility. These patterns differ qualitatively from those reported in previous studies on learners using standard *Pīnyīn*, suggesting that local romanization practices can fundamentally reshape the emergent L2 phonological system rather than merely attenuate existing contrasts. These findings are best read as an intensive, case-bound account of one institutional context rather than as a statistical generalization to the broader IMW population in Taiwan, and their value lies in the depth of insight they offer into how a specific training design can shape a specific interlanguage system.

The analysis further indicates that these phonetic characteristics arise from the interaction of three main factors: pedagogical interference from modified romanization that visually aligns Mandarin affricates with Indonesian /tʃ/ and /dʒ/, transfer from an L1 phonological system lacking aspirated and retroflex segments, and psycholinguistic strategies of hyper articulation developed under communicative pressure in the workplace. Within the framework of the Speech Learning Model and orthographic effects research, the findings highlight how orthography can bias equivalence classification and contribute to interlanguage fossilization when not carefully designed.

Suggestions

Future research could test the generalizability of these findings by examining other LPKS that use different romanization practices, comparing IMW in different destination countries, or implementing controlled pedagogical interventions that replace modified romanization with *Pīnyīn* and introduce explicit articulatory instruction. Such studies would help determine to what extent the patterns documented here are locally contingent or representative of a broader phenomenon in migrant-worker language training. At a practical level, revising romanization schemes and pronunciation teaching in vocational contexts may not only improve IMW's segmental accuracy but also enhance their communicative confidence and professional integration in Mandarin-speaking environments.

REFERENCES

- Afrina, U., & Cleveresty, T. B. (2021). Analisis penggunaan Hanyu Pinyin dalam pelafalan bahasa Mandarin berdasarkan unsur pembentuknya. *Metahumaniora*, 11(2), 135–144. <https://jurnal.fib.unpad.ac.id/index.php/metahumaniora/article/view/32776>
- Alwi, H., Moeliono, A. M., Dardjowidjojo, S., & Lapoliwa, H. (2000). *Tata bahasa baku bahasa Indonesia* (3rd ed.). Balai Pustaka.
- Bagaskara, A. S. D. (2026). *Produksi fonetik konsonan afrikat bahasa Mandarin pekerja migran Indonesia di Taiwan: Studi kasus LPKS Eka Management Cipta Sejahtera* [Undergraduate thesis, Universitas Brawijaya]. <https://repository.ub.ac.id/id/eprint/260010>
- Bassetti, B. (2007). Effects of Hanyu Pinyin on pronunciation in learners of Chinese as a foreign language. In A. Guder, X. Jiang, & Y. Wan. (Eds.), *The Cognition, Learning and Teaching of Chinese Characters*. Beijing Language and Culture University Press. <https://eprints.bbk.ac.uk/id/eprint/528>
- Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 191–206). Multilingual Matters. <https://doi.org/10.21832/9781847691118-013>
- Bassetti, B. (2024). Orthographic effects in the phonetics and phonology of second language learners and users. In *The Cambridge handbook of bilingual phonetics and phonology* (pp. 699–720). Cambridge University Press.
- Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (6.4.07).
- BP2MI. (2025). *Laporan tahunan publikasi data pekerja migran Indonesia tahun 2024*. https://kp2mi.go.id/uploads/statistik/images/data_06-02-2025_Laporan_Tahunan_Publikasi_Data_PMI_2024.pdf

- Campbell, L., & Grondona, V. (2008). Ethnologue: Languages of the world. In G. Raymond (Ed.), *Language*, 84(3), 636-641. SIL International. <https://doi.org/10.1353/lan.0.0054>
- Chang, Y.-H. (2012). *Variability in cross-dialectal production and perception of contrasting phonemes: The case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin* [Dissertation]. University of Illinois.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design* (4th ed.). SAGE.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345–360. <https://doi.org/10.1016/j.wocn.2007.11.002>
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Farizi, F., & Herwiana, S. (2022). A study of teaching speaking techniques used at Peace English Course Pare, Kediri. *Pioneer: Journal of Language and Literature*, 14(1), 43-57. <https://doi.org/10.36841/pioneer.v14i1.1382>
- Fitriyah, Z., & Subandi, S. (2020). Analisis kesalahan pelafalan huruf konsonan (声母 shēngmǔ) oleh para tokoh dalam drama Thailand yang berjudul Love Beyond Frontier. *Jurnal Bahasa Mandarin*, 5(2), 99. <https://ejournal.unesa.ac.id/index.php/manadarin/article/view/45181>
- Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Flege, J. E., & Bohn, O.-S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901>
- Goh, H. L., Woon, F. T., Moisi, S. R., & Styles, S. J. (2024). Contrastive alveolar/retroflex phonemes in Singapore Mandarin bilinguals: Comprehension rates for articulations in different accents, and acoustic analysis of productions. *Language and Speech*, 67(4), 924–944. <https://doi.org/10.1177/00238309231205012>
- Han, Zhaohong. (2004). *Fossilization in adult second language acquisition* (D. Singleton, Ed.; 1st ed.). Multilingual Matters.
- Irawan, Y. (2017). *Fonetik akustik: Sebuah pengantar telaah wujud akustik bahasa* (Vol. 1). Angkasa.
- Lee-Kim, S.-I. (2014). Revisiting Mandarin ‘apical vowels’: An articulatory and acoustic study. *Journal of the International Phonetic Association*, 44(3), 261–282. <https://doi.org/10.1017/S0025100314000267>

- Li, H., Baryadi, I. P., & Wijana, I. D. P. (2019). Sound pattern of Indonesian plosives. *Linguistik Indonesia*, 37(1), 1–12. <https://doi.org/10.26499/li.v37i1.84>
- Li, J., Xu, R. B., & Mok, P. P. K. (2014). Effect of orthography on L2 production of Mandarin tones. *Proc. 4th International Symposium on Tonal Aspects of Languages (TAL 2014)*, 147–151. https://www.isca-archive.org/tal_2014/li14_tal.html#
- Li, S., & Gu, W. (2015). Acoustic analysis of Mandarin affricates. *Proc. Interspeech*, 1680–1684. <https://doi.org/10.21437/Interspeech.2015-387>
- Liu, D. D., Artawa, K., Budiarsa, M., Suastra, I. M., & Septevany, E. (2021). Pronunciation errors of Indonesian Mandarin learners. *TIJOLAC: The International Journal of Language and Cultural*, 3(1), 66–82. <https://www.growingscholar.org/journal/index.php/TIJOLAC/article/view/111>
- LPKS. (n.d.). *Modul Bahasa Mandarin*. LPKS Eka Management Cipta Sejahtera.
- Narhan, R., Sholihatun, P., & Syarfina, T. (2023). Analisis frekuensi, intensitas, dan durasi pada bahasa Turki oleh native speaker dan non-native speaker menggunakan Praat. *LINGUA: Jurnal Bahasa, Sastra, Dan Pengajarannya*, 20(2), 351–372. <https://doi.org/10.30957/lingua.v20i2.840>
- Purba, Y. A., Aini, Y. N., Asiati, D., & Ngadi, N. (2020). Kesenjangan kepesertaan jaminan sosial tenaga kerja di Indonesia: Analisis data Sakernas 2018. *Jurnal Ketenagakerjaan*, 15(2). <https://doi.org/10.47198/naker.v15i2.72>
- Reetz, H., & Jongman, A. (2020). *Phonetics: transcription, production, acoustics, and perception* (2nd ed.). John Wiley & Sons, Inc.
- Reranta, R. C., & Huntley, M. L. (2022). Acoustic parameters giving the angry impression in Lampungese neutral speech. *Linguistik Indonesia*, 40(1), 49–61. <https://doi.org/10.26499/li.v40i1.272>
- RTI. (2024, January 26). Taiwan's Labor Ministry to launch digital learning program for migrant workers. *Ministry of Foreign Affairs, Republic of China (Taiwan)*. https://roc-taiwan.org/th_en/post/6482.html
- Saville-Troike, M., & Barto, K. (2016). *Introducing second language acquisition* (3rd ed.). Cambridge University Press.
- Sinurat, B. and Herman. (2019). The typology of the first year students' pronunciations at FKIP-Universitas HKBP Nommensen Pematangsiantar on English voiceless plosive consonants. *Journal of English Language and Culture*, 10(1), 1-9. <http://dx.doi.org/10.30813/jelc.v10i1.1903>
- Wang, M., Lin, C. Y., & Yang, C. (2013). Contributions of phonology, orthography, and morphology in Chinese-English biliteracy acquisition: A one-year longitudinal study. In X. Chen, Q. Wang, & Y. C. Luo (Eds.), *Reading development and difficulties in monolingual and bilingual Chinese children* (pp. 191–211). Springer. https://doi.org/10.1007/978-94-007-7380-6_10

- Widodo, H., & Belgradoputra, R. J. (2019). Perlindungan pekerja migran Indonesia. *Binamulia Hukum*, 8(1), 107–116. <https://ejournal.hukumunkris.id/index.php/binamulia/article/view/343>
- Wu, S., & Sang, Y. (2026). Production of Mandarin affricates by Thai learners: An acoustic study. *Acoustical Science and Technology*, 47(3), 221–231. <https://doi.org/10.1250/ast.e25.72>
- Zhou, M. (2019). *Language ideology and order in rising China* (1st ed.). Palgrave Macmillan Singapore.