



Intelligibility and Comprehensibility of Thai-Accented English: The Role of Listener L1 in ELF Communication

Suthinee Peerachachayanee^{a,*}, Sujinat Jitwiriyanont^b

^a suthinee_p@payap.ac.th, English as an International Language, Chulalongkorn University
^b sujinat.j@chula.ac.th, Department of Linguistics and Center of Excellence in Southeast Asian Linguistics, Faculty of Arts, Chulalongkorn University, Thailand
* Corresponding author, suthinee_p@payap.ac.th

APA Citation: Peerachachayanee, S., & Jitwiriyanont, S. (2026). Intelligibility and comprehensibility of Thai-accented English: The role of listener L1 in ELF communication. <i>LEARN Journal: Language Education and Acquisition Research Network</i> , 19(2), 879-912. https://doi.org/10.70730/TNQB3068	
Received 21/10/2025	ABSTRACT This study investigates how listeners from six first language (L1) backgrounds: Thai, Burmese, Chinese (Mandarin), Japanese, Korean, and English, perceive Thai-accented English in terms of intelligibility and comprehensibility. Using a mixed-methods design, participants completed two tasks: (1) a sentence transcription task measuring objective intelligibility, and (2) a rating task evaluating subjective comprehensibility of unscripted speech. Results show a strong effect of listener L1 on both outcomes. Thai listeners demonstrated the highest intelligibility and comprehensibility, confirming the Interlanguage Speech Intelligibility Benefit (ISIB) in matched-L1 contexts. Burmese listeners performed comparably to native English speakers in intelligibility and rated Thai-accented English as highly comprehensible, likely due to phonological similarities and regional exposure. Mandarin Chinese listeners achieved moderate scores in both intelligibility and comprehensibility, with notable individual
Received in revised form 16/03/2026	
Accepted 27/05/2026	

	variation. Japanese and Korean listeners performed lower in both tasks, which reflects greater phonological distance from both Thai and English. Native English listeners achieved a moderately high score on the intelligibility task, yet rated Thai-accented English as the least comprehensible, illustrating the native listener paradox. These findings add to English as a Lingua Franca (ELF) research by underscoring the listener-dependent nature of intelligibility and questioning the continued reliance on native-speaker norms in speech assessment and pedagogy. Keywords: Thai-accented English, intelligibility, comprehensibility, listener L1, English as a Lingua Franca
--	---

Introduction

English today functions as a global lingua franca, serving as the primary medium of communication among speakers from diverse linguistic and cultural backgrounds. In many of these interactions, no native speaker of English is present, and the communicative success depends less on adherence to native norms and more on whether interlocutors can understand one another. Within this paradigm, the concepts of intelligibility (the degree to which listeners accurately decode and identify individual words in the speech) and comprehensibility (the perceived ease of understanding) have gained prominence as key constructs for analyzing spoken communication in multilingual settings (Munro & Derwing, 1995; Smith & Nelson, 1985). Research in English as a Lingua Franca (ELF) has consistently highlighted that intelligibility, rather than native-like accuracy, constitutes the most critical measure of communicative effectiveness (Jenkins, 2000; Seidlhofer, 2004). This orientation challenges traditional views of accentedness, which position deviation from a native norm as a deficit, and instead emphasizes the importance of mutual understanding across diverse first language (L1) backgrounds. Despite this theoretical reorientation, empirical evidence remains uneven across regions. While European varieties of English and their intelligibility have been studied extensively, Southeast Asian varieties remain comparatively underexplored, even though the region is characterized by frequent cross-linguistic encounters involving English.

Research in cross-language speech perception has long demonstrated that the listener's L1 exerts a powerful influence on how non-native speech is perceived (Best, 1995). Phoneme categories, phonotactic constraints, and suprasegmental expectations shaped by the L1 can bias recognition, sometimes facilitating perception but often hindering it when the target

contrasts do not exist in the L1. For instance, Japanese listeners have difficulty perceiving /r/-/l/ contrasts (Aoyama et al., 2004), while Cantonese listeners often struggle with voiced syllable-final plosives (Chan & Li, 2000). Such asymmetries extend to the evaluation of accented speech, that is, what may be intelligible for one listener group may not be so for another. This asymmetry is further complicated by findings from Interlanguage Speech Intelligibility Benefit (ISIB) research, which suggests that listeners often understand speech from talkers who share their L1 more easily than speech from native speakers or other non-native speakers (Bent & Bradlow, 2003). For example, a Thai listener may find Thai-accented English more intelligible than native English, due to shared expectations and overlapping error patterns. However, evidence for ISIB in mismatched contexts where the listener does not share the speaker's L1 but comes from another non-native background is mixed, and the degree to which phonological distance between languages predicts intelligibility remains unresolved.

Thailand provides a particularly compelling context for investigating these questions. English holds an important role in business, education, and regional cooperation under frameworks such as ASEAN, yet exposure to native English input remains relatively limited for many Thai speakers. Thai-accented English, therefore, emerges as a common variety in local and regional ELF interactions. While studies have documented segmental and suprasegmental features of Thai English, such as consonant cluster simplification, substitution of consonants, and differences in stress placement (e.g. Peerachachayanee, 2022; Suntornsawet, 2022), few studies have examined the extent to which such features affect intelligibility for diverse listener groups. Even less is known about how the listener's own L1 background shapes perception of Thai-accented speech. This gap in the literature is problematic because the communicative success of Thai speakers in ELF contexts depends on how they produce English, but equally on how interlocutors from other linguistic backgrounds process and evaluate that speech. If intelligibility is jointly constructed between speaker and listener (Pickering, 2006), then focusing exclusively on the production side neglects half of the interaction. Previous studies have tended to focus on either production features of Thai English or comparisons between Thai and native English listeners, leaving a significant gap regarding broader cross-L1 comparisons within the Asian context.

The present study addresses this gap by investigating the role of listener L1 background in shaping the intelligibility and comprehensibility of Thai-accented English. The research brings together listener groups from six L1 backgrounds that represent diverse phonological systems and typological characteristics.

- *Thai*, sharing the same L1 as the speakers representing matched ISIB;
- *Burmese* and *Mandarin Chinese*, both tonal languages, differ in structure but share some key features with Thai, such as restricted coda inventories and limited consonant clusters.
- *Japanese* and *Korean*, both non-tonal languages, exhibit strong phonotactic constraints. Japanese lacks most final consonants and consonant clusters, while Korean allows only unreleased stops and nasals in coda position; in this respect, Korean shows a coda restriction similar to Thai.
- *English*, as the target language and the only native L1 in the sample, provides a reference point for how deviations from native norms are perceived. Native English listeners are equipped to recognize all standard phonological contrasts but may be less tolerant of non-native variations.

The study employs both objective measures (transcription accuracy in a sentence-level task) and subjective measures (comprehensibility ratings of unscripted speech). By triangulating these approaches, the research provides a multidimensional account of listener experience and contributes to a fuller understanding of intelligibility in ELF contexts, with implications for language pedagogy and assessment practices in multilingual Asian settings.

Research Questions

The following research questions guide the investigation:

1. How does listener L1 affect the intelligibility and comprehensibility of Thai-accented English?
2. What is the relationship between intelligibility and comprehensibility across different listener L1 groups?

Literature Review

Intelligibility and Comprehensibility in Second Language Speech

A central distinction in the study of L2 speech concerns the concepts of intelligibility and comprehensibility. Intelligibility refers to the degree of success in decoding speech, typically measured through objective transcription tasks where listeners reproduce exactly what they hear (Munro & Derwing, 1995). Comprehensibility, in contrast, captures listeners' subjective perceptions of ease or difficulty in understanding speech, usually assessed through rating scales (Derwing & Munro, 1997). While closely related, the two constructs are not the same. A listener may correctly recognize a word (high intelligibility) while simultaneously perceiving the

speech as effortful to process (low comprehensibility). Conversely, listeners may rate speech as easy to follow due to contextual cues, even if individual word recognition is incomplete.

This conceptual distinction has significant implications for both theory and practice. Intelligibility aligns with the processing outcome of whether communication succeeds at the level of word recognition, while comprehensibility reflects processing effort, indexing cognitive load in real time (Isaacs, 2008).

The Role of Listener Background

A substantial body of research demonstrates that listener characteristics exert a powerful influence on both intelligibility and comprehensibility. Among these characteristics, listener L1 background is especially consequential, as it shapes the perceptual system through phonemic inventories, phonotactic constraints, and prosodic expectations. For example, Thai listeners might struggle with voiced fricatives, while Mandarin listeners find final consonant clusters challenging, often perceiving them inaccurately or failing to recognize them at all (Wei & Zhou, 2002). This means that intelligibility is not simply a property of the speaker's production but a joint construct determined by the alignment or misalignment between speaker and listener repertoires (Pickering, 2006).

Building on these insights, Bent and Bradlow (2003) introduced the concept of the Interlanguage Speech Intelligibility Benefit (ISIB). ISIB posits that listeners understand speech produced by non-native talkers who share their L1 more easily than speech from native talkers. The explanation is that shared L1 backgrounds create overlapping patterns of transfer and error, which render speech more predictable. For instance, Thai listeners may anticipate the substitution of /θ/ with /t/ in Thai-accented English, which could enhance recognition accuracy.

Research has extended ISIB beyond matched contexts to test mismatched conditions, where the listener and talker do not share the same L1. The evidence here is mixed. Some studies report a mismatched benefit, where non-native listeners outperform native listeners in recognizing foreign-accented speech (Hayes-Harb et al., 2008; Stibbard & Lee, 2006). Other studies, however, find no advantage or even a disadvantage for mismatched non-native listeners compared to native listeners (Imai et al., 2005; Major et al., 2002).

These discrepancies can be accounted for within a phonological distance framework. Mismatched ISIB is predicted to emerge when listener and talker L1s share overlapping phonological constraints (e.g., similar coda inventories, comparable syllable templates, or parallel transfer strategies), but

not when they are typologically distant. When a talker's L1-conditioned substitution (e.g., devoicing of final obstruents) corresponds to a constraint that also exists in the listener's L1, the listener can assimilate the non-native variant without perceptual cost, yielding a mismatched benefit. When no such overlap exists, the substitution falls outside the listener's category structure and is more likely to be misidentified.

Phonological Profiles of Listener L1 Groups

A fuller understanding of cross-L1 perception requires consideration of the phonological characteristics of the listener groups included in this study. The segmental inventories and key phonological features of all six languages are summarized in Table 1, which enables direct cross-linguistic comparison of the phonological features most relevant to Thai English intelligibility.

Thai

Thai is a tonal, syllable-timed language belonging to the Tai-Kadai language family. Standard Thai has approximately 21 consonant phonemes and 9-18 vowel phonemes (depending on analysis), including long and short vowel distinctions that are phonemically contrastive (Tingsabadh & Abramson, 1993). Codas are restricted to voiceless unreleased stops /p, t, k/, nasals /m, n, ŋ/, and glides /j, w/, with no voiced obstruents or fricatives permitted in final position. Thai employs a five-tone system serving lexical functions. These constraints produce systematic L1 transfer in English: final consonant devoicing (/b, d, g/ → [p̚, t̚, k̚]), cluster simplification, and substitution of absent consonants such as /v/ → [w], /z/ → [s], /θ/ → [t] (Peerachachayanee, 2022).

Burmese

Burmese, a member of the Tibeto-Burman branch of the Sino-Tibetan family, is a tonal language with a CV(C) syllable structure and syllable-timed rhythm (Okell, 1969). Burmese codas are highly restricted to unreleased voiceless stops (/p, t, k/ often realized as glottal stops) and nasals (/m, n, ŋ/); no consonant clusters occur in native vocabulary, and a three-tone system (high, low, creaky) serves lexical functions. These coda restrictions parallel Thai, creating structural familiarity with Thai-accented English. Burmese English typically exhibits consonant cluster reduction, epenthetic vowel insertion, and substitutions of fricatives with stops (Kirkpatrick, 2010).

Chinese (Mandarin)

Mandarin Chinese is a tonal, syllable-timed language that belongs to the Sinitic branch of the Sino-Tibetan language family. It has approximately 21 consonants and five to six core vowels (Duanmu, 2002). Mandarin prohibits consonant clusters and restricts codas to nasals (/n/, /ŋ/), with no final obstruents permitted. This severely limits exposure to the types of complex syllables found in English. Mandarin uses four lexical tones plus a neutral tone. The absence of final obstruents hinders accurate perception of English coda consonants (Hao, 2012). These constraints contribute to recurring misperceptions when Mandarin listeners process non-native English varieties.

Japanese

Japanese is a non-tonal, mora-timed language from the Japonic family, with a relatively simple phonological system. According to Vance (2008), it possesses 21 consonant phonemes and 5 vowels with phonemic length contrasts. The syllable structure is strictly (C)V, with only moraic nasal /N/ or geminate /Q/ in coda position. No consonant clusters occur in native words; speakers insert epenthetic vowels to break up English clusters. Japanese uses a pitch accent system (not lexical tone) and mora-timed rhythm, promoting systematic vowel epenthesis and coda simplification in English perception and production (Masuda & Arai, 2008). Key interlanguage features include vowel epenthesis, /r/-/l/ neutralization via the alveolar tap [ɾ], and systematic final consonant modification (Nishio & Tsuzuki, 2014).

Korean

Korean is a syllable-timed language belonging to the Koreanic language family. It has 19 consonant phonemes with a three-way laryngeal contrast in stops: tense, lax, and aspirated (e.g., /p/, /p̚/, /pʰ/) (Ahn & Iverson, 2004). According to Sohn (1999), coda consonants are restricted to nasals (/m, n, ŋ/) and unreleased stops (/p̚, t̚, k̚/). No consonant clusters occur in native phonology (Lee & Ramsey, 2001). A key phonological feature is final obstruent neutralization, in which all obstruents in coda position are devoiced and unreleased (Cho & Park, 2006). These L1 constraints contribute to common interlanguage repair strategies in both perception and production, including consonant cluster simplification, vowel epenthesis, and final consonant deletion.

English

As the target language, English represents the baseline with full phonological inventory including all consonant clusters, fricative contrasts, and complex syllable structures. Native English speakers possess complete access to phonemic distinctions and prosodic patterns, though they may lack familiarity with systematic patterns of L2 English variation and may interpret deviations from native norms as errors rather than predictable interlanguage features.

Table 1

Cross-Linguistic Comparison of Phonological Features Relevant to Thai English Intelligibility

	Tone System	Vowel Inventory	Consonant Inventory	Syllable Structure	Rhythm Type
Thai	5 tones: Mid, low, falling, high, rising	9–18 vowels: Distinct phonemic length contrast (short vs. long)	~21 consonants: Lacks aspirated and voiced obstruents in final position; limited clusters	(C)(C)V(V)(C): Clusters restricted to /r, l, w/; codas restricted to unreleased stops, nasals, or glides	Syllable-timed
Burmese	3 tones: High, low, creaky (voice quality linked)	7 oral + 4 nasal: Vowels often realized with nasalization in checked syllables	~34 consonants: Three-way stop contrast (voiced, voiceless, aspirated)	C(G)V(C): Onset clusters allowed (glides /w, y/); final stops debuccalized to /ʔ/; final nasals realized as vowel nasalization	Syllable-timed
Mandarin Chinese	4 lexical tones: + neutral tone	~5–6 core vowels: Rich inventory of diphthongs and triphthongs	~21 consonants: No final obstruents; retroflex series present	(C)(G)V(N): No consonant clusters; codas restricted to nasals /n, ŋ/ (and /ɹ/ in erhua)	Syllable-timed
Japanese	Pitch accent: (High/Low patterns)	5 vowels: Phonemic length contrast (short vs. long is distinct)	~14-21 consonants: Lacks /r/-/l/ contrast; Palatalized onsets common	(C)(j)V(C): Coda restricted to moraic nasal /N/ or geminate obstruent /Q/	Mora-timed

Korean	None: (Intonational pitch only)	8–10 vowels: Length contrast largely lost in modern Seoul dialect	~19 consonants: Three-way laryngeal contrast (plain, aspirated, tense); final obstruents unreleased	(C)(G)V(C): No onset clusters; complex coda clusters exist orthographically but simplify in pronunciation (C)(C)(C)V(C)(C) (C): Extensive complex clusters allowed in both onset and coda	Syllable- timed
English	None: (Stress & Intonation)	~12–20 vowels: High dialectal variation (monophthons & diphthongs)	~24 consonants: Large set of fricatives (/f, v, θ, ð, s, z, ʃ, ʒ, h/) and affricates (/tʃ, dʒ/); /r/ and /l/ are distinct		Stress- timed

Methodology

This study investigated how listeners from different first language (L1) backgrounds perceive Thai-accented English in terms of intelligibility and comprehensibility. The methodology involved careful selection of speakers and listeners, controlled elicitation of speech samples, and a combination of objective and subjective perception tasks.

Speakers

Speech samples were produced by Thai speakers of English as a second language (L2). To represent typical Thai-accented English, speakers were excluded if they had lived in English-speaking countries for over three months or held degrees in English-related fields. Eligibility was confirmed through a background questionnaire collecting demographic data, English proficiency, and exposure to English in daily life.

Ten speakers (five males, five females) recorded five scripted sentences designed to elicit common Thai-accented features. Three raters, including a native English speaker, a non-Thai non-native English speaker, and a Thai English speaker, assessed the accentedness of each speaker using a 7-point Likert scale. Based on these ratings, six speakers (three males and three females) representing a range of accent strengths were selected. Gender balance was maintained to ensure ecological validity, although it was not hypothesized to affect outcomes.

Listeners

Thirty listeners were recruited from six L1 backgrounds: English, Japanese, Korean, Mandarin Chinese, Burmese, and Thai. Each group included five participants, ensuring balanced representation of both native and non-native English speakers from tonal and non-tonal language backgrounds.

Eligibility required at least intermediate English proficiency (B1 CEFR level), determined via self-report. Listeners were excluded if they had significant exposure to Thai-accented English e.g., living in Thailand for over a year, studying Thai formally, having close Thai contacts, or working as English teachers in Thailand. A background questionnaire collected demographic data and prior exposure details.

Speech Materials

Two types of materials were used: scripted and unscripted speech. Scripted sentences highlighted key pronunciation features and formed the basis of the intelligibility task. Unscripted utterances were elicited through prompts on familiar topics (e.g., hobbies) to capture more natural speech and were used in the comprehensibility task. This dual approach allowed the study to measure both word recognition accuracy and ease of understanding.

Tasks

Intelligibility was assessed through a sentence transcription task. Listeners heard each scripted sentence (with a maximum of two plays) and transcribed what they heard using standard English orthography. This provided an objective measure based on correct word recognition.

Comprehensibility was assessed via listener ratings of the unscripted speech on a 7-point Likert scale (1 = “*very difficult to understand*”; 7 = “*very easy to understand*”). This reflected perceived processing ease rather than accuracy.

Data Analysis

Transcription responses were scored binarily: 1 point for correct words (allowing minor spelling variants), and 0 for incorrect words or misperceptions. Scores were calculated as percentages of correctly transcribed words per listener. All vocabulary items were drawn from high-frequency word lists to minimize lexical unfamiliarity.

Comprehensibility scores were treated as continuous variables, with mean ratings computed per utterance. Descriptive statistics (means, SDs) were calculated for each L1 group and the overall sample.

Comparative analysis focused on group differences across the six L1s. Although not the primary aim, the inclusion of tonal and non-tonal language groups enabled exploratory investigation of ISIB-related effects. Statistical tests assessed the significance of any group differences in perception scores.

Results

Intelligibility Performance Across Listener Groups

This section examines intelligibility outcomes for Thai-accented English across six listener groups: Thai, English, Burmese, Mandarin Chinese, Japanese, and Korean. Intelligibility was measured through a sentence-level transcription task in which participants transcribed target words spoken by Thai L1 speakers of English. Scores represent the number of correctly transcribed words out of 33 possible targets. Table 2 presents mean scores, accuracy percentages, and score ranges for each group.

Table 2

Intelligibility Scores by L1 Group

L1 Group	Mean Score (out of 33)	Percentage Accuracy	Score Range
Thai	24.20	73.30%	22 - 25
English	18.80	56.97%	15 - 22
Burmese	18.00	54.55%	16 - 19
Chinese (Mandarin)	16.60	50.30%	10 - 19
Japanese	14.40	43.64%	12 - 18
Korean	13.80	41.82%	12 - 17

The results demonstrate a clear intelligibility hierarchy across L1 groups. Thai listeners, sharing the same L1 as the speakers, achieved substantially higher accuracy than all other groups. English and Burmese listeners performed at similar levels in the second tier, while Mandarin Chinese listeners occupied an intermediate position. Japanese and Korean participants demonstrated the lowest intelligibility scores. The subsections below examine each group's performance, focusing on how specific L1 backgrounds influenced listener outcomes.

Thai L1 Listeners

Thai listeners achieved the highest performance with a mean score of 24.20 out of 33 (73.3% accuracy) and showed strong consistency, with scores ranging from 22 to 25. While other groups exhibited score ranges extending up to nine points, Thai listeners clustered within a narrow three-point spread. This consistency indicates that their advantage stemmed from deep familiarity with Thai-accented English phonology rather than individual proficiency differences.

Several pronunciation features demonstrate this phonological alignment. Thai English frequently exhibits final devoicing of obstruents (e.g., *bag* realized as [bæk]) and substitution of affricates (/tʃ/ → [tɕ]). While such deviations can compromise intelligibility for unfamiliar listeners, Thai participants interpreted them effortlessly. For instance, the word *chair*, commonly produced with a palatalized onset by Thai speakers, was consistently identified correctly by all Thai listeners. Similarly, the word *nature*, typically realized in Thai English as [neɪtɕəː], contains a second-syllable affricate that is unaspirated ([tɕ]). This reflects a two-stage L1 transfer process in which Thai /tɕʰ/ serves as the default mapping for English /tʃ/, but aspiration is often neutralized in non-initial position, yielding [tɕ]. The substitution results from the interaction of perceptual assimilation and positional deaspiration, and is perceived as phonetically natural by Thai listeners. Their familiarity with unaspirated affricates enables them to interpret the Thai-accented *nature* with ease, despite potential confusion for listeners from other L1 backgrounds.

These findings confirm that Thai listeners' advantage derived not from general English proficiency but from their ability to map expected phonological deviations onto intended English lexical targets. The data highlight the central role of phonological familiarity in supporting intelligibility, even when segmental accuracy deviates from native English norms.

English L1 Listeners

English listeners ranked second overall with a mean score of 18.80 (56.97%), yet their broader score range (15 - 22) suggests considerable individual variation in processing Thai-accented input. While their native English competence provided strong lexical and phonological foundations, these listeners lacked familiarity with Thai phonological transfer patterns, resulting in errors when segmental deviations exceeded familiar accent variation boundaries.

Several pronunciation features illustrate this perceptual mismatch. A common issue involved the substitution of voiced obstruents with voiceless counterparts. For example, the word *zip* was frequently misheard as *sip*, *steep*, or *seek*, which reflects the absence of the voiced fricative /z/ in Thai phonology. Thai speakers often substitute /z/ with /s/, producing *zip* as [sɪp], which native English listeners interpreted as a different lexical item. In this case, their reliance on voicing as a contrastive cue which is absent in the input led to incorrect lexical mapping.

A similar issue arose with the word *rock*, which was misperceived as *lock*. This misperception reflects the neutralization of the onset /r/-/l/ distinction in Thai English, where Thai speakers often merge /r/ with [l] in colloquial speech. While Thai listeners could interpret this variation based on familiarity, English L1 listeners were more likely to misidentify the intended word, particularly when combined with limited context or ambiguous preceding verbs (e.g., “*found that rock*” misheard as “*found that lock*”).

These findings demonstrate that English L1 listeners’ performance was limited by their expectations of native segmental norms. In the absence of phonological familiarity with Thai English, they struggled to interpret predictably altered features.

Burmese L1 Listeners

Burmese listeners achieved a mean score of 18.00 (54.55%), performing similarly to English native speakers, but demonstrated notably less individual variation with scores ranging from 16 to 19. This uniformly strong performance likely reflects a combination of phonological alignment between Burmese and Thai languages and significant regional exposure to Thai-accented English varieties.

Phonological similarities between Burmese and Thai could facilitate intelligibility. Both languages impose restrictions on syllable codas and prohibit complex consonant clusters in word-final positions. Specifically, Burmese permits only nasal consonants and glottal stops in final position (Ankura, 2021), while Thai allows nasal consonants and unreleased voiceless stops (Slayden, 2009). Although the phonetic realization of the stops differs (glottal vs. unreleased oral), the structural constraint, prohibiting complex clusters and open fricatives in the coda, creates a phonotactic familiarity. These structural constraints mean that when Thai speakers simplify or omit final consonant clusters in English words, Burmese listeners encounter phonological patterns that align similarly with their own language’s syllable structure preferences. However, intelligibility was still compromised in several instances, particularly when segmental deviations conflicted with phonemic categories or contrasts in Burmese.

Two examples highlight how voiceless dental fricative /θ/, absent in both Thai and Burmese phonologies, posed consistent challenges. First, in the case of *theme*, while only one listener successfully identified the phrase (“*This theme fit well*”), others produced responses such as “*This team fix well*,” “*Just thin, fit well*,” or “*It been fixed well*.” These variations suggest that Thai speakers’ substitution of /θ/ with /t/ (e.g., *theme* pronounced as [ti:m]) led to confusion for Burmese listeners, who similarly lack /θ/ in their native phoneme inventory and did not recognize this substitution. As a result, they interpreted the initial /θ/ as a more familiar onset, resulting in perceptual drift toward *team*, *thin*, or other plausible alternatives. Similarly, the word *bath* showed a parallel pattern with /θ/ in final position. Burmese listeners often misheard it as *bus*, *bat*, *box*, or even *bird*, which indicates difficulty processing the final interdental fricative.

These findings indicate that Burmese listeners may benefit from structural phonological similarities with Thai; however, the absence of shared segmental categories remains a significant barrier to intelligibility. Their relatively consistent performance suggests a stable perceptual strategy informed not only by phonological alignment but also by increased exposure to Thai-accented English. Geographical and cultural proximity to Thailand likely enhanced this exposure, allowing Burmese listeners to develop familiarity with regional English varieties despite limited formal instruction. In addition, other extralinguistic variables such as regional proximity, cross-border communication, and informal immersion through Thai media may have contributed to this performance independently of purely phonological factors. However, this familiarity alone is insufficient when specific phonemic mismatches persist between the speaker’s and listener’s L1s.

Mandarin Chinese L1 Listeners

Mandarin Chinese listeners obtained a mean score of 16.60 (50.30%) accompanied by the widest score range of all groups (10 - 19 points). This substantial individual variability indicates marked differences in participants’ ability to process Thai-accented English.

One difficulty involved voicing contrasts in obstruents, particularly in word-initial position. In the case of *zip*, many listeners failed to identify the intended word and instead produced alternatives such as *sit*, *say*, *save*, *sleep*, or even *receive*. These responses indicate a consistent confusion between the voiced fricative /z/ and its voiceless counterpart /s/, a distinction that is non-phonemic in Mandarin. Thai-accented English, which often replaces /z/ with /s/, likely presented the word *zip* as [sɪp], which Mandarin Chinese listeners then interpreted through the lens of their L1 phoneme inventory, leading to substitutions with /s/-initial words such as *sit* or *say*. Without exposure to this

patterned substitution, Mandarin Chinese listeners were unable to map the intended lexical target accurately.

The word *kid* showed another problem, especially due to the final voiced stop /d/, which is also not permitted in Mandarin coda position. Responses like *key* and *tea* suggest that listeners either missed the final consonant or substituted it with more familiar structures. Thai English typically neutralizes final voiced stops, so *kid* was produced as [kɪt̚] which lacks a clear final voiced closure. As a result, Mandarin Chinese listeners tended to interpret the input as phonetically similar but semantically unrelated words. These substitutions indicate a breakdown in segmental mapping due to the absence of direct phonemic equivalence and unfamiliarity with the typical patterns of Thai-accented English.

However, the exceptionally wide score range suggests that some Mandarin Chinese participants were able to compensate for phonological mismatches through higher levels of English proficiency or greater exposure to diverse English accents in international contexts. In contrast, others struggled significantly to map Thai-accented English forms onto their intended lexical targets. The high variability within this group illustrates that listener-level factors, such as exposure to diverse accents and individual adaptability, can moderate the effects of L1 phonological constraints.

Japanese L1 Listeners

Japanese listeners recorded a mean score of 14.40 (43.64%) with scores ranging from 12 to 18, placing them among the lowest-performing groups. This relatively consistent low performance across participants reflects phonological incompatibilities between Japanese and English sound systems.

Japanese phonology diverges fundamentally from English in ways that correspond directly to observed transcription error patterns. These patterns align with known phonotactic constraints in Japanese, where consonant inventories are more restricted and syllable structure is predominantly CV (Tsujiura, 2007), limiting exposure to English word-final segments or clusters. Japanese listeners often reinterpreted Thai-accented English forms through the lens of their L1 phonology, which led to segmental substitutions or omissions.

A clear example of this is seen with the word *half*. While some listeners successfully identified the phrase “*Only half is okay*,” others produced responses such as “*Only have it ok*” or “*Only has okay*.” In Thai English, *half* was realized as [ha:p̚], which deviates from English targets. Japanese listeners who were unfamiliar with these segmental variations and lacked a strong representation of word-final fricatives, appeared to mishear the word as morphologically similar or functionally plausible alternatives such as *have* or

bas. These substitutions indicate that Japanese listeners actively assimilated the final voiceless fricative /f/ to the nearest available L1 segment, preserving manner or voicing features (e.g., *have* via voiced counterpart; *bas* via shared manner/voicing), rather than resorting to deletion.

The word *beach* also revealed segmental confusion, with some Japanese listeners correctly recognizing the target word, while others produced responses of “*bee*”. In Thai-accented English, *beach* is often realized as [bi:tɕ] or [bi:t̚], reflecting either the substitution of the English affricate /tʃ/ with the Thai palatal affricate [tɕ], or the reduction of the final affricate to a plain voiceless stop. Japanese phonotactics strictly prohibit word-final obstruents (except for the geminate /Q/), and borrowed words typically end in a vowel through epenthesis (e.g., *beach* → *biichi*). Compounding this, Japanese phonotactics strongly favor open CV syllables, and listeners may have employed a coda deletion strategy, stripping the final affricate to conform to familiar syllable templates. Thus, they failed to recover the final consonant. The misperceptions observed here point to a broader issue of segmental filtering shaped by L1 phonological constraints and unfamiliarity with substitutions found in Thai-accented English.

Despite these substantial challenges, Japanese participants successfully identified high-frequency, phonetically transparent items such as *far* and *city*, suggesting that familiar vocabulary with clear phonological mapping could overcome structural barriers. However, their consistently low overall performance reflects both the considerable phonological distance between Japanese and English systems and their limited exposure to Thai-accented English varieties in their linguistic environment.

Korean L1 Listeners

Korean listeners achieved the lowest scores with a mean of 13.80 (41.82%) and a narrow range of 12–17, which indicates uniformly poor performance across the group. While Korean phonology permits more consonantal contrasts than Japanese, final position restrictions are notable, with only a limited set of consonants allowed (Kang, 2003).

In the case of *ball*, Korean L1 listeners produced a variety of responses including *bone*, *bar*, and *bun*, with only one correctly identifying the intended word. Korean phonology does in fact contrast final /l/ and /n/ (e.g., *kal* ‘knife’ vs. *kan* ‘liver’). The listeners accurately perceived the phonetic input [n] provided by the Thai speaker and mapped it faithfully to lexical items ending in nasals (*bone*, *bun*). The perceptual shift also involved the vowel dimension: the target vowel /ɔ/ in *ball* was reinterpreted as /o:/ in *bone* or /ɑ/ in *bun*, indicating that both segmental dimensions (coda consonant and vowel quality) contributed to the misperception. Here, the listeners’

phonological competence in distinguishing final liquids from nasals actually hindered intelligibility, as it forced them to commit to the speaker's error rather than reconstructing the intended target /l/.

The word *guard* also proved challenging, with responses including *God*, *dot*, *cap*. Korean phonology lacks a voicing distinction in word-initial obstruents and does not permit voiced obstruents like /d/ or /g/ in final position. Thai speakers produced *guard* as [ka:t̚], neutralizing both voicing and final obstruent aspiration. Korean listeners, unfamiliar with these realizations and without a robust /d/–/t/ contrast in final position, appeared to interpret the word through the lens of semantically and phonetically proximate alternatives like *God* or *dot*. The consistent confusion of final /d/ with /t/, and of initial /g/ with /k/, highlight the role of both L1 constraints and the accented input in shaping intelligibility.

Overall, the consistently low performance across all Korean participants indicates both significant phonological misalignment between Korean and English sound systems and limited exposure to Thai-accented input varieties.

Table 3 organizes Thai English deviations into two categories: (i) those that impaired intelligibility and (ii) those that did not. Under Group (i), two to three representative misperceptions per listener L1 group illustrate how each mistranscription traces to that group's phonological transfer patterns. Under Group (ii), words that remained largely intelligible appear alongside the reasons the deviations did not interfere with recognition.

Table 3

L1-Specific Transfer Patterns and Their Impact on Intelligibility

Listener L1	Target	Transcriptions	L1 Transfer Explanation	Acc.
(i) Deviations Affecting Intelligibility				
Burmese	bath	<i>bus, bat, bird, box</i>	No /θ/ in Burmese; listeners substituted stops /t/ or /s/. Coda inventory (nasals and glottal stop only) blocked recovery of the final fricative.	17%
	rock	<i>rod, block, lock</i>	Thai /r/–/l/ merger yielded [ɔk]. Limited Burmese coda options added further onset-level confusion.	47%
	bag	<i>back, bank</i>	Thai devoiced /g/ to [k]. Since Burmese does not permit oral obstruents in coda position, the voiced target was unrecoverable.	7%
Mandarin Chinese	guard	<i>car, god, guy</i>	Thai speakers devoiced both onset /g/ → [k] and final /d/. Mandarin allows only	23%

			nasal codas (/n, ŋ/), leaving listeners unable to recover the obstruent endings.	
	kid	<i>key, tea</i>	Thai realized final /d/ as [t̚]. Without final stops in Mandarin, listeners tended to delete the coda altogether.	53%
	zip	<i>sit, receive, sleep</i>	Thai substituted /z/ with [s]. Mandarin does not distinguish /z/ from /s/ by voicing, removing the cue needed to identify the intended word.	13%
Japanese	beach	<i>bee</i>	Thai speakers realized beach as [bi:t̚]. The strong CV syllable template in Japanese promoted coda deletion, so the final consonant was lost.	53%
	half	<i>have, has</i>	Word-final /f/ does not exist in Japanese. Listeners remapped it through voicing (/f/ → /v/ → have) or place shift (/f/ → /s/ → has).	77%
	bend	<i>bin, brain, ben</i>	Thai reduced the /nd/ cluster to [n]. Japanese allows only moraic /N/ in coda position, making the full cluster unrecoverable.	13%
Korean	ball	<i>bone, bun, bar</i>	Thai realized /l/ as [n]. Because Korean maintains a final /l/-/n/ contrast, listeners mapped the [n] input directly to nasal-final words. The vowel was also reinterpreted.	53%
	bath	<i>bat, bag, bed</i>	Korean has no /θ/. L1 coda neutralization rules produced diverse final-consonant interpretations across listeners.	17%
	guard	<i>God, god, dot, cap</i>	Korean uses a three-way laryngeal contrast rather than a voicing distinction in onset position; this, together with final neutralization, disrupted recognition.	23%
English	zip	<i>sip, seek</i>	Thai replaced /z/ with [s]. Native English listeners depend heavily on voicing; without it, they selected the wrong lexical item.	13%
	garage	<i>car, carriage, garden</i>	Shifted stress placement combined with /dʒ/ → [tɕ] pushed the word beyond native listeners' accent-tolerance threshold.	57%
	tent	<i>thing</i>	The unaspirated Thai [t] in tent was heard as /θ/ by English listeners, who lacked familiarity with Thai aspiration patterns and perceived <i>thing</i> instead.	20%

Thai	vest	<i>vest, word</i>	/v/ is absent from Thai; all speakers realized it as [w]. Shared-L1 familiarity (ISIB) could not compensate for full phonemic neutralization.	0%
	bag	<i>back</i>	Devoicing of final /g/ to [k̚] follows Thai phonotactic rules restricting voiced obstruents in coda position.	7%
	cub	<i>cup, cap</i>	Devoicing of final /b/ to [p̚] likewise reflects the Thai prohibition on voiced stops in coda position.	3%
(ii) Deviations Not Affecting Intelligibility				
All groups	shop	<i>97% overall</i>	[tʰ] is perceptually similar to [ʃ]; high word frequency aided recognition.	97%
	city	<i>87% overall</i>	Unaspirated [t] preserved the syllable template; high frequency supported recognition.	87%
	there	<i>87% overall</i>	/ð/ → [d]; word frequency and contextual cues offset the substitution.	87%
	far	<i>87% overall</i>	Non-rhotic realization is widespread across English varieties; no critical consonant contrast was affected.	87%
	land	<i>87% overall</i>	[nd] simplified to [n]; no competing lexical item was activated.	87%
	gift	<i>83% overall</i>	[ft] reduced to [f], yet the core CVC shape was preserved; accuracy ranged 60–100% across groups.	83%
	kiwi	<i>83% overall</i>	Unaspirated /k/ did not trigger any competing lexical candidate.	83%
	fail	<i>77% overall</i>	Final /l/ realized as [w]; no lexical competitor was activated by the change.	77%

Note. Acc. = overall recognition accuracy across all 30 listeners. Transcription examples are representative responses selected from each L1 group.

The table reveals three broad patterns. When a phonemic contrast is fully neutralized, accuracy drops to near-floor levels irrespective of listener L1: *vest* (0%, /v/ entirely absent from Thai phonology) and *bag* and *cub* (7% and 3%, reflecting final devoicing of /g/ → [k̚] and /b/ → [p̚]) show that shared-L1 familiarity cannot offset the loss of a target contrast. At the same time, a single Thai deviation can produce divergent misperceptions depending on the listener's L1: *bath* (/θ/ → [t]) was heard as *bus* by Burmese listeners, whose coda constraints blocked fricative recovery, but as *bat* or *bag* by Korean listeners, whose own neutralization patterns channelled perception differently. Likewise, *guard* yielded *car/god/guy* from Mandarin listeners, *god/dot/cap* from Korean listeners, and *rod/block/lock* from Burmese listeners,

each group filtering the same [k]-initial, coda-deleted output through its own phonological system. Group (ii), by contrast, shows that high-frequency words with intact syllable shapes remain broadly intelligible even with non-native segmental realization: *shop*, *there*, *far*, and *land* all exceeded 87% accuracy, indicating that lexical frequency and preserved syllable structure can absorb minor substitutions such as /ð/ → [d] or cluster simplification when no plausible lexical competitor exists.

Statistical Analysis

The intelligibility patterns were confirmed by statistical analysis. A one-way ANOVA assessed the effect of listener L1 background on intelligibility scores, using correctly transcribed words (out of 33) as the dependent variable. The analysis revealed a significant main effect of listener L1 group, $F(5, 24) = 10.73$, $p < .001$, with a large effect size ($\eta^2 = .691$), which indicates that approximately 69% of variance in intelligibility scores was attributable to differences in listeners' native languages.

Post hoc comparisons using Tukey's HSD test clarified these differences. Thai listeners significantly outperformed all other groups, with large effect sizes ranging from $d = 1.13$ (vs. English) to $d = 2.18$ (vs. Korean). All pairwise comparisons between Thai listeners and other groups reached significance ($p < .05$ or lower), supporting a clear Interlanguage Speech Intelligibility Benefit (ISIB) driven by phonological familiarity and shared L1 features.

Additional comparisons among non-Thai groups explored potential advantages of tonal over non-tonal backgrounds. Burmese listeners performed significantly better than Korean listeners ($p = .037$, $d = 0.88$) and marginally better than Japanese listeners ($p = .054$, $d = 0.75$). Other comparisons (e.g., Mandarin Chinese vs. Japanese or Korean) did not reach significance, which indicates that shared typological features such as tonal systems alone do not fully explain performance differences. Phonological overlap in coda restrictions, exposure to Thai-accented English through regional contact, and educational background may play more critical roles than tonal background per se. Overall, the statistical findings confirm that intelligibility is primarily shaped by the degree of phonological alignment between speaker and listener L1s, with Thai listeners' advantage representing a clear case of the Interlanguage Speech Intelligibility Benefit.

Comprehensibility Evaluations of Thai-Accented English

This section examines how listeners from six L1 backgrounds evaluated the comprehensibility of Thai-accented English. Comprehensibility

was assessed using a 7-point Likert scale (1 = very difficult to understand, 7 = very easy to understand), with each participant rating three unscripted speech samples. The results reveal significant patterns of convergence and divergence across listener groups, which reflect the complex interplay of linguistic proximity, perceptual experience, and attitudes toward accented English varieties.

Overall Comprehensibility Patterns

Table 4 presents mean comprehensibility ratings by L1 group. The results indicate generally favorable perceptions of Thai-accented English, with most groups rating above the scale midpoint. However, meaningful variation exists across groups, ranging from Thai listeners' highest rating (6.600) to English listeners' lowest rating (4.467).

Table 4

Mean Comprehensibility Ratings by L1 Background

L1 Group	Mean Comprehensibility
Thai	6.600
Burmese	6.333
Chinese (Mandarin)	6.067
Japanese	5.600
Korean	5.467
English	4.467
Overall Mean	5.760

The comprehensibility ratings mirror the intelligibility hierarchy closely, which reinforces the view that the two measures, though conceptually distinct, track the same underlying variable of phonological alignment in this dataset. Thai listeners' ratings ($M = 6.600$) were the highest, consistent with the ISIB, since listeners who share the speaker's L1 tend to process phonological deviations not as unexpected noise but as predictable interlanguage features. Processing load is reduced not because the speech is acoustically clearer, but because Thai listeners possess an implicit "error grammar" for Thai-accented English that enables anticipatory perception. A deviation such as final devoicing of /g/ to [k] does not require effortful recovery; it is already part of the listener's perceptual expectation, and the lexical candidate is activated with minimal competition.

Burmese and Mandarin Chinese listeners also rated Thai-accented English highly, suggesting that both regional proximity and tonal language experience may enhance comprehensibility. Burmese listeners likely benefited

from exposure to Thai English through cross-border interactions, tourism, and media. Similarly, Mandarin Chinese listeners may have been aided by their shared structural features with Thai, such as restricted coda inventories and syllable-timed rhythm, which could facilitate processing of Thai-accented English. Their relatively high ratings, despite moderate intelligibility performance, highlight that comprehensibility and intelligibility are distinct constructs. Even when word recognition is imperfect, listeners may still process meaning with ease.

Japanese and Korean listeners provided moderately positive ratings, still above the neutral midpoint, though lower than other L2 groups. These results indicate a greater perceptual burden, likely due to greater phonological distance and less exposure to Thai-accented English.

In contrast, English L1 listeners reported the lowest comprehensibility ratings, despite relatively strong intelligibility scores. This apparent paradox, with accurate word recognition paired with low processing ease, underscores the influence of native speaker expectations. Native listeners may be more attuned to deviations from standard English norms, such as vowel shifts, cluster reduction, or suprasegmental variation, which can increase perceived difficulty even when meaning is clear. Additionally, limited exposure to diverse L2 Englishes may lower tolerance for accent variation, which reinforces bias toward standard forms. These results reaffirm that comprehensibility is a subjective judgment shaped by listener expectations, familiarity, and sociolinguistic attitudes as much as by the sound signal itself. The data also suggest a cumulative effect. When stress deviations, consonant substitutions, and vowel shifts co-occur in the same utterance, they may compound to lower comprehensibility ratings even when listeners can still decode individual words successfully. This cumulative burden likely explains why comprehensibility ratings can diverge from intelligibility scores, particularly for listener groups with limited exposure to Thai-accented English.

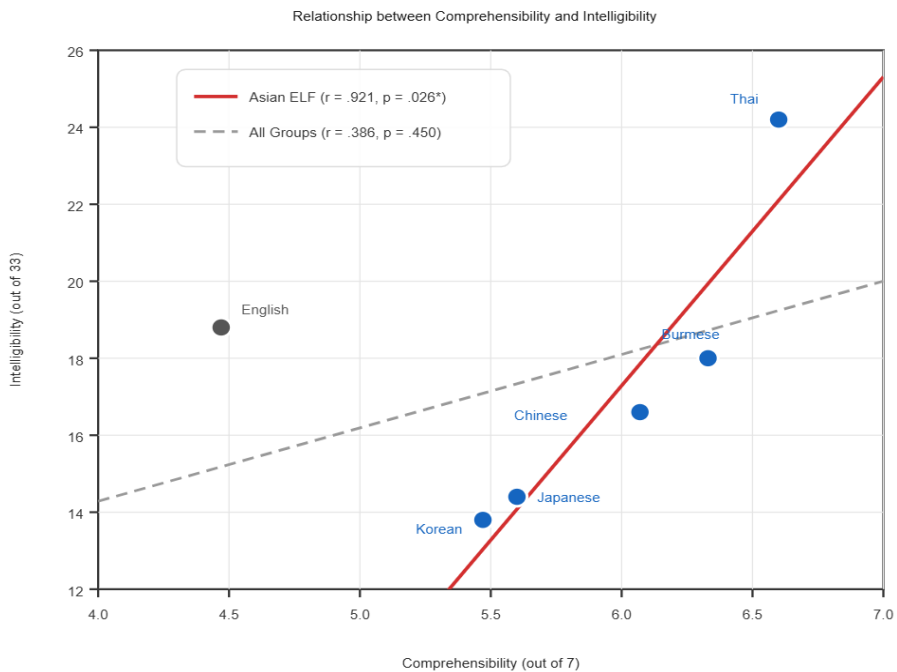
The divergence between English listeners' high intelligibility but low comprehensibility suggests that native speakers may experience greater cognitive effort even when achieving accurate word recognition, possibly due to increased sensitivity to accent features.

Relationship Between Comprehensibility and Intelligibility

To examine the relationship between comprehensibility and intelligibility, a Pearson correlation analysis was conducted using mean scores for each L1 group.

Figure 1

Intelligibility and Comprehensibility Relationship



A Pearson correlation coefficient was computed to assess the relationship between comprehensibility and intelligibility across all listener groups. Initial analysis including all six L1 groups yielded a moderate, though non-significant, positive relationship ($r = .386$, $p = .450$). However, closer inspection of the data distribution revealed that the native English speaker group functioned as a significant outlier, exhibiting a distinct dissociation pattern: they reported the lowest comprehensibility (4.467) yet achieved a moderate-to-high intelligibility score (18.80). This suggests that while native speakers found the Thai-accented English difficult to process (low comprehensibility), their familiarity with the target language's phonology allowed them to reconstruct the intended words with relatively high accuracy.

When the English group was excluded to focus on the performance of L2 (non-native) and L1 (Thai) listeners, the relationship strengthened

considerably. For these five groups, the Pearson correlation coefficient ($r = .921$, $p = .009$) indicates a very strong positive linear relationship between the two constructs. This strength falls into the “very strong” category according to conventional interpretation guidelines, suggesting that for these listeners, the two measures are intrinsically linked despite capturing different facets of speech processing.

Further analysis through the coefficient of determination ($r^2 = .848$) reveals that approximately 84.8% of the variance observed in intelligibility scores can be statistically explained by variations in comprehensibility ratings. In practical terms, listener groups who reported greater ease in processing Thai-accented English also demonstrated significantly higher success rates in accurately identifying words. While Thai listeners anchored the high end of this spectrum (6.600/24.20), the overall trend confirms that as perceived ease of understanding increases, so does objective word recognition.

The strength of this correlation suggests that comprehensibility ratings serve as reliable predictors of actual word recognition performance across diverse L1 backgrounds, despite the different cognitive processes underlying each measure.

Discussion

The results converge on three theoretical claims that extend and refine current models of ELF intelligibility. The large and significant effect of listener L1 on intelligibility ($F(5, 24) = 10.73$, $p < .001$, $\eta^2 = .691$) confirms that phonological background is a primary determinant of cross-linguistic speech perception, not simply one variable among many. The three subsections below interpret this central finding in relation to phonological distance theory, the ISIB framework, and the implications of the native listener paradox for assessment in ELF contexts.

Phonological Influence on Intelligibility and Comprehensibility

The results provide strong evidence for the influence of L1 phonological structure on both intelligibility and comprehensibility outcomes. The large effect size ($\eta^2 = .691$) indicates that L1 differences account for approximately 69% of the variance in intelligibility performance, which demonstrates that phonological background is a primary determinant of cross-linguistic speech perception.

- Thai listeners, who share the same phonological system as the speakers, demonstrated significantly higher intelligibility and comprehensibility (Cohen's $d = 1.13$ - 2.18 compared to other groups). This exemplifies perfect phonological alignment, where systematic L1 transfer patterns

(final devoicing, cluster simplification, consonant substitutions) are predictable rather than deviant. This predictability enables effortless lexical mapping despite surface-level deviation from native English norms.

- Burmese and Mandarin Chinese listeners, while not sharing the same L1, exhibited moderate performance levels, outperforming listener groups from typologically more distant language backgrounds. Though inferential statistics did not consistently reach conventional significance levels, descriptive statistics reveal a noteworthy trend suggesting both groups performed more successfully than listeners from non-tonal, non-Southeast Asian backgrounds. Burmese listeners significantly outperformed Korean participants ($d = 0.88$, $p = .037$), and Mandarin Chinese listeners, despite greater variability, maintained relatively high comprehensibility ratings. These patterns may reflect partial phonological compatibility: both Burmese and Mandarin Chinese share certain structural features with Thai, specifically syllable-timed rhythm and restricted coda inventories, which may ease perceptual mapping. Burmese listeners, in particular, achieved intelligibility scores nearly identical to native English listeners (18.00 vs. 18.80), suggesting regional familiarity may approximate native speaker advantage in processing L2 English varieties. Geographical and cultural proximity between Myanmar and Thailand likely provides sustained exposure to Thai-accented English through cross-border trade, tourism, and media, allowing Burmese listeners to develop implicit knowledge of Thai English phonological transfer patterns. This combination of structural phonological compatibility and extensive regional exposure demonstrates that accent familiarity and phonological alignment can be as influential as native linguistic competence in determining intelligibility outcomes. Additionally, familiarity with regional English varieties in multilingual contexts may further support successful meaning extraction, pointing to the influence of both phonological proximity and ELF-related listening experience in shaping comprehension.
- Japanese and Korean listeners, whose L1 phonologies differ from both Thai and English in key respects (though Korean shares some features with Thai such as unaspirated stops and coda neutralization), struggled in both tasks with no significant differences between them ($d = 0.13$, $p = .930$). This suggests a threshold effect where phonological distance beyond a certain point results in pervasive recognition failure. Their L1 constraints (obligatory open syllables in Japanese, restricted final consonant inventories in Korean) create fundamental incompatibilities with both Thai interlanguage patterns and English phonological structures, which led to severe recognition failures.

Interlanguage Speech Intelligibility Benefit (ISIB)

The results strongly support the ISIB hypothesis (Bent & Bradlow, 2003), but only in the matched-L1 condition. Thai listeners who shared L1 with the speakers consistently outperformed all other groups, not only in accuracy but also in perceptual consistency. Regarding this, Thai listeners appear to access a shared phonological grammar that renders Thai English interlanguage features transparent.

The mismatched-L1 conditions revealed a more complex picture. While Burmese listeners showed some advantages over more distant L1 groups, these benefits were inconsistent and much smaller in magnitude than matched-L1 effects. Mandarin Chinese listeners' moderate intelligibility but high comprehensibility ratings suggest compensatory processing strategies where global meaning extraction compensates for segmental recognition difficulties. This also suggests a potential phonological proximity benefit, where shared typological traits (e.g., tonal systems, coda limitations) approximate the perceptual advantage of a shared L1. Similar claims have been made in prior studies (Hayes-Harb et al., 2008; Stibbard & Lee, 2006) which found intelligibility benefits in typologically related L2-L2 interactions.

By contrast, Japanese and Korean listeners' consistently low scores indicate that phonological distance was sufficiently large to negate any mismatched benefit. Thai-accented English features that reflect Thai phonotactic constraints such as devoiced final obstruents, absent /z/, /v/, and /θ/, and reduced consonant clusters fall outside the assimilation categories available to Japanese and Korean listeners. Rather than facilitating recognition through shared transfer predictions, these features introduce an additional layer of perceptual opacity: the listener must first decode a non-native substitution without the advantage of shared L1 expectations, then map it onto a lexical target, and do so without the scaffolding that phonological proximity would otherwise provide. The resulting double filtering once through the speaker's interlanguage and again through the listener's L1 phoneme categories accounts for the near-floor intelligibility observed in these two groups. Taken together, the three-tier pattern of intelligibility (Thai > Burmese/Mandarin > Japanese/Korean) constitutes the principal theoretical contribution of this study: it demonstrates that ISIB is not a binary matched-versus-unmatched distinction but a gradient function of inter-L1 phonological proximity, with regional exposure and shared typological constraints acting as additional moderating variables.

The Native Listener Paradox: Challenging ELF Assessment Paradigms

The native English listeners' performance reveals a fundamental paradox in ELF intelligibility research that has profound implications for language assessment and pedagogy. Despite achieving relatively high intelligibility scores (18.80/33, 56.97%), native listeners provided the lowest comprehensibility ratings (4.467/7). This indicates that accurate word recognition does not necessarily correlate with subjective ease of understanding, especially when the speech departs from native phonological norms.

Several factors likely contribute to this dissociation between objective accuracy and subjective difficulty. Native speakers possess highly automatized phonological representations calibrated primarily to native or near-native input. When confronted with L2 English that systematically deviates from these representations through features such as final devoicing, consonant substitutions, or non-native prosody, they may experience increased cognitive load even when successfully recovering lexical targets. Native listeners also approach speech perception with strong expectations about phonetic realization, stress patterns, and prosodic contours. When Thai-accented English violates these expectations, for example, through unexpected stress placement in words like *garage*, they may experience momentary processing disruption. Although lexical access often succeeds, the accumulated effect of repeated expectation violations may contribute to perceived difficulty. Many native speakers, particularly those in monolingual environments, may have limited experience with L2 varieties that deviate substantially from native norms. Without a strong perceptual framework for interpreting predictable transfer patterns, they may treat each deviation as an individual error requiring conscious correction, rather than recognizing them as predictable features of a coherent accent system.

These results caution against over-reliance on native speakers as evaluators of speech intelligibility and comprehensibility in ELF settings, where interactions predominantly involve non-native users. If native speakers consistently underestimate the communicative effectiveness of L2 English varieties that function successfully in ELF contexts, then native-normed assessments may penalize effective ELF communication. Evaluations from native speakers may reflect expectation-driven bias rather than actual communicative breakdowns.

Intelligibility vs. Comprehensibility: Related but Distinct

The strong correlation between intelligibility and comprehensibility among ELF listeners ($r = .921$) confirms that the two measures are closely aligned, yet the results also illustrate their non-equivalence. The English listeners are a clear outlier: despite scoring high in intelligibility, they rated Thai-accented English as relatively difficult to understand.

This reinforces the conceptual distinction proposed by Munro and Derwing (1995) and expanded by Isaacs and Trofimovich (2012) that intelligibility reflects outcome accuracy, while comprehensibility indexes cognitive effort during processing. In ELF communication, both measures matter, demonstrating that communication succeeds when listeners can decode individual words and can do so without excessive strain or ambiguity.

Pedagogical and Theoretical Implications

The findings carry implications for both theory and language pedagogy.

- Reframing “Accent” in ELF: The findings reinforce that native-likeness is not a prerequisite for intelligibility or comprehensibility in international English communication. Instead, shared phonological expectations and exposure to diverse accent varieties are more predictive of communicative success.
- Accent Familiarity Training: Given the perceptual advantages linked to experience, pedagogical models should emphasize exposure to global English accents rather than promoting a single target variety (e.g., General American or Received Pronunciation). This can help build adaptability and perceptual flexibility, especially for learners in multilingual regions like Southeast Asia.
- Priority-Based Syllabus Development: The transcription data allow identification of specific phonological features that cause intelligibility breakdowns versus those that do not. Features causing breakdowns include: (a) final obstruent devoicing ($/b, d, g/ \rightarrow [p̥, t̥, k̥]$), which neutralizes voicing contrasts; (b) initial obstruent devoicing and fricative substitution ($/v/ \rightarrow [w]$, $/z/ \rightarrow [s]$, $/g/ \rightarrow [k]$) which eliminate critical voicing contrasts absent in Thai; (c) liquid modifications and $/r/-/l/$ neutralization, which collapse a phonemic distinction in English; and (d) Final manner-changing substitutions (e.g. $/ʃ, tʃ, dʒ/ \rightarrow [t]$). These features should be prioritized in pronunciation instruction as they directly disrupt word recognition. In contrast, several Thai English features do not compromise intelligibility and need not be targeted in instruction: (a) unaspirated stops (e.g., Thai $[t]$ for English aspirated $[t^h]$), (b) non-rhoticity such as *far*, (c) final cluster simplification where most items

remained intelligible; and (d) fricative/affricate substitutions (e.g. /tʃ/ → [tʃʰ]) as it preserves key perceptual cues. Teachers and curriculum developers should focus instructional time on features that create genuine phonemic neutralizations while accepting features that preserve the core segmental contrasts of English.

- **Listener-Responsive Instruction:** Teachers and curriculum developers should consider listener variables, including likely L1 backgrounds of interlocutors, when designing pronunciation instruction or assessing speech. The use of native-speaker models as sole benchmarks is increasingly untenable in global English communication.
- **Reevaluating Standard Norms:** The native listener paradox demonstrated here provides further support for moving away from native-speaker norms in intelligibility research and assessment. Non-native listeners, especially those from similar phonological backgrounds, may be more appropriate judges of intelligibility in ELF contexts.

Conclusion

This study makes three interrelated theoretical contributions to ELF intelligibility research. It provides empirical evidence that ISIB is not a binary, matched-versus-unmatched distinction but a gradient function of inter-L1 phonological proximity: listener groups whose L1s share structural constraints with Thai (restricted codas, limited fricative inventories, syllable-timed rhythm) demonstrated measurably higher intelligibility and comprehensibility than groups whose L1s are typologically distant. This gradient pattern, spanning Thai > Burmese/Mandarin > Japanese/Korean, supports a theoretical account in which mismatched perceptual benefits scale with the degree to which listeners can assimilate talker-L1-conditioned substitutions into their own phoneme categories. The study also documents a three-way interaction between phonological distance, regional exposure, and intelligibility outcomes: Burmese listeners' high intelligibility scores suggest that extended contact with a regional English variety can approximate the perceptual advantage that matched-L1 familiarity confers, pointing to the role of accent exposure as an independent moderator of ISIB. Finally, the native listener paradox observed here (high intelligibility paired with the lowest comprehensibility ratings) challenges the default use of native-speaker evaluators in ELF assessment, demonstrating that native phonological competence and tolerance for L2 variation are empirically dissociable.

Accordingly, these findings contribute to ongoing efforts to understand how intelligibility is co-constructed in multilingual environments. In contexts where English functions as a lingua franca, the success of communication depends less on adherence to native norms and more on

mutual intelligibility shaped by shared linguistic experience and perceptual adaptability. As global English continues to evolve, research and pedagogy must evolve with it, emphasizing accommodation, variation, and listener diversity as central components of communicative competence.

Limitations and Future Research

While the current study offers important insights into the role of listener L1 in shaping the intelligibility and comprehensibility of Thai-accented English, several limitations should be acknowledged. The relatively small sample size within each L1 group ($N = 5$) limits the generalizability of the findings. Although clear patterns emerged, future studies should expand the participant pool to include a broader and more diverse range of listeners to strengthen statistical power and to explore intra-group variability in more detail.

In addition, the study relied on controlled listening tasks involving sentence-level transcription and comprehensibility ratings of isolated speech samples. While these tasks offer methodological control and comparability across groups, they do not fully capture the dynamic, interactive nature of ELF communication, where co-construction of meaning, accommodation strategies, and contextual support play significant roles. Incorporating spontaneous dialogue data, listener-speaker interaction, or real-time comprehension monitoring could provide deeper insights into how intelligibility is negotiated in practice.

The present study also did not fully control for extralinguistic variables that may have influenced listener performance independently of L1 phonological background. Factors such as length of residence in Thailand, frequency of contact with Thai speakers, exposure to Thai media, and socioeconomic circumstances affecting linguistic immersion were collected via the background questionnaire and informed the exclusion criteria, but were not formally operationalized as controlled variables. Future studies should implement more rigorous screening, including shorter residency thresholds and formal measures of accent exposure, to disentangle linguistic from extralinguistic influences on intelligibility.

Since the present study focused on consonantal features, future research should extend the investigation to include suprasegmental features such as tone, stress, and intonation patterns, which constitute a prominent aspect of Thai-accented English. Examining how these prosodic features interact with intelligibility and comprehensibility would offer a more complete account of Thai English in ELF contexts.

Despite these limitations, the study provides useful insights into how listener L1 background affects both the intelligibility and comprehensibility

of Thai-accented English. The observed differences across groups highlight the influence of phonological familiarity, prior exposure, and listener expectations in shaping how speech is decoded and evaluated. These findings suggest that intelligibility involves more than accurate word recognition, and comprehensibility reflects more than acoustic clarity alone; both are shaped by how listeners process and adapt to non-native speech patterns. Future research should extend beyond controlled, sentence-level tasks to examine how intelligibility and comprehensibility operate in interactive, real-world communication. This shift would offer a deeper understanding of how Thai English functions within English as a Lingua Franca (ELF) contexts and how meaning is collaboratively negotiated across diverse listener-speaker pairings.

Acknowledgements

This research article is part of a doctoral dissertation titled “*The Impact of Consonants’ Functional Load on Intelligibility and Comprehensibility of Thai-Accented English in English as a Lingua Franca Contexts*” at Chulalongkorn University. The research was funded by Payap University through a PhD scholarship.

About the Authors

Suthinee Peerachachayanee: A PhD candidate in English as an International Language at Chulalongkorn University and a lecturer at the International College, Payap University. Her research focuses on English phonetics and phonology, specifically examining Thai-English phonological interactions and the pedagogical implications for language instruction and international communication contexts.

Sujinat Jitwiriyant: An Assistant Professor and a researcher at the Center of Excellence in Southeast Asian Linguistics, Faculty of Arts, Chulalongkorn University. His research interests encompass Experimental Phonetics and Foreign Language Speech.

References

- Ahn, S., & Iverson, G. (2004). Dimensions in Korean laryngeal phonology. *Journal of East Asian Linguistics*, 13(4), 345–379.
<https://doi.org/10.1007/s10831-004-4256-x>
- Ankura. (2021). *A phonological study of Burmese language spoken in Magway Division of Myanmar* [Doctoral dissertation, Mahachulalongkornrajavidyalaya University]. MCU e-Thesis.
<https://e-thesis.mcu.ac.th/thesis/864>

- Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32(2), 233–250. [https://doi.org/10.1016/S0095-4470\(03\)00036-6](https://doi.org/10.1016/S0095-4470(03)00036-6)
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114(3), 1600–1610. <https://doi.org/10.1121/1.1603234>
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). York Press.
- Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, 13(1), 67–85. <https://doi.org/10.1080/07908310008666590>
- Cho, J., & Park, H.-K. (2006). A comparative analysis of Korean-English phonological structures and processes for pronunciation pedagogy in interpretation training. *Meta*, 51(2), 229–246. <https://doi.org/10.7202/013253ar>
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19(1), 1–16. <https://doi.org/10.1017/S0272263197001010>
- Duanmu, S. (2002). *The phonology of Standard Chinese*. Oxford University Press.
- Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40(2), 269–279. <https://doi.org/10.1016/j.wocn.2011.11.001>
- Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, 36(4), 664–679. <https://doi.org/10.1016/j.wocn.2008.04.002>
- Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *The Journal of the Acoustical Society of America*, 117(2), 896–907. <https://doi.org/10.1121/1.1823291>
- Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *Canadian Modern Language Review*, 64(4), 555–580. <https://doi.org/10.3138/cmlr.64.4.555>
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility

- ratings. *Studies in Second Language Acquisition*, 34(3), 475–505.
<https://doi.org/10.1017/S0272263112000150>
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, 20(2), 219–273.
<https://doi.org/10.1017/S0952675703004524>
- Kirkpatrick, A. (2010). *English as a lingua franca in ASEAN: A multilingual model*. Hong Kong University Press.
- Lee, I., & Ramsey, S. R. (2001). *The Korean language*. State University of New York Press.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173–190.
<https://doi.org/10.2307/3588329>
- Masuda, H., & Arai, T. (2008). Perception and production of consonant clusters in Japanese-English bilingual and Japanese monolingual speakers. *Interspeech 2008*, 1988–1991.
<https://doi.org/10.21437/Interspeech.2008-523>
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Nishio, Y., & Tsuzuki, M. (2014). Phonological features of Japanese EFL speakers from the perspective of intelligibility. *JACET Journal*, 58, 57–78.
- Okell, J. (1969). *A reference grammar of colloquial Burmese*. Oxford University Press.
- Peerachachayanee, S. (2022). Towards the phonology of Thai English. *Academic Journal of Humanities and Social Sciences Burapha University*, 30(3), 64–92. <https://so06.tci-thaijo.org/index.php/husojournal/article/view/257710>
- Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics*, 26, 219–233.
<https://doi.org/10.1017/S0267190506000110>
- Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, 24, 209–239.
<https://doi.org/10.1017/S0267190504000145>
- Slayden, G. (2009). *Central Thai phonology* [Unpublished manuscript]. University of Washington. <http://www.thai-language.com/resources/slayden-thai-phonology.pdf>

- Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes*, 4(3), 333–342.
<https://doi.org/10.1111/j.1467-971X.1985.tb00423.x>
- Sohn, H. (1999). *The Korean language*. Cambridge University Press.
- Stibbard, R. M., & Lee, J.-I. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *The Journal of the Acoustical Society of America*, 120(1), 433–442.
<https://doi.org/10.1121/1.2203595>
- Suntornsawet, J. (2022). A systematic review of Thai-accented English phonology. *PASAA*, 63, 348–370.
<https://doi.org/10.58837/CHULA.PASAA.63.1.13>
- Tingsabadh, K., & Abramson, A. S. (1993). Thai. *Journal of the International Phonetic Association*, 23(1), 24–28.
<https://doi.org/10.1017/S0025100300004746>
- Tsujimura, N. (2007). *An introduction to Japanese linguistics* (2nd ed.). Blackwell Publishing.
- Vance, T. J. (2008). *The sounds of Japanese*. Cambridge University Press.
- Wei, Y., & Zhou, Y. (2002). *Insights into English pronunciation problems of Thai students*. ERIC Clearinghouse on Languages and Linguistics.
<https://files.eric.ed.gov/fulltext/ED476746.pdf>