JoAL
Journal of Applied Linguistics
JOURNAL OF APPLIED LINGUISTICS
Vol. 5 No. 2 , January 2026
e- ISSN 2809 - 1922 , p-ISSN 2809 - 1914
Available online at:
https://journal.eltaorganization.org/index.php/joal/index

The Performance of OpenL as an AI Recognition of Interdental Fricatives [θ] and [ð] in Indonesian-Accented English

M. Iqbal Farabi Sufi[1*], Tengku Syarfina[2]
[1,2]Faculty of Cultural Sciences, Universitas Sumatera Utara, Indonesia
Corresndence Email : iqbalfarabi20@gmail.com

ARTICLE INFO

Article history:
Received : December 15, 2025
Reviewed : December 18, 2025
Revised : December 20, 2025
Accepted : December 25, 2025
Available online : December 26, 2025

Keywords:
Interdental fricatives; Indonesian-accented English; OpenL

Abstract

This study investigates how Artificial Intelligence (AI) systems recognize the English interdental fricatives /θ/ and /ð/ in Indonesian-accented speech. Because these sounds are absent from the Indonesian phonological inventory, Indonesian learners often experience difficulty producing them, which may affect intelligibility and AI recognition. Using a qualitative phonetic analysis with AI-based comparison, speech data from six Indonesian learners of English and one native speaker were collected. The recordings were analyzed using Praat to examine acoustic characteristics and OpenL to generate speech-to-IPA transcriptions. The results show that many learner productions lacked sustained fricative turbulence, indicating non-target realizations of interdental fricatives. OpenL generally reflected these acoustic deviations rather than producing canonical forms, suggesting limited sensitivity to subtle fricative cues. Overall, the findings reveal an intersection between human phonetic challenges and technological limitations in current ASR systems, highlighting the need for accent-inclusive training data and focused pronunciation instruction to improve both intelligibility and AI speech recognition performance.

INTRODUCTION

The accurate recognition of speech sounds is essential for effective communication and language learning. The interdental fricatives /θ/ (as in *think*) and /ð/ (as in *this*) are widely recognized as problematic phonemes for learners of English whose native language does not include these sounds (Weinberger, 2011). In the case of Indonesian language, the absence of equivalent interdental fricatives frequently leads learners to substitute them with more familiar consonants such as /t/ or /d/. This substitution pattern represents a form of negative transfer, whereby the phonological structure of the first language shapes second language articulation, producing systematic deviations that may hinder intelligibility in international communication (Kurniawan, 2016).

Although pronunciation challenges have increasingly received scholarly attention, limited research has explored their interaction with Artificial Intelligence (AI), especially Automatic Speech Recognition (ASR) technologies. Because most ASR models are trained primarily on native English datasets, they frequently struggle with accented speech, misidentifying phoneme substitutions in particular (Shi et al., 2021). Such difficulties have been documented across various non-native

Published by English Lecturers and Teachers Association (ELTA)

DOI: 10.52622/joal.v5i2.608

English varieties, underscoring the constraints of current AI systems in accommodating phonological variation (Farooq et al., 2025).

One platform that provides a more detailed approach to speech recognition is OpenL, an AI-based transcription system capable of converting audio inputs into written and, in some cases, phonetic outputs. Designed for multilingual processing, OpenL supports more than one hundred languages and accepts a variety of input formats, including audio, image, and document files (OpenL, 2025). Its capacity to generate precise transcriptions makes it especially useful for analyzing how non-native speech is interpreted by AI, particularly in cases involving segmental substitutions.

A critical issue arises when learners replace interdental fricatives with alternative consonants, as speech recognition systems may interpret these substitutions as entirely different lexical items, resulting in communication breakdowns. For example, when an Indonesian learner produces *think* as [tink], an ASR system may render it as *tink*, thereby compromising intelligibility and comprehension. (Fadila Firdaus et al., 2020) Such misrecognitions not only create frustration for users but also reveal systemic inequities in AI technologies, where native-like articulation is privileged over accented speech (Chan et al., 2022). Confronting this limitation is essential to ensure that ASR systems remain accessible, equitable, and effective for a diverse global user base.

The interaction between Indonesian-accented English and AI-based recognition represents a distinct case due to the systematic nature of phoneme substitutions. Unlike random articulation errors, the substitutions of /θ/ with /t/ and /ð/ with /d/ occur in a consistent and predictable manner, suggesting that ASR systems could, in principle, be adapted to account for them. Yet, in the absence of focused investigations on these phonological patterns within Indonesian contexts, current models remain unable to distinguish accented variation from genuine mispronunciation, thereby constraining their utility in authentic communicative environments (Sirwan et al., 2022).

Accordingly, this study aims to examine the production and recognition of the interdental fricatives /θ/ and /ð/ in Indonesian-accented English using an IPA-based phonetic framework. Specifically, the study seeks to address the following research questions:

1. How are the interdental fricatives /θ/ and /ð/ produced by Indonesian learners of English across different word positions?
2. How does the AI-based transcription system OpenL recognize these interdental fricatives in Indonesian-accented speech compared to native pronunciation?

By addressing these questions, the study seeks to bridge linguistic research on second language pronunciation with technological evaluations of AI-based speech recognition, contributing insights to both language pedagogy and the development of more accent-inclusive ASR systems.

METHOD

This study employs a qualitative descriptive. Creswell (1994) characterizes qualitative methods as investigative strategies designed to understand social or human problems by developing a holistic and interconnected representation through the use of language. Qualitative description is particularly suitable for examining how specific phonological features are produced and perceived, rather than measuring recognition accuracy numerically. This research is rooted in phonological theory and speech perception frameworks to provide a basis for analyzing learners' pronunciation patterns. The Contrastive Analysis Hypothesis (CAH) posits that phonological disparities between a speaker's first and second languages can serve as predictors of pronunciation difficulties (Wardhaugh, 1970).

From a technological perspective, the study is informed by ASR and IPA error analysis frameworks, which emphasize how phonological variation and accented input affect speech recognition performance (Jurafsky & Martin, 2016). This dual theoretical basis ensures that the study addresses both human phonetic challenges and AI system limitations in handling accented English.

Participants

The respondents of this study consisted of six female Indonesian learners of English, aged 18 to 25, who were enrolled in an English language program. All participants were native speakers of the Indonesian language with lower intermediate English proficiency, ensuring that their speech represented naturalistic productions of interdental fricatives. In addition, a native English speaker served as a control group, allowing comparison between accented and standard pronunciation. The materials for data collection included a word list of seven English words containing interdental fricatives in initial, medial, and final positions (e.g., *think, mother, bath, this, birthday, and smooth*). All participants were informed about the purpose of the study and provided verbal consent before recording.

Data collection was conducted through individual recording sessions. Each participant read a prepared word list in a quiet setting, with high-quality audio equipment used to capture the recordings. The data were analyzed in two complementary ways. First, a trained phonetician transcribed the recordings using the International Phonetic Alphabet (IPA) with support from Praat software, focusing on whether /θ/ and /ð/ were realized accurately or substituted with other consonants. Second, the recordings were then processed using OpenL's speech-to-IPA and ASR system, which directly outputs phonetic transcriptions. This approach allows a more linguistically detailed comparison between AI recognition and human phonetic transcription.

Data analysis procedures

The analysis proceeded in several stages. On the human phonetic side, the researchers inspected Praat spectrograms to verify acoustic features and document substitutions, such as /θ/ realized as [t] or /ð/ as [d]. On the AI side, recognition outputs were compared with both the intended phonemes and the human transcriptions, enabling the identification of learner-driven substitutions versus system-specific misclassifications. The results were summarized in descriptive tables and confusion matrices, mapping correct recognitions against errors (Miles & Huberman, 1994). Finally, findings from both human and AI analyses were qualitatively interpreted to draw insights into the overlap between Indonesian learners' phonetic difficulties and the recognition limitations of current ASR and IPA systems.

FINDINGS

Production of the Voiced Interdental Fricative /ð/

Table 1. Production of Voiced Interdental Fricative /ð/

Word	Position	Learners' Production (Praat)	Native Production	OpenL IPA Output
this	Initial	[dɪs], [dis], [ðɪs]	[ðɪs]	[dɪs], [ðɪs]
mother	Medial	[mʌdər], [mʌðər]	[mʌðər]	[ˈmeidər], [mʌðər]
smooth	Final	[smu:d], [smut], [smuːð]	[smuːð]	[smut], [smuːð]

Table 1 shows that learners demonstrate variable accuracy in producing /ð/. In *this* (initial position), some learners substitute /ð/ with [d], while others produce the target [ð]. The native speaker consistently produces [ðɪs]. OpenL reflects this variation by outputting both [dis] and [ðɪs], indicating sensitivity to learners' actual acoustic realizations.

Published by English Lecturers and Teachers Association (ELTA)

In *mother* (medial position), learners again show mixed productions, with both [d] substitution and accurate [ð] realization. The native speaker produces [mʌðər], and OpenL outputs both stop and fricative forms, sometimes aligning with the target when learners produce sufficient frication.

For *smooth* (final position), learners produce both accurate [ð] and substituted stop realizations. OpenL mirrors these patterns rather than uniformly correcting them, suggesting that the system largely follows phonetic input.

1. Spectrogram of Voiced Interdental Fricative

a) Initial Position

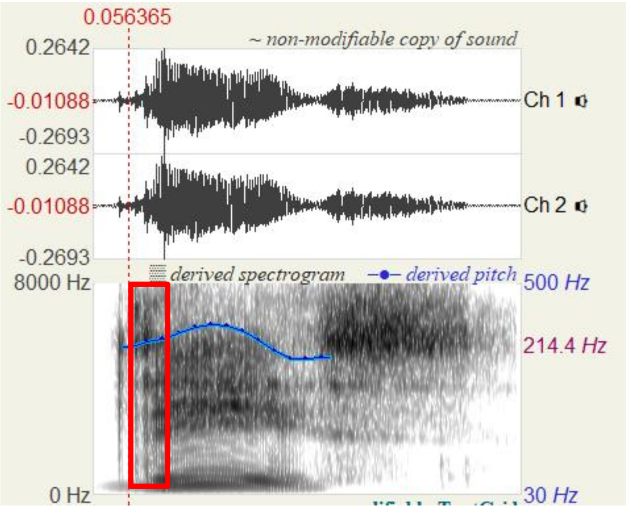

Figure 1. *Spectrogram of P03 (“this”)*

Figure 1 presents the spectrogram of P03’s production of *this*. The voiced segment shows continuous voicing but lacks high-frequency turbulence, indicating realization as a voiced alveolar stop [d]. The spectrogram displays a brief closure followed by a release burst, confirming stop articulation rather than fricative airflow. This substitution reflects the influence of the Indonesian phonological system, which does not include interdental fricatives.

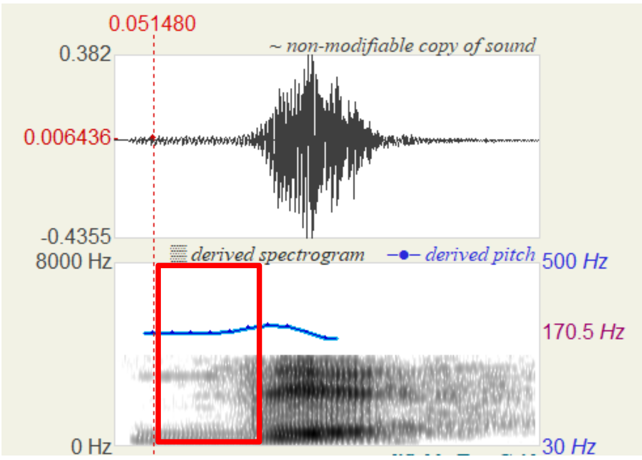

Figure 2. *Spectrogram of the native speaker (“this”)*

Figure 2 shows the native speaker’s production of *this*. In contrast, the spectrogram exhibits continuous voicing accompanied by diffuse high-frequency frication throughout the segment. The absence of a stop closure and the presence of sustained turbulence indicate accurate production of the voiced interdental fricative [ð].

b) Medial Position

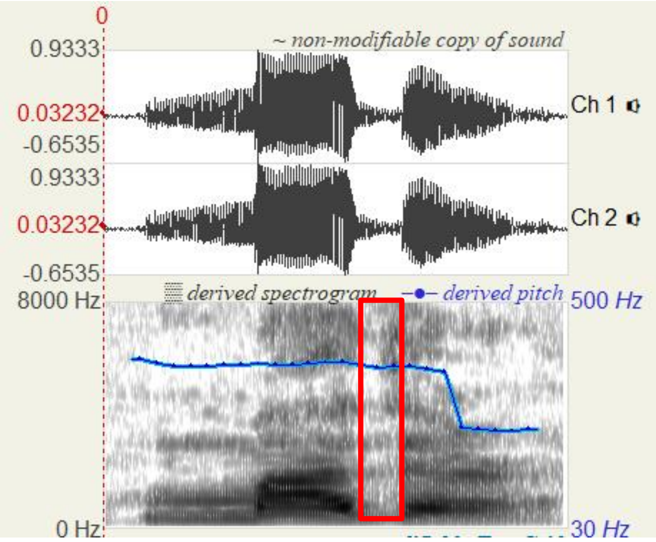

Figure 3. *Spectrogram of P06 ("mother")*

Figure 3 illustrates P06's production of *mother*. The spectrogram reveals voicing during the medial segment; however, fricative noise is weak and intermittent. The segment lacks the sustained turbulence characteristic of [ð], resulting in a stop-like realization closer to [d]. Although partial frication is present, it is insufficient to be categorized as a full interdental fricative.

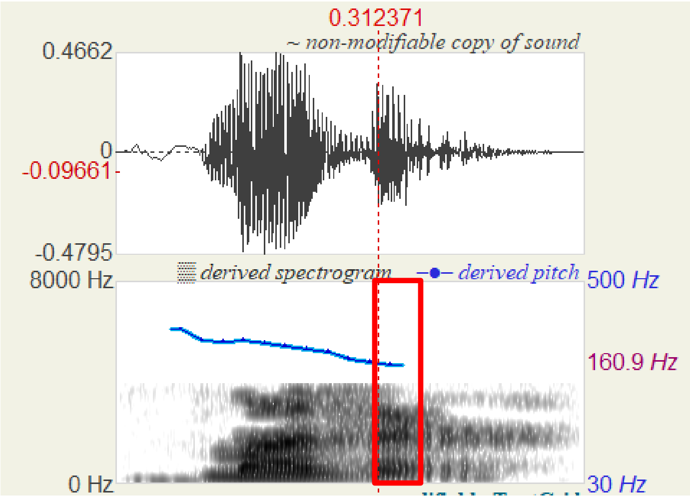

Figure 4. *Spectrogram of the native speaker ("mother")*

Figure 4 presents the native speaker's production of *mother*. The spectrogram displays stable voicing combined with clear fricative energy in the mid-to-high frequency range. This pattern confirms accurate articulation of [ð] in medial position and contrasts sharply with the learner's reduced frication.

c) Final Position

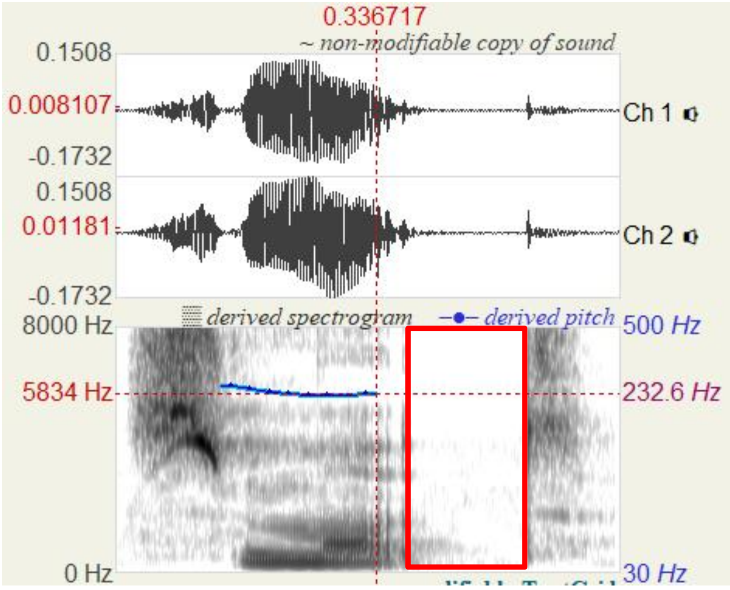

Figure 5. *Spectrogram of P05 ("smooth")*

Figure 5 shows P05's production of *smooth*. The final segment exhibits continuous voicing but lacks a visible frication band in the high-frequency range, followed by a rapid decrease in amplitude, indicating realization as a voiced alveolar stop [d]. The absence of sustained turbulence suggests that airflow was prematurely obstructed, resulting in stop substitution.

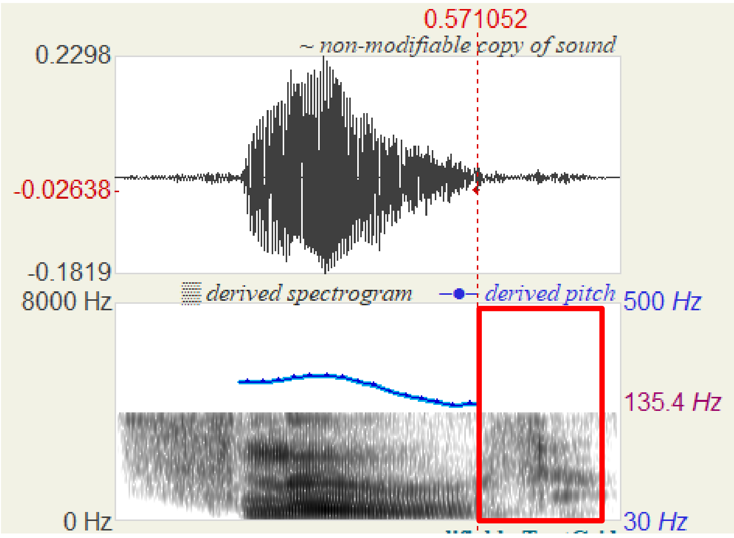

Figure 6. *Spectrogram of the native speaker ("smooth")*

Figure 6 depicts the native speaker's production of *smooth*. In contrast, the spectrogram displays continuous voicing accompanied by clear fricative noise extending to the end of the segment. This sustained turbulence confirms accurate production of the voiced interdental fricative [ð] in word-final position.

Production of the Voiceless Interdental Fricative /θ/

Table 2 presents the learners' production of the voiceless interdental fricative /θ/ and compares it with native speaker data and OpenL IPA outputs.

Published by English Lecturers and Teachers Association (ELTA)

Table 2. Production of Voiceless Interdental Fricative /θ/

Word	Position	Learners' Production (Praat)	Native Production	OpenL IPA Output
think	Initial	[tɪŋk], [θɪŋk]	[θɪŋk]	[tɪŋk], [θɪŋk]
birthday	Medial	[bɜːrtdeɪ], [ˈbɜːθdeɪ]	[ˈbɜːθdeɪ]	[bɜːrtdeɪ], [ˈbɜːθdeɪ]
bath	Final	[bæt], [bɑːθ]	[bɑːθ]	[bæs], [bɑːθ]

As shown in Table 2, learners' production of /θ/ also varies across positions. In *think* (initial position), learners produce both [θɪŋk] and [tɪŋk], while the native speaker consistently produces [θɪŋk]. OpenL outputs both forms, reflecting learner variation.

In *birthday* (medial position), learners alternate between accurate [θ] and stop substitution [t]. OpenL similarly produces both forms, indicating partial alignment with learners' pronunciation.

In *bath* (final position), learners produce both [θ] and [t]. The native speaker produces [θ], and OpenL occasionally outputs a target-like form even when learners' production deviates, suggesting limited normalization toward canonical pronunciation.

1. Spectrogram of Voiceless Interdental Fricative

a) Initial Position

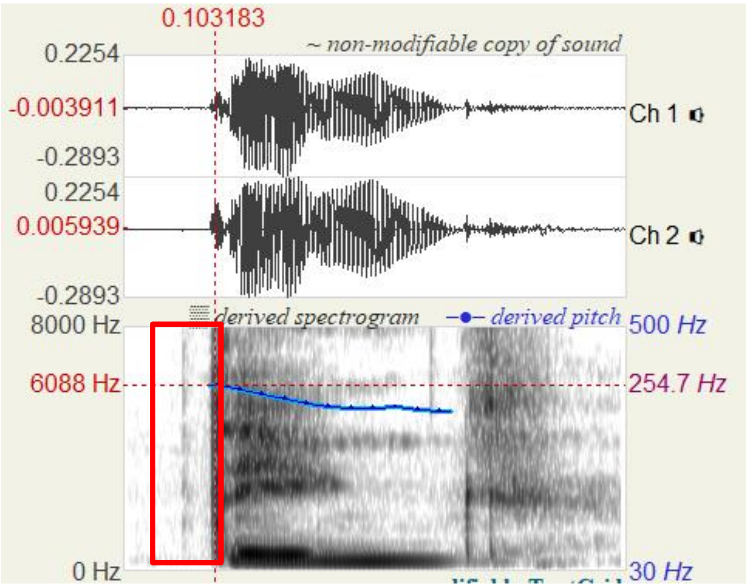

Figure 7: *Spectrogram of P05 ("think")*

Figure 7 presents the spectrogram of P05's production of *think*. The initial segment shows a complete closure followed by a brief release burst, with no visible high-frequency frication. This acoustic pattern indicates realization of the voiceless interdental fricative /θ/ as the voiceless alveolar stop [t]. The absence of sustained turbulence confirms stop articulation rather than fricative airflow, reflecting influence from the Indonesian phonological system, which does not include interdental fricatives.

Published by English Lecturers and Teachers Association (ELTA)

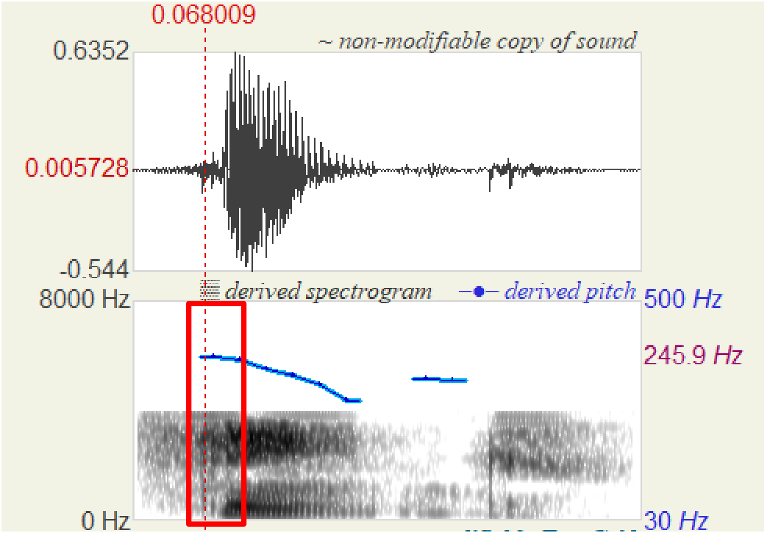

Figure 8: *Spectrogram of the native speaker ("think")*

Figure 8 shows the native speaker's production of *think*. In contrast, the spectrogram exhibits diffuse high-frequency turbulence extending throughout the initial segment without evidence of stop closure. This continuous frication confirms accurate production of the voiceless interdental fricative [θ].

b) Medial Position

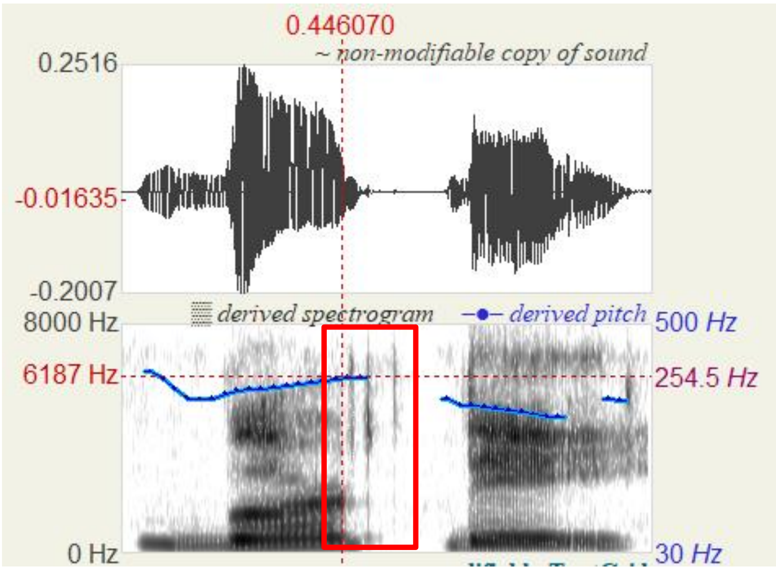

Figure 9. *Spectrogram of P01 ("birthday")*

Figure 9 presents the spectrogram of P01's production of *birthday*. The medial segment is characterized by a brief closure and a release burst, with minimal fricative noise. This pattern indicates substitution of the voiceless interdental fricative /θ/ with the voiceless alveolar stop [t]. The lack of sustained turbulence suggests that airflow was momentarily obstructed rather than continuously released.

Published by English Lecturers and Teachers Association (ELTA)

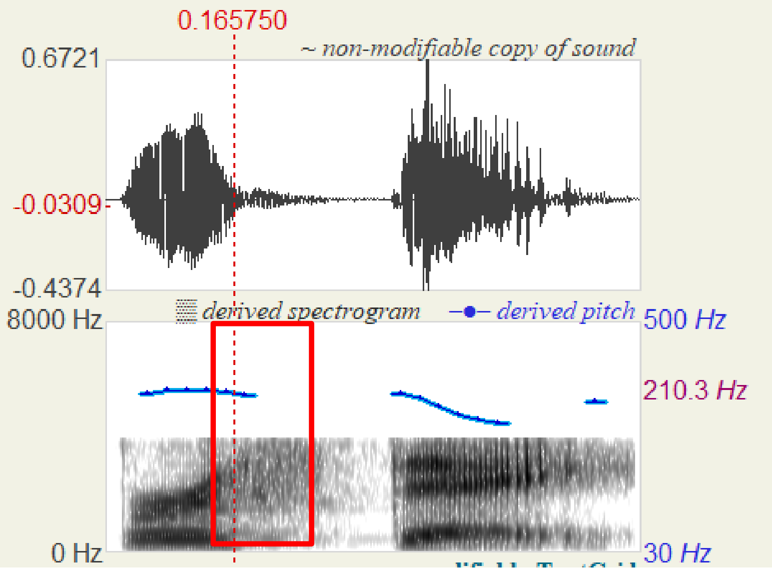

Figure 10. *Spectrogram of the native speaker ("birthday")*

Figure 10 shows the native speaker's production of *birthday*. In contrast, the spectrogram displays continuous high-frequency frication during the medial segment, without a complete closure. This acoustic profile confirms accurate realization of [θ] in intervocalic position.

c) Final Position

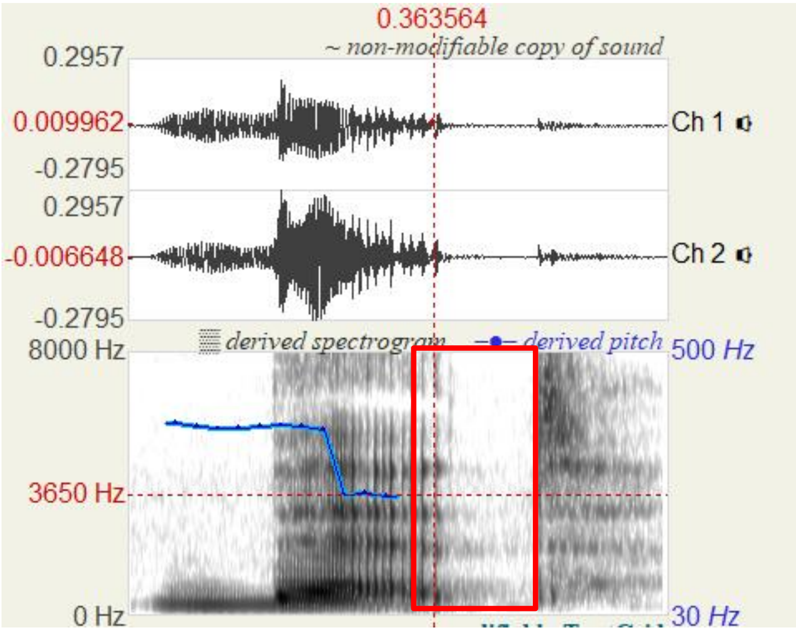

Figure 11. *Spectrogram of P05 ("bath")*

Figure 11 presents the spectrogram of P05's production of *bath*. The final segment exhibits a rapid decrease in amplitude with no visible high-frequency frication, followed by a stop-like release, indicating realization of the voiceless interdental fricative /θ/ as the voiceless alveolar stop [t]. The absence of sustained turbulence suggests premature obstruction of airflow at the end of the word.

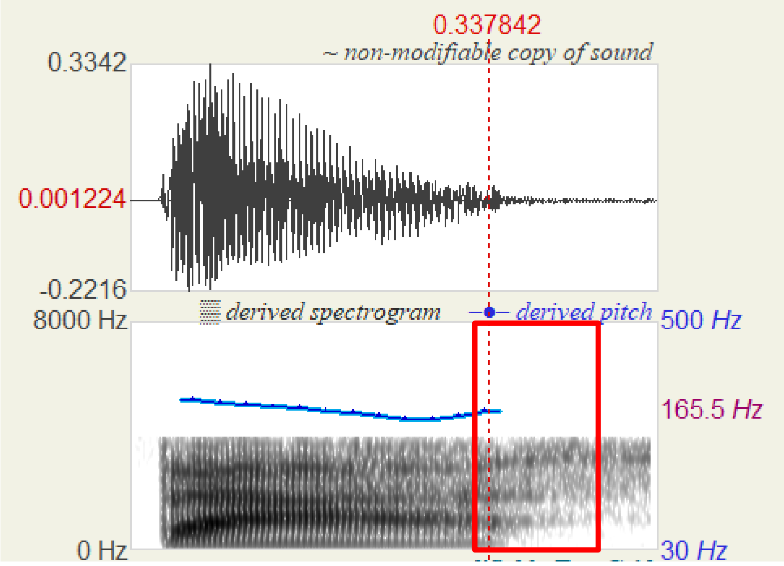

Figure 12. *Spectrogram of the native speaker ("bath")*

Figure 12 shows the native speaker's production of *bath*. In contrast, the spectrogram displays sustained high-frequency frication extending to the end of the segment, without evidence of stop closure. This sustained turbulence confirms accurate production of the voiceless interdental fricative [θ] in word-final position.

3.2 AI (OpenL) recognition as IPA

OpenL's speech-to-IPA transcription generally reflected the learners' actual productions rather than their intended phonemic targets. When learners produced /θ/ as [t] or /ð/ as [d], OpenL frequently returned [t] or [d] in its IPA output. This indicates that the model is responsive to the same acoustic features identified in the Praat analysis, such as closure, burst, and the absence of frication in stop consonants. In some tokens with weak but perceptible frication, particularly medial /θ/ in *birthday* and medial /ð/ in *mother*, OpenL occasionally recognized the intended fricative. However, stop-like realizations were more consistently transcribed as stops, reflecting a pattern similar to human perception.

One notable exception was observed in the native control data, where OpenL failed to produce an IPA transcription for *bath* despite the production being native-like. This isolated error highlights that tool reliability can also influence analytical outcomes independently of the phonetic content itself.

Published by English Lecturers and Teachers Association (ELTA)

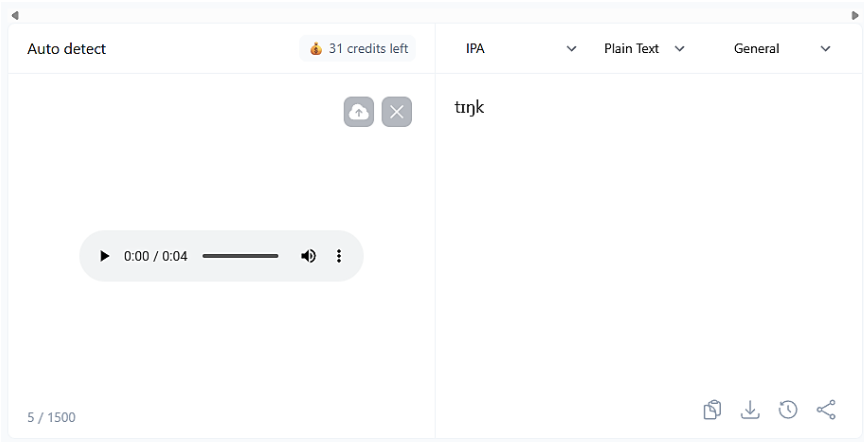

Figure 13. *OpenL IPA output for P05 (“think”) – [tɪŋk]*

The OpenL system transcribed P05’s production of *think* as [tɪŋk], which precisely matched the human transcription. This result indicates that the AI recognized the same stop-like acoustic cues, including the brief closure and strong release burst, and interpreted them as [t] rather than [θ]. The consistency between OpenL and human perception suggests that the system relies on the absence of frication to identify a stop consonant.

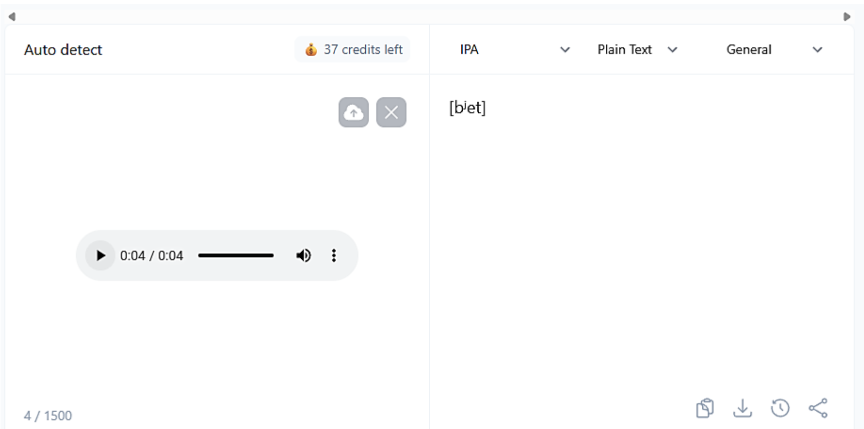

Figure 14. *OpenL IPA output for P05 (“bath”) – [bɛt]*

The OpenL transcription for P05’s production of *bath* displays [bɛt], confirming the system’s classification of the final /θ/ as [t]. The AI evidently responded to the absence of high-frequency fricative noise and the presence of a stop-burst cue at the word’s end. This aligns with the human Praat transcription, indicating that OpenL, like human listeners, perceived the learner’s voiceless interdental fricative as an alveolar stop. The convergence between human and AI analyses suggests that both rely on similar acoustic indicators when processing low-frication tokens.

Published by English Lecturers and Teachers Association (ELTA)

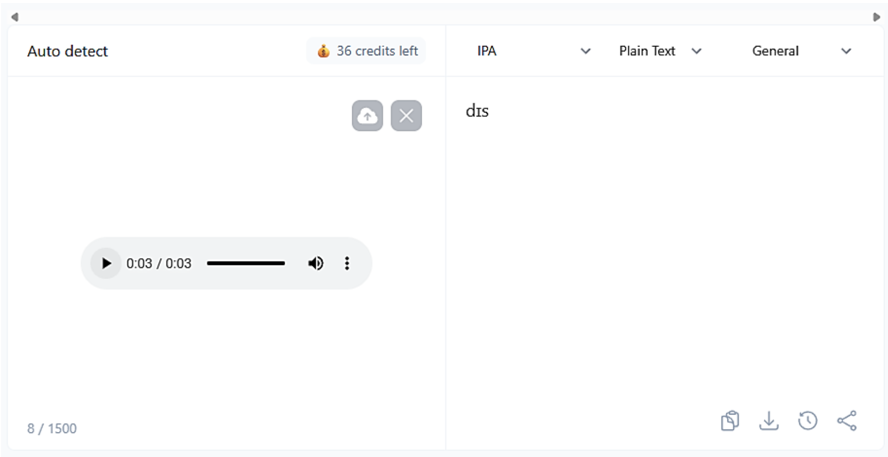

Figure 15. *OpenL IPA output for P04 (“this”) – [dɪs]*

In *this*, OpenL identified the initial segment as [d], consistent with the learner’s substitution of /ð/→[d]. The model captured the voiced closure without fricative energy, showing sensitivity to the same low-frequency voicing cues noted in Praat. Both human and AI analyses classify the token as a voiced alveolar stop rather than a fricative.

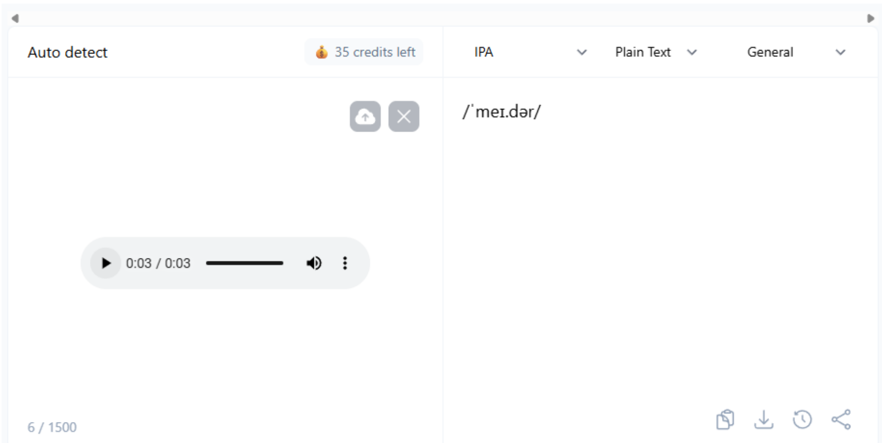

Figure 16. *OpenL IPA output for P06 (“mother”) – [ˈmeidər]*

The OpenL system transcribed P06’s production of *mother* as [ˈmeidər], which constitutes a clear transcription error. The AI not only missed the intended word but also failed to represent the voiced interdental fricative /ð/, interpreting it instead as [d]. In contrast, the Praat analysis of P06’s recording revealed partial frication in the medial position, suggesting that the learner produced a weak but perceptible [ð]. This comparison indicates that OpenL was overly influenced by vowel formant cues and the closure element of the segment, overlooking the subtle fricative noise. Such a mismatch demonstrates the system’s limitation in distinguishing low-intensity interdental fricatives in connected vocalic contexts, while human auditory analysis can still recognize these as approximations of [ð].

Published by English Lecturers and Teachers Association (ELTA)

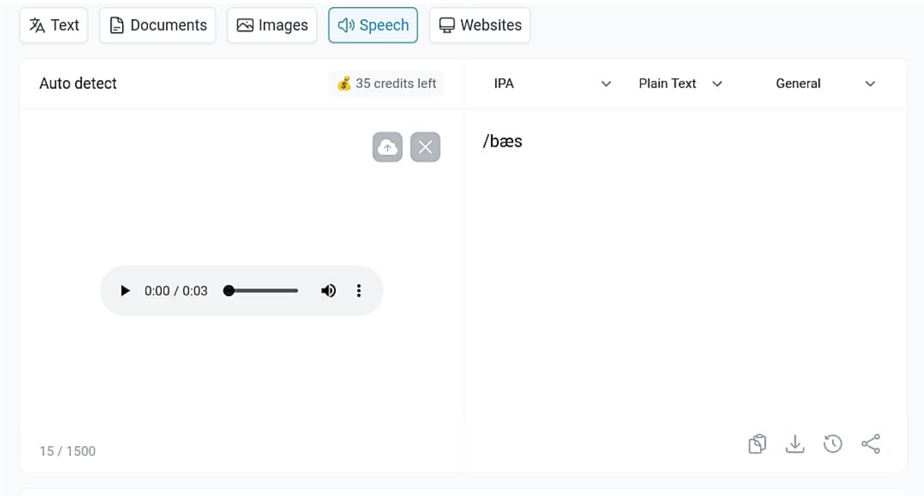

Figure 17. *OpenL IPA output for the native speaker ("bath") – [bæs]*

OpenL transcribed the native token as [bæs], indicating it perceived strong frication but labeled it with [s] rather than the intended [θ]. Praat for the native control shows a clear interdental fricative (continuous, diffuse high-frequency noise without a stop burst), so the [s] label reflects a place-of-articulation misclassification by the AI (alveolar vs interdental), not a pronunciation error. This contrasts with learners' tokens (e.g., [bæt]) where both human and AI identified a stop rather than frication.

DISCUSSION

The findings of this study confirm earlier research indicating that interdental fricatives /θ/ and /ð/ pose significant challenges for English learners whose first language lacks these sounds. Consistent with Weinberger (2011) and Kurniawan (2016), Indonesian learners in this study frequently substituted /θ/ with [t] and /ð/ with [d]. These substitutions directly reflect the central prediction of the Contrastive Analysis Hypothesis (CAH), which posits that phonemes absent from the learner's first language will be replaced by the closest available articulatory equivalents in the target language (Wardhaugh, 1970). The systematic nature of these substitutions demonstrates negative transfer rather than random error, thereby providing empirical support for CAH at the segmental level.

However, the present findings extend previous studies by showing that learner production is not uniformly inaccurate. Some participants produced target-like interdental fricatives, particularly in initial and medial positions, suggesting partial acquisition despite the absence of explicit phonetic training. This supports earlier claims that interdental fricatives are difficult but not unattainable for EFL learners. Acoustic analysis using Praat further confirms that accurate productions display sustained frication and, in the case of /ð/, continuous voicing, in line with established phonetic descriptions.

From a technological perspective, the findings further extend CAH implications to AI-based speech recognition. Because ASR systems are trained predominantly on native phonological patterns, they tend to mirror the same contrastive biases that challenge human learners. As observed in previous studies (Farooq et al., 2025; Shi et al., 2021), OpenL frequently reflected learners' stop substitutions when fricative cues were absent, effectively treating these productions as legitimate realizations rather than deviations. However, in some cases, OpenL produced target-like IPA outputs despite non-native pronunciation, particularly for high-frequency words, suggesting lexical normalization rather than phonetic sensitivity (Chan et al., 2022).

Published by English Lecturers and Teachers Association (ELTA)

The systematic nature of Indonesian learners' substitutions, as highlighted by (Sirwan et al., 2022), is clearly evident in the present data. Rather than random errors, substitutions of interdental fricatives follow consistent patterns, suggesting that ASR systems could be adapted to better accommodate such variation. Nonetheless, the comparison between Praat-based human analysis and OpenL output demonstrates that current AI systems still struggle to distinguish accented variation from phonemic error, limiting their reliability in pronunciation assessment and real-world communication.

CONCLUSION

This study demonstrates that Indonesian learners of English commonly substitute interdental fricatives /θ/ and /ð/ with alveolar stops [t] and [d], a pattern confirmed through both human phonetic analysis and AI-based transcription. OpenL's IPA outputs largely mirrored learners' articulatory realizations, indicating that the system relies on salient acoustic cues such as closure and release while showing limited sensitivity to weak fricative turbulence. These findings reinforce the observation that current ASR systems, similar to human listeners, are influenced by accent-related phonological variation, which can result in misrecognition when non-native pronunciations diverge from canonical targets.

From a pedagogical perspective, the results highlight the need for focused pronunciation instruction on interdental fricatives for Indonesian learners of English. Since learners consistently rely on alveolar stops as substitutes, pronunciation training should emphasize articulatory placement and sustained airflow associated with /θ/ and /ð/. The use of acoustic visualizations, such as spectrograms, may help learners develop greater awareness of fricative features and distinguish them from stop consonants.

From a technological perspective, the findings underscore the limitations of current ASR systems in handling predictable accent-based phonological variation. Because OpenL reflects learners' productions rather than accounting for systematic substitutions, ASR outputs may fail to distinguish accented realizations from canonical forms. These results suggest that incorporating accented speech data and phonologically informed modeling strategies could improve ASR performance for non-native speakers without altering the intended lexical output.

Future research should expand the dataset to include more diverse Indonesian accents and other non-native varieties of English. Comparative testing with multiple ASR systems (e.g., Whisper, Google Speech-to-Text) could provide broader insight into AI sensitivity to accented fricatives. Moreover, integrating acoustic modeling or deep-learning approaches trained on accented input could improve AI recognition robustness. Pedagogically, future studies may explore the use of AI-assisted pronunciation training tools to help learners perceive and produce interdental fricatives more accurately.

REFERENCES

Chan, M. P. Y., Choe, J., Li, A., Chen, Y., Gao, X., & Holliday, N. (2022). Training and Typological Bias in ASR Performance for World Englishes. *Interspeech 2022*, 1273–1277. https://doi.org/10.21437/Interspeech.2022-10869

Creswell, J. W. (1994). Research Design: Qualitative and Quantitative Approaches. Sage Publications.

Fadila Firdaus, S., Maulia Indrayani, L., & Soeria Soemantri, Y. (2020). The Production of Interdental Fricatives by English as a Foreign Language Students in English Course Bandung. *Linguistics and English Language Teaching Journal*, *8*(1), 2339–2940.

Farooq, M. A., Unsa Hafiz, & Hussain, M. A. (2025). Speech Recognition and Phonetic Variation: Understanding the Impact of Accents and Dialects on AI-Based Speech Systems. *Qualitative Research Journal for Social Studies, 2*(3), 296–311. https://doi.org/10.63878/qrjs295

Jurafsky, D., & Martin, J. H. (2016). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition Second Edition. Pearson.

Kurniawan, D. (2016). The Error Analysis of the Pronunciation of Dental Education Study Program Faculty of Teacher Training and Education Sriwijaya University. *Jounal of English Literacy Education, 3*(2), 157–163.

Miles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis: an Expanded Sourcebook. CA: Sage Publications.

OpenL. (2025). OpenL Translate: Accurate AI Translation in 100+ Languages. https://openl.io/

Shi, X., Yu, F., Lu, Y., Liang, Y., Feng, Q., Wang, D., Qian, Y., & Xie, L. (2021). The Accented English Speech Recognition Challenge 2020: Open Datasets, Tracks, Baselines, Results and Methods. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 6918–6922. https://doi.org/10.1109/ICASSP39728.2021.9413386

Sirwan, A., Thama, K. A., & Suyanto, S. (2022). Indonesian Automatic Speech Recognition Based on End-to-end Deep Learning Model. *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 410–415. https://doi.org/10.1109/CyberneticsCom55287.2022.9865253

Wardhaugh, R. (1970). Teachers of English to Speakers of Other Languages, Inc. (TESOL). *Source: TESOL Quarterly, 4*(2), 123–130.

Weinberger, S. H. (2011). Minimal Segments in Second Language Phonology. In *Second-Language Speech* (pp. 263–312). De Gruyter Mouton. https://doi.org/10.1515/9783110882933.263

Published by English Lecturers and Teachers Association (ELTA)