International Journal of English Learning and Applied Linguistics (IJELAL)
6(2), 2026, 300–313
ISSN: 2775-4359 (Online)
DOI: http://dx.doi.org/10.21111/ijelal.v4i2.15467

INTERLANGUAGE IN SOUND: SYLLABLE-STRUCTURE MODIFICATION IN ENGLISH PRONUNCIATION AMONG INDONESIAN PESANTREN LEARNERS

Dian Fajrina Mubarokah[1], Faiz Rizki Muhammad[2]*, Dinar Dipta[3], Haida Umiera Hashim[4]

[1,2,3]Universitas Darussalam Gontor, Jl. Raya Siman, Demangan - Siman - Ponorogo, Jawa Timur, 63471, Indonesia
[4]Universiti Teknologi MARA, Jl. Ilmu 1/1, 40450 Shah Alam, Selangor, Malaysia
[2] faizrizkimuhammad@unida.gontor.ac.id

Abstract

This study investigates syllable-structure modification in the English pronunciation of 11th-grade students at one of pesantrens institution in East Java. It aims to identify modification types in monosyllabic, disyllabic, and trisyllabic words, determine the most frequently affected syllable positions (onset, nucleus, and coda), and examine these modifications from a Second Language Acquisition (SLA) perspective. Using a descriptive qualitative design, data were collected through pronunciation tasks and semi-structured interviews. The findings show that pronunciation modification occurs systematically, reflecting learners' developing interlanguage phonology. Four major modification types were identified: substitution, epenthesis, deletion, and cluster reduction. Substitution was the most dominant type, while nucleus modification occurred most frequently, indicating persistent difficulty in vowel production. From an SLA perspective, these patterns are influenced by first-language transfer, phonotactic constraints, and learners' tendency to assimilate unfamiliar English sounds into existing phonological categories. Increased modification in longer words also suggests the role of cognitive processing limitations. Despite substantial English exposure in the pesantren environment, the persistence of these patterns indicates that input alone is insufficient without explicit phonological awareness and corrective feedback. This study highlights the importance of integrating phonological instruction with SLA-informed approaches to support pronunciation development in EFL contexts.

Keywords: Interlanguage Phonology, Pesantren, Pronunciation Modification, Syllable Structure

INTRODUCTION

English has become a global language used in education, business, politics, religion, and digital communication, making English proficiency increasingly important for learners in English as a Foreign Language (EFL) contexts (Chenenje, 2024). Among the components of communicative competence, pronunciation plays a crucial role because it directly affects intelligibility and successful communication. Effective pronunciation does not necessarily require a native-like accent; rather, it enables speakers to convey messages clearly and confidently (Srakaew, 2021). In contrast, poor pronunciation may reduce learners' confidence and willingness to participate in oral communication activities, limiting opportunities for language development (Datu, 2025; Gumelar & Lestari, 2024).

Despite its importance, pronunciation remains one of the most difficult aspects of second language learning. These difficulties are largely influenced by differences between the phonological systems of learners' first language (L1) and the target language

(L2). English contains complex vowel distinctions, consonant clusters, stress patterns, and syllable structures that are often absent in Indonesian phonology (Larasati & Pammu, 2023). Consequently, Indonesian EFL learners frequently rely on their L1 phonological knowledge when producing English sounds. From a second language acquisition (SLA) perspective, this phenomenon can be explained through interlanguage theory, which views learner language as a dynamic linguistic system shaped by both L1 influence and ongoing exposure to the target language (Selinker, 1972). One of the most influential processes within interlanguage development is language transfer, particularly negative transfer, where learners apply L1 phonological rules to L2 pronunciation (Odlin, 1989).

Research on pronunciation in second language acquisition has consistently emphasized the importance of phonological awareness and syllable structure in developing intelligible speech among EFL learners. According to Selinker (1972), learners construct an interlanguage system that reflects a transitional linguistic competence influenced by both the target language and the native language. Within this framework, pronunciation errors are viewed not as random deviations but as systematic manifestations of learners' developing phonological knowledge. Similarly, Odlin (1989) argues that language transfer plays a central role in shaping second language pronunciation, particularly when learners apply the phonotactic constraints of their first language to the target language. This phenomenon is especially visible in syllable-structure modification, where learners simplify unfamiliar consonant clusters or alter vowel patterns to conform to familiar L1 structures. Previous studies by Khoirunnisa and Misnadin (2022) as well as Li (2023) further demonstrate that Indonesian EFL learners frequently experience difficulties in producing English syllable patterns due to differences between Indonesian and English phonological systems. These findings support the view that pronunciation development is strongly connected to learners' awareness of syllable organization and phonological constraints.

In addition to language transfer, several scholars highlight the cognitive and perceptual dimensions of pronunciation acquisition. Flege (1995), through the Speech Learning Model, explains that learners tend to categorize unfamiliar L2 sounds into the closest equivalents available in their native phonological inventory, leading to substitution and simplification processes. This theory is particularly relevant in explaining why Indonesian learners often modify English syllable structures through epenthesis, deletion, or cluster reduction. Furthermore, Krashen (1985) emphasizes in the Input Hypothesis that comprehensible input is necessary for language acquisition, although exposure alone does not guarantee accurate pronunciation development. Complementing this perspective, Schmidt (1990) argues through the Noticing Hypothesis that learners must consciously recognize the discrepancy between their own production and the target form in order to achieve phonological improvement. Likewise, Swain (1985) stresses the role of output and corrective feedback in encouraging learners to restructure their interlanguage system. In pesantren contexts, where students actively use English in daily communication yet continue to experience pronunciation difficulties, these theories collectively suggest that persistent errors may result from limited phonological noticing, insufficient feedback, and the stabilization of non-target-like forms. Therefore, analyzing syllable-structure modification provides a valuable approach for understanding how learners negotiate phonological challenges within multilingual EFL environments.

Pronunciation difficulties are particularly visible in syllable structure production. A syllable consists of three main components: onset, nucleus, and coda, which together form the basic phonological structure of words (Shahid & Mahmood, 2022). English allows more complex syllable patterns and consonant clusters than Indonesian, causing learners to modify unfamiliar sound combinations. These modifications commonly appear in the form of substitution, deletion, epenthesis, and cluster reduction. Such patterns reflect learners' attempts to simplify difficult phonological structures while maintaining intelligibility. According to Flege's Speech Learning Model, learners tend to assimilate unfamiliar L2 sounds into the closest categories available in their native phonological system, resulting in systematic pronunciation modification (Flege, 1995). Previous studies have shown that Indonesian learners frequently experience pronunciation problems involving vowel production and consonant clusters (Khoirunnisa & Misnadin, 2022; Li, 2023). However, many studies focus mainly on segmental errors without examining how these errors relate to syllable structure and broader SLA processes.

This study was conducted in Pondok Pesantren Al-Ikhsan Beji, an Islamic boarding school where students are exposed to multiple languages, including Indonesian, local languages, Arabic, and English (Muhammad & Ashadi, 2019). English is used in various formal and informal activities such as classroom interaction, public speaking, and daily communication (Deyanti & Hindun, 2024). Despite this exposure, pronunciation difficulties remain evident among students. This condition suggests that exposure alone is insufficient for accurate pronunciation acquisition. Krashen (1985) explains that comprehensible input is necessary for language acquisition, but learners must also internalize the input effectively. Similarly, Schmidt (1990) argues that learners need to consciously notice the gap between their own production and the target form to improve their language performance. Without sufficient feedback and phonological awareness, learners may continue producing non-target-like pronunciation forms, leading to fossilization within their interlanguage system (Selinker, 1972; Swain, 1985).

Based on these considerations, this study investigates syllable-structure modification in the English pronunciation of 11th-grade students at Pondok Pesantren Al-Ikhsan Beji. Specifically, the study aims to identify the types of syllable-structure modification found in monosyllabic, disyllabic, and trisyllabic words, determine which syllable positions (onset, nucleus, and coda) are most frequently modified, and explore the factors influencing students' pronunciation within the pesantren context. By integrating phonological analysis with SLA perspectives, this study is expected to provide a deeper understanding of pronunciation development among Indonesian EFL learners.

METHOD

This study employed a descriptive qualitative design to investigate syllable-structure modification in students' English pronunciation. The focus of the research was to identify the types of pronunciation modification and to analyze the specific syllable positions (onset, nucleus, and coda) affected by those modifications. A qualitative approach was chosen because the study aimed to provide detailed descriptions and interpretations of phonological phenomena rather than statistical measurements.

The primary source of data was the spoken pronunciation of 11th-grade students at one of Modern Pesantrens in Mojokerto, East Java. The research was conducted in the

context of an Islamic boarding school where students receive formal English instruction and participate in a daily language system that encourages the use of English in certain situations. This educational context provided relevant linguistic exposure, making the students suitable participants for examining pronunciation patterns. The main data consisted of audio recordings from a pronunciation task, supported by data from semi-structured interviews with selected students and English teachers.

Instruments and Procedures

The data were collected through two instruments. First, a pronunciation task in which required students to read aloud a list of English words representing different syllable structures (CV, CVC, CCVC) and word lengths (monosyllabic, disyllabic, and trisyllabic). The target pronunciations were taken from the Oxford Learner's Dictionary and transcribed using the International Phonetic Alphabet (IPA) as the standard reference. Second, semi-structured interviews were conducted to obtain contextual information regarding students' language background, exposure to English, and perceived pronunciation difficulties.

Data Collection

The data collection process was conducted in several stages to ensure the accuracy, consistency, and ethical integrity of the study. This study involved human participants and followed ethical standards for research involving human subjects. Prior to the data collection, ethical approval was obtained from the relevant institutional authority in accordance with institutional guidelines for academic research conducted in educational settings. All participants were informed about the purpose of the study, the research procedures, and the use of audio recordings during the data collection process. Informed consent was obtained from all participants before their involvement in the study, and anonymity was ensured through the use of participant codes instead of real names. In addition, all collected data were handled confidentially and used solely for academic purposes.

Before conducting the pronunciation task, the researcher explained the instructions and procedures clearly to ensure that participants understood the activity. Each student was then asked to read aloud the prepared word list individually in a quiet environment to minimize background noise and distractions. The pronunciation performances were recorded using a digital audio recorder or smartphone recording application to preserve the authenticity of the students' speech production. During the reading activity, the researcher did not interrupt or correct the participants in order to obtain natural pronunciation data. After the recordings were collected, the audio files were transcribed phonologically and compared with the target IPA transcriptions to identify patterns of syllable-structure modification such as substitution, deletion, epenthesis, and cluster reduction.

Furthermore, the researcher conducted the interview to explore main challenges and gain additional information on the analyzed case. To facilitate clearer communication and more detailed responses, the interviews were conducted in Indonesian. With the participants' consent, all interview sessions were audio-recorded and subsequently transcribed for analysis. The interview findings were used to support the interpretation of

the pronunciation data and to provide deeper insights into the factors influencing students' syllable-structure modification in English pronunciation within the pesantren context.

Data Analysis Procedures

The data analysis was conducted systematically in several stages. All audio recordings were first organized and labeled according to participant codes. Each pronunciation was then transcribed using IPA symbols to represent the students' actual spoken forms. The transcriptions were compared with the target pronunciations to identify deviations in syllable structure. Identified modifications were classified into deletion, epenthesis, substitution, or cluster reduction. Finally, each modification was localized within the syllable structure to determine whether it occurred in the onset, nucleus, or coda. To enhance credibility, source triangulation was applied by comparing pronunciation data with interview findings, ensuring consistency between phonological patterns and contextual explanations

From the data gathered in the interview session, the researcher used thematic analysis proposed by Braun and Clarke (2006) to identify, analyze, and present the data in the form of patterns of meaning from the explored coding. This thematic analysis is particularly useful because it allows researchers to explore participants' experiences, perceptions, and language practices in depth.

FINDINGS

The findings reveal that pronunciation modification does not occur randomly. Instead, it follows systematic and predictable patterns influenced by word length, syllable structure, and phonotactic differences between English and Indonesian. From a second language acquisition perspective, these patterns can be understood within the framework of interlanguage development. According to interlanguage theory, learners construct a dynamic linguistic system that is neither fully their first language nor the target language, but a rule-governed system in its own right (Selinker, 1972). Therefore, the observed pronunciation modifications should not be viewed merely as errors, but as evidence of learners' evolving phonological competence.

Types of Pronunciation Modification

The analysis demonstrates that students employed four primary strategies in modifying English pronunciation: substitution, epenthesis, deletion, and cluster reduction. Although all four types appeared across different word lengths, their frequency and distribution varied depending on structural complexity as it is showed in the following table.

Table 1. Types of Pronunciation Modification Cases

Pronunciation Modification	Monosyllable	Disyllable	Trisyllable
Substitution	63	77	90
Deletion	0	31	13
Epenthesis	7	19	28
Cluster reduction	3	9	0

Substitution

Substitution refers to the replacement of one sound with another. Kenworthy (1987) explains that substitution occurs when learners replace unfamiliar target sounds with more familiar ones from their first language (Khoirunnisa & Misnadin, 2022). This explanation aligns with the present findings. For example, in the word "breath" /breθ/, students pronounced it as *brith* /brið/. In this case, the vowel /e/ in the nucleus of syllable 1 was replaced by /ɪ/, and the consonant /θ/ in the coda of syllable 1 was substituted with /ð/. These substitutions indicate strong first-language influence, as Indonesian does not contain the interdental fricative /θ/.

In disyllabic words, substitution frequently occurred in stressed syllables. For instance, in the word "nation" /ˈneɪ.ʃən/, students pronounced it as *nasyen* /naʃən/. The diphthong /eɪ/ in the nucleus of syllable 1 was simplified into the monophthong /a/. This pattern suggests that students avoided complex vowel sequences and preferred stable vowel sounds that resemble Indonesian vowel patterns.

In trisyllabic words, substitution often involved vowel strengthening in unstressed syllables. In the word "collected" /kə.lɛk.tɪd/, students pronounced it as *kolekt* /kolɛkt/. The schwa /ə/ in the nucleus of syllable 1 was replaced by /o/, indicating difficulty in producing reduced vowels. Since Indonesian does not employ vowel reduction in the same way English does, students tended to replace weak vowels with full vowels.

Deletion

Deletion was strongly influenced by word length. No deletion cases were identified in monosyllabic words. However, deletion occurred 31 times in disyllabic words and 13 times in trisyllabic words. Deletion, particularly in longer words, can be associated with cognitive and processing constraints in second language production. From a psycholinguistic perspective, this phenomenon can be explained by cognitive processing limitations in second language production. Learners with limited automaticity tend to simplify linguistic output to reduce articulatory and cognitive load (de Jong, 2023). For the example in the word "flower" /ˈflaʊ.ər/, students deleted the final /r/ in the coda of syllable and in trisyllabic words such as "collected", students deleted the entire third syllable /tɪd/, reducing the word to two syllables.

Epenthesis

Epenthesis was the second modification strategy identified. It occurred 7 times in monosyllabic words, 19 times in disyllabic words, and 28 times in trisyllabic words. Epenthesis refers to the insertion of an additional sound into a word. For the example in the word "great" /greɪt/, students produced *greint*, inserting /n/ before the final /t/. The insertion occurred between the nucleus and the coda of syllable. Other example is in the word "crowded" /ˈkraʊ.dɪd/where students pronounced it as *karoudid*, inserting a vowel /a/ after the onset /k/ in syllable. This insertion created an additional syllabic element and reduced the complexity of the consonant cluster /kr/.

In trisyllabic words, epenthesis frequently occurred within medial clusters. In the word "collected", some students produced *kolekstid*, inserting /s/ before /t/ in syllable

two. This suggests that insertion functioned as a compensatory strategy to manage articulatory difficulty.

Cluster Reduction

Cluster reduction was the least frequent modification type. It occurred 3 times in monosyllabic words, 9 times in disyllabic words, and did not occur in trisyllabic words. Cluster reduction further illustrates learners' sensitivity to phonotactic constraints. Like in the example of word "pump" /pʌmp/, students pronounced it as *pam*, deleting the final /p/ in the coda cluster. Another example is in the word "project" /prodʒɛkt/where the students produced *projet*, deleting /k/ in the final cluster. These examples show simplification of consonant clusters that are less common in Indonesian phonology

Modification Based on Syllable Structure

The analysis in this finding is categorizing modification three main aspects: onset, nucleus, and coda positions across the syllables. The findings showed the modification case as follows.

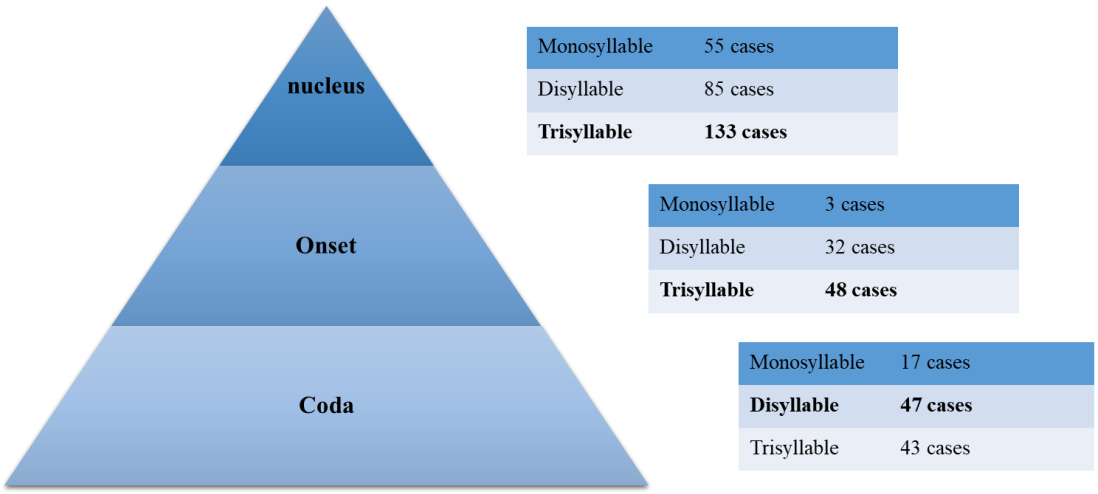

Figure 1. Modification Cases Based on Syllable Structure

The findings show that pronunciation modification occurred across all syllable positions, namely the nucleus, onset, and coda, with different frequencies in monosyllabic, disyllabic, and trisyllabic words. The nucleus position recorded the highest number of modification cases among all syllable components. Specifically, there were 55 cases in monosyllabic words, 85 cases in disyllabic words, and 133 cases in trisyllabic words. These results indicate that nucleus modification increased as the syllable length became more complex, with trisyllabic words showing the greatest number of occurrences. The data demonstrate that vowel-related modifications were the most dominant pattern identified in students' pronunciation production.

The onset position showed fewer modification cases compared to the nucleus. The findings reveal 3 cases in monosyllabic words, 32 cases in disyllabic words, and 48 cases in trisyllabic words. Similar to the nucleus findings, the number of onset modifications

increased in longer words, particularly in trisyllabic forms. This pattern indicates that onset-related pronunciation modifications were more frequent when students produced words with more complex syllable structures and consonant arrangements.

In the coda position, the findings identified 17 modification cases in monosyllabic words, 47 cases in disyllabic words, and 43 cases in trisyllabic words. Unlike the nucleus and onset positions, coda modification was highest in disyllabic words rather than trisyllabic words. Overall, the data show that the nucleus consistently represented the most frequently modified syllable component, followed by the coda and onset positions. The findings also demonstrate that pronunciation modification occurred more frequently in disyllabic and trisyllabic words than in monosyllabic words across all syllable positions**.**

SLA Perspectives toward the Modification Cases in Pesantren Context

The interview analysis generated several major themes related to students' pronunciation difficulties, namely: (a) spelling–pronunciation mismatch, (b) limited exposure to authentic English, (c) limited vocabulary mastery, and (d) affective factors in speaking performance.

Spelling-Pronunciation Mismatch

The first theme, spelling–pronunciation mismatch, emerged from students' statements indicating confusion between written English forms and their actual pronunciation. Many participants reported that they tended to pronounce words based on spelling because they were unsure of the correct phonological forms. This issue was particularly evident when students encountered unfamiliar vowel sounds, silent letters, or multisyllabic words. It could be interpreted from the following excerpts:

> STD 5
>
> *"Sometimes I read English words exactly like the spelling because I do not know the correct pronunciation."*
>
> STD 9
>
> *"I get confused when the written word is different from how it is pronounced."*

From a second language acquisition (SLA) perspective, this phenomenon reflects learners' dependence on existing linguistic knowledge when processing unfamiliar L2 forms. Because Indonesian has a more consistent relationship between spelling and pronunciation, students transferred this orthographic expectation into English pronunciation. As a result, learners produced interlanguage forms influenced by L1 phonological and orthographic patterns, particularly in vowel production and syllable realization.

Limited Exposure to Authentic English

The second theme identified was limited exposure to authentic English. Although students used English regularly in the pesantren environment, the interviews revealed that most communication occurred among fellow learners with similar linguistic backgrounds and pronunciation patterns. Students stated that they rarely listened to authentic spoken English from native speakers or highly proficient users outside classroom activities.

STD 2

"I usually practice English only with friends in the dormitory, so we often make the same pronunciation mistakes."

STD 4

"I rarely listen to native speakers, so I do not really know how some words should sound."

As a result, exposure to accurate pronunciation models remained limited. From the SLA perspective, this condition illustrates the distinction between input and intake proposed in language acquisition theories. Although learners received English input daily, not all input was successfully internalized into accurate phonological competence. The limited availability of authentic pronunciation models also reduced opportunities for noticing the gap between students' own pronunciation and target-like forms, as explained in Schmidt's Noticing Hypothesis. In addition, the repeated interaction among learners with similar pronunciation patterns may have reinforced non-target-like phonological forms within their interlanguage system.

Limited Vocabulary Mastery

The theme of limited vocabulary mastery also appeared throughout the interviews. Several students explained that unfamiliar vocabulary made them uncertain about pronunciation, especially when producing longer words with complex syllable structures.

STD 1

"When I find a difficult word that I do not know, I feel unsure how to pronounce it."

STD 3

"Long English words are difficult because I do not know where to stress or divide the syllables."

This uncertainty reflects the cognitive demands involved in second language production, where learners simultaneously process vocabulary, pronunciation, and syllable organization during speech production.

Affective Factors in Speaking Performance

Another important theme emerging from the data was affective factors in speaking performance. Many participants expressed feelings of nervousness, embarrassment, and fear of making mistakes when speaking English publicly.

STD 10

"I am afraid my friends will laugh if my pronunciation is wrong."

STD 8

"I feel nervous when speaking English in front of the class because I think my pronunciation is bad."

In SLA theory, these emotional conditions are closely related to Krashen's Affective Filter Hypothesis, which suggests that anxiety and low self-confidence can inhibit language acquisition by limiting learners' ability to process linguistic input effectively. The findings indicate that students often avoided spontaneous oral practice because they feared producing incorrect pronunciation. Consequently, opportunities for meaningful output and pronunciation improvement became more limited. This condition also relates to Swain's Output Hypothesis, which emphasizes that language production plays an important role in helping learners test and restructure their linguistic knowledge. Therefore, the findings demonstrate that pronunciation difficulties among students were shaped not only by phonological interference but also by cognitive processing, environmental exposure, and affective dimensions within the second language acquisition process.

DISCUSSION

The findings indicate that pronunciation modification among Indonesian pesantren learners occurs systematically and reflects the development of learners' interlanguage phonology rather than random pronunciation errors. According to Selinker (1972), interlanguage represents a transitional linguistic system constructed by learners during the process of second language acquisition. In this study, the dominance of substitution across monosyllabic, disyllabic, and trisyllabic words suggests that learners relied heavily on familiar phonological categories when producing unfamiliar English sounds. This finding supports Odlin's (1989) theory of negative transfer, which explains that learners' first language strongly influences second-language pronunciation, especially when the target language contains phonological features absent from the learners' native language. The replacement of English interdental fricatives such as /θ/ with more familiar sounds demonstrates learners' attempts to approximate unfamiliar phonemes using Indonesian phonological patterns.

The frequent occurrence of vowel modification in nucleus positions further indicates that vowel production remains one of the major difficulties for Indonesian EFL learners. English vowels involve distinctions in vowel quality, stress, and reduction that differ significantly from Indonesian phonology. This finding aligns with the Speech Learning Model proposed by Flege (1995), which explains that learners tend to assimilate unfamiliar second-language sounds into the closest phonological categories available in

their first language. The simplification of diphthongs into monophthongs and the replacement of schwa vowels with full vowels indicate that learners preferred stable vowel articulations that resemble Indonesian vowel systems. This phenomenon is consistent with Muñoz Mallén and Vázquez (2019), who argued that learners from syllable-timed language backgrounds often experience difficulty producing reduced vowels in English. Therefore, the dominance of nucleus modification reflects the influence of first-language phonological constraints on English vowel production.

The findings also reveal that epenthesis and cluster reduction were frequently employed to simplify complex English consonant clusters. Indonesian phonology generally favors simple consonant-vowel syllable structures, while English allows more complex consonant combinations. As a result, learners inserted additional vowels or consonants to reduce articulatory difficulty and adapt English syllable structures into forms that fit Indonesian phonotactic patterns. This supports Eckman's (1977) Markedness Differential Hypothesis, which states that structurally marked forms in the target language are more difficult for learners to acquire. The insertion of vowels within consonant clusters demonstrates learners' attempts to repair unfamiliar syllable structures, while cluster reduction indicates the simplification of consonant sequences that are uncommon in Indonesian phonology. These findings suggest that pronunciation modification functions as a phonological adaptation strategy rather than as random deviation.

Another significant finding is the increase of pronunciation modification in longer words. Deletion occurred more frequently in disyllabic and trisyllabic words, indicating that learners experienced greater articulatory and cognitive difficulty when processing more complex lexical items. This finding supports the psycholinguistic perspective proposed by Baddeley et al. (1975), who explained that longer linguistic forms impose heavier demands on working memory. Consequently, learners simplified pronunciation by omitting segments or entire syllables in order to reduce cognitive load during speech production. The deletion of codas and weak syllables demonstrates that learners prioritized intelligibility and fluency over phonological accuracy. In this context, deletion reflects a developmental strategy within interlanguage phonology rather than careless pronunciation. Furthermore, de Jong (2023) argued that limited automaticity in second-language production often leads learners to simplify linguistic output in order to maintain fluency.

Within the pesantren context, the findings suggest that exposure to English alone does not automatically lead to accurate pronunciation development. Although students were exposed to English through muhadharah, classroom discussions, and daily communication practices, non-target-like pronunciation patterns persisted. This phenomenon can be explained through Krashen's (1985) distinction between input and intake. Learners may receive substantial linguistic input, yet not all input is successfully internalized without active cognitive processing. In addition, Schmidt's (1990) Noticing Hypothesis emphasizes that learners must consciously notice the gap between their own pronunciation and the target form in order to improve. The absence of consistent phonological feedback may therefore prevent learners from recognizing inaccurate pronunciation patterns, allowing such forms to persist over time.

The findings also support Swain's (1985) Output Hypothesis, which argues that language production activities alone are insufficient without corrective feedback that encourages learners to modify and restructure their interlanguage system. Although students actively used English in communication, the lack of explicit pronunciation correction may have contributed to fossilization, where inaccurate pronunciation forms become stable and resistant to change (Selinker, 1972). This indicates that pronunciation development requires not only frequent exposure and practice but also focused phonological instruction, corrective feedback, and awareness-building activities. Therefore, effective pronunciation teaching in pesantren contexts should integrate communicative exposure with explicit phonological support to facilitate more accurate pronunciation development.

Overall, the findings demonstrate that pronunciation modification among Indonesian pesantren learners is influenced by multiple interacting factors, including first-language transfer, phonotactic constraints, perceptual assimilation, syllable structure complexity, and cognitive processing limitations. These modifications reflect adaptive strategies within learners' developing interlanguage systems rather than isolated pronunciation errors. By integrating phonological analysis with second language acquisition theory, this study provides a more comprehensive understanding of English pronunciation development in multilingual EFL contexts

CONCLUSION

This study demonstrates that pronunciation modification among Indonesian EFL learners occurs in systematic and predictable patterns, reflecting the development of learners' interlanguage phonology rather than random error production. The findings show that substitution, epenthesis, deletion, and cluster reduction function as adaptive strategies that enable learners to cope with unfamiliar phonological structures in English. Among these, substitution emerged as the most dominant modification type, while nucleus modification was particularly frequent, indicating that vowel production remains a central challenge for learners. From a second language acquisition perspective, these patterns are strongly influenced by first-language transfer, phonotactic constraints, and the assimilation of unfamiliar L2 sounds into existing L1 phonological categories. In addition, the increased occurrence of modification in longer words suggests that cognitive processing limitations play a significant role in shaping learners' pronunciation. These findings confirm that pronunciation development is not only a linguistic process but also a cognitive and developmental one.

Furthermore, the pesantren context reveals that frequent exposure to English does not automatically result in accurate pronunciation. The persistence of non-target-like forms suggests that input alone is insufficient without explicit phonological instruction and corrective feedback, which are necessary to facilitate noticing and restructuring in learners' interlanguage system. Without such support, certain pronunciation patterns may become stabilized over time.

Overall, this study underscores the importance of analyzing pronunciation at the level of syllable structure while integrating insights from SLA theory. By doing so, it provides a more comprehensive understanding of how learners construct and modify their phonological systems, and offers pedagogical implications for improving pronunciation instruction in EFL contexts, particularly in multilingual environments such as pesantren.

REFERENCES

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behavior, 14(6), 575–589.

Chenenje, S. L. (2024). Phonological basis of mispronunciation in spoken English of Kabarasi people in social setting. Journal of Linguistics, Literature, and Language Teaching (JLLLT), 3(2), 60–69. https://doi.org/10.37249/jlllt.v3i2.704

Datu, Y. A. (2025). Achieving speaking proficiency for professionals: Segmental and suprasegmental challenges. IOSR Journal of Humanities and Social Science, 30(1), 40–48. https://doi.org/10.9790/0837-3001024048

de Jong, N. H. (2023). Fluency in speaking as a dynamic construct. Language Teaching Research Quarterly, 37, 179–187. https://doi.org/10.32038/ltrq.2023.37.09

Deyanti, F., & Hindun, H. (2024). Pengaruh muhadarah terhadap kemampuan public speaking di pondok pesantren cendekia amanah. Jurnal Pendidikan dan Sastra Inggris, 4(1), 110–116. https://doi.org/10.55606/jupensi.v4i1.3216

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. Language Learning, 27(2), 315–330. https://doi.org/10.1111/j.1467-1770.1977.tb00124.x

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233–277). York Press.

Gumelar, W. S., & Lestari, T. S. (2024). The phonics method to increase English pronunciation for young learners. Jurnal Simki Pedagogia, 7(2), 350–362. https://doi.org/10.29407/jsp.v7i2.687

He, Y. (2014). Production of English syllable final /l/ by Mandarin Chinese speakers. Journal of Language Teaching and Research, 5(4), 742–750. https://doi.org/10.4304/jltr.5.4.742-750

Inka, A., Lubis, P., Nadia, C., Rahmah, M., & Perangin-angin, A. B. (2024). Analysis of English diphthongs pronunciation errors of senior high school EFL students. Journal of Social Science Research, 4, 6774–6781.

Khoirunnisa, A., & Misnadin, M. (2022). An analysis of pronunciation errors produced by Indonesian learners of English: A case study of English department students of Trunojoyo University. Indonesian EFL Journal, 8(1), 57–70.

Krashen, S. D. (1985). The input hypothesis: Issues and implications. Longman.

Kurniati, N. U. S., Koh, Y. H., & Kim, D. H. (2024). Phonological challenges in Indonesian syllabification: Insight from Korean learners. Language Literacy, 8(2), 517–524. https://doi.org/10.30743/ll.v8i2.10071

Larasati, S., & Pammu, A. (2023). Students' perspective on pronunciation using YouTube during COVID-19 pandemic. ELS Journals on Interdiscplinary Studies in Humanities, 6(4), 676–683.

Muhammad, F. R., & Ashadi. (2019). Behaviorism or Mentalism? the Case of Foreign Language Learning Process in Modern Islamic Boarding Schools. International Journal of Indonesian Education and Teaching, 3(1), 58-66.

Maisyaroh Hasibuan, M., Tanjung, N., & Lubis, Y. (2023). An in-depth analysis of syllable formation and variations in linguistic phonology. Bhinneka: Jurnal Bintang Pendidikan dan Bahasa, 1(3), 220–231. https://doi.org/10.59024/bhinneka.v1i3.225

Mila Adinda, & Ahmad, Z. (2024). The importance of teaching English pronunciation using English in Mind textbook. In Proceedings of the 22nd Atma Jaya Annual Linguistics Conference (KOLITA) (pp. 233–238). https://doi.org/10.25170/kolita.22.5979

Muñoz Mallén, A. M., & Vázquez, V. P. (2019). The improvement of intelligibility in the oral production of standard English: A study on vowel quality in stressed and unstressed syllables. English Language Teaching, 12(4), 115. https://doi.org/10.5539/elt.v12n4p115

Odlin, T. (1989). Language transfer: Cross-linguistic influence in language learning. Cambridge University Press.

Schmidt, R. (1990). The role of consciousness in second language learning. Applied Linguistics, 11(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Selinker, L. (1972). Interlanguage. International Review of Applied Linguistics, 10(1–4), 209–231. https://doi.org/10.1515/iral.1972.10.1-4.209

Shahid, M., & Mahmood, A. (2022). An analytical study of English syllabic structure by EFL learners. Hayatian Journal of Linguistics and Literature, 6(1), 184–206.

Srakaew, P. N. (2021). The importance of pronunciation in English communication. Journal of Teaching English, 2(2), 11–18.

Sulaiman, M., Syahri, I., & Saputri, K. (2022). A contrastive analysis between English and Indonesian copulative verbs. English Community Journal, 6(1), 36. https://doi.org/10.32502/ecj.v6i1.4342

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), Input in second language acquisition (pp. 235–253). Newbury House.