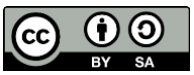
IDEAS

Journal of Language Teaching and Learning, Linguistics and Literature

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Volume 13, Number 2, December 2025
pp. 9468 - 9482

Copyright © 2025 The Author
IDEAS is licensed under CC-BY-SA 4.0 License



Issued by English study program of IAIN Palopo

Vowel and Inflected-Form Realisation in the Recorded Read-Aloud Performance of Indonesian First-Year English Education Students

Mustika Aprilia[1], Andi Naniwarsih[2], Ernitasari Mulyadi[3], Laila Kurniawaty[4]
[1,2,3]Universitas Muhammadiyah Palu
Corresponding E-Mail: andinaniwarsih@unismuhpalu.ac.id

Received: 2025-10-21 Accepted: 2025-12-31
DOI: 10.24256/ideas. v13i2.13065

Abstract

This qualitative descriptive study examined how first-year English Education students at an Indonesian university realised vowels and inflected word forms in recorded read-aloud submissions, and how these realisations related to their reading fluency. The question matters because pre-service teachers' read-aloud performance will serve as a pronunciation model for their future pupils, and errors produced fluently may go unnoticed by both reader and listener. Twelve recordings of a 109-word narrative passage, submitted through Google Classroom, were analysed using error analysis and descriptive fluency profiling. Fluency tended to run ahead of accuracy: six readers were placed higher on fluency than on accuracy, none showed the reverse, and most deviations were produced smoothly without self-correction. Deviations clustered in two areas: word endings carrying inflectional meaning, which were deleted (months → month) or replaced by the base form (left → leave), and vowels in high-frequency words, which were pronounced according to spelling (came → cam, seems → sems). Only the two weakest readers combined low accuracy with word-by-word reading. Teacher education programmes should teach inflected forms and spelling-to-sound vowel patterns explicitly, and assess accuracy and fluency separately so that fluent errors are not overlooked.

Keywords: *read-aloud; pronunciation errors; inflectional morphology; reading fluency; Indonesian EFL learners*

Introduction

Oral reading competence occupies a distinctive place in the preparation of English teachers in Indonesia. Unlike school pupils, who read aloud mainly to show that they can decode print, pre-service teachers will read aloud to their own learners: they will model dialogues, narrate stories and voice instructions, and their

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

realisation of English sounds, stress and intonation will become input for pupils who may have few other spoken English models (Celce-Murcia et al., 2010); (Derwing, T. M. & Munro, 2015). For first-year students in a Faculty of Teacher Training and Education (*Fakultas Keguruan dan Ilmu Pendidikan*, *FKIP*, *Universitas Muhammadiyah Palu*), oral reading is therefore both a language skill and an early professional competence.

Analyzing this competence well is harder than it appears. Live read-aloud activity consumes class time in proportion to class size, depends on a single fleeting hearing by the researcher, and leaves no record that can be revisited for moderation or feedback. Reading aloud in front of peers is also a recognised source of foreign language classroom anxiety (Horwitz et al., 1986), which can depress performance for reasons unrelated to underlying competence. A researcher who hears a reading once can form an overall impression but can rarely pinpoint which segments or phrase boundaries caused difficulty.

Learning management systems alter these conditions. In the study program examined here, Google Classroom is used to set assignments to which students attach an audio file recorded on their own phone or laptop. Students can choose when to record, listen to their attempt, and re-record before submitting, while the researcher receives a durable file that can be replayed, transcribed and analysed segment by segment.

Multimodal social semiotics holds that communication rarely relies on a single mode. Speech, writing, image and gesture each offer particular affordances, and meaning emerges from how they are combined (Jewitt et al., 2016); (Kress, 2010); (Kress, G. & Van Leeuwen, 2001) distinguishes *modes*, the culturally shaped resources for making meaning, from *media*, the material means through which meaning is produced and disseminated. Applied to an audio submission, the written script supplies the linguistic resources, the reader's voice realises them through segmental and prosodic choices, and the recording device and platform constitute the medium. A weak final fricative may reflect a learner's articulation or a low-quality microphone and lossy compression; an assessor who ignores the medium risks attributing one to the other. Table 1 sets out the analytic framework derived from this literature.

Table 1 - Analytic framework for recorded oral reading submissions

Dimension	Analytic focus	Indicators examined	Theoretical anchor
Linguistic mode	Decoding the written text into speech	Grapheme–phoneme mapping; segmental accuracy; inflectional endings; word omissions, insertions and substitutions	Error analysis; L1 phonological transfer
Audio/vocal mode	Vocal realisation of meaning	Phrasing and pausing; intonation and stress; smoothness; pace	Automaticity theory; reading prosody
Media/technology dimension	Conditions of production and transmission	File format and duration; audibility; background noise; traces of editing or restarts; submission compliance	Mode–medium distinction; computer-assisted language learning

Adapted from (Kress, 2010); (Kuhn et al., 2010); (Corder, 1981)

(Corder, 1967; 1981) reframed learner errors as systematic evidence of a developing system rather than as random failure, a system (Selinker, 1972) termed *interlanguage*. In second-language phonology, the Speech Learning Model (Flege, 1995) explains that learners initially perceive new sounds through the categories of their first language: L2 sounds judged similar to L1 sounds are assimilated, while genuinely new sound patterns are slow to be established.

Three features of Indonesian make particular English targets vulnerable. First, Indonesian permits few word-final consonant clusters, so English codas such as /-nθs/ in *months*, /-ndz/ in *friends* or /-ndʒd/ in *changed* invite simplification through deletion or vowel insertion (Jasuli, Sari, I. D. P. & Astuti, 2025); (Prasatyo, 2024). Second, Indonesian marks neither plurality nor past tense through inflectional suffixes, so the final /-s/, /-z/, /-t/ and /-d/ that carry grammatical meaning in English lack a morphological anchor in the learner's L1. Third, Indonesian spelling is relatively transparent, whereas English grapheme–phoneme correspondences are irregular (Celce-Murcia et al., 2010); readers used to a transparent orthography tend to pronounce vowels as they are spelled and to give full vowels to syllables that English reduces.

Not every deviation carries equal weight. (Munro, M. J. & Derwing, 1995) showed that a strong foreign accent does not necessarily reduce intelligibility, and pronunciation teaching has shifted from nativeness towards intelligibility as its benchmark (Levis, 2005). Likewise (Jenkins, 2000) that some features matter more than others for mutual intelligibility among speakers of English as an international language. Final inflectional consonants deserve particular attention in teacher education not because they mark a native accent, but because their omission removes grammatical information that pupils need in order to hear plurality and tense in the teacher's model.

Oral reading fluency is commonly defined as the integration of accurate decoding, automatic word recognition reflected in an appropriate rate, and prosody (Kuhn et al., 2010); (Rasinski, 2004). Automaticity theory (LaBerge, D., & Samuels, 1974) proposes that attention is limited: when word recognition is effortful, it absorbs the resources needed to group words into phrases and project meaning, producing slow, word-by-word reading with flat intonation. Fluency has accordingly been described as a bridge between decoding and comprehension (Pikulski, J. J. & Chard, 2005). Prosody covers pitch variation, stress, timing and, crucially, phrasing, the grouping of words into meaningful units that are marked by pauses and pitch movement at syntactic boundaries. According to (Zutell, J. & Rasinski, 1991) that Multidimensional Fluency Scale operationalises these qualities through the dimensions of phrasing, smoothness and pace, and remains a practical tool for classroom assessment.

For second-language readers, an additional explanation must be considered. Flat or evenly timed delivery may not result only from decoding load; it may also reflect L1 rhythm. Indonesian learners have been reported to carry a syllable-timed

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

tendency into stress-timed English (Jasuli, Sari, I. D. P. & Astuti, 2025), which would produce evenly weighted syllables even when words are recognised quickly. Distinguishing decoding-driven from rhythm-driven prosodic difficulty matters, because the two call for different instructional responses.

Research on the English pronunciation of Indonesian learners converges on three findings. First, segmental difficulty is concentrated in features absent from Indonesian, especially word-final consonant clusters, vowel length, and certain vowels and diphthongs (Firdaus, M. R. & Milal, 2021); (Jasuli, Sari, I. D. P. & Astuti, 2025); (Prasatyo, 2024); among first-year English Education students in Palu, for example, the dental fricatives /θ/ and /ð/ were replaced by /t/ and /d/, an error attributed to mother-tongue interference (Yulianti et al., 2021). Second, this difficulty persists beyond the early stages of learning and into continuous reading: read-aloud recordings show segmental and silent-letter errors attributed to first-language interference and fossilization (Firdaus, M. R. & Milal, 2021); (Senowarsito & Ardini, 2019), and even postgraduate English Education students who already teach in schools mispronounce words in read-aloud texts, most often in consonant clusters (Purnama, S., Pawiro, M. A. & Azis, 2023). Third, inflected forms are a persistent weak point: the simple past was the most frequent grammatical error in the spoken English of Indonesian English majors (Najla, N. & Fatimah, 2020), and speakers of Malay, a language closely related to Indonesian, frequently omitted or misrealised the *-ed* ending (Lotfie et al., 2017).

In parallel, recorded reading has proved a workable medium for assessing and developing fluency, both among Indonesian undergraduates (Sinambela, 2017) and in other EFL contexts (Rodríguez-Fuentes, R. A. & Calle-Díaz, 2023); (Rodríguez-Fuentes et al., 2025). The state of the art is therefore well developed on each side separately: it is known which segments and forms Indonesian learners find difficult, and it is known that recorded reading can capture fluency and prosody.

What these studies have not done is examine the two together. Studies of segmental difficulty, including those based on read-aloud recordings (Firdaus, M. R. & Milal, 2021); (Purnama, S., Pawiro, M. A. & Azis, 2023), describe what learners get wrong but not whether they get it wrong hesitantly or fluently, while fluency studies rate pace and prosody without asking which forms are being produced fluently. Studies of recorded oral tasks in computer-mediated settings, meanwhile, have concentrated on anxiety and learner perceptions (McNeil, 2014) rather than on the linguistic content of the recordings.

The distinction matters in teacher education: an inaccurate form produced hesitantly signals difficulty to both reader and listener, whereas one produced fluently is less likely to be noticed and may become part of the model a future teacher offers pupils. That practising teachers continue to mispronounce words in read-aloud texts (Purnama, S., Pawiro, M. A. & Azis, 2023) suggests that such forms are not reliably corrected during teacher training. What remains to be done is to describe which vowels and inflected forms pre-service teachers realise inaccurately at the

start of their training, and whether these inaccuracies occur in fluent or disfluent reading.

The originality of the present study lies in three features. It analyses segmental accuracy and fluency in the same recordings rather than separately; it focuses on first-year pre-service English teachers, whose read-aloud performance will later serve as a model for pupils; and it draws on routine Google Classroom submissions rather than elicited test data, showing how ordinary coursework can serve diagnostic purposes.

Accordingly, the objective of this study is to describe how first-year English Education students realise vowels and inflected word forms in recorded read-aloud submissions, and how these realisations relate to their reading fluency. To meet this objective, twelve Google Classroom recordings of the same 109-word passage were analysed qualitatively: deviations were identified and categorised through error analysis, each reading was profiled for fluency, and the resulting accuracy and fluency levels were compared across readers. The study was guided by one research question: How do first-year English Education students realise vowels and inflected word forms in recorded read-aloud submissions, and how do these realisations relate to their reading fluency?

Method

Research design

The study adopted a qualitative descriptive design (Miles et al., 2014) organised by the analytic framework in Table 1. The aim was to describe how first-year students realised vowels and inflected forms and how these realisations related to their fluency, not to test causal relationships; the recording medium was examined as a condition shaping what could be heard and judged. Counts are reported only to show how widespread a pattern was within the analysed set; they are not intended for statistical generalisation.

Context and Participants

The study took place in the English Education Study Program of the FKIP at Universitas Muhammadiyah Palu in the first semester of the 2025/2026 academic year. All 12 students in the first-year cohort submitted a self-recorded read-aloud via Google Classroom, and all 12 recordings were analysed. The participants comprised 7 females and 5 males, aged 18–20 years. All used Indonesian as the language of instruction at the university; 9 identified Indonesian as their main language of daily communication, while 3 regularly spoke a regional language (*Kaili*, *Buginese* and *Balinese*) alongside Indonesian at home. To preserve anonymity, participants are identified by the codes SS1–SS12 throughout this paper.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Reading Instrument

All participants read the same 109-word passage: paragraph A of the narrative reading text *University Six Months On* from (Harrison et al., 2006), read as printed in the textbook. The passage contains features relevant to the analysis: word-final consonant clusters (*months*, *friends*, *left*, *changed*); regular and irregular past-tense and plural forms (*learned*, *left*, *came*, *studies*, *differences*); vowels governed by a final silent *e* (*came*, *made*); other common words whose vowels are not predictable from spelling (*seems*, *since*, *fitting*); and a short exclamative (*How time flies!*) that invites marked intonation.

Data Collection

Students received written instructions in Google Classroom to download the text, read it aloud while recording their voice, and upload the recording as an audio file. Because students recorded offline on their own phones or laptops, they could choose when to record, listen to their attempt, and re-record before uploading their preferred version. Files were downloaded from Google Classroom and stored in a password-protected folder.

Data Analysis

Analysis followed the three concurrent activities described by (Miles et al., 2014). During data condensation, each recording was listened to three times by the third author while following the script. For the linguistic mode, every audible deviation from the script was noted with the target word and the realised form, and then coded as a vowel deviation, consonant deviation, final-consonant omission, or word-level miscue (whole-word substitution, insertion or boundary error). Whole-word substitutions in which an inflected word was replaced by its base form (e.g., *leave* for *left*) were additionally flagged as base-form substitutions, because they bear on inflectional morphology.

Target forms are given in broad IPA. Realised forms were transcribed impressionistically, that is, on the basis of the coder's auditory judgement rather than instrumental measurement and are reported as perceived approximations in orthographic form. For the audio/vocal mode, each reading was described in terms of pace, smoothness (hesitations, repetitions and self-corrections), phrasing and intonation, using the dimensions of the Multidimensional Fluency Scale (Zutell, J. & Rasinski, 1991) as sensitising categories rather than as numerical scores. For the media dimension, the file format, audibility, background noise and recording volume of each submission were noted.

Result

Table 2 refers to summarise the evidence for each recording across the three dimensions, together with the accuracy and fluency levels assigned in the analysis. The sections that follow describe the patterns in each dimension and then the relationships among them.

Table 2 - Analysis matrix and assigned levels for the twelve recordings

ID	Linguistic mode (segmental accuracy)	Audio/vocal mode (prosody and fluency)	Media/technology dimension	Accuracy Level	Fluency Level
SS1	Minor omission of final /-s/, /-z/ (*months → month*, *friends → friend*)	Excellent fluency, natural intonation, appropriate chunking	.m4a; clean, no background noise	High	High
SS2	Vowel deviations (*came → cam*, *fitting → fighting*); final sound omission	Steady pace, smooth rhythm, no disruptive hesitation	.mp3; clear	Moderate	Moderate
SS3	Vowel deviations (*since → science*; *made → mady*)	Stable speed, clear articulation, moderate expression	.m4a; good quality	Moderate	Moderate
SS4	Base-form substitution (*left → leave*); *has → hels*	Confident opening, steady pace	.aac; slight background noise, readable	Basic	Moderate
SS5	Accurate; plural and final sounds realised	Natural pitch variation, expressive prosody, consistent chunking	.mp3; high fidelity	High	High
SS6	Consonant and vowel substitution (*months → mounts*); *major → mya-jor*	Calm, consistent tempo	.m4a; clear, minor room echo	Basic	Moderate
SS7	Inflected-form misreadings (*left → lifes*, *learned → learn*)	Regular tempo, clear basic articulation	.mp3; standard, no glitches	Basic	Moderate
SS8	Initial consonant voicing (*six → ziks*); one word inserted	Unhurried, relaxed delivery	.mp3; good volume	Basic	Moderate
SS9	Word-final truncation (*left → leave*; *meant → mean)*	Expressive intonation, natural sentence stress	.m4a; clear	Moderate	High
SS10	Multiple deviations (*months → mats*, *flies → fals*)	Word-by-word reading, frequent hesitation, self-correction	.oga; low volume, slight ambient noise	Low	Emerging
SS11	Vowel deviations in common words (*since*, *came*)	Slow, deliberate, word-by-word	.m4a; clean	Low	Emerging
SS12	Vowel deviation (*seems → sems*); word-boundary confusion	Stable pace, good sentence cadence	.m4a; clear	Basic	Moderate

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

On the accuracy scale, two readings were rated High (SS1, SS5), three Moderate (SS2, SS3, SS9), five Basic (SS4, SS6, SS7, SS8, SS12) and two Low (SS10, SS11). On the fluency scale, three readings were rated High (SS1, SS5, SS9), seven Moderate (SS2, SS3, SS4, SS6, SS7, SS8, SS12) and two Emerging (SS10, SS11). The two distributions differ in shape: fluency clustered at the Moderate level, whereas accuracy was spread more widely and leaned towards the lower levels. Six readers were placed higher on fluency than on accuracy (SS4, SS6, SS7, SS8, SS9, SS12), and none showed the reverse pattern.

The remaining six received matching levels, including the two readers at the lower end of both scales (SS10, SS11). In this set, therefore, control of pace tended to run ahead of segmental accuracy, except among the weakest readers, for whom difficulties in both dimensions coincided. Only SS5 was essentially free of deviations, realising plural and final sounds precisely, and SS1 showed just two minor omissions. Across the other recordings, deviations fell into two broad groups of similar size that often co-occurred within the same reading.

The first group concerned word endings, especially endings that carry inflectional meaning, and appeared in six recordings (SS1, SS2, SS4, SS7, SS9, SS10). It took two forms. Some readers deleted the final consonant: SS1 read *months* /mʌnθs/ as *month* and *friends* /frendz/ as *friend*, SS10 reduced *months* to something close to *mats*, and SS2 omitted the final consonant in *world* /wɜːld/ (realised as *worl*). Others replaced the inflected word with its base form: SS4 and SS9 read *left* /left/ as *leave*, and SS7 read *learned* /lɜːnd/ as *learn*; SS7 also misread *left* as *lifes*. Two words were clear hotspots: *months* was misread in three recordings (SS1, SS6, SS10) and *left* in three (SS4, SS7, SS9). Base-form substitution is not a phonetic simplification. It suggests that readers retrieved the lemma and produced its uninflected form, although the visual similarity of *left* and *leave* may also have contributed; either way, the difficulty lies in the inflected form as much as in pronunciation.

The second group consisted of vowel deviations, found in six recordings (SS2, SS3, SS6, SS10, SS11, SS12). Strikingly, they occurred mostly in short, frequent words rather than in long or unfamiliar ones. *Came* /keɪm/ was read close to *cam* (SS2, SS11), apparently ignoring the final silent *e* that signals the long vowel; *made* /meɪd/ was produced with a final vowel, close to *mady* (SS3), as if the silent *e* were pronounced; *seems* /siːmz/ became *sems* (SS12), with *ee* read as a short /e/; *fitting* /ˈfɪtɪŋ/ was read with the letter-name vowel of *fighting* (SS2); and *since* /sɪns/ was produced as *science* (SS3, with SS11 also showing difficulty on this word).

These are patterns of spelling-to-sound conversion: the readers knew the words, but applied a letter-by-letter reading of the vowel graphemes. The silent *e* was mishandled in both directions, ignored in *cam* and voiced in *mady*. Consonant substitutions were less frequent. SS6 produced *months* as *mounts*, replacing /θ/ with /t/, and SS8 voiced the initial consonant of *six* (*ziks*). Word-level miscues other than base-form substitution were few: SS8 inserted one word and SS12 showed a word-boundary confusion. Only SS10 audibly self-corrected.

Comparing the audio/vocal descriptions yielded three profiles. The *high* profile (SS1, SS5, SS9) was distinguished by explicit expressive features: natural pitch variation, sentence stress placed on content words, and phrase-level chunking. SS5 combined these with consistent chunking throughout; SS9 was notable for expressive intonation and natural sentence stress despite word-final truncation in the linguistic mode.

The *moderate* profile was the largest group (SS2, SS3, SS4, SS6, SS7, SS8, SS12). Its common feature was control of pace: readings were described as steady, stable, calm, regular or unhurried, without disruptive hesitations. What separated these readers from the high profile was the absence of clearly reported expressive variation rather than the presence of disfluency. Expression was described as moderate in SS3 and sentence cadence as good in SS12, while for the other five readers the descriptions concerned pace and articulation only. Brief within-phrase breaks still occurred, such as SS7's hesitation pause between the verb and preposition in *fitted… in* (SS7, 00:09.12).

The *emerging* profile (SS10, SS11) showed word-by-word delivery. SS10's reading included frequent hesitations and self-corrections, and SS11 read slowly and deliberately, articulating words individually rather than grouping them into phrases.

Media/technology dimension: device defaults and uneven recording conditions

Students submitted files in four formats: .m4a (six recordings), .mp3 (four), .aac (SS4) and .oga (SS10). All submissions were audible enough to identify deviations from the script. Nine were clean, while three carried recording problems: slight background noise (SS4), minor room echo (SS6), and low volume with ambient noise (SS10). The weakest reading therefore coincided with the weakest recording, whereas SS11, the other Low/Emerging reader, submitted a clean file.

Relationships across dimensions

Segmental accuracy and fluency did not move in lockstep. Two readers were strong on both scales (SS1, SS5) and two were weak on both (SS10, SS11), but SS9 combined high fluency with moderate accuracy, and five readers with a steady, controlled pace (SS4, SS6, SS7, SS8, SS12) were placed in the Basic accuracy band. Recording quality conditioned the evidence available without determining the outcome: mild noise in SS4 and SS6 did not prevent analysis, and the two Low/Emerging readers submitted, respectively, a poor recording (SS10) and a clean one (SS11).

Discussion

The deviations cluster around two sources that both reflect contrasts between Indonesian and English. The first is inflectional morphology. Final consonant deletion in *month* and *friend* is consistent with the L1 transfer of a syllable structure that permits few final clusters (Flege, 1995); (Jasuli, Sari, I. D. P. & Astuti, 2025);

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

(Prasatyo, 2024). The base-form substitutions (*leave* for *left*, *learn* for *learned*) point further: because Indonesian marks neither tense nor number on the word, readers who recognised the lemma may have produced it in the uninflected form.

Error analysis treats such regularities as evidence of an interlanguage rule rather than as carelessness (Corder, 1981); (Selinker, 1972). That these errors appeared in fluent readers (SS1, SS9) suggests that they will not disappear as reading becomes faster and may need explicit attention before they stabilise.

The second source is orthography. Readers accustomed to Indonesian's relatively transparent spelling read English vowel graphemes letter by letter, giving *came* a short vowel and *seems* a short /e/ (Celce-Murcia et al., 2010). That these deviations occurred in high-frequency words rather than difficult vocabulary indicates a gap in grapheme–phoneme knowledge, not in word knowledge. Consonants absent from Indonesian were a smaller source of difficulty: SS6's /t/ for /θ/ in *months* matches the substitution pattern reported among first-year English Education students in Palu (Yulianti et al., 2021), and also the difficulties of junior high school learners in the same city (Naniwarsih & Muniar, 2019).

The two clusters of difficulty identified here are consistent with earlier studies of Indonesian learners. As in (Firdaus, M. R. & Milal, 2021), vowel and silent-letter errors appeared in read-aloud data and are plausibly explained by first-language interference; the present study adds that these errors occurred mainly in high-frequency words and in both directions of the silent *e* rule, which points to grapheme–phoneme knowledge rather than vocabulary. The inflection findings parallel the frequent omission or misrealisation of *-ed* among Malay speakers (Lotfie et al., 2017) and the dominance of simple-past errors in the speech of Indonesian English majors (Najla, N. & Fatimah, 2020).

They differ, however, in that the present readers had the inflected form in front of them: base-form substitutions such as *leave* for *left* occurred even though the task did not require students to produce the form themselves. Unlike (Purnama, S., Pawiro, M. A. & Azis, 2023), whose postgraduate participants erred most often in consonant clusters, cluster simplification here was only one of several sources of difficulty, possibly because first-year students also lack secure spelling-to-sound knowledge for common vowels. The main contribution of this study lies elsewhere: whereas these earlier studies report what learners get wrong, the present analysis shows that most errors were produced fluently and without self-correction, a pattern that earlier error analyses could not reveal because they did not examine fluency alongside accuracy.

Three further points qualify this interpretation. First, regional languages used at home may shape these patterns alongside Indonesian, as (Senowarsito & Ardini, 2019) showed that for university students whose first language is Javanese. Although Indonesian is the medium of instruction for all twelve participants, the three students who regularly speak *Kaili*, *Buginese* or *Balinese* at home may show variations influenced by the phonology of those languages, and factors other than

transfer, such as anxiety when reading English, may also play a role. Second, some final-consonant losses may reflect connected-speech elision, which fluent speakers produce in consonant clusters before another consonant (Roach, 2009), although base-form substitutions such as *leave* for *left* cannot be explained this way. Third, because realised forms were recorded impressionistically rather than transcribed acoustically, fine distinctions, such as whether *cam* reflects a vowel quality or a vowel length difference, cannot be settled from these data. Such forms rarely impede intelligibility in context (Munro, M. J. & Erwing, 1995); their importance lies in the model future teachers provide to pupils.

The emerging profile fits automaticity theory: effortful word recognition absorbed attention, and phrasing broke down into word-by-word delivery (LaBerge, D., & Samuels, 1974); (Pikulski, J. J. & Chard, 2005). The more instructive pattern is the moderate group, whose members read at a controlled pace while producing Moderate or Basic segmental accuracy, together with the high-profile readers who combined expressive delivery with inflectional errors. In these readings, deviant forms were produced fluently and, except for SS10, without self-correction. This suggests that the deviations are not momentary decoding failures but forms that have already become automatic. In (Kuhn et al., 2010) terms, these readers have rate without full accuracy. For teacher education this is the more serious finding, because a smoothly delivered *leave* for *left* is less likely to be noticed by either the student or a listener than a hesitant one.

The data do not allow strong claims about intonation in the moderate group. Their descriptions concern pace and articulation, so the absence of reported expressiveness cannot be equated with flat pitch. Whether some of these readers carry a syllable-timed rhythm into English, as reported for Indonesian learners (Jasuli, Sari, I. D. P. & Astuti, 2025), is a hypothesis for acoustic analysis rather than a finding of this study.

The recording medium mattered more than a technical check would suggest. Most students used device-default formats that were adequate for analysis, but the weakest reading came with the weakest recording. Poor conditions may have made segments harder to judge, or less engaged students may have taken less care over both; these explanations cannot be separated here, and word-final sounds in low-volume or noisy files should be treated with caution. Because students were permitted to re-record, each file is best read as a curated best performance rather than a sample of typical reading. Calm delivery should not be read as evidence of reduced anxiety, which was not measured; (McNeil, 2014) found that asynchronous oral environments both relieve and create anxiety. The clearest contribution of the medium is evidential: without a durable record, the case-by-case comparison of accuracy and fluency reported here would not have been possible.

Four implications follow for English teacher education. First, pronunciation instruction for first-year English Education students should treat inflected forms explicitly, linking sound and meaning through pairs such as *month–months* and

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

learn–learned–left/leave. Second, spelling-to-sound rules for vowels, including the final silent *e* (*cam–came*, *mady–made*) and vowel digraphs (*sems–seems*), deserve direct teaching even at university level, using the high-frequency words in which errors actually occurred. Third, rubrics for recorded reading should score accuracy and fluency separately, since a holistic grade would have hidden the difference between SS1 and SS9, who read equally expressively but with different levels of accuracy. Fourth, Google Classroom's private comments allow time-stamped feedback on specific words, and asking students to listen to their own recording and mark their deviations before receiving feedback could make fluent errors visible to the students who produce them; simple recording guidance would also make the evidence more reliable. The models learners hear matter here: in Palu, vocational students who practical pronunciation with songs from YouTube videos improved more than a control group (Irmayanti et al., 2025) which illustrates how strongly learners' pronunciation is shaped by the models they imitate.

For the future teachers or lecturers, whose own read-aloud performance will become such a model, accurate realisation of vowels and inflected forms is therefore a professional rather than merely personal concern.

Conclusion

This study examined how first-year English Education students realised vowels and inflected word forms in recorded read-aloud submissions, and how these realisations related to their reading fluency. The central finding is that fluency tended to run ahead of accuracy. Two readers were strong on both scales (SS1, SS5) and two were weak on both (SS10, SS11), but six were placed higher on fluency than on accuracy, and none showed the reverse pattern. Deviations occurred in two areas: endings of inflected words, which were deleted or replaced by the base form, and vowels in common words, which were pronounced according to spelling. Because most of these deviations were produced smoothly and without self-correction, they appear to be established forms rather than momentary decoding failures.

For English teacher education, this means that a fluent-sounding reading can conceal inaccurate pronunciation models. Inflected forms and spelling-to-sound vowel patterns should be taught explicitly, and recorded reading tasks should assess accuracy and fluency separately. Recorded submissions are useful for this purpose: they provide a durable record that allows fluent errors to be identified, and students can be asked to listen to their own recordings and locate their deviations before receiving feedback.

Several limitations qualify these conclusions. The analysis covered a single cohort of twelve students reading one short passage, so the patterns are not generalisable. Recordings were coded by one researcher, and realised forms were transcribed impressionistically rather than acoustically, so fine distinctions such as vowel quality versus vowel length cannot be resolved. Because students could re-record, each file represents a curated best performance rather than typical reading.

Future studies should use independent second coding, acoustic analysis, and both read-aloud and spontaneous speech tasks.

References

Celce-Murcia, M., Brinton, D. M., &, & Goodwin, J. M. (2010). Teaching pronunciation: A course book and reference guide (2nd ed.). Cambridge University Press.

Corder, S. P. (1981). The significance of learners' errors. International Review of Applied Linguistics in Language Teaching, 5(4), 161–170.

Derwing, T. M., &, & Munro, M. J. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. (J. Benjamins. (ed.)).

Firdaus, M. R., &, & Milal, A. D. (2021). Segmental and suprasegmental mispronunciations made by EFL learners in Indonesia. Vivid: Journal of Language and Literature, 10(2), 52–62. https://doi.org/https://doi.org/10.25077/vj.10.2.52-62.2021

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, 233-277.

Harrison, L., Cushen, C., & Hutchinson, S. (2006). Achieve IELTS 2 English for Onternational Education (Workbook). Marshall Cavendish Education.

Horwitz, E. K., Horwitz, M. B., &, & Cope, J. (1986). Foreign language classroom anxiety. The Modern Language Journal, 70(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Irmayanti, I., Mukrim, M., Wahyudin, W., &, & Hasyim, Z. (2025). The Effectiveness of Songs of Youtube Videos in Improving Pronunciation skill of the Grade Eleventh Students at SMK Negeri 2 Palu. E-Journal of ELTS (English Language Teaching Society, 12(2), 746–757. https://doi.org/10.22487/elts.v12i2.3825

Jasuli, Sari, I. D. P., &, & Astuti, E. S. (2025). Challenges in English phonology for Indonesian EFL learners with a focus on consonant clusters, vowel length, and rhythm. Prosiding Seminar Nasional IKIP Budi Utomo, 6(1), 25–35. https://ejurnal.uibu.ac.id/index.php/prosiding_penelitian/article/view/2456

Jenkins, J. (2000). The phonology of English as an international language. Oxford University Press.

Jewitt, C., Bezemer, J., &, & O'Halloran, K. (2016). Introducing multimodality. Routledge.

Kress, G., &, & Van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. Arnold.

Kress, G. (2010). Multimodality: A social semiotic approach to contemporary communication. Routledge.

Kuhn, M. R., Schwanenflugel, P. J., &, & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. Reading Research Quarterly, 45(2), 230–251.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

processing in reading. Cognitive Psychology, 6(2), 293–323.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. TESOL Quarterly, 39(3), 369–377.

Lotfie, M. M., Abdul Kadir, M., &, & Pilus, Z. (2017). –Ed allomorphs and linguistic knowledge of Malay speakers of English: A descriptive and correlational study. Indonesian Journal of Applied Linguistics, 7(2), 445–454. https://doi.org/10.17509/ijal.v7i2.8353

McNeil, L. (2014). Ecological affordance and anxiety in an oral asynchronous computer-mediated environment. Language Learning & Technology, 18(1), 142–159.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE Publications.

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.

Munro, M. J., &, & Erwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. Language Learning, 45(1), 73–97.

Najla, N., &, & Fatimah, S. (2020). Common grammatical errors in oral communication made by Indonesian EFL learners. Journal of English Language Teaching, 9(4), 740–749. https://doi.org/10.24036/jelt.v9i4.110341

Naniwarsih, A. &, & Muniar. (2019). Improving Students' Pronunciation through Reading Aloud Strategy. Journal of Foreign Language and Educational Research, 2(July), 25–33. https://doi.org/10.31934/jofler.v2i2.1350

Pikulski, J. J., &, & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. The Reading Teacher, 58(6), 510–519.

Prasatyo, B. A. (2024). Analyzing mispronunciation patterns of alveolar and velar-alveolar consonant clusters: A comprehensive investigation. Beyond Linguistika, 6(2). https://doi.org/doi.org/10.36448/bl.v6i2.3584

Purnama, S., Pawiro, M. A., &, & Azis, A. (2023). Detecting mispronunciations of non-native (L2) post-graduate students of English language education in Indonesia. International Journal of Language Education, 7(1), 134–142. https://doi.org/10.26858/ijole.v1i1.36368

Rasinski, T. V. (2004). Assessing reading fluency. Pacific Resources for Education and Learning.

Roach, P. (2009). English phonetics and phonology: A practical course. In Cambridge University Press. ((4th ed.).).

Rodríguez-Fuentes, R. A., &, & Calle-Díaz, L. (2023). Fluency development beyond speech rate: A study on the effects of read aloud protocols on EFL learners. SSRN Electronic Journal, 11. https://doi.org/10.1016/j.amper.2023.100133

Rodríguez-Fuentes, R. A., Ferrer Ariza, E., &, & Linero, J. (2025). Digital reading aloud protocols and its effects on EFL fluency. System, 133,.

https://doi.org/10.1016/j.system.2025.103712

Selinker, L. (1972). Interlanguage. International Review of Applied Linguistics in Language Teaching, 10(3), 209–231.

Senowarsito, &, & Ardini, S. N. (2019). Phonological fossilisation of EFL learners: The interference of phonological and orthographic system of L1 Javanese. 3L: The Southeast Asian Journal of English Language Studies, 25(2), 74–85. https://doi.org/https://doi.org/10.17576/3L-2019-2502-06

Sinambela, S. E. (2017). Prosody as a tool for assessing reading fluency of adult ESL students. Advances in Language and Literary Studies, 8(6), 83–87. https://doi.org/10.7575/aiac.alls.v.8n.6p.83

Yulianti, P., Rita, F., & Hasyim, Z. (2021). An Analysis of Error in Pronunciation Dental Fricative Sounds by the First Year Students of Engish Education Study Program. Journal of ELTS (English Language Teaching Society), 9(2), 172–182. https://doi.org/10.22487/elts.v9i2.1874

Zutell, J., &, & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. Theory Into Practice, 30(3), 211–217.