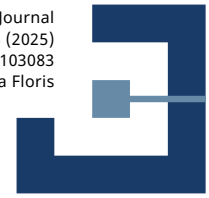
The JALT CALL Journal
ISSN 1832-4215, 21(3), 103083 (2025)
https://doi.org/10.29140/jaltcall.v21n3.103083



Castledown

From words to pixels: Artificial intelligence struggles with world Englishes

Flora Debora Floris
debora@petra.ac.id
Petra Christian University, Indonesia

This study examines how DALL-E 3 interprets English descriptions written by Indonesian university students. Sixteen descriptive texts were submitted to the artificial intelligence (AI) tool, and the resulting images were compared to original photos. Most outputs showed clear mismatches. The analysis found that misinterpretations originated from two main sources: grammatical and vocabulary patterns reflecting Indonesian English and broader stylistic choices, such as the use of vague, emotional, or abstract language. The study also found that a high level of concrete detail could often mitigate the negative effects of non-standard grammar. The findings suggest that current AI tools are not yet equipped to fairly process the full range of human linguistic variation, from local English features to the stylistic patterns of human-centric writing. To support more inclusive use of AI in education, this study adapts the established concept of intelligibility into the idea of "digital intelligibility," and recommends improving training data and creating classroom space for open discussions about AI bias toward language diversity and stylistic choices.

Keywords: World Englishes, language variation, AI bias, text-to-image generator, intelligibility

Introduction

The rapid development of artificial intelligence (AI) has changed many aspects of people's daily life. One important development is the rise of generative AI systems that can turn written text into images. A well-known example is OpenAI's DALL-E 3, which creates detailed images based on text prompts. While the system appears to be highly advanced, its output depends heavily on how closely the user's input matches the language patterns found in its training data.

As these tools become more widely used in education, concerns have emerged about their ability to understand the many different English varieties used around the world. From a sociolinguistic standpoint, these varieties are part of what scholars call World Englishes or localized forms that reflect

the cultural, social, and linguistic backgrounds of non-native speakers (Kachru et al., 2009). One such emerging variety is Indonesian English, which shows distinct patterns in vocabulary, grammar, and pronunciation shaped by local linguistic patterns. However, most AI models are trained mainly on "standard" English, such as American English (Payne et al., 2024), which may limit their ability to process localized inputs effectively.

Some existing studies (e.g. Bailey et al., 2025; Jackson et al., 2024) have confirmed that AI systems often struggle with non-standard English varieties. However, despite these findings, little attention has been given to how such AI systems respond to non-standard Englishes in real classroom contexts, especially in multilingual settings where students regularly use localized forms of English.

Many studies in English Language Teaching (ELT) focus on the practical pedagogical use of AI in classrooms, for example, as tools for writing support, conversational practice, and language learning enhancement. Some recent studies, as listed in Floris (2025), have examined both the potential benefits and limitations of implementing such tools in ELT contexts. However, these studies do not address how well AI tools understand the kinds of English students actually use.

To address this gap, the present study explores how DALL-E 3 interprets English descriptions written by Indonesian university students. It focuses on whether features such as local vocabulary, cultural references, and Indonesian-influenced word order affect the AI tool's ability to generate accurate visual representations. This study treats AI tools as "reader" and asks what happens when it encounters English that departs from standard norms.

Importantly, this study does not treat the students' English as "incorrect." Instead, it follows the view that many of these features reflect natural patterns in emerging localized varieties such as Indonesian English, even if this variety is not yet formally codified. In this sense, the goal is not to correct student writing but to examine how AI systems respond to variation that falls outside dominant norms.

This study investigates two key questions:

- How accurately does DALL-E 3 interpret descriptions written in Indonesian English?
- Which language features of Indonesian English lead to misinterpretation in the generated images?

To frame this research within current academic discussions, the next section reviews relevant literature on World Englishes, Indonesian English, and AI's engagement with English language diversity.

Literature review

World Englishes

World Englishes refers to the many different varieties of English used around the world. The concept, introduced by Kachru (1992), challenges the dominance of native-speaker norms by recognizing localized varieties, such as Indian English, Nigerian English, and Singaporean English. These varieties follow their own linguistics patterns and rules to meet the communicative needs of their users, reflecting their local identities, values, cultures and social settings. The World Englishes concept encourages people to see English as a global language with many legitimate forms, not just one correct version, such as American or British English.

Scholars working within the World Englishes framework often draw on the concepts of intelligibility, comprehensibility, and interpretability to explain how communication functions across different English varieties. Intelligibility refers to one's basic ability to recognize words or utterances, focusing mainly on pronunciation and phonological clarity (Smith & Nelson, 1985). Comprehensibility involves understanding the meaning of those words in context (Matsuura, 2007). Interpretability refers to the ability to grasp the intended meaning, which includes pragmatic and cultural cues (Kachru, 2008). These concepts help explain why speakers of different Englishes can still communicate successfully, even if their grammar, vocabulary, or pronunciation differs. This success is often made possible through shared context, willingness to adjust, and cultural awareness.

However, such human flexibility does not easily translate to AI systems. Unlike people, AI tools rely on fixed training data and statistical patterns to interpret language. As a result, they may struggle to process English varieties that do not match the norms they were trained on, which are typically based on forms like American English (Payne et al., 2024). Understanding this limitation sets the stage for examining how DALL-E 3 handles Indonesian English.

Indonesian English

Indonesian English has developed its own unique characteristics due to Indonesia's diverse language environment and education system, where English is taught as a compulsory subject in high schools. Although this variety has not yet been formally codified as a recognized variety, recent studies (e.g., Daeng & Weda, 2019; Endarto, 2020; Hamsa & Weda, 2019; Mahmud & Slabakova, 2020; Martin-Anatias, 2018; Mulder & Marentek, 2017; Percillier, 2015; Pipit & Rahyono, 2020; Syam et al., 2024) have shown that it displays consistent linguistic features shaped by local usage and sociolinguistic context. It has distinct features in vocabulary, grammar, and pronunciation.

Vocabulary. One of the most notable aspects of Indonesian English is its distinctive vocabulary. According to Endarto (2020), the lexical features of this variety can be categorized into four groups, namely: (1) loanwords of Indonesian

origin, (2) words shaped by semantic shifts, (3) words influenced by morphological changes, and (4) distinctive collocations.

Many Indonesian or local terms have been borrowed directly into English use. These borrowed words often describe cultural concepts, objects, or daily practices that do not have exact matches in standard English (Endarto, 2020). For example, cultural and religious terms such as *Galungan* (Balinese Hindu festival), *Nyepi* (Balinese Hindu Day of Silence), *Waisak* (Vesak or Buddha's Day) are widely used in Indonesian English. Titles and addressing terms like *Bu* (Madam) or *Pak* (Sir) are also included, as well as local food names such as *tumpeng* (a cone-shaped rice dish served with side dishes), *rendang* (a slow-cooked spicy beef dish), and *rawon* (a black beef soup from East Java). Some loanwords, like *batik*, *gamelan*, and *orangutan*, are already integrated into international English dictionaries (Endarto, 2020).

In addition to borrowing, Indonesian English often involves changes in meaning. Words familiar to native English speakers may carry different interpretations or cultural connotations in Indonesian contexts (Endarto, 2020). For instance, the word *terminal* is used to mean *bus station* in Indonesian English, while in standard English it usually refers to a specific part of an airport or station.

Morphological adaptations are also common. Indonesian English speakers may modify English word forms by adding local prefixes or suffixes, aligning them more closely with Indonesian linguistic norms (Endarto, 2020). For example, *staff* is frequently used as a countable noun referring to *one employee*, whereas in standard English *staff* is typically uncountable, meaning *all employees*.

Furthermore, Indonesian English has creative lexical blending and hybrid word combinations, especially in informal communication. Speakers frequently mix Indonesian and English elements, especially to express emotions or interpersonal feelings (Endarto, 2020; Martin-Anatias, 2018). Distinctive collocations also appear, such as *discuss about* or *explain about*, which directly reflect Indonesian usage where verbs like *berdiskusi* (discuss) or *menjelaskan* (explain) are normally followed by the preposition *tentang* (about).

Grammar. In addition to vocabulary, Indonesian English shows distinct grammatical features. One of the most noticeable differences is how speakers express time through verbs. Since Indonesian lacks verb conjugation to show when something happened (past, present, or future), Indonesian English users often encounter difficulties with English verb tenses. They might say I go yesterday instead of I went yesterday or use tenses inconsistently because these patterns are absent in Indonesian (Pipit & Rahyono, 2020). In contrast, standard English uses specific verb inflections, such as adding -ed or -s, to clearly mark time and subject agreement.

Sentence structure in Indonesian English generally follows the basic subject-verb-object (SVO), which is similar to English. However, the influence of Indonesian grammar can result in subtle shifts in word order or clause construction (Daeng & Weda, 2019). For example, in Indonesian the head of a noun

phrase usually precedes the modifier, while in English the modifier normally comes before the head. This structural difference can influence Indonesian English sentence patterns, making them diverge from standard English norms (Daeng & Weda, 2019). While these differences might not affect basic understanding, they can make the sentence sound unusual to other speakers of English or AI systems trained on standard structures.

Code-switching is one feature of Indonesian English, referring to the mixing of Indonesian and English words in the same sentence. This practice is widely observed among Indonesians using English, particularly youngsters and public figures (Endarto, 2020). For example, someone might say *I suka this movie*, blending English with the Indonesian word *suka* (meaning *like*). This mixing is especially common with nouns and often reflects social or emotional expression (Martin-Anatias, 2018). Standard English, by contrast, tends to maintain language separation, switching between complete sentences rather than mixing within sentences.

Indonesian English also shows non-standard grammar patterns. Percillier (2015) reports frequent verb deletion, including copular, lexical, and even modal verbs, as in *The beach so far away from here* where the copula *is* is omitted. Indonesian English also displays relative pronoun deletion, producing forms such as *the first table Ø will finish it.* Mahmud and Slabakova (2020) note that Indonesian English often shows omission of inflections, low use of the third-person singular *-s*, and omission of copula or auxiliary be. They also observe overuse of bare verb forms in contexts that require finite verbs. Examples include *she go to school, she eat a mango, he Ø happy, the dog want come*, and *she say something*. These features can differ significantly from other Southeast Asian Englishes like Singaporean or Malaysian English (Percillier, 2015). They also contrast with standard English, which generally has verb endings, copula, and auxiliary verbs.

Another significant feature involves number marking and plural formation. Mulder and Marentek (2017) report that Indonesian English speakers often leave the head noun in a noun phrase unmarked for plurality when other elements in the phrase, the discourse, or the context already indicate plurality. This reflects the influence of Indonesian, where plurality is usually implied rather than marked. Endarto (2020) adds that Indonesians transfer morphological features from their first language into English, as in the case of *staff* being used as a countable noun (*a staff, staffs*), while in standard English it is normally uncountable. Standard English, by comparison, has rules that distinguish between count and mass nouns, and generally adds *-s* or *-es* to mark plurals, although irregular forms also exist.

Another pattern involves adjective placement. In Indonesian, adjectives typically come after the noun, such as *house big* instead of *big house*. This word order directly reflects Indonesian grammar, where modifiers follow the noun they describe (Hamsa & Weda, 2019). Standard English, in contrast, places adjectives before the main noun in a fixed order.

Indonesian English also shows some differences in the use of prepositions. Mulder and Marentek (2017) explain that this comes from the influence of

Indonesian grammar. Sometimes prepositions like on, in, and at are used in the same way, because their Indonesian equivalents are also used interchangeably. In other cases, a preposition may be left out because its meaning is already clear from the verb or noun, which is common in Indonesian. There are also cases where one preposition is used instead of another, such as *with* instead of *from* (e.g. *separated with*), following the Indonesian preposition *dengan* (with). Mulder and Marentek (2017) also note that Indonesian English users may add a preposition even when standard English would not. For example, *on* may be added after *discuss* (e.g. *discuss a lot on this subject*) to emphasize the topic. Finally, prepositions can be used with a different meaning from standard English, such as *of* in *elected of the people* where standard English would use *by*, or *on* in setting ambitious targets on emission reduction where standard English would normally use *for*.

Pronunciation. Indonesian English also has distinct pronunciation patterns, such as simplified consonant clusters, reduced diphthongs, and flat intonation (Syam et al., 2024). While these features are part of the variety, they are not included in this study because the focus is on how written language, not speech, influences AI interpretation.

To sum up, Indonesian English features unique vocabulary, grammar, and pronunciation patterns shaped by the local language and cultural context. These features make Indonesian English noticeably different from other English varieties.

AI and English language diversity

AI models are often trained in a wide range of languages, which helps them understand and process different linguistic structures and vocabularies (Nicholas & Bhatia, 2023). However, this training typically emphasizes distinct languages, such as Spanish, Arabic, or Mandarin, rather than the varieties within a single language such as Nigerian English, Singaporean English, or Indonesian English.

Non-standard English varieties are often missing because the training data is mainly taken from the Web, where certain dominant forms of English appear more often. As Caines et al. (2025) point out, the training methods used in these systems tend to prioritize frequent language patterns. This means that language forms used more often are treated as more acceptable, while less common varieties, such as regional or localized Englishes, are underrepresented or overlooked.

This gap in representation has practical consequences. AI tools often misinterpret or ignore features of English varieties and treat them as errors that should be corrected. For instance, Jackson et al. (2024) found that GPT-4 performed inconsistently when handling Trinidadian English Creole (TEC). While it could translate pronouns and basic grammar fairly well, it failed to recognize many vocabulary items. Only 39% of TEC-generated sentences were rated fully

grammatical by GPT-4. In many cases, the AI system classified TEC as incorrect English and attempted to revise it into standard forms.

Similar results appear in writing contexts. Lee et al. (2025) showed that ChatGPT-4 defaulted to Native English norms even when asked to produce essays in Korean English. It ignored typical features such as borrowings like "hand phone" (commonly used in Korean English) and topic-prominent word order or selective subject omission, which are natural to Korean discourse. The tool produced simplified standardized forms; and when giving feedback on learner texts, it focused only on standard grammar and vocabulary. ChatGPT-4 did not consider whether the text was communicatively effective in its local context or reflected Korean cultural identity.

Similar bias appears in speech-based systems. Bailey et al. (2025) reported that speech-based AI systems, such as ASR (automatic speech recognition), tend to underperform when interacting with English varieties like African American English (AAE), especially in educational assessments. These shortcomings are often due to differences in pronunciation, word choice, or syntax as these features that were not adequately represented in the training data. As a result, students who use AAE received inaccurate feedback or low scores in AI-supported learning platforms.

The "communication" problems between users and AI tools highlight a deeper issue. Unlike human listeners, AI tools cannot ask for clarification or use context to figure out what a speaker means when the language is unfamiliar. Instead, AI expects the input to match what it already knows. As Harris et al. (2024) explain, this puts the burden on users, especially those who use non-standard English, to change how they speak or write in order to be understood, rather than the AI systems adapting to them.

People often try to overcome these limits by adjusting how they speak to AI. Thomas et al. (2020) found that users tend to simplify their language, change word choices, or modify their tone to match what the system can understand. These adjustments help avoid confusion but also show how much users must work to make the system function smoothly.

In educational settings, these misunderstandings can affect learning outcomes. Students may receive poor or inaccurate feedback because their English does not match the model's expectations. When these patterns are framed as "mistakes," it risks reinforcing the belief that only standard English is acceptable. This study pushes back against that view by framing localized English as legitimate language use.

As Sartori and Theodorou (2022) argue, AI tools are not neutral; they reflect the choices and assumptions built into them. When used in classrooms, these tools can increase the risk of unfair treatment. Students whose English reflects local norms may receive inaccurate feedback or lower scores, not because their writing is unclear, but because the AI system does not recognize their way of using English. This is particularly relevant in multilingual settings like Indonesia, where students' English often reflects local influences.

These challenges emphasize the need for more inclusive AI design. Systems that aim to support language learning must be able to recognize and respond to

the many forms of English in use around the world. Without this ability, there is a risk that AI tools will continue to misunderstand students and reinforce narrow ideas of what English should look like.

Method

Participants

This study was conducted during a regular Writing 1 course at a private university in Indonesia. All sixteen participants were first-year English-major students enrolled in the course. The participants' English proficiency levels were determined by their university entrance exam results. Based on these results, ten students were at the intermediate level, four at the pre-intermediate level, and two at the advanced level. The university used these exam results to divide students into classes, and the present Writing 1 class was one of the two parallel classes formed in this way. The researcher did not determine the grouping of students. Therefore, the varying proficiency levels are reported here as background information only, not as a factor analyzed in relation to their writing performance.

Data collection

The Writing 1 course focused on paragraph writing, and students were introduced to both descriptive and narrative paragraph structures. This research was integrated into one of the required assignments, i.e. writing a descriptive paragraph.

At the time of the study, the researcher was also the students' teacher. This dual role is common in classroom-based research and allows for deeper insight into learners' behaviors, language use, and task engagement (Hopkins, 2014). It enables the researcher to observe naturally occurring learning processes and respond to students' needs during the task while ensuring consistency in instruction and classroom support.

To address ethical considerations, students of the present Writing 1 class were clearly informed that their writing might be used for research purposes, that their participation in the research was voluntary, and that it was unrelated to their grades. All gave verbal agreement and written consent. In this study, all student names were replaced with pseudonyms to protect their identity.

The assignment asked each student to take a photograph of a familiar location on campus, using their smartphones. In the following class sessions, these students used the photo to write a 100–150-word descriptive paragraph. Students were instructed to write for a human audience and were not informed that their texts would be used as AI prompts. The writing process took place over two 120-minute class meetings. All stages of the writing task were completed in class, including outlining, drafting, editing, and submitting the final version.

As part of the classroom practice, the students were not allowed to use

digital devices during the writing sessions but were permitted to use printed dictionaries. This decision was supported by research showing that printed dictionaries can lead to better vocabulary retention than electronic alternatives, particularly when learners spend more time actively engaging with word forms and meanings (Chiu & Liu, 2013). This approach was also consistent with the course's pedagogical focus on building foundational writing skills and promoting language awareness.

The class was designed to be a supportive space where these students could learn comfortably, ask questions, and work with each other without fear of making mistakes. In addition to peer support, the students were also welcome to ask the teacher for clarification and guidance as they worked through each stage of the task.

Data analysis

This study's analysis was designed to explore how DALL-E 3, treated as a non-human "reader", interpreted descriptive paragraphs written by English language learners. Once the final descriptions were collected, each student writing was entered verbatim by the teacher (the researcher) into the DALL-E 3 text-to-image generator using the default web interface. No edits or additions were made to the input, so the writing stayed in its original classroom form.

This procedure avoided researcher bias. However, this also created a key methodological consideration. Because students were instructed to write a descriptive paragraph for a human audience and were unaware of the AI, it was anticipated that any mismatches could originate from various sources within the texts, such as the documented features of Indonesian English, and the broader stylistic choices that are a natural part of writing for people.

Once the AI-generated images were produced, they were printed out to be compared with the students' original photographs. During the comparison session, each student's original photo was displayed alongside the corresponding AI-generated image. To evaluate how accurately the AI interpreted the written descriptions, two independent raters reviewed each image pair. Both raters were experienced writing teachers, each with over 15 years of teaching experience, and were colleagues of the researcher from the same department and university. The researcher facilitated the process but did not participate in assigning outcome labels. Neither of the raters was involved in guiding the writing task.

Each rater independently assigned one of the following outcome labels:

- Close match – the AI image captured the setting, key objects, and general mood.
- Partial match – the scene was recognizable but included one or two noticeable inaccuracies.
- Mismatch – the AI image showed major errors or unrelated content.

Each rater was also asked to provide a brief justification for their decision. In cases where the assigned label differed, the raters would have discussed their

reasoning to reach agreement. However, in this study, no major disagreements occurred during the process.

The labels and their descriptions were developed by the researcher to support a straightforward, non-numerical evaluation of visual similarity. In line with the qualitative nature of this study, these categorical labels were designed to focus on patterns of interpretation rather than measurement. As Sandelowski (2000) argues, qualitative research emphasizes meaning over metrics, and the use of simple descriptive categories allows for flexible, context-sensitive judgment without reducing interpretation to numbers. Before rating all the image pairs, both raters joined a short practice session using a few sample images. This helped ensure that both raters shared a common understanding of the categories and applied them in a consistent way.

After the evaluations were completed, the results were shared with the students. They were invited to comment if they disagreed with the assessments, but no objections were raised. The students' agreement helped confirm that the evaluations made by the raters were reasonable and fair.

The student-written descriptions were uploaded into ChatGPT-4 for analysis. The researcher used it to help identify recurring vocabulary and grammar features in the texts. Phonological features were not included, as they were not relevant to the written data.

These AI-generated suggestions served as a starting point, but all linguistic features were manually reviewed and refined by the researcher. The researcher also consulted the literature on Indonesian English to guide the classification and naming of these features. To support validation, some of the student texts were shared with the two external raters mentioned earlier. The raters reviewed the features alongside these samples to confirm that the categories were clearly defined, consistently applied, and grounded in both the data and relevant theory.

With the evaluation process complete, the next section presents the key findings and explores how features of the students' English may have influenced the way the AI interpreted their descriptions.

Findings

While the research questions focused on Indonesian English, the analysis also revealed that stylistic choices in descriptive writing contributed to mismatches. These are reported alongside Indonesian English features to provide a fuller account of the data.

When DALL-E 3 misunderstands: how language features affect image accuracy

This section addresses the research question of how features of Indonesian English influence DALL-E 3's image outputs. The findings reveal a consistent pattern of misinterpretation. Out of the sixteen student descriptions analyzed, the majority (fourteen) produced clear mismatches when compared to the

original photographs. The remaining two (S14 and S2) resulted in only partial matches, and none of the descriptions generated an image that was considered a close match. The accuracy of the AI output seems to depend more on the level of concrete visual detail than on grammatical correctness.

For example, S14's description of the rooftop garden was written in non-standard English, with some grammar patterns likely influenced by his first language. S14's writing showed features typical of locally shaped English use, such as missing plural endings, expressions like "look-like," and phrasing such as "the garden that located at 8th floor." However, he included concrete details such as "a grey bridge with white fences," "six white block of stairs," and "three oval-shaped pool of grass and yellow tulips." The specific visual cues gave DALL-E 3 enough guidance to generate an image that was more accurate than most others. The two images in Figure 1, showing S14's original photo and the AI-generated image, illustrate how concrete visual details reduced the effect of non-standard grammar and allowed the AI to interpret the scene more successfully.

Figure 1. Original photo of the rooftop garden (top) and AI-generated image based on S14's description (below)

In contrast to S14's partial match, the descriptions by S4, S8, and S13 resulted in clear mismatches between the students' intended scenes and the AI-generated images. S4 described one of the campus canteens with general phrases like "very spacious," "lots of table and bench," and "lots of option when it comes to food choices." These words suggest the place is large and active, but he did not explain how the canteen actually looks. He did not mention the layout, the kind of tables and benches, the ceiling, lighting, or other details that could help the AI picture the space more clearly. Because of this, DALL-E 3 filled in the gaps using what it knows from other images of food courts. It imagined a big, modern place with rows of food stalls, and tall glass ceilings. However, this did not match the real canteen, which is smaller, simpler, and more casual. This case shows how vague expressions, such as "lots of," reduce clarity for AI systems. These are not recognized features of Indonesian English in the literature, but in this study, they were found in student writing and contributed to mismatches.

S8 described the campus Love Garden using emotional phrases like "makes me feel good" and "nice properties." Although she mentioned a few visual details, such as "25 accessories like a lock heart shape red" and "a small stage around 1×1 meter", the description did not clearly explain what the garden looked like overall. She did not say where the locks were placed, how they were arranged, or how the space was set up. Because these visual details were vague, DALL-E 3 filled in the missing parts by guessing. The result was a colorful, decorative garden full of large heart-shaped locks hanging from trees and walls. This is something that looked very different from the simple display in the real photo. This shows that when a description relies more on feelings than on clear visual explanation, the AI may end up creating something more imagined than real.

In S13's case, her night-time description of the main campus lobby focused strongly on mood and feeling rather than physical details. She used vivid phrases: "eerie," "ominous shadows," "desolate voids," and "untrustful noises" to describe the atmosphere, but she did not explain the layout, shape, or materials of the space. There was no mention of the building's structure, lighting fixtures, or how the tables and plants were arranged. Because of this, DALL-E 3 relied heavily on her emotional tone and filled in the missing details with its own visual interpretation of "eerie" and "ominous." The result was a dramatic, almost horror-style image with dim lighting, heavy shadows, and an abandoned look. This is far from the actual setting, which is simply a quiet, open-air campus area at night. The two images in Figure 2, showing the original photo of the campus lobby at night and the AI-generated image, illustrate how abstract and emotional language led the AI to imagine a horror-like scene rather than reproduce the real setting.

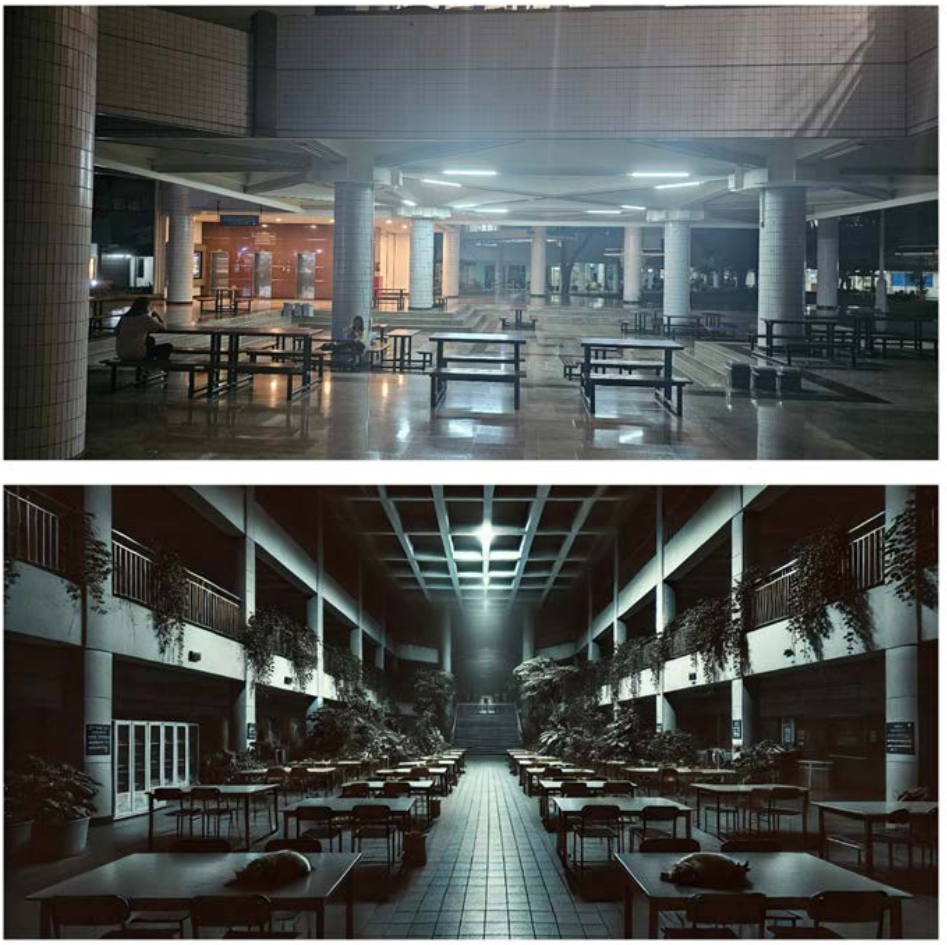


Figure 2. Original photo of the main campus lobby at night (top) and AI-generated image based on S13's description (below)

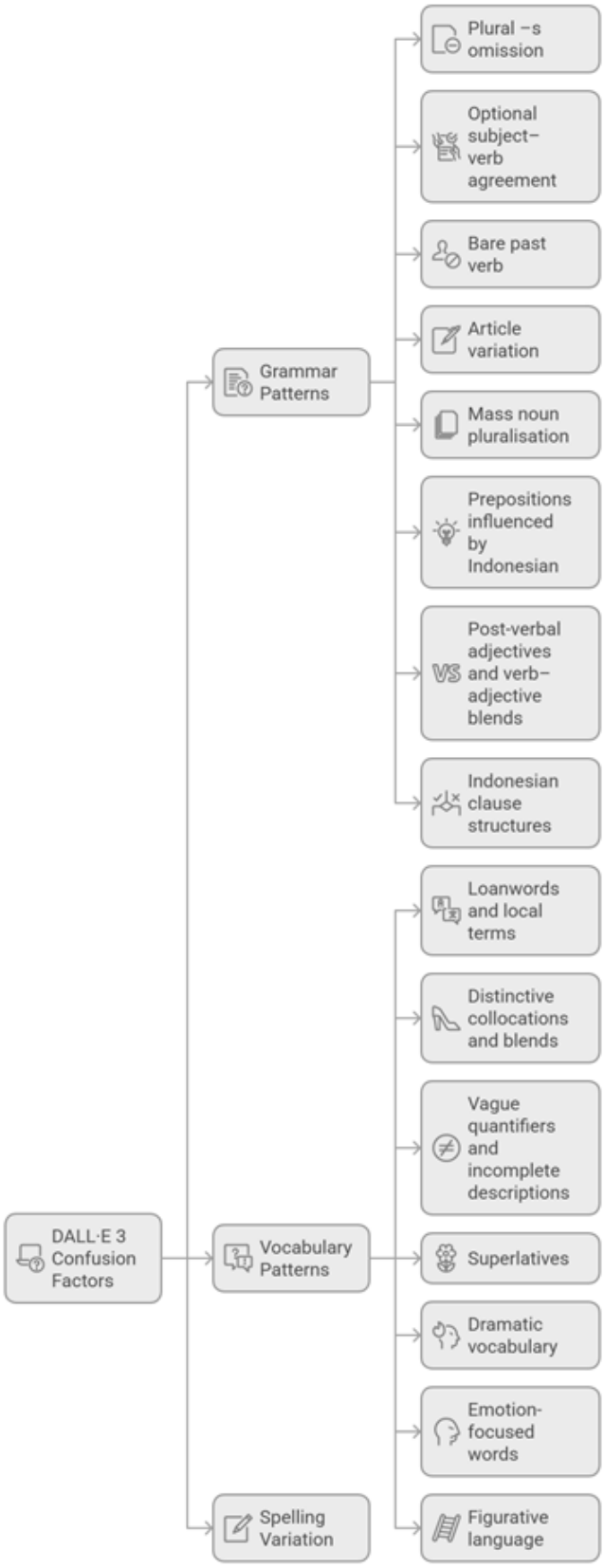
Overall, the findings show that DALL-E 3 does not consistently interpret student-written descriptions with high accuracy when the text contains vague expressions, non-standard grammar, or lacks specific visual information. Some of the patterns observed, such as plural *–s* omission, unusual verb forms, or distinctive prepositional use, have been documented in the literature as features of Indonesian English (e.g., Mahmud & Slabakova, 2020; Mulder & Marentek, 2017). Others, such as vague quantifiers or heavy use of emotional language, are not considered as features of Indonesian English but appeared in this dataset. Both types of patterns reduced clarity for the AI and contributed to mismatches. However, the AI's accuracy improved when descriptions included concrete visual details, such as colors, shapes, and numbers. This was true even when non-standard grammar was used, which suggests that for DALL-E 3, a high level of detail was more important than the other language patterns.

DALL-E 3 often misunderstood the students' descriptions, producing images that did not closely match the original photographs. The language patterns responsible for these mismatches appeared to originate from two main sources. The first source is a set of grammatical patterns that are consistent with documented features of Indonesian English found in existing studies. The second source is a range of other stylistic choices that appeared to be more specific to the context of this study. All of them, however, affected how DALL-E 3 interpreted the scene described in the text.

The key language patterns that caused misalignment between student texts and image outputs are summarized in Figure 3 and discussed in more detail below.

Grammar patterns.

- Plural *–s* omission: The students often omitted plural markers. For example, S1 wrote "3 portion", and S2 described "more than 20 table and 80 chair". This aligns with research on Indonesian English, which notes that speakers often leave nouns unmarked for plurality due to the influence of Indonesian grammar, where plurality is often implied by context (Mulder & Marentek, 2017). However, this led DALL-E 3 to misjudge quantity and either under- or overestimate how many objects should appear in the image.
- Optional subject–verb agreement: Examples include such as "canteen Q have much space" (S2), "each table have six planted power outline" (S5), and "this place also have a very nice view" (S7). This is a documented feature of Indonesian English, linked to the low use of the third-person singular *–s* and omission of auxiliary verbs (Mahmud & Slabakova, 2020). Such variation may have made it harder for DALL-E 3 to match subjects with actions or descriptions, especially when there were multiple objects mentioned in the sentence.
- Bare past verb: Instead of using past tense forms, the students often used base verbs, as in "I wait for my friends" (S7) and "I visit here with 10 of my friends" (S2). Indonesian English has been shown to rely heavily on bare verbs due to the lack of verb conjugation to show when something happened (Mahmud & Slabakova, 2020; Pipit & Rahyono, 2020). This may have led DALL-E 3 to interpret or treat the events as present or general, not something that happened in the past, which could affect how the scene was shown.
- Article variation: Articles were sometimes added where not needed ("a beautiful sceneries", S7), omitted ("Entrance at Q building", S16), or mismatched ("has a good facilities", S10). While not widely documented as a specific feature of Indonesian English in the literature, this was a recurring pattern in the student writing. These variations likely made it harder for the AI to recognize noun types or to understand what kind of object or scene to show.



Note: Image generated by Napkin AI (2025)

Figure 3. Key language patterns that affected DALL-E 3 image accuracy

- Mass noun pluralisation: Words like "cool furnitures" (S5), "the foods in canteen W" (S4), and "beautiful sceneries" (S7, S14) were treated as countable. This is consistent with Indonesian English usage, where words like staff, food, and furniture are often pluralized due to the influence of the first language (Endarto, 2020). As AI image generation often depends on detecting known object categories, these words may have caused it to show excessive or mismatched visual elements.
- Prepositions influenced by Indonesian: Phrases such as "different with library" (S6) and "comfortable to stayed long at here" (S2), "sit at there for a long time" (S10), and "compared than the other hallway" (S7) reflect direct translation from Indonesian. This pattern aligns with established research on Indonesian English, which notes that distinctive prepositional uses (e.g., "different with") are directly shaped by the grammar of the local language (Mulder & Marentek, 2017). These may have led DALL-E 3 to misplace objects or create layouts that did not match the intended scene.
- Post-verbal adjectives and verb–adjective blends: Some students combined verbs with adjectives in a way influenced by Indonesian structure, such as "makes me relaxing" (S8), "makes me can be relax" (S9), or "want to relaxing" (S15). While not documented with these exact words, these non-standard verb constructions are consistent with documented features of Indonesian English, such as unconventional verb patterns (e.g., Mahmud & Slabakova, 2020; Percillier, 2015). These blended forms may have made it unclear whether the sentence described an action or a feeling, which could affect how the AI generated the scene.
- Indonesian clause structures: Some students used sentence patterns that follow Indonesian grammar, e.g., "What the different this garden and the others is …" (S8) or "There are three reason why I like this canteen" (S2), and "the canteen only have 2 rows than canteen Q" (S3). These patterns align with what Daeng and Weda (2019) describe as the subtle shifts in word order or clause construction found in Indonesian English. These patterns deviate from standard English, which may have reduced prompt clarity and made the meaning less clear for the AI.

Vocabulary patterns.

- Loanwords and local terms: Although nearly all of the students kept to English, a few slipped in Indonesian food names that have no clear English equivalents. S2, for instance, mentioned "ayam lalapan" and "nasi lemak", while S4 referred to "pangsit mie". The use of such loanwords for local food and cultural items is a documented feature of Indonesian English (Endarto, 2020). Because DALL-E 3's training data does not always map these local dishes to clear visuals, the AI either ignored the words or replaced them with random cafeteria food.
- Distinctive collocations and blends: Some students used English word combinations that were grammatically correct but unusual in global usage, likely influenced by Indonesian phrasing or direct translation.

This aligns with research identifying distinctive collocations and lexical blends as features of Indonesian English (Endarto, 2020; Martin-Anatias, 2018). For example, S5 wrote "lift up my mood," a phrase that makes sense to local readers but is not commonly used in standard English, where expressions like "cheer me up" are more typical. S8 described a garden as having "nice properties," which may have been intended to mean "nice features", but the word "properties" in this context sounds vague and slightly off. These word combinations may have confused DALL-E 3 and led it to focus more on mood or style instead of showing what the place actually looked like.

- Vague quantifiers and incomplete descriptions: Many students used general terms like "lots of," "many," or "several" without giving clear amounts or specific details. While not a documented feature of Indonesian English, this was a common pattern in the student data. For example, S14 described "lots of green grass and several orange tulips," which may have led DALL-E 3 to exaggerate the size of the garden or overfill the scene with plants. S2 used phrases like "many table with bright color chair" and "many food outlet," which may have encouraged DALL-E 3 to insert more tables and larger food counters than were actually there. Some other students omitted clear spatial details such as where objects were placed or what size they were. S7, for example, wrote, "I can see a lot of house from several kilometres," without saying from which vantage point. S3, who described one of the canteens available, said that the canteen "only have 2 rows" and "only have 3 food stands," but did not explain how these were arranged or where they were located. These details hint at the space's simplicity and emptiness, but without layout cues, DALL-E 3 had to guess and filled in the gaps using its own assumptions.
- Superlatives: The use of unsupported superlatives is not an established feature of Indonesian English but was found in this dataset. Some students used superlatives to express strong opinions or comparisons, but they did not include concrete measurements or physical limits to support these claims. For example, S2 wrote that "canteen Q have the highest ceiling than other canteen" and S3 stated that "canteen T is the best canteen ever." These kinds of statements may be meaningful to the students, but without more detail, such as actual ceiling height or reasons why one space is "better", the AI responded by creating exaggerated scenes. For instance, DALL-E 3 created images with extra-tall buildings, fancy lighting, or modern designs that did not appear in the original photos.
- Dramatic vocabulary: Some students used strong or dramatic language that may have caused the AI to imagine the place as more luxurious or futuristic than it really was. These expressions showed how the students felt, but they did not give clear details about what the place looked like. Such uses of dramatic vocabulary are not documented features of Indonesian English; they reflect writing patterns found in

this group of students. For example, S11 described one of the campus' lobbies as "a futuristic architecture which looks like some kind of structure from a science fiction literature or from the future." Although the space may feel special to the student, this kind of dramatic description likely made the AI imagine a high-tech building with shiny white walls, bright lights, or futuristic features that did not actually exist in the real lobby. S10, who was describing one of the campus' auditoriums, also used dramatic phrases like "how amazing things that this place can offer" and "a luxury lightning with a various different colours." This kind of language may have led the AI to create spotlights, colorful lighting effects, or expensive-looking materials. As a result, the AI picture of this particular auditorium looked more glamorous than the real space. These descriptions reflected the students' excitement and personal connection to the place. However, the lack of clear and concrete details gave the AI room to interpret freely, which often resulted in images that looked more extravagant than the real setting.

– Emotion-focused words: Many students described how the place made them feel rather than what it physically looked like. For example, S3 wrote that the canteen "makes me feel calm," S6 mentioned the "comfy vibes" of the campus bookstore, and S16 said the entrance made him feel "calm and nostalgic." These emotional expressions gave the AI little concrete visual detail to work with. As a result, DALL-E 3 focused on atmosphere, like adding warm lights, cozy colors, or dreamy backgrounds, while ignoring layout or materials. These emotional expressions showed how much the students liked the place, but they did not give enough visual details for the AI to picture the place accurately. These emotional expressions are not documented features of Indonesian English but are found in the students' writing.
– Figurative language: Some students used metaphors or playful comparisons in their writing, and the AI tools interpreted literally. This led to image outputs that looked visually striking but strayed far from the real locations. For example, S13 described the main campus lobby at night by writing, "the classrooms… become desolate voids waiting to swallow you." This poetic metaphor was meant to express a gloomy atmosphere. However, the AI made an image that looked like a dark tunnel or a spooky hallway with lots of shadows and plants. Similarly, S14 wrote that the garden's "coral-shaped pathway… look-like Spongebob's coral." He likely intended this as a light-hearted way to describe the curve of the path. However, DALL-E 3 created an elaborate spiral design in a manicured, symmetrical rooftop garden. The original garden is much simpler and asymmetrical, with natural circular plant beds and straightforward walkways. The AI's image reflected a more stylized interpretation of "coral-shaped" rather than the actual physical layout. These cases show how figurative expressions, while vivid and creative, may confuse image-generation tools like DALL-E 3. The use of figurative language is not an established

feature of Indonesian English but a creative writing choice found in this dataset.

Spelling variation. Spelling differences in the students' writing made it harder for the AI to understand certain words, which likely led to some of the mismatches in the generated images. DALL-E 3 depends on accurate spelling to match words with the right objects or scenes. When it encounters unfamiliar or non-standard spellings, it may guess the meaning or substitute something else. For example, S10 described a "polcadot" wall (likely meant to be polka dot), which caused the model to create decorative or unusual wall patterns. S6's phrase "transparant wall" (for transparent) may have changed how the AI showed the lighting or wall material. S5 wrote "great lightning" when she meant lighting, which may have confused the AI and caused it to show something like stormy light or dramatic effects instead of normal lamps. People can usually understand these kinds of spelling differences from context, but AI systems are not as flexible. Although spelling variation is not a documented feature of Indonesian English, in this study it made it harder for DALL-E 3 to follow what the students meant and likely contributed to visual mismatches, such as incorrect colors, decorations, or objects.

In short, this analysis suggests a consistent pattern between the AI's misinterpretations and the presence of specific linguistic patterns in the student texts. These patterns included both documented features of Indonesian English and other stylistic choices that emerged from this study's context. Both types of patterns seemed to make the text less clear for the AI and led to mismatches.

Validating the findings: A comparative analysis

To further clarify how Indonesian English features may have shaped the mismatches in this study, two representative cases (S14 and S13) were revisited for evaluative comparison. Both texts were revised into a more standard English format that prioritizes concrete, literal, and visually descriptive language over emotional or vague phrasing. These revised prompts were then submitted to DALL-E 3 to evaluate if a change in language would produce a more accurate image.

Case 1: S14's Rooftop Garden. S14's original description achieved a partial match because it included specific visual details like "a grey bridge with white fences" and "three oval-shaped pool of grass". However, it also contained grammatical features of Indonesian English, such as missing plural endings and non-standard phrasing like "look-like" and "the garden that located at 8th floor," which likely contributed to inaccuracies in the AI-generated image. The revised prompt preserves the original concrete details while correcting the grammatical variations to align with standard English: "A rooftop garden on the 8th floor of a university building. The garden features a coral-shaped pathway, three oval-shaped pools of grass planted with yellow tulips, and small trees around. A white fence and stairs connect the garden to the building. In the background,

there are other buildings and a clear blue sky. The right side of the picture shows a tall building with yellow and gray panels and hanging plants." The resulting image, presented in Figure 4, corresponds more closely to the original rooftop garden photo.

Figure 4. Original photo of the rooftop garden (top) and AI-generated image created from a revised prompt (below).

Case 2: S13's Lobby. S13's original text focused heavily on abstract and figurative language, using phrases like "eerie," "ominous shadows," and "desolate voids waiting to swallow you". This led the AI to generate a dramatic, horror-style image rather than the actual quiet, open-air lobby. This revised prompt removes all emotional and figurative language and replaces it with precise, physical details about the lighting, materials, objects, and layout: "A wide, open-air university lobby at night. The floor is shiny and reflects the bright ceiling lights. Several thick round columns support the roof. There are many long brown tables and benches arranged in rows. A glass-walled elevator is visible in the back. The place looks quiet and spacious." The effect of this revision is

illustrated in Figure 5, where the AI-generated image more accurately represents the actual open-air lobby.

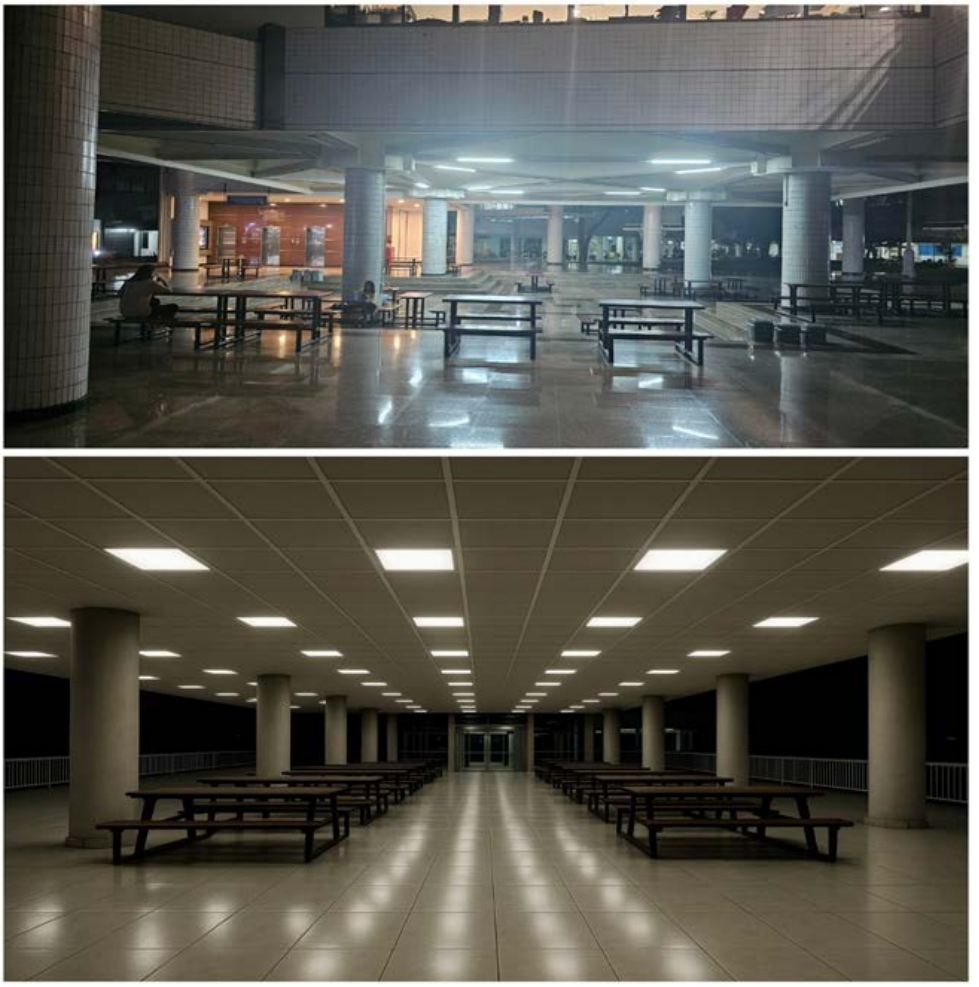


Figure 5. Original photo of the main campus lobby (top) and AI-generated image created from a revised prompt (below)

The images generated from the revised prompts were a closer match to the original photographs. This outcome suggests that the AI's initial misinterpretations were influenced by the grammatical features of Indonesian English and the use of human-centric stylistic choices. It reinforces the argument that a bias toward standard English limits the digital intelligibility of current AI systems.

Discussion

The findings confirm that several common features of Indonesian English did appear in the student texts and contributed to misinterpretations in the DALL-E 3 image outputs. The analysis also revealed other features not previously discussed in the literature which also contributed to mismatches between the students' intended meaning and the AI-generated images. This interpretation

is strengthened by the comparative analysis, which showed that revising the prompts into standard English led to clearer and more accurate images.

From a World Englishes perspective, the students' writing reflects legitimate variation. As Kachru (1992) emphasizes, English exists in many localized forms that meet the communicative needs of their users. Indonesian English, in particular, carries distinct vocabulary choices (e.g., distinctive collocations and blends) or grammar patterns (e.g., zero plurals, subject–verb mismatches) shaped by Indonesian language structures and cultural preferences (Daeng & Weda, 2019; Endarto, 2020). These patterns, clearly present in student texts, align closely with previous descriptions of Indonesian English (e.g., Hamsa & Weda, 2019; Pipit & Rahyono, 2020).

In human interactions, such features do not necessarily block understanding. According to Smith and Nelson (1985), intelligibility, comprehensibility, and interpretability allow communication across English varieties even when surface-level grammar or vocabulary differ. For example, when a student wrote, "This canteen have many food outlet," a human reader could infer the intended meaning, even if the grammar is non-standard. This is because human listeners use shared context, pragmatic cues, and a willingness to accommodate unfamiliar forms. However, DALL-E 3 cannot replicate this human flexibility. It struggles with language forms that fall outside dominant norms.

This problem is not unique to Indonesian English. Evidence from other English varieties shows a similar pattern of misrecognition and bias. Commercial speech-recognition tools record about twice as many mistakes with African American English as they do with mainstream American English (Martin & Wright, 2023). Bailey et al. (2025) found that speech-based AI systems, like automatic speech recognition (ASR), often give inaccurate feedback to students who use African American English, especially during assessments. GPT-4 was able to recognize Trinidadian English Creole in only 21% of test prompts and often changed it into standard English (Jackson et al., 2024). Similarly, a study on ChatGPT-4 found that the model consistently defaulted to Native English Speaker (NES) norms when asked to perform writing and feedback tasks using Korean English (Lee et al., 2025). A large study of language models also found that prompts written in African American English were more likely to be linked with low-status jobs or harsher penalties compared to the same prompts in standard English (Hofmann et al., 2024). Other studies show that writing produced by AI often reflects the language style of more powerful social groups, which creates an unfair idea of what "good" English looks like (Alvero et al., 2024). Earlier research on word embeddings found the same pattern in smaller language models where the system favored certain types of language and identities over others (Caliskan et al., 2017).

These studies show that AI systems often struggle with English that is different from the standard forms they were trained on. This helps explain why Indonesian English may face similar challenges in AI processing. To this, Bailey et al. (2025), Caines et al. (2025), Jackson et al. (2024), (Lee et al., 2025), and Payne et al. (2024) point out that AI tools rely heavily on dominant linguistic norms and that they typically misinterpret unfamiliar grammar and cultural

expressions. The underlying issue might reflect standard language ideology or the belief that one variety of English (often standard American or British) is more correct or desirable than others. As a result, non-standard forms, including Indonesian English, are underrepresented in AI training datasets. AI tools treat standard forms as defaults and view non-standard varieties as errors to be corrected. As a result, students who use natural, local forms of English face additional pressure: to communicate clearly with AI, they must adjust their language toward more standardized forms. This shifts the responsibility for intelligibility onto the users rather than onto AI developers. It also goes against the ideas of the World Englishes approach, which promotes recognition and respect for local norms.

This reliance on standard English raises broader ethical and educational concerns. As Sartori and Theodorou (2022) note, AI tools are not neutral as they reflect and reproduce social biases. Blodgett et al. (2020) warn that without careful design, natural language processing (NLP) systems, including tools like ChatGPT, DALL-E, and other language-based AI models, can reinforce existing inequalities. When the AI repeatedly misinterprets local language forms, it sends the message that varieties like Indonesian English are incorrect or incompatible with AI. In classrooms, this may inspire students that only standardized forms are acceptable in digital spaces.

This study identifies the gap as an issue of "digital intelligibility," meaning how well digital technologies can process and understand people's natural, non-standard language without requiring them to change their writing. While this is a broad concern, the focus here is specifically on AI. Digital intelligibility connects to earlier work on algorithmic intelligibility and transparency, which looked at whether people can understand how algorithms work (Diakopoulos & Koliska, 2017). It is also linked to AI intelligibility, defined as the epistemological sense of how an AI works and how accurate it is (Ashok et al., 2022), and to Kamiura (2025), who notes that digital intelligibility has usually been used to describe how humans understand AI systems. In this study, however, the term is adapted to highlight the reverse issue, i.e. whether AI systems can understand human language. The findings indicate that current AI models, especially DALL-E 3, still lack this kind of digital intelligibility.

Improving digital intelligibility will require changes in how AI systems are trained. Developers need more inclusive training strategies to reduce misinterpretation of local English varieties. A key step is to diversify the training data. Future training data should include real examples of Indonesian English and other local varieties, not just American or British English. This includes authentic texts with real-world examples of grammatical variation, creative phrasing, local collocations, and culturally specific vocabulary. Such variety helps AI systems better understand and respond to different forms of English. A study by Hidayatullah et al. (2025) found that adding code-mixed Indonesian–Javanese–English text improved performance across multiple NLP tasks. This shows that diverse language data makes the AI systems more reliable.

Beyond English varieties, training data also needs to capture stylistic diversity. Student writing often includes vague quantifiers, emotional tone, or

figurative phrasing. These are natural choices in descriptive paragraph tasks but can confuse AI systems that rely on concrete input. Including authentic examples of such stylistic variation in training corpora would help AI models handle real-world student writing more effectively.

AI developers should also regularly evaluate how well AI systems handle different English varieties and writing styles. One way to do this is by using tests such as the Vernacular Language Understanding Evaluation (VALUE). Using this test, Ziems et al. (2022) found that AI tools perform worse when the input is written in African American English. This reveals that the tools are biased toward standard English. A similar test could be made for image-generation tools like DALL-E to see how well they handle prompts written in different English varieties. These evaluations can help AI developers identify weaknesses early and improve fairness and accuracy.

Educational institutions also have a role to play. Schools and universities should acknowledge and discuss these linguistic biases with students. Teachers can encourage students to view AI-generated output with caution and remind them that misunderstandings often happen due to system limitations. Classroom activities that ask students to compare their original prompts with the AI-generated images, identify where meaning shifts, and discuss which version of English is treated as the standard can help build awareness of bias in both AI and language norms. Similar strategies have been used in teacher-education programs, where sessions focused on identifying bias in generative tools have helped future teachers develop a more reflective attitude toward technology (Dilek et al., 2025).

Conclusion

This study examined how DALL-E 3 interpreted descriptive texts written in Indonesian English and found that the AI's misinterpretations appeared to originate from two sources. The first was the use of documented features of Indonesian English, such as the omission of plural markers and the use of local vocabularies. These features align with the World Englishes perspective, which views such local variations as legitimate. The second source was the students' other writing patterns, such as the use of emotional or figurative language. These patterns are not specific to Indonesian English but they appeared in the students' texts.

These findings show that current AI systems are not yet able to handle linguistic diversity fairly. Improving digital intelligibility requires responses at both the technical and educational levels. Technically, AI systems need more diverse training and evaluation. At the same time, classroom activities that explore how AI handles different English varieties can help students recognize bias and build critical digital literacy.

This study had several limitations. First, the students were not told that their writing would be used as prompts for an AI. They were instructed to write a standard descriptive paragraph for a human audience and were not aware their text would be used by an AI. Because of this, the findings reflect how the

AI interprets writing from a natural classroom setting, not how it might perform with a text specifically designed for it. The findings provide authentic data on how the AI handles real-world student language.

Second, this study did not clearly isolate the cause of the AI's misinterpretations. A comparative analysis of two cases, however, showed that revising the student texts into standard English resulted in more accurate images. While limited to two examples, this finding supports the interpretation that the features of Indonesian English and other stylistic patterns might be a primary cause of the misinterpretations.

Third, the study was limited to written texts from sixteen students at a single institution and used only one AI model. Pronunciation and intonation as other features of spoken Indonesian English were not included. Future research should explore how AI systems handle spoken texts, especially when processed through speech recognition tools that often struggle with non-standard English.

Finally, future studies could also compare how different image-generation models handle localized English. These strategies would help move toward AI systems that are more inclusive, more aware of linguistic diversity, and better able to support fair communication in education and beyond.

References

Alvero, A. J., Lee, J., Regla-Vargas, A., Kizilcec, R. F., Joachims, T., & Antonio, A. L. (2024). Large language models, social demography, and hegemony: comparing authorship in human and synthetic text. *Journal of Big Data, 11*(1), 138. https://doi.org/10.1186/s40537-024-00986-7

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management, 62,* 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Bailey, A. L., Johnson, A., Shankar, N. B., Veeramani, H., Washington, J. A., & Alwan, A. (2025). Addressing bias in spoken language systems used in the development and implementation of automated child language-based assessment. *Journal of Educational Measurement.* https://doi.org/10.1111/jedm.12435

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. In D. Jurafsky, J. Chai, N. Schluter, J. Tetreault (Eds.), *Proceedings of the 58th annual meeting of the Association for Computational Linguistics* (pp. 5454–5476). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.485

Caines, A., Buttery, P., & Seed, G. (2025). Towards equitable, diverse and inclusive language models in ai-enhanced education technology for language learners. In W. Wei & D. Chao (Eds.), *The Routledge handbook of the sociopolitical context of language learning* (pp. 401–416). Routledge. https://doi.org/10.4324/9781003398172-29

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Chiu, L.-L., & Liu, G.-Z. (2013). Effects of printed, pocket electronic, and online dictionaries on high school students' English vocabulary retention. *The Asia-Pacific Education Researcher, 22*(4), 619–634. https://doi.org/10.1007/s40299-013-0065-1

Daeng, K., & Weda, S. (2019). Contrastive analysis of Makassarese, Indonesian, and English syntax. *Asian EFL Journal, 25*(5.2), 111–129.

Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. *Digital Journalism, 5*(7), 809–828. https://doi.org/10.1080/21670811.2016.1208053

Dilek, M., Baran, E., & Aleman, E. (2025). AI literacy in teacher education: Empowering educators through critical co-discovery. *Journal of Teacher Education, 76*(3), 294–311. https://doi.org/10.1177/00224871251325083

Endarto, I. T. (2020). A corpus-based lexical analysis of Indonesian English as a new variety. *Indonesian Journal of Applied Linguistics, 10*(1), 95–106. https://doi.org/10.17509/ijal.v10i1.24993

Floris, F. D. (2025). Exploring shared repertoire in virtual communities of practice: Integration of artificial intelligence in English language teaching. *The JALT CALL Journal, 21*(2), 102420. https://doi.org/10.29140/jaltcall.v21n2.102420

Hamsa, A., & Weda, S. (2019). Comparative study in Indonesian and English: Identifying linguistic units of comparison. *Asian EFL Journal, 25*(5.2), 96–110.

Harris, C., Mgbahurike, C., Kumar, N., & Yang, D. (2024, November). Modelling gender and dialect bias in automatic speech recognition. In Y. Al-Onaizan, M. Bansal, & Y. Chen (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 15166–15184). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-emnlp.890

Hidayatullah, A. F., Apong, R. A., Lai, D. T. C., & Qazi, A. (2025). Pre-trained language model for code-mixed text in Indonesian, Javanese, and English using transformer. *Social Network Analysis and Mining, 15*(1), 1–17. https://doi.org/10.1007/s13278-025-01444-9

Hofmann, V., Kalluri, P. R., Jurafsky, D., & King, S. (2024). AI generates covertly racist decisions about people based on their dialect. *Nature, 633*(8028), 147–154. https://doi.org/10.1038/s41586-024-07856-5

Hopkins, D. (2014). *A teacher's guide to classroom research* (5th ed.). McGraw-Hill Education.

Jackson, S., Beekhuizen, B., Zhao, Z., & McEwen, R. (2024). GPT-4-Trinis: Assessing GPT-4's communicative competence in the English-speaking majority world. *AI & Society, 40*(3), 1–17. https://doi.org/10.1007/s00146-024-01945-9

Kachru, B. B. (1992). *The other tongue: English across cultures* (2nd ed.). University of Illinois Press.

Kachru, B. B., Kachru, Y., & Nelson, C. L. (2009). *The handbook of world Englishes.* John Wiley & Sons.

Kachru, Y. (2008). Cultures, contexts, and interpretability. *World Englishes, 27*(3-4), 309–318. https://doi.org/10.1111/j.1467-971x.2008.00569.x

Kamiura, M. (2025). The four fundamental components for intelligibility and interpretability in AI ethics. *American Philosophical Quarterly, 62*(2), 103–112. https://doi.org/10.5406/21521123.62.2.01

Lee, S., Jeon, J., McKinley, J., & Rose, H. (2025). Generative AI and English language teaching: A global Englishes perspective. *Annual Review of Applied Linguistics,* 1–24. https://doi.org/10.1017/s0267190525100184

Mahmud, M., & Slabakova, R. (2020). A longitudinal investigation of morpheme acquisition by Indonesian learners of L2 English. *International Journal of Language Studies, 14*(4), 19–38.

Martin, J. L., & Wright, K. E. (2023). Bias in automatic speech recognition: The case of African American language. *Applied Linguistics, 44*(4), 613–630. https://doi.org/10.1093/applin/amac066

Martin-Anatias, N. (2018). Language selection in the Indonesian novel: Bahasa gado-gado in expressions of love. *South East Asia Research, 26*(4), 347–366. https://doi.org/10.1177/0967828x18809592

Matsuura, H. (2007). Intelligibility and individual learner differences in the EIL context. *System, 35*(3), 293–304. https://doi.org/10.1016/j.system.2007.03.003

Mulder, J., & Marentek, A. (2017). Emergent use of prepositions by Indonesian speakers of English in EFL interactions. *Asian English Studies, 19,* 87–105. https://doi.org/10.50875/asianenglishstudies.19.0_87

Napkin AI. (2025). *Napkin AI: Get visuals from your text* (Beta-0.11.0 version). https://app.napkin.ai/

Nicholas, G., & Bhatia, A. (2023). Lost in translation: Large language models in non-English content analysis. *arXiv preprint arXiv:2306.07377.* https://doi.org/10.48550/arXiv.2306.07377

Payne, A. L., Austin, T., & Clemons, A. M. (2024). Beyond the front yard: The dehumanizing message of accent-altering technology. *Applied Linguistics, 45*(3), 553–560. https://doi.org/10.1093/applin/amae002

Percillier, M. (2015). Postcolonial and learner Englishes in Southeast Asia: Implications for international communication. *Communicating with Asia: The Future of English as a Global Language,* 135–152. https://doi.org/10.1017/cbo9781107477186.010

Pipit, M., & Rahyono, F. X. (2020). The past tense expression of Indonesian learners: A morphosyntactic review and its implication toward teaching field. *Asian EFL Journal, 27*(2), 120–143.

Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing & Health, 23*(4), 334–340. https://doi.org/10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g

Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: Narratives, inequalities, and human control. *Ethics and Information Technology, 24*(1). https://doi.org/10.1007/s10676-022-09624-3

Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes, 4*(3), 333–342. https://doi.org/10.1111/j.1467-971x.1985.tb00423.x

Syam, A. R., Gardner, S., & Cribb, M. (2024). Pronunciation features of Indonesian-accented English. *Languages, 9*(6), 222. https://doi.org/10.3390/languages9060222

Thomas, P., McDuff, D., Czerwinski, M., & Craswell, N. (2020). Expressions of style in information seeking conversation with an agent. In A. An & L. Xia (Eds.), *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval* (pp. 1171–1180). The Association for Computing Machinery. https://doi.org/10.1145/3397271.3401127

Ziems, C., Held, W., Yang, J., Dhamala, J., Gupta, R., & Yang, D. (2022). VALUE: Understanding dialect disparity in NLU. In S. Muresan, P. Nakov, A. Villavicencio (Eds.), *Proceedings of the 60th annual meeting of the Association for Computational Linguistics* (Volume 1: Long papers) (pp. 3701–3720). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.258