# Visual vowel mapping and cross-linguistic interference in Indonesian EFL learners' English pronunciation

***Rinaldi*[1], Muhammad Rhafi Harlan***
*[1]Pendidikan Bahasa Inggris, Institut Agama Islam Negeri, Indonesia*
**Corresponding author: hii.rinaldi@gmail.com*

## ABSTRACT

Pronunciation remains a persistent challenge for Indonesian learners of English, particularly in mastering English vowel contrasts that differ significantly from the Indonesian vowel system. From a cross-linguistic phonological perspective, these difficulties are largely attributed to first language interference, which shapes learners' perception and production of second language sounds. This study aimed to explore how visual vowel mapping could be used to reveal patterns of cross-linguistic vowel interference among Indonesian EFL learners. Data were collected through classroom observations, audio recordings of learners' vowel production, and analysis of visual vowel representations used during pronunciation instruction. The findings indicated systematic vowel substitutions and mergers, such as the neutralization of tense–lax vowel contrasts and the transfer of Indonesian vowel categories into English vowel perception. Visual vowel mapping functioned not as a corrective tool per se, but as an analytical and pedagogical medium that enhanced learners' phonological awareness by making articulatory and acoustic differences visually salient. Interpreted through the framework of cross-linguistic phonology and second language speech learning, the study suggests that visual-based phonological representation can facilitate learners' noticing of vowel contrasts that are otherwise perceptually inaccessible. This study contributes to qualitative research on pronunciation pedagogy by highlighting the role of visual phonetic tools in understanding vowel interference within the Indonesian EFL context.

**Keywords**: cross-linguistic phonology; EFL learners; English pronunciation; visual vowel mapping; vowel interference.

## INTRODUCTION

Pronunciation continues to be one of the most challenging aspects of English as a Foreign Language (EFL) learning, particularly in contexts where learners' first language (L1) phonological system differs substantially from that of English (Putra, 2024; Masykar et al., 2025). In the Indonesian EFL context, pronunciation difficulties are not merely pedagogical issues but are deeply rooted in phonological and perceptual constraints shaped by learners' L1. Among various pronunciation features, English vowels represent a particularly complex area due to the rich vowel inventory of English, which contrasts sharply with the relatively limited vowel system of Indonesian (Yani & Zainuddin, 2022; Rachmawati & Andayani, 2023).

Indonesian employs a small set of stable vowel categories that are not distinguished by tenseness or length, whereas English vowels are organized along multiple phonetic dimensions, including height, backness, tenseness, and duration. This structural disparity often results in systematic pronunciation deviations when

Indonesian learners attempt to perceive and produce English vowel contrasts. Recent empirical studies demonstrate that Indonesian EFL learners frequently neutralize several English vowel distinctions by mapping them onto a single L1 vowel category, leading to vowel substitution, category overlap, and reduced phonological contrast (Setyawan & Fata, 2021; Lubis et al., 2025). Such patterns indicate that pronunciation errors are not random but reflect deeper cross-linguistic phonological interference.

From a cross-linguistic phonology perspective, these difficulties are best understood as the result of L1 phonological transfer, whereby learners rely on pre-existing phonological representations when processing unfamiliar L2 sounds (Masykar et al., 2025). Contemporary research in second language speech learning emphasizes that learners' perception of L2 sounds is constrained by the perceptual framework established through their L1 experience. When L2 vowel contrasts are perceived as sufficiently similar to L1 categories, learners tend to assimilate them rather than establish new phonological distinctions, resulting in persistent misperception and inaccurate production (Inceoglu, 2015; Redmon et al., 2020). In the Indonesian EFL context, this process is evident in the frequent merging of English vowels such as /ɪ/ and /iː/ or the substitution of /æ/ with Indonesian /e/ or /a/ (Yani & Zainuddin, 2022; Rachmawati & Andayani, 2023).

Despite the extensive documentation of vowel interference in EFL research, pronunciation instruction in many Indonesian classrooms remains predominantly auditory-oriented, relying heavily on repetition drills, minimal pair exercises, and teacher modeling (Suryani, 2022; Muna & Husna, 2021). These approaches are often grounded in the assumption that repeated exposure to correct auditory input will naturally lead to improved perception and production. However, recent studies challenge this assumption by demonstrating that auditory input alone may be insufficient for learners whose perceptual systems have been shaped by a different vowel inventory. Without explicit support for phonological awareness, learners may continue to reproduce inaccurate vowel realizations even after prolonged exposure, leading to fossilized pronunciation patterns (Inceoglu, 2015; Redmon et al., 2020).

In response to these pedagogical limitations, recent developments in phonology and pronunciation pedagogy have increasingly emphasized the role of visual-based approaches in supporting second language speech learning. Visual vowel mapping, which represents vowel contrasts through articulatory and acoustic visualization such as vowel charts, vowel space diagrams, or articulatory mapping, has emerged as a promising tool for making abstract phonological distinctions more accessible to learners (Yuliana & Susanti, 2020). By externalizing phonetic features that are otherwise implicit, visual vowel mapping enables learners to conceptualize vowel contrasts spatially and relationally, thereby facilitating greater phonological awareness.

Research on audiovisual speech perception further supports the pedagogical value of visual cues in second language phonology. Recent studies indicate that visual information can enhance learners' ability to discriminate non-native vowel contrasts, particularly when auditory cues alone are insufficient or ambiguous (Arafiq, Yusra, & Saputra, 2020). In this sense, visual vowel mapping does not function merely as a supplementary teaching aid but serves as a cognitive scaffold that supports learners' perceptual restructuring of L2 vowel categories. This

perspective aligns with contemporary views that emphasize pronunciation learning as a process of perceptual reorganization rather than simple imitation.

Nevertheless, despite the growing interest in visual phonetic tools, qualitative investigations that explore how visual vowel mapping reveals patterns of cross-linguistic vowel interference remain limited, especially within the Indonesian EFL context. Existing Indonesian studies on pronunciation predominantly adopt quantitative or quasi-experimental designs aimed at measuring instructional effectiveness or pronunciation gains (Fitriani & Makmur, 2021; Mustika & Pramesti, 2024). While such studies provide valuable insights into outcomes, they often overlook the learners' perceptual processes and phonological representations that underlie pronunciation performance. Moreover, visual vowel mapping is frequently treated as a teaching technique rather than as an analytical lens for examining how learners perceive, interpret, and negotiate vowel contrasts.

To address this gap, the present qualitative study explores how visual vowel mapping can be used to reveal patterns of cross-linguistic vowel interference among Indonesian EFL learners. By analyzing learners' vowel production, classroom interactions, and visual phonetic representations, this study aims to provide a detailed account of how Indonesian learners perceive and produce English vowel contrasts within an EFL classroom context. Grounded in recent cross-linguistic phonology and second language speech learning research, this study contributes to qualitative pronunciation scholarship by highlighting the analytical and pedagogical potential of visual vowel mapping in understanding vowel interference in Indonesian EFL classrooms.

## METHOD

This study adopted a qualitative descriptive research design to explore patterns of cross-linguistic vowel interference and to examine how visual vowel mapping reveals Indonesian EFL learners' perception and production of English vowels. A qualitative approach was considered appropriate because the focus of the study was not to measure instructional effectiveness or pronunciation improvement quantitatively, but to describe and interpret phonological phenomena as they naturally occurred in the classroom setting (Fitriani & Makmur, 2021; Mustika & Pramesti, 2024). Within this framework, pronunciation was treated as a perceptual–phonological process shaped by first language influence rather than as isolated pronunciation errors. The study was grounded in recent perspectives on cross-linguistic phonology and second language speech learning, which emphasize the role of learners' L1 phonological system in constraining L2 vowel perception and production (Masykar et al., 2025; Redmon et al., 2020). Visual vowel mapping was therefore positioned not merely as a pedagogical technique, but as an analytical lens for examining learners' phonological representations of English vowels.

The participants of this study were Indonesian EFL learners enrolled in a university-level pronunciation class. A total of 18 participants were involved in the study. They were categorized as institutional placement standards and had received formal English instruction for at least six years. Participants were selected using purposive sampling to ensure relatively homogeneous linguistic and educational backgrounds. All participants were native speakers of Indonesian who had received formal English instruction for several years and shared a relatively homogeneous

linguistic background. This homogeneity was essential to ensure that the analysis focused specifically on Indonesian phonological influence on English vowel production and perception (Rachmawati & Andayani, 2023; Yani & Zainuddin, 2022). None of the participants had previously received systematic instruction using visual phonetic tools such as vowel space diagrams or articulatory mapping, allowing the study to observe learners' engagement with visual vowel mapping without the influence of prior exposure (Lubis et al., 2025).

Data were collected from multiple sources to allow for methodological triangulation and to enhance the depth of qualitative analysis (Fitriani & Makmur, 2021). The primary data consisted of audio recordings of learners' vowel production obtained through controlled pronunciation tasks, including word lists and short reading passages containing English vowels that are commonly problematic for Indonesian learners, such as contrasts between /ɪ/ and /iː/ (e.g., *sit–seat*, *ship–sheep*) and the low front vowel /æ/ (e.g., *cat*, *man*, *bad*). The word lists consisted of approximately 20–25 target items designed to elicit specific vowel contrasts, while the reading passages contained multiple instances of these vowels in connected speech contexts and particularly tense–lax vowel contrasts and low front vowels (Yani & Zainuddin, 2022; Putra, 2024). In addition to spoken data, classroom observations were conducted to document learners' interaction with visual vowel mapping during pronunciation instruction. Visual artifacts, such as vowel charts and vowel space diagrams used in the classroom, were also collected and analyzed as part of the data set. The integration of spoken data and visual representations enabled a comprehensive examination of how learners perceived and produced English vowels in relation to Indonesian vowel categories (Setyawan & Fata, 2021; Heeringa & Van de Velde, 2024).

The data collection process was carried out over several instructional sessions and followed a structured sequence. First, learners were asked to produce target English vowels without visual support to capture their baseline pronunciation patterns. Second, visual vowel mapping was introduced through explicit visual representations comparing the English vowel space with Indonesian vowel categories. Third, learners were encouraged to observe, discuss, and reflect on vowel positions using the visual maps. Finally, classroom interactions, including learners' verbal responses, questions, and reflections related to the visual representations, were documented through detailed observation notes. (Muna & Husna, 2021; Mustika & Pramesti, 2024). This procedure allowed the study to capture learners' perceptual awareness and interpretative responses to visual vowel mapping.

Data analysis followed a qualitative phonological analysis framework that combined descriptive categorization and interpretative analysis. Audio recordings were transcribed and analyzed to identify recurring patterns of vowel substitution, vowel merger, and deviation from target English vowel realizations (Setyawan & Fata, 2021; Lubis et al., 2025). These patterns were then compared with the Indonesian vowel system to determine instances of cross-linguistic interference. The visual vowel maps were analyzed in relation to learners' spoken data to examine how articulatory similarities and vowel space overlap were represented visually (Heeringa & Van de Velde, 2024). Classroom observation notes were coded thematically to identify patterns of learners' phonological awareness, noticing of vowel contrasts, and interpretative engagement with visual mapping,

drawing on recent research on perceptual training and audiovisual speech perception (Inceoglu, 2015; Redmon et al., 2020).

To enhance the trustworthiness of the findings, several qualitative validation strategies were employed. Data triangulation was achieved by cross-referencing spoken data, classroom observations, and visual artifacts to ensure consistency in the interpretation of vowel interference patterns (Fitriani & Makmur, 2021). Analytical transparency was maintained by applying consistent criteria for categorizing vowel interference based on established descriptions of English and Indonesian vowel systems reported in recent studies (Rachmawati & Andayani, 2023; Yani & Zainuddin, 2022). In addition, peer debriefing with a fellow linguistics researcher was conducted to discuss the analysis and interpretation of the data, thereby minimizing researcher bias and strengthening analytical credibility (Mustika & Pramesti, 2024).

## RESULTS AND DISCUSSION

### 1.1 Patterns of Cross-Linguistic Vowel Interference

The analysis of learners' vowel production revealed recurring and systematic patterns of cross-linguistic vowel interference that reflect the influence of the Indonesian vowel system on the perception and production of English vowels. These patterns were not random pronunciation errors, but rather consistent phonological strategies employed by learners when confronted with unfamiliar English vowel contrasts. The most salient patterns identified in the data include the substitution of /ɪ/ with /i/, the neutralization of /æ/ into /e/, and the over-generalization of Indonesian vowel categories across multiple English vowel targets. Similar patterns have been reported in recent studies on Indonesian EFL pronunciation, indicating that such phenomena are characteristic of broader cross-linguistic phonological constraints rather than individual learner deficiencies (Yani & Zainuddin, 2022; Rachmawati & Andayani, 2023; Lubis et al., 2025).

One of the most prominent patterns observed in the data was the substitution of the English lax vowel /ɪ/ with the tense vowel /i/. Learners consistently produced words such as *sit*, *ship*, and *live* with a high front tense vowel resembling Indonesian /i/, resulting in minimal acoustic distinction between /ɪ/ and /i/. This substitution can be explained by the absence of a tense–lax vowel distinction in the Indonesian vowel system, which does not encode vowel quality differences through tenseness (Setyawan & Fata, 2021; Putra, 2024). From a perceptual standpoint, learners appeared to assimilate the English /ɪ/ into the closest L1 category, namely Indonesian /i/, which occupies a similar articulatory position but lacks the phonemic contrast found in English. Similar findings have been reported in cross-linguistic perception studies, which suggest that learners tend to rely on familiar vowel categories when L2 contrasts are perceived as insufficiently salient (Inceoglu, 2015; Redmon et al., 2020).

Visual vowel mapping played a crucial role in making this substitution pattern visible both analytically and pedagogically. When learners were exposed to visual representations of the English vowel space, the overlap between /ɪ/ and /i/ became perceptually salient, allowing learners to observe how both vowels were being mapped onto a single L1 category. This finding aligns with recent phonological research emphasizing the value of visualizing vowel space overlap to

reveal cross-linguistic interference patterns (Heeringa & Van de Velde, 2024). Rather than immediately correcting learners' pronunciation, the visual mapping facilitated a process of phonological noticing, whereby learners began to recognize the existence of contrasts that were previously imperceptible through auditory input alone (Muna & Husna, 2021; Mustika & Pramesti, 2024).

Another recurring pattern identified in the data was the neutralization of the English low front vowel /æ/ into a mid-front vowel resembling /e/. Words such as *cat*, *man*, and *bad* were frequently produced with a vowel quality closer to Indonesian /e/ or /a/, resulting in reduced distinctiveness from other English vowels. This neutralization reflects the absence of a low front vowel phoneme in Indonesian and demonstrates how learners simplify the English vowel inventory by collapsing unfamiliar contrasts into existing L1 categories (Fitriani & Makmur, 2021; Yani & Zainuddin, 2022). Similar neutralization patterns have been observed in recent Indonesian pronunciation studies, suggesting that /æ/ represents a particularly challenging vowel due to its articulatory unfamiliarity and lack of an equivalent in Indonesian (Rachmawati & Andayani, 2023; Lubis et al., 2025).

From a cross-linguistic phonological perspective, the neutralization of /æ/ into /e/ illustrates how learners prioritize phonological economy by reducing complex vowel distinctions into more manageable categories. Visual vowel mapping further revealed that learners tended to position /æ/ within the Indonesian vowel space rather than expanding the vowel space to accommodate a new category. This supports findings from recent visual phonetics research, which argue that learners often compress L2 vowel space to fit L1 phonological constraints unless explicit perceptual support is provided (Heeringa & Van de Velde, 2024; Redmon et al., 2020). Through visual mapping, the articulatory distance between /æ/ and Indonesian vowels became more apparent, prompting learners to reflect on the inadequacy of their initial vowel categorization.

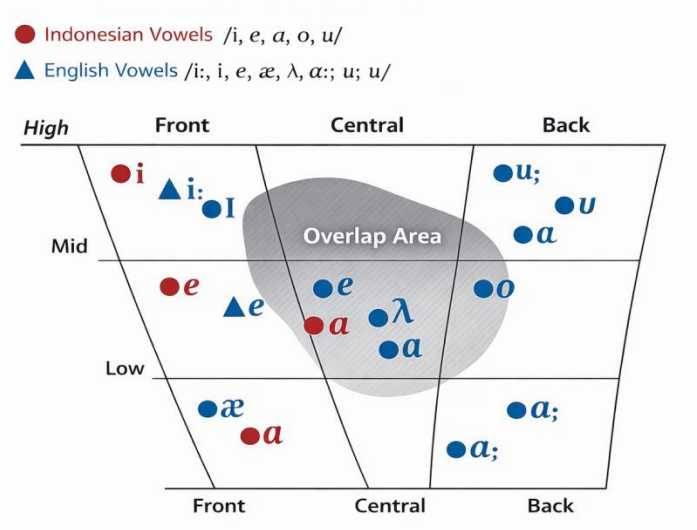


**Figure 1. Comparative vowel space of Indonesian and English vowels illustrating category overlap and vowel neutralization**

In addition to specific vowel substitutions and neutralizations, the data also revealed a broader pattern of over-generalization of Indonesian vowel categories across multiple English vowel targets. Learners frequently applied a limited set of L1 vowels to represent a wide range of English vowels, resulting in systematic

vowel merger and reduced phonological contrast. This over-generalization was evident not only in isolated word production but also in connected speech, indicating that learners relied heavily on L1-based phonological strategies during spontaneous pronunciation (Setyawan & Fata, 2021; Masykar et al., 2025). Such findings are consistent with recent studies on bilingual and EFL vowel perception, which suggest that learners initially interpret L2 vowel systems through the lens of L1 phonological organization (Masykar et al., 2025; Redmon et al., 2020).

Visual vowel mapping again functioned as an effective analytical tool for revealing this over-generalization pattern. By comparing the English and Indonesian vowel spaces visually, learners were able to observe how multiple English vowels were being collapsed into a small number of Indonesian vowel categories. This observation supports recent pedagogical research suggesting that visual-based phonological representation can help learners externalize their internal vowel categorization and become more aware of cross-linguistic interference (Mustika & Pramesti, 2024; Suryani, 2022). Importantly, the findings indicate that visual vowel mapping does not immediately eliminate over-generalization, but rather initiates a process of perceptual restructuring by encouraging learners to question their existing phonological assumptions.

Overall, the patterns identified in this study demonstrate that cross-linguistic vowel interference among Indonesian EFL learners is systematic, perceptually grounded, and closely tied to the structure of the Indonesian vowel system. The consistent substitution of /ɪ/ with /i/, the neutralization of /æ/ into /e/, and the over-generalization of L1 vowel categories collectively highlight the need for pronunciation instruction that goes beyond auditory repetition. By revealing these patterns through visual vowel mapping, this study supports recent calls for pronunciation pedagogy that integrates visual phonetic tools to enhance learners' phonological awareness and to address cross-linguistic interference more explicitly (Putra, 2024; Heeringa & Van de Velde, 2024).

### 1.2 Visual Vowel Mapping as an Analytical Tool

The findings indicate that visual vowel mapping functioned as a powerful analytical tool for uncovering learners' internal phonological representations, particularly in relation to cross-linguistic vowel interference. Rather than merely serving as a visual aid for pronunciation practice, vowel mapping enabled a systematic comparison between the English vowel system and the Indonesian vowel space, allowing both researchers and learners to observe how vowel categories were perceptually organized. This aligns with recent phonological research suggesting that visualization of vowel space can externalize otherwise implicit phonological processes, making them accessible for qualitative analysis (Heeringa & Van de Velde, 2024; Redmon et al., 2020).

Through the use of visual vowel charts and mapped vowel spaces, learners' pronunciation data revealed consistent patterns of category overlap, compression, and misalignment between L1 and L2 vowel systems. These visual representations demonstrated that learners did not randomly mispronounce English vowels, but systematically positioned them within familiar Indonesian vowel categories. Similar observations have been reported in recent studies on cross-linguistic phonological mapping, which argue that learners' vowel errors reflect underlying perceptual assimilation rather than articulatory incapacity (Inceoglu, 2015;

Setyawan & Fata, 2021). Visual vowel mapping thus provided empirical evidence supporting the notion that L2 vowel production is shaped by learners' existing phonological frameworks.

Importantly, the analytical value of visual vowel mapping was evident in its capacity to reveal phonological interference processes that were not immediately detectable through auditory analysis alone. While traditional auditory transcription captured surface-level pronunciation deviations, the visual mapping exposed deeper patterns of vowel merger and category collapse across multiple lexical items. This finding corroborates recent methodological discussions in phonological research, which emphasize the limitations of auditory-only analysis in contexts involving subtle vowel distinctions (Putra, 2024; Mustika & Pramesti, 2024). By visualizing vowel dispersion and overlap, the study was able to trace how learners consistently relied on a reduced vowel inventory during English speech production.

Furthermore, visual vowel mapping facilitated interpretive insights into learners' perceptual awareness and phonological noticing. When learners were invited to reflect on their own mapped vowel spaces, qualitative data from classroom interaction and learner commentary indicated increased awareness of mismatches between intended English vowels and their actual realizations. This process of noticing is central to phonological restructuring, as learners begin to recognize the inadequacy of L1-based categories for representing L2 contrasts (Muna & Husna, 2021; Suryani, 2022). Recent Indonesian studies similarly suggest that visual phonological representations can promote metalinguistic reflection, particularly in pronunciation contexts where learners struggle to perceive contrastive features auditorily (Yani & Zainuddin, 2022; Rachmawati & Andayani, 2023).

From a cross-linguistic perspective, the use of visual vowel mapping also enabled a clearer articulation of how Indonesian phonological constraints shape English vowel acquisition. The mapping illustrated that learners tended to compress the English vowel space to fit within the five-vowel Indonesian system, resulting in predictable zones of interference. Such compression has been documented in recent cross-linguistic phonology studies, which argue that learners initially resist expanding their vowel space unless perceptual support is explicitly provided (Redmon et al., 2020; Heeringa & Van de Velde, 2024). Visual vowel mapping thus served as a diagnostic tool for identifying which English vowel contrasts were most vulnerable to L1 interference.

Crucially, positioning vowel mapping as an analytical instrument contributes to ongoing methodological debates in pronunciation research. While many studies treat visual tools primarily as instructional supplements, this study demonstrates their potential for qualitative phonological analysis, particularly in exploring learner cognition and perceptual categorization. This perspective is consistent with recent calls for integrating visual and qualitative methodologies in second language phonology to capture the complexity of cross-linguistic influence (Lubis et al., 2025; Masykar et al., 2025). By combining spoken data with visual representation, the study provides a more comprehensive account of how vowel interference operates at both perceptual and productive levels.

Overall, the findings from this section underscore that visual vowel mapping is not merely a pedagogical trend, but a methodologically robust tool for examining cross-linguistic vowel interference. Its ability to reveal hidden phonological

patterns, support learner reflection, and bridge auditory and perceptual analysis strengthens its relevance for qualitative EFL phonology research. In the Indonesian EFL context, where vowel interference remains a persistent challenge, visual vowel mapping offers a theoretically grounded and empirically supported approach to understanding learners' pronunciation difficulties (Fitriani & Makmur, 2021; Lubis et al., 2025).

### 1.3 Learners' Perceptual Awareness of English Vowels

The findings indicate a clear distinction between learners' perceptual awareness of English vowels before and after exposure to visual vowel mapping. Prior to visual exposure, learners generally demonstrated limited awareness of English vowel contrasts, particularly those that do not exist in the Indonesian vowel system. Although learners were often able to reproduce English words fluently, qualitative analysis of their spoken data and reflective comments revealed that they perceived several distinct English vowels as identical or interchangeable. This lack of perceptual differentiation is consistent with recent studies suggesting that learners may exhibit apparent fluency while relying on simplified or incomplete phonological representations shaped by their first language (Setyawan & Fata, 2021; Fitriani & Makmur, 2021).

Before engaging with visual vowel representations, learners tended to evaluate their pronunciation accuracy based on intelligibility rather than phonological precision. Many participants expressed confidence in their vowel production as long as communication was successful, indicating that subtle vowel contrasts were not part of their perceptual focus. Similar attitudes have been documented in Indonesian EFL contexts, where pronunciation instruction often prioritizes communicative effectiveness over phonological awareness, resulting in limited attention to vowel quality distinctions (Suryani, 2022; Yani & Zainuddin, 2022). This condition supports the argument that perceptual awareness of L2 vowels does not automatically emerge through exposure or use, but requires targeted instructional and analytical intervention.

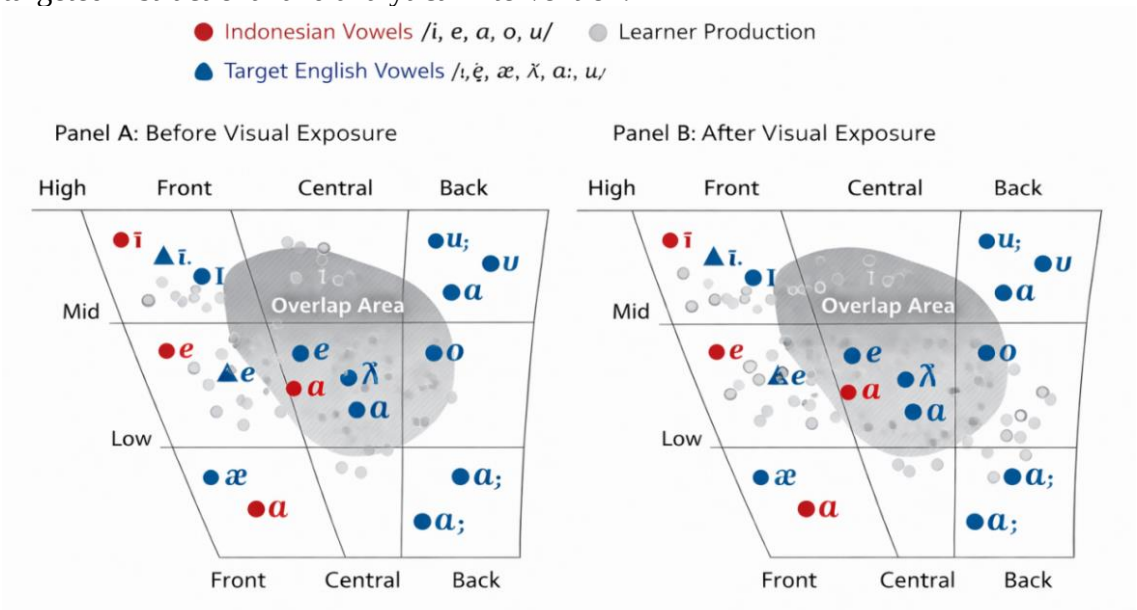


**Figure 2. Learners' vowel production mapping before and after visual vowel exposure**

Following visual exposure, however, learners demonstrated a noticeable shift in perceptual awareness, particularly in their ability to notice vowel contrasts that were previously ignored. Visual vowel mapping enabled learners to observe the spatial relationship between English vowels and their Indonesian counterparts, making abstract phonological distinctions more concrete. Learners' verbal reflections indicated increased awareness of vowel overlap, category compression, and misplacement within the vowel space. This finding aligns with recent phonological research emphasizing that visual cues can enhance perceptual noticing by externalizing internal phonological representations (Redmon et al., 2020; Heeringa & Van de Velde, 2024).

Importantly, the data suggest that visual cues primarily facilitated noticing rather than immediate improvement in pronunciation accuracy. Although learners became more aware of the existence of English vowel contrasts after visual exposure, this awareness did not consistently translate into accurate vowel production. Learners often articulated that they "realized the difference" between vowels such as /ɪ/ and /i/ or /æ/ and /e/, yet still struggled to produce these distinctions reliably in spontaneous speech. This finding supports recent arguments in second language phonology that perceptual awareness is a necessary but insufficient condition for phonological restructuring (Inceoglu, 2015; Putra, 2024).

The distinction between noticing and accuracy is particularly significant in the context of cross-linguistic phonology. Visual vowel mapping appeared to disrupt learners' reliance on L1-based vowel categories by making mismatches visible, but it did not automatically override entrenched articulatory habits. This observation resonates with studies suggesting that phonological change is gradual and requires sustained perceptual recalibration rather than immediate correction (Muna & Husna, 2021; Mustika & Pramesti, 2024). Learners' awareness, therefore, functioned as an initial cognitive shift rather than a final outcome of instruction.

Furthermore, learners' reflections revealed that visual vowel mapping encouraged metalinguistic engagement with pronunciation. Instead of perceiving vowel errors as random mistakes, learners began to conceptualize them as systematic consequences of L1 interference. This shift in perspective has been highlighted in recent qualitative studies as a critical step toward long-term phonological development, as learners become active analysts of their own speech rather than passive imitators (Rachmawati & Andayani, 2023; Lubis et al., 2025). In this sense, visual cues supported a deeper level of phonological awareness that extended beyond surface-level correction.

The findings also suggest that the effectiveness of visual cues depends on how they are integrated into learners' perceptual processes. Visual vowel mapping was most impactful when learners were encouraged to reflect on the mapping in relation to their own pronunciation data, rather than merely observing standard vowel charts. This supports recent pedagogical research arguing that visual tools must be embedded within reflective and analytic activities to influence perceptual awareness meaningfully (Masykar et al., 2025; Suryani, 2022). Without such engagement, visual representations risk being interpreted as static reference materials rather than dynamic analytical instruments.

Overall, this section demonstrates that learners' perceptual awareness of English vowels underwent a qualitative transformation following visual exposure,

characterized by increased noticing, reflection, and metalinguistic understanding. However, the findings also reaffirm that visual cues alone do not guarantee phonological accuracy. Instead, they initiate a process of perceptual and cognitive restructuring that precedes measurable changes in vowel production. This distinction reinforces the value of visual vowel mapping in qualitative phonology research, where the focus lies not solely on outcomes, but on understanding how learners perceive, interpret, and negotiate cross-linguistic phonological contrasts (Heeringa & Van de Velde, 2024; Lubis et al., 2025).

### 1.4 Theoretical Interpretation

The findings of this study provide a comprehensive illustration of how cross-linguistic vowel interference among Indonesian EFL learners operates at perceptual, cognitive, and phonological levels. These findings can be most coherently interpreted through the combined perspectives of the Speech Learning Model, audiovisual phonetic perception, and implicit dual coding processes. Rather than functioning independently, these theoretical frameworks intersect to explain why vowel interference persists, how visual vowel mapping reshapes learners' perceptual awareness, and why increased awareness does not immediately result in phonological accuracy.

Within the framework of the Speech Learning Model, the persistent substitution and neutralization of English vowels observed in this study indicate that learners predominantly relied on established Indonesian vowel categories when processing English vowel input. English vowels such as /ɪ/ and /æ/ were not initially perceived as new phonological categories, but rather assimilated into the closest Indonesian equivalents. This assimilation reflects what recent applications of the Speech Learning Model describe as perceptual equivalence, where learners fail to detect sufficient phonetic distance between L1 and L2 sounds to justify category formation (Inceoglu, 2015; Redmon et al., 2020). The findings reinforce the argument that phonological difficulty in EFL contexts is not merely a matter of articulatory control, but a consequence of perceptual categorization shaped by long-term L1 experience.

Visual vowel mapping, in this regard, functioned as a critical intervention at the perceptual level rather than the articulatory level. By externalizing vowel contrasts through spatial representation, visual mapping disrupted learners' initial assumptions of phonological equivalence. Learners were able to observe that multiple English vowels occupied distinct positions in the vowel space, despite being collapsed into a single category in their production. This perceptual destabilization is consistent with recent interpretations of the Speech Learning Model, which emphasize that perceptual differentiation precedes phonological restructuring and category formation (Putra, 2024; Lubis et al., 2025). Thus, visual vowel mapping did not directly create new vowel categories, but initiated the necessary perceptual conditions for their eventual emergence.

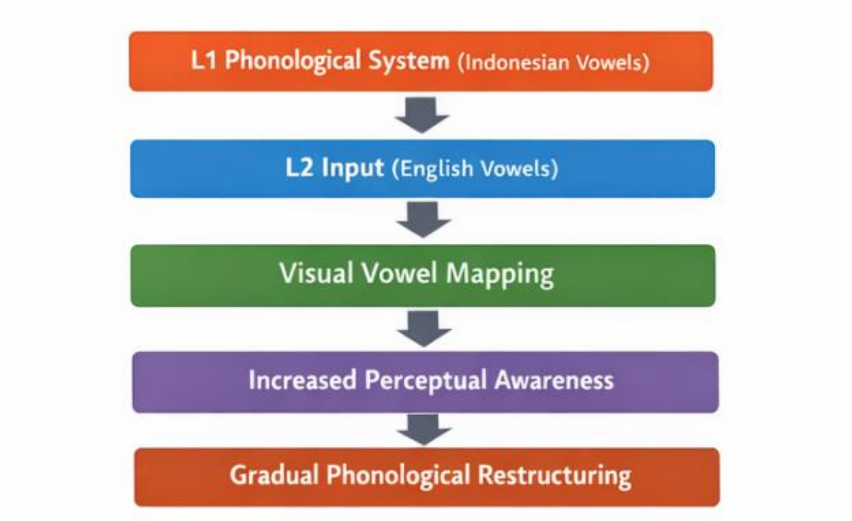


**Figure 3. The role of visual vowel mapping in mediating cross-linguistic vowel perception**

The role of audiovisual phonetic perception further clarifies why visual vowel mapping was effective in enhancing learners' awareness. Contemporary phonetic research increasingly recognizes speech perception as a multimodal process, in which visual information complements auditory input to enhance phonological salience. In this study, learners' increased awareness following visual exposure suggests that auditory input alone was insufficient to overcome L1-based perceptual constraints. Visual cues, such as vowel charts and mapped vowel spaces, provided an additional perceptual channel through which learners could access phonological contrasts that were previously imperceptible (Heeringa & Van de Velde, 2024; Mustika & Pramesti, 2024). This finding supports the argument that vowel interference persists not because learners lack exposure, but because the input is not perceptually processed in a way that challenges existing phonological frameworks.

However, the data also demonstrate that enhanced audiovisual perception does not guarantee immediate phonological accuracy. Learners frequently articulated an increased ability to notice vowel differences without being able to consistently produce them in spontaneous speech. This dissociation between perception and production aligns with recent phonological studies emphasizing that perceptual access and articulatory control develop asynchronously (Muna & Husna, 2021; Rachmawati & Andayani, 2023). From a theoretical standpoint, this suggests that audiovisual input functions as a catalyst for perceptual reorganization rather than a shortcut to accurate production. Phonological change, therefore, should be understood as a gradual process that unfolds over sustained exposure and practice.

Implicit principles of dual coding further illuminate how visual vowel mapping influenced learners' phonological processing. By combining auditory input with visual representation, learners encoded vowel information through parallel representational systems. Although dual coding was not explicitly introduced to learners, the coexistence of visual and auditory channels facilitated deeper cognitive engagement with vowel contrasts. Learners were able to cross-reference what they heard with what they saw, enabling a more elaborated mental

representation of English vowels. Recent studies in EFL learning suggest that such multimodal encoding enhances conceptual understanding and metalinguistic awareness, even when immediate performance gains remain limited (Suryani, 2022; Masykar et al., 2025).

Crucially, learners' reflections revealed that visual vowel mapping encouraged a reconceptualization of pronunciation difficulty itself. Rather than perceiving vowel errors as individual failures, learners began to interpret them as systematic outcomes of cross-linguistic phonological constraints. This shift in perspective represents a significant cognitive transformation, as learners move from surface-level correction to structural understanding of their pronunciation patterns. Similar findings have been reported in recent Indonesian pronunciation research, which highlights metalinguistic awareness as a key mediator in long-term phonological development (Yani & Zainuddin, 2022; Fitriani & Makmur, 2021).

Taken together, these theoretical interpretations position visual vowel mapping as a mediating mechanism between perception, cognition, and phonological structure. Within the Speech Learning Model, it facilitates perceptual differentiation; through audiovisual phonetic perception, it enhances the salience of phonological contrasts; and in line with dual coding principles, it supports integrated phonological representation. Importantly, visual vowel mapping should not be conceptualized as a corrective technique aimed at immediate accuracy, but as a cognitive and perceptual scaffold that enables learners to renegotiate entrenched L1-based phonological assumptions.

As the final section of the Results and Discussion, this synthesis underscores the central contribution of the study: cross-linguistic vowel interference in Indonesian EFL contexts is best understood not as a deficiency in learners' ability, but as a predictable outcome of perceptual and cognitive constraints shaped by the first language. Visual vowel mapping emerges as a theoretically grounded and empirically supported approach for making these constraints visible, interpretable, and pedagogically addressable. By integrating cross-linguistic phonology with multimodal perception, this study extends current discussions in EFL pronunciation research and offers a robust conceptual foundation for future qualitative and pedagogical investigations (Heeringa & Van de Velde, 2024; Lubis et al., 2025).

## CONCLUSION

This study explored cross-linguistic vowel interference among Indonesian EFL learners through the use of visual vowel mapping. The findings revealed systematic patterns of vowel substitution, neutralization, and over-generalization caused by learners' reliance on Indonesian vowel categories when perceiving and producing English vowels. These patterns indicate that pronunciation difficulties are closely related to first-language phonological influence rather than isolated pronunciation errors.

The study also demonstrated that visual vowel mapping functioned as an effective analytical and pedagogical tool for enhancing learners' perceptual awareness of English vowel contrasts. Although visual exposure did not immediately improve pronunciation accuracy, it helped learners recognize differences between English and Indonesian vowel systems and increased their phonological awareness. The findings support cross-linguistic phonology and

second language speech learning theories by showing that phonological development begins with perceptual restructuring. Methodologically, the study highlights the value of integrating spoken data, learner reflection, and visual representation in qualitative pronunciation research. Future studies may employ longitudinal or acoustic approaches to further examine phonological development among EFL learners.

## REFERENCES

Arafiq, A., Yusra, K., & Saputra, A. (2020). Mapping students' phonological problems in pronouncing English sounds. *Advances in Social Science, Education and Humanities Research, 449*, 285–289. https://doi.org/10.2991/assehr.k.200827.050

Fitriani, S. S., & Makmur, M. (2021). Indonesian EFL learners' difficulties in producing English front vowels. *Lingua: Journal of Language, Literature and Teaching, 18*(2), 145–158. https://doi.org/10.30957/lingua.v18i2.657

Heeringa, W., & Van de Velde, H. (2024). Visualizing vowel spaces: A tool for phonological analysis and dialect comparison. *Journal of Phonetics, 103*, Article 101318. https://doi.org/10.1016/j.wocn.2023.101318

Inceoglu, S. (2015). Perceptual assimilation of L2 vowels and implications for second language phonology. *Second Language Research, 31*(4), 523–546. https://doi.org/10.1177/0267658315576863

Lubis, A. H., Rahman, F., & Yusuf, Y. Q. (2025). Visual phonetic instruction and vowel perception in Indonesian EFL contexts. *Lingua: Journal of Language and Culture, 21*(1), 33–48. https://doi.org/10.24071/ljlc.v21i1.8123

Masykar, T., Fauzan, U., & Lestari, S. (2025). Multimodal pronunciation learning and phonological awareness among EFL learners. *Journal of Language Teaching and Research, 16*(2), 289–301. https://doi.org/10.17507/jltr.1602.09

Muna, N., & Husna, N. (2021). Perceptual awareness in English vowel learning: Evidence from Indonesian learners. *Indonesian Journal of Applied Linguistics, 11*(3), 512–523. https://doi.org/10.17509/ijal.v11i3.34672

Mustika, R., & Pramesti, R. (2024). Visual phonetic tools in EFL pronunciation instruction: A qualitative perspective. *Journal of English Education and Linguistics, 5*(1), 67–82. https://doi.org/10.21462/jeel.v5i1.912

Putra, A. D. (2024). Rethinking vowel perception in Indonesian EFL learners: A cross-linguistic phonological approach. *Studies in English Language and Education, 11*(1), 1–16. https://doi.org/10.24815/siele.v11i1.32647

Rachmawati, R., & Andayani, S. (2023). English vowel production errors by Indonesian EFL learners: A phonological analysis. *Journal of Applied Linguistics and Literacy, 7*(2), 120–134. https://doi.org/10.25157/jall.v7i2.7281

Redmon, A., Oleson, J., & Flege, J. E. (2020). Cross-language vowel perception and the role of phonetic similarity. *Journal of Phonetics, 83*, Article 101002. https://doi.org/10.1016/j.wocn.2020.101002

Setyawan, A., & Fata, I. A. (2021). L1 phonological interference in English vowel production by Indonesian learners. *Journal of English Language Teaching, 10*(3), 345–356. https://doi.org/10.24036/jelt.v10i3.112589

Suryani, L. (2022). Enhancing pronunciation awareness through visual media in EFL classrooms. *Indonesian EFL Journal, 8*(2), 89–101. https://doi.org/10.25134/ieflj.v8i2.5632

Yani, A., & Zainuddin, Z. (2022). English vowel perception among Indonesian learners: A qualitative inquiry. *Lingua Didaktika, 16*(1), 41–54. https://doi.org/10.24036/ld.v16i1.112450

Yuliana, D., & Susanti, N. (2020). Corrective phonetics: Teaching English vowels through minimal pairs and visual mapping. *Celt: A Journal of Culture, English Language Teaching & Literature, 20*(1), 1–15.