Volume 6, Issue 1, February 2025

J-LELC
Journal of Language Education, Linguistics, and Culture
P-ISSN 2774-6003 E-ISSN 2775-099X

# An Analysis Of EFL Students' Difficulties In Speech Perception: A Qualitative Study On Misheard Minimal Pairs

**Puspa Triyani Waruwu[1], Yunita Tafonao[2], Merlin Waruwu[3], Arniwati Gulo[4], Afore Tahir Harefa[5]**

Universitas Nias
waruwupuspatriani@gmail.com[1], yunitatafonao2005@gmail.com[2], merlinwaruwu8@gmail.com[3]
arniguloarnigulo@gmail.com[4], aforetahirharefa@unias.ac.id[5]

| Received: January 2026 | Revised: February 2026 | Accepted: February 2026 |
|---|---|---|

**Abstract**

Perception of speech is a fundamental skill in EFL learning that is very challenging for Indonesian learners due to the limitations of their L1 vocal system. Although previous studies have mostly used quantitative accuracy measurements, qualitative exploration of authentic experiences in the classroom is still limited, especially for university-level EFL students in the context of psycholinguistics. This study investigates the difficulties of fifth-semester English Education students at Nias University in perceiving minimal vowel pairs through a descriptive qualitative design. Data from 15 purposive sampling participants were collected through a minimal pairs listening test (20 items), semi-structured interviews, and classroom observations. The results show the highest error rates of 87% for /ɪ/–/iː/ (ship-sheep) and 80% for /ʊ/–/uː/ (full-fool), caused by the Indonesian vowel system, which has only five monophthongs without length/tension contrast. Internal factors included L1 phonological interference and cognitive load, while external factors included poor classroom audio quality and environmental noise. Students relied more on context-based top-down strategies than bottom-up phonetic analysis. These findings reveal patterns of vowel perception difficulties characteristic of Indonesian EFL in an authentic learning context.

**Keywords:** Speech Perception, Minimal Pairs, Indonesian EFL,Vowel Contrast, Phonological Awareness



Citation:
Waruwu, P. T., Tafonao, Y., Waruwu, M., Gulo, A., Harefa, A. T. 2026. An Analysis Of EFL Students' Difficulties In Speech Perception: A Qualitative Study On Misheard Minimal Pairs. *J-LELC: Journal of Language Education, Linguistics, and Culture*, 6(1), 65-71. https://doi.org/10.25299/j-lelc.2026.27015

**UIR PRESS/J-LELC: Journal Language Education, Linguistics, and Culture**

## INTRODUCTION

Speech perception serves as the most fundamental foundation in the process of second language or foreign language acquisition. Theoretically, this ability is not merely a passive auditory process but a fundamental skill involving the accurate recognition and discrimination of speech sounds. In the discipline of psycholinguistics, speech perception is viewed as the primary gateway to effective oral communication comprehension. Without the ability to distinguish precise phonemes, a learner will remain trapped in continuous semantic ambiguity. This poses a very real challenge for English as a Foreign Language (EFL) learners in Indonesia, where the native language's vowel system often conflicts with the English phonological system.

One of the most difficult manifestations of speech perception is the discrimination of minimal pairs. Minimal pairs are defined as pairs of words that differ by only a single phoneme, yet this slight difference drastically alters the entire meaning of the word. Classic examples frequently encountered include the distinction between the words pull and pool, or full and fool. Although the acoustic difference appears very subtle, failure to detect it can lead to fatal errors in message interpretation. This phenomenon is reinforced by the view of Sailuddin et al. (2025), who assert that minimal pairs are often the primary source of perceptual errors because subtle phonemic differences frequently go undetected by the ears of phonologically untrained learners. This lack of phonological sensitivity contributes directly to comprehension failure during listening exercises.

Cognitively, speech perception is a complex process. Zhang et al. (2023) explain that this mechanism involves a two-way integration: bottom-up processing, which relies on raw acoustic information from incoming sound, and top-down processing, which utilizes the listener's linguistic knowledge and context. For EFL learners, an imbalance between these two processes often occurs. If the target language's phonological system has not been fully absorbed into long-term memory, learners tend to perform incorrect sound categorization; they will perceive two acoustically different sounds as the same if those sounds lack a corresponding equivalent in their native language's phonological categories.

This difficulty is further exacerbated by the factor of cognitive load. Shaojie et al. (2022) state that complex or unfamiliar linguistic input creates a high cognitive load, which in turn hinders the learner's ability to capture subtle phonetic details. When the brain is overly occupied with understanding the overall sentence structure, attention to vowel duration or vocal fold tension is often neglected. This condition explains why perceptual errors remain frequent even at the higher education level, where students should theoretically possess sufficient understanding. As emphasized by Kajiura et al. (2023), failure to recognize basic sounds will lead to systemic failure in understanding words, phrases, and spoken messages in their entirety.

This gap between theory and practice is vividly observed among fifth-semester students of Class A in the English Education Study Program at Nias University, specifically within the Psycholinguistics course. Based on initial classroom observations, it was found that students still experience significant difficulty in distinguishing English vowel contrasts, as seen in the identification errors of word pairs like pull-pool and full-fool. This phenomenon is quite ironic, considering that students at this level should have completed foundational courses such as Phonetics and Phonology. This proves that theoretical mastery of the material does not automatically improve speech perception abilities in actual communication situations.

A review of previous research indicates a methodological and theoretical gap that needs to be filled. Most prior studies have focused more on quantitative assessments of sound perception accuracy. However, qualitative studies that explore perceptual challenges in depth based on authentic classroom experiences remain very limited. Particularly within the Nias University environment, research specifically dissecting how students process vowel minimal pairs in the context of a Psycholinguistics class is virtually non-existent. This research aims to bridge that gap.

Through a qualitative approach, this study intends not only to count the number of errors but to delve deeper into the patterns of minimal pair perception errors and identify the underlying cognitive and linguistic elements. Furthermore, this research seeks to uncover the strategies used by students to overcome these perceptual barriers in natural learning situations. By understanding these challenges from a more holistic perspective, it is hoped that the results of this study can provide a tangible contribution to the development of teaching materials and instructional strategies that are more precisely targeted for EFL learners in Indonesia. Overall, the urgency of this research lies in its

effort to reconnect students' declarative knowledge (theory) with their procedural ability (practice) in speech perception. By mapping specific difficulties, educators can design more effective exercises to enhance students' phonological awareness, ensuring they do not merely "hear" but truly "perceive" English sounds accurately according to the correct phonemic categories.

## RESEARCH METHOD

### Research Design

This study examines students' challenges in identifying English vowel minimal pairings using a descriptive qualitative research design. Because speech perception in a second or foreign language includes intricate cognitive and perceptual processes that cannot be well explained by numerical data alone, this design was selected. According to recent research, individual cognitive characteristics including phonological short-term memory, attention, and listening experience have a significant impact on errors in L2 vowel contrast perception(Georgiou 2023). These variables necessitate meaning-based, in-depth investigation, which is best served by a qualitative method. Studies on the hearing abilities of EFL learners reveal challenges in differentiating vocal contrasts associated with exposure to English input and subjective listening (Lee and Ahn 2025). Learners' perceptions of similar vowel sounds in real-world situations are captured using descriptive qualitative designs. When paired with qualitative information that explains perceptual errors, minimal pair listening tasks are valid (Uchihara,et., al 2025). In the natural setting of the participants, researchers are the main tool.

### Research Participants

The participants consisted of 15 fifth-semester students in class A of the English Education Study Program at Nias University who were studying psycholinguistics. Selection was conducted through purposive sampling with the following criteria: (1) actively attending the Psycholinguistics course, (2) demonstrating minimal difficulty in perceiving vowel pairs (full-fool, pull-pool) based on initial observations, (3) willing to voluntarily participate in listening tests, interviews, and observations. A total of 15 participants were selected to achieve thematic saturation in in-depth qualitative analysis**.**

### Research Instruments

In qualitative research, the researcher acts as the main instrument. Four supporting instruments ensure systematic data collection:

1. The Minimal Pairs Listening Test (20 items) identifies patterns of vocal perception errors (/ɪ-iː/, /ʊ-uː/, /æ-e/) that are problematic for Indonesian EFL learners. Each item is played twice; participants write down the words they hear. Valid for L2 phonetic perception (Uchihara et al. 2025).
2. Semi-Structured Interview Guide explores listening experiences, causes of errors, strategies (e.g., “How to distinguish ship-sheep or full-fool?”). Captures subjective factors beyond test scores (Lee and Ahn 2025).
3. Observation Checklist/Field Notes document contextual factors (audio quality, concentration, noise) affecting perception (Georgiou 2023)
4. Documentation (audio recordings, transcripts, answer sheets) supports data credibility and triangulation.

### Data Analysis

1. Data Selection: Selecting, grouping, and coding test and interview data.
2. Data Presentation: Frequency tables of misperceptions based on vocal contrasts; interview excerpts and thematic findings.
3. Conclusion Drawing & Verification: Formulating themes that answer the research questions. Verified through: (1) triangulation of techniques (tests, interviews, observations), (2) member checking with participants.

## RESULTS AND DISCUSSION

### Results

The research results from minimal pairs listening tests, semi-structured interviews, and classroom observations are presented in this part using the methods outlined in the Research Methodology. In order to address the research questions about the kinds of challenges EFL students face when perceiving English minimal pairs, the variables that affect sound misperceptions, and the students' experiences and methods for perceiving phonologically similar sounds, the results are methodically arranged.

The results of the minimal pairs listening test show that all students (n=15) had difficulty perceiving English vowel minimal pairs with a relatively consistent pattern among participants. The vowel contrast /ɪ/–/iː/ was most often misperceived by 13 students (87%), followed by /ʊ/–/uː/ with 12 students (80%). The most errors occurred in the word pairs ship–sheep, bit–beat, full–fool, and pull–pool. Students tended to perceive long and short vowels as the same sound when pronounced at normal speed without special emphasis.

The contrast /æ/–/e/ in the pairs bet–bat and peck–pack showed a moderate level of difficulty, with 9 students (60%) misidentifying them. In contrast, the contrasts /ɒ/–/ʌ/ (dog–dug) and /ɪ/–/e/ (still–steel) were relatively easier, with 5 students (33%) and 4 students (27%) making mistakes, respectively. This variation in difficulty level shows that not all vowel contrasts pose the same challenge for students.

**Table 1. Patterns of Misheard English Vowel Minimal Pairs (n = 15)**

| No | Vowel Contrast | Example Minimal Pairs | Students with Errors | Difficulty Level |
|---|---|---|---|---|
| 1. | /ɪ/–/iː/ | ship–sheep, bit–beat | 13 Students (87%) | Very high |
| 2. | /ʊ/–/uː/ | full–fool, pull–pool | 12 Students (80%) | Very high |
| 3. | /æ/–/e/ | bet–bat, peck–pack | 9 Students (60%) | Moderate |
| 4. | /ɒ/–/ʌ/ | dog–dug | 5 Students (33%) | Low |
| 5. | /ɪ/–/e/ | still–steel | 4 Students (27%) | Low |

The data presented in Table 1 provides an empirical overview of the complex phonological challenges faced by English Education students at Nias University. Substantially, these error patterns indicate a non-random hierarchy of perceptual difficulty, rooted in the typological differences between the phonological systems of the native language (L1) and the target language (L2). The following is an in-depth analysis of these findings:

**The Perceptual Crisis in High Vowels (/ɪ/–/iː/ and /ʊ/–/uː/)**

The findings reveal extreme error rates regarding high vowel contrasts, specifically 87% for the /ɪ/–/iː/ pair and 80% for /ʊ/–/uː/. Scientifically, this phenomenon is categorized as "perceptual assimilation," where listeners fail to detect distinctive vowel features. The Indonesian language possesses only five monophthongs (/a,i,u,e,o/) which do not recognize differences in duration (length) or tenseness as semantic dividers. Consequently, the students' cognitive systems automatically categorize both English sounds into a single vowel "slot" within their memory. This is further supported by qualitative data showing significant student confusion during the sound discrimination process.

> *"During the test, to be honest, I was confused distinguishing between 'ship' and 'sheep'. To my ears, they sounded exactly the same, like the ordinary 'i' sound in Indonesian. I didn't realize that one should be pronounced longer or with more tension."*

This failure has implications for broader oral communication failure, as words such as pull and pool are perceived as homophones. When placed incorrectly within a sentence lacking strong contextual cues, these words obscure the intended message.

**Gradation of Difficulty in Mid and Low Vowels**

Interestingly, the data indicates a decrease in the error rate for the /æ/–/e/ contrast, which stands at 60% (Moderate). Articulatory speaking, mid-front vowels possess a wider acoustic distance compared to high vowels; thus, although students still experience interference, they are more capable of capturing the differences in sound quality. Meanwhile, for low vowel contrasts such as /ɒ/–/ʌ/ (33%) and /ɪ/–/e/ (27%), the difficulty level falls into the "Low" category. This confirms the theory that the greater the acoustic distance and the difference in physical features between two phonemes, the lower the likelihood of perceptual overlap. Listeners are able to utilize clearer bottom-up signals to distinguish these sounds without having to rely too heavily on contextual guessing.

**Dissociation Between Declarative and Procedural Knowledge**

These findings reveal a fundamental issue in language education: the gap between theory and practice. Although these fifth-semester students have theoretically studied phonetics and phonology, their ability to apply that knowledge automatically (in real-time) during listening remains significantly weak. This aligns with the research assertion that speech perception is not a passive process but rather a cognitive integration influenced by the brain's workload. The following interview excerpt illustrates how this cognitive load hinders phonetic analysis:

> "The lecturer or the audio plays it at a normal speed, but to me, it feels too fast. Just as I am trying to recall whether that was a long /iː/ sound, the audio has already moved on to the next word. In the end, I feel pressured and lose focus on the sound itself."

This condition forces students to rely on top-down strategies (guessing meaning based on the sentence) because they are no longer able to accurately discriminate the incoming acoustic signals.

**The Impact of Contextual and External Variables**

The interpretation of the data in Table 1 cannot be separated from the external factors identified during classroom observations. It was noted that inconsistent audio quality and environmental noise at Nias University contributed significantly to the students' perceptual failures. Under noisy conditions, subtle vowel features such as duration become increasingly difficult to discern, which in turn drives the error rates for high vowels above the 80% threshold. Students also expressed a sense of insecurity resulting from these environmental barriers:

> "I feel a lack of confidence whenever there is a minimal pairs listening task. Especially if the speakers aren't clear or if there's noise from outside the classroom, I become even more hesitant. I'm afraid of mishearing."

Overall, Table 1 is not merely a collection of figures but clear evidence of systemic challenges in EFL vowel acquisition in Indonesia. These data suggest the need for a radical shift in teaching methods moving from simple theoretical delivery to intensive and repetitive sound discrimination exercises, such as High Variability Phonetic Training (HVPT). The integration of in-depth sound analysis (bottom-up) and contextual understanding (top-down) must become the central axis of the Psycholinguistics curriculum to ensure students are capable of decoding English phonology with precision.

**Discussion**

This study found a clear pattern of misperception among fifth-semester EFL students majoring in English Education at Nias University. The highest error rate occurred in the contrast between the vowels /ɪ/–/iː/ (87%) and /ʊ/–/uː/ (80%). These figures indicate the great difficulty EFL students have in distinguishing between long and short vowels in English. The main cause is the Indonesian vowel system, which has only five simple vowels without any distinction in vowel length. Therefore, EFL students naturally consider the pairs ship-sheep and full-fool to be the same sounds because there is no similar contrast in their first language.

These findings are consistent with other Indonesian EFL studies. (Shaojie, et.,al 2022) also found that high vowel contrasts pose the greatest challenge for Indonesian learners due to the strong

influence of their first language. The moderate error rate of 60% for the /æ/–/e/ contrast indicates that mid-front vowels are easier to recognize because the sound difference is clearer and the acoustic distance is greater. Lower error rates of below 35% for the contrasts /ɒ/–/ʌ/ and /ɪ/–/e/ further reinforce that vowel pairs with greater sound differences are easier for EFL students to understand, even though there are still no direct equivalents in the Indonesian vowel system (Firdaus 2019).

A thorough picture of the real EFL learning environment can be obtained through classroom observation. Classroom learning demonstrates a variety of environmental elements that impact EFL students' sound perception skills, in contrast to laboratory studies, where conditions are rigorously controlled. Real difficulties are caused by the classroom speakers' uneven sound quality, outside noise, and regular speaking speed without repetition. These circumstances are very comparable to the everyday communication scenarios that EFL students in Indonesia frequently encounter, where extrinsic factors like these frequently interfere with the ability to recognize small sounds.

In this study, EFL students' listening strategies place more emphasis on comprehending sentences' overall meaning than on focusing on particular vowel sounds. This top-down method works well for comprehending general communication signals, but it is less suitable for minimal pair discriminating tasks that need for accuracy in vowel quality and duration. When taking listening examinations, only a tiny percentage of EFL students intentionally identify and pay attention to vowel length discrepancies. This suggests that pupils have a great chance to increase their phonological awareness with more focused and organized practice.

The findings of this study have practical implications that can be directly applied in English language learning for EFL students. Specialized exercises targeting problematic high vowel contrasts such as /ɪ/–/iː/ and /ʊ/–/uː/ need to be the main focus in related courses. A teaching approach that combines explicit phonological awareness exercises with effective listening strategies will be very helpful. This strategy involves integrating detailed sound analysis (bottom-up processing) with sentence context comprehension (top-down processing). In addition, improving classroom audio facilities and simulating listening exercises in real-life conditions will better prepare EFL students to face everyday communication challenges.

This study was successful in identifying particular patterns of difficulty Indonesian EFL students encountered when trying to differentiate English vowel sounds. The findings indicate that high vowel contrasts are far more difficult than mid and low vowels. Both the students' objective performance from the hearing test results and their subjective experiences in the setting of true and actual classroom learning were successfully disclosed by the qualitative approach employed in this study (Tang, et,al 2025) .

The ability to capture the complexity of EFL learning as it occurs in students' real lives is the primary benefit of data gathering in a classroom context. In contrast to laboratory studies, which are frequently excessively controlled, this study examines the real-world learning challenges that EFL students encounter on a regular basis. Compared to quantitative approaches alone, the combination of three data collection methods minimal pairs tests, in-depth semi-structured interviews, and field observations offers a more comprehensive and comprehensive explanation of the phenomenon of speech perception. These findings' ecological validity guarantees that the research findings have a direct bearing on the creation of English language curricula and instructional strategies.

This study also emphasizes the critical connection between EFL students' practical skills and their comprehension of psycholinguistic theory. It is still difficult for students to apply phonological ideas in real-time listening settings at regular speaking speeds, even when they have studied them theoretically in class. These results highlight the necessity of building a more robust bridge in EFL instruction between theoretical knowledge (declarative knowledge) and practical abilities (procedural knowledge). The speech perception abilities of EFL students in English language education will increase more significantly when planned, intensive practice sessions are combined with conceptual comprehension.

## CONCLUSION

This study concludes that fifth-semester EFL students majoring in English Education at Nias University experience the highest difficulty in perceiving the contrast between the vowels /ɪ/–/iː/ (87%) and /ʊ/–/uː/ (80%). The main findings from the minimal pairs listening test, semi-structured interviews, and classroom observations indicate that the limited vowel system in Indonesian causes

the ship-sheep and full-fool pairs to be grouped as similar sounds. Contributing factors include classroom audio quality and normal speaking speed. Students' strategies relied more on contextual meaning than detailed phonetic analysis. The results of this classroom-based study confirm the pattern of vowel perception difficulties typical of Indonesian EFL and support the development of contextual minimal pairs exercises.

## REFERENCES


Firdaus, Meisya Ivone. 2019. *Listening Comprehension Strategies Used by EFL Students in State University of Surabaya*. Vol. 7.

Flege, James Emil, Katsura Aoyama, and Ocke-Schwen Bohn. n.d. *The Revised Speech Learning Model (SLM-r) Applied*.

Georgiou, Georgios P. 2023. "Nonverbal Intelligence Does Matter for the Perception of Second Language Sounds." *Psych* 5(1):123–32. doi:10.3390/psych5010012.

Holt, Lori L. 2025. "Speech Perception Is Speech Learning." *Current Directions in Psychological Science* 34(4):218–24. doi:10.1177/09637214251318726.

Kajiura et., al. 2023. *Fast-Rate Multimodal Training Improves L2 Fast-Rate Multimodal Training Improves L2 Listening and Fast-Speech Adjusting Skills Names and Contact Information for All Authors Fast-Rate Multimodal Training Improves L2 Listening and Fast-Speech Adjusting Skills 2 3*. https://ssrn.com/abstract=4467038.

Kogan, Vita V, and Joan C. Mora. 2022. *The Effects of Individual Differences in Native Perception on Discrimination of a Novel Non-Native Contrast*. https://doi.org/10.

Lee, Kyunga, and Hyunkee Ahn. 2025. "Blended Phonetic Training with HVPT Features for EFL Children: Effects on L2 Perception and Listening Comprehension." doi:10.3390/languages.

Mackey et., al. 2022. *Second Language Research*.

Rakat, Bilqisti, Ika Handayani, and Delsa Miranty. 2024. "The Effectiveness of Using Minimal Pairs towards Students' Pronunciation Mastery." *Journal of English Language Teaching and Cultural Studies* 7(1):1–13. doi:10.48181/jelts.v7i1.22262.

Rocabado, Francisco, Gianna Schmitz, and Jon Andoni Duñabeitia. 2025. "You Can Stand Under My Umbrella: Cognitive Load in Second-Language Reading." *Behavioral Sciences* 15(8). doi:10.3390/bs15081051.

Sailuddin, Sartika Putri, Zainal Rafli, Muhammad Kamal, and Abdul Hakim. 2025. "Audio versus Multimodal Input: A Case Study of Speech Perception among Learners in English as a Foreign Language (EFL) Context." *Journal on Interdisciplinary Studies in Humanities* 8:2025. doi:10.34050/elsjish.v8i4.49140.

Segretin, María Soledad, Gonzalo Daniel, and Tan Shaojie. n.d. *Systematic Literature Review on Audio-Visual Multimodal Input in Listening Comprehension*.

Tang, Chuanhong, Danfeng Huang, and Andrew K. F. Cheung. 2025. "Cognitive Constraints in Bilingual Processing an Entropy-Based Discrimination between Translation and Second Language Production." *Frontiers in Psychology* 16. doi:10.3389/fpsyg.2025.1503147.

Uchihara, Takumi, Michael Karas, and Ron I. Thomson. 2025. "High Variability Phonetic Training (HVPT): A Meta-Analysis of L2 Perceptual Training Studies." *Studies in Second Language Acquisition* 47(3):794–827. doi:10.1017/S0272263125100879.

Yani Lubis, Afrohul Fadhillah Batubara, and Jihan Aldie Fathani. 2023. "Understanding English Minimal Pairs of Vowel." *CENDEKIA: Jurnal Ilmu Sosial, Bahasa Dan Pendidikan* 3(3):164–75. doi:10.55606/cendikia.v3i3.1424.

Zhang, Ye, Rong Ding, Diego Frassinelli, Jyrki Tuomainen, Sebastian Klavinskis-Whiting, and Gabriella Vigliocco. 2023. "The Role of Multimodal Cues in Second Language Comprehension." *Scientific Reports* 13(1). doi:10.1038/s41598-023-47643-2.