


English Diphthong Pronunciation: A Quantitative Evidence of Articulatory Difficulties Among EFL Learners

 <https://doi.org/10.31004/jele.v11i4.2751>

*Anggun Angriani, Konder Manurung, Hastini^{abc} 

¹²³Tadulako University, Indonesia

Corresponding Author: rianianggun@gmail.com

A B S T R A C T

This study investigated students' difficulties in pronouncing English diphthongs and the factors contributing to those difficulties among second-semester students of the English Education Study Program at Tadulako University. A quantitative descriptive design was employed. Data were collected through a pronunciation test and a questionnaire. The pronunciation test was used to identify students' ability in pronouncing English diphthongs, while the questionnaire examined factors affecting their pronunciation difficulties. The findings revealed that students experienced varying levels of difficulty in producing English diphthongs, particularly in maintaining smooth vowel glides and accurate articulatory movements. Several students tended to simplify diphthongs into monophthongs or produce incomplete vowel transitions, reflecting the influence of their first language. Questionnaire results showed that learning environment was the most influential factor ($M = 3.12$), followed by articulatory difficulty ($M = 2.95$) and learning media ($M = 2.90$), all categorized as high-level factors. Teaching materials were perceived as moderately influential, with mean scores of 2.68 and 2.79. These findings indicate that students' pronunciation difficulties are influenced by a combination of articulatory complexity, first-language interference, learning conditions, instructional resources, and limited opportunities for pronunciation practice. Overall, the study highlights the importance of creating supportive learning environments and providing systematic pronunciation instruction to improve EFL learners' mastery of English diphthongs.

Keywords: *English Diphthongs, Pronunciation Difficulties, EFL Learners, Learning*

Article History:

Received 12th May 2026

Accepted 19th June 2026

Published 20th June 2026



INTRODUCTION

English pronunciation is a central component of effective communication, particularly in contexts where English is used as a foreign language. Within pronunciation, diphthongs play a crucial role because they occur in many high-frequency English words, influence word meaning, and strongly affect speech intelligibility. Diphthongs require a smooth glide from one vowel quality to another, making them more complex than simple vowels. When these sounds are not produced accurately, the listener may misunderstand the message or perceive the speech as unnatural. Because of this, mastering English diphthongs is not only important for achieving clear communication but also for developing overall phonological competence that supports learners' confidence and fluency in speaking.

The importance of diphthong mastery is particularly recognized by those involved in English language education. English Education students, as future teachers, are expected to possess accurate pronunciation so they can serve as reliable pronunciation models for their own learners. Lecturers emphasize this skill because pronunciation is closely linked to

teaching quality, especially in speaking and listening classes. Teacher-training programs and curriculum planners also consider accurate diphthong production essential, as it reflects students' phonetic awareness and readiness to teach. When teacher candidates demonstrate weak pronunciation, especially in fundamental sounds like diphthongs, it raises concerns about their ability to teach English effectively in the future. Therefore, multiple stakeholders— students, lecturers, and institutions—view diphthong mastery as an important competency in teacher preparation.

Despite its importance, many English Education students at Tadulako University still experience difficulties in producing English diphthongs accurately. Through informal observations in classrooms, student presentations, and recorded speaking tasks, the researcher noticed recurring issues such as reduced vowel glides, substitution with Indonesian monophthongs, and inconsistent articulation. These problems appear even among students who have completed phonetics or pronunciation courses, suggesting that exposure and practice during coursework may not be sufficient for full mastery. In addition, similarities and differences between Indonesian and English vowel systems likely contribute to these difficulties. Since Indonesian does not use diphthongs in the same way English does, students may struggle to transition smoothly between two vowel qualities within a single syllable. This situation indicates that diphthong pronunciation remains a persistent challenge that deserves systematic examination.

Research on English diphthong pronunciation among Indonesian EFL learners has been conducted in various educational contexts, and most studies indicate that diphthongs remain one of the most challenging vowel categories for learners. One recent study by (Suciati et al., 2024) examined university students' pronunciation of English diphthongs and found recurring difficulties in producing accurate vowel glides, particularly for /eɪ/, /aɪ/, and /oʊ/. Many students reduced these diphthongs into monophthongs, which suggests difficulties in coordinating articulatory movement during pronunciation.

Another study conducted by Alimin, (2022) investigated English diphthong pronunciation problems among fourth-semester English Education students at Qomaruddin University. The findings revealed that diphthongs such as /aʊ/ and /ɔɪ/ were frequently mispronounced due to first language interference. In addition, the study highlighted that the teaching materials used in class provided limited practice and lacked sufficient audio models, which restricted students' exposure to accurate diphthong pronunciation.

A further investigation by Cahyaningrum, (2023) focused on diphthong pronunciation in a microteaching class involving English Education students preparing for teaching practice. The results showed that many students produced weak or inconsistent diphthong transitions in their spoken English. The study emphasized that limited speaking opportunities, minimal pronunciation-focused feedback, and low exposure to English outside the classroom contributed to students' pronunciation difficulties, pointing to the importance of a supportive learning environment.

This study is grounded in theories of phonology and second language acquisition that explain why learners experience difficulties in pronouncing English diphthongs and how learning conditions influence those difficulties. English diphthongs are phonologically complex sounds that require a smooth glide from one vowel quality to another within a single syllable. According to phonetic theory, producing diphthongs involves coordinated movement of the tongue, lips, and jaw, which can be challenging for learners whose first language does not use similar vowel glides (Ladefoged & Johnson, 2015; Roach, 2009). This articulatory complexity provides a foundational explanation for why diphthongs are often perceived as difficult by EFL learners.

From a contrastive perspective, Contrastive Analysis Theory proposed by Lado, (1957) suggests that learners tend to experience difficulty when the sound system of the target language differs from that of their first language. Indonesian has a vowel system that is largely composed of monophthongs and does not require the dynamic vowel movement

found in English diphthongs. As a result, Indonesian learners may transfer pronunciation patterns from their first language and simplify English diphthongs into single vowels. This theory helps explain the linguistic source of students' difficulties in pronouncing English diphthongs.

Based on these observations, the researcher intends to investigate students' difficulties more deeply by focusing on eight English diphthongs in American English, namely /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /oʊ/, /ɪr/, /ɛr/, and /ʊr/. This study aims to quantitatively measure how difficult each diphthong is for English Education students at Tadulako University by analyzing their pronunciation performance and their self-perceived difficulty. A pronunciation test will be used to evaluate actual production, while a questionnaire will collect students' opinions on how difficult each diphthong feels to them. By combining these two instruments, the researcher hopes to provide comprehensive, data-based insights that can help lecturers refine teaching methods, design more targeted pronunciation practice, and support students in developing clearer and more accurate diphthong pronunciation. Ultimately, this research seeks to contribute to improving pronunciation learning within the English Education Study Program and preparing students to become confident and competent English teachers.

METHOD

This study employed a quantitative descriptive research design. A quantitative approach was considered appropriate because the objective of this study was to measure students' pronunciation performance numerically and to examine their perceptions using structured questionnaire items. According to (Creswell, 2014), quantitative research enables researchers to collect numerical data and apply statistical procedures to describe patterns and trends within a population. This study focused on two main aspects: (1) identifying the most difficult English diphthong for second-semester students to pronounce, and (2) examining the factors that contributed to students' difficulties, including articulatory challenges, teaching materials, learning media, and the learning environment.

The descriptive nature of this design means that the study did not aim to manipulate variables or test cause-and-effect relationships. Instead, it sought to provide an accurate description of the existing condition. The data obtained from the pronunciation test and questionnaire were analyzed using frequencies, percentages, and mean scores to present a clear picture of students' pronunciation performance and perceived difficulties. As noted by (Derwing & Munro, 2015), descriptive approaches are commonly used in pronunciation research to identify patterns of pronunciation problems without implementing experimental treatment.

This design was appropriate for the purpose of the present study, as it did not aim to compare groups or evaluate the effectiveness of instructional methods, but rather to document and describe the pronunciation difficulties experienced by English Education students at Tadulako University.

Respondents

The research took place at the English Education Study Program, Teacher Training and Education Faculty, Tadulako University. This study focused on second-semester students because they were in the early stage of phonological development. At this level, students had begun learning basic pronunciation but were still developing their control over English vowel sounds, including diphthongs. Investigating their pronunciation at this stage was important to prevent long-term fossilization and to provide early instructional support. At the same time, preliminary observations and informal discussions with lecturers indicated that pronunciation difficulties, especially in diphthongs, still persisted among students at this level. The participants were selected using purposive sampling, a method that selects individuals who are most relevant to the research focus (McMillan & Schumacher, 2010). Approximately 15–17 students participated in the study. This sample size was suitable for

descriptive statistics because it was large enough to represent class-level tendencies while still manageable for detailed pronunciation scoring. Purposive sampling was justified in this research because the participants had prior exposure to English diphthongs, were in the stage where pronunciation development was still very active, and learned in relatively standardized learning environments and materials. The participants were informed about the purpose of the study, and their participation was voluntary.

Instruments

This study employed two research instruments to collect the data required to answer the research questions. The instruments were designed to obtain comprehensive information regarding students' pronunciation performance and their perceptions of the factors contributing to their difficulties. The first instrument was a pronunciation test, which was used to assess students' accuracy in producing the eight English diphthongs in American English. The second instrument was a questionnaire, which was used to explore students' perceptions of the factors influencing their pronunciation difficulties.

The pronunciation test was designed to assess second-semester students' ability to pronounce eight English diphthongs in American English, namely /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /oʊ/, /ɪr/, /ɛr/, and /ʊr/. The test aimed to measure students' pronunciation accuracy and identify which diphthong was the most difficult for them to pronounce. To obtain reliable data, the pronunciation test consisted of a sentence reading task containing 24 sentences. Each diphthong was represented by three different target words embedded in meaningful sentences. The use of sentence-level context was intended to provide a controlled yet natural environment for pronunciation production. During the test, students read the sentences aloud individually, and their pronunciation was recorded using a digital audio recorder to ensure accuracy and reliability. The recordings served as the primary data source for scoring and analysis. Students' pronunciation was evaluated using a scoring rubric that measured the accuracy of diphthong production. Each correctly pronounced diphthong received a score of 1, while incorrect pronunciation received a score of 0.

In addition to the pronunciation test, a questionnaire was used to collect data regarding students' perceptions of the factors contributing to their difficulties in pronouncing English diphthongs. While the pronunciation test measured students' actual performance, the questionnaire provided insight into how students perceived the sources of their pronunciation challenges. The questionnaire was designed in the form of close-ended statements using a four-point Likert scale, allowing students to indicate the degree to which they agreed or disagreed with each statement. The questionnaire consisted of 16 statements distributed across four perception categories, namely difficulty in articulating English diphthongs, difficulties related to teaching materials, difficulties related to learning media, and difficulties related to the learning environment. Before being administered, the questionnaire was reviewed to ensure clarity, relevance, and alignment with the research objectives. All participants were instructed to respond honestly based on their personal learning experiences, and the responses were later coded numerically and analyzed quantitatively to identify trends and patterns in students' perceptions.

Data Collection Techniques

This study used a systematic sequence of data collection to ensure accuracy and consistency. First, the researcher prepared the pronunciation test materials and questionnaire before administering the instruments to the participants. The participants were informed about the purpose and process of the study, and their participation was voluntary. After the preparation stage, the pronunciation test was administered individually in a quiet classroom to minimize distractions and ensure clear recordings. Students were asked to read the sentence items aloud, and all readings were recorded clearly using a digital audio recorder. To maintain natural pronunciation performance, students were not allowed to rehearse the words during the test. After completing the pronunciation test, students

completed the questionnaire either in printed or digital form. Participants were instructed to respond honestly based on their learning experiences and perceptions regarding pronunciation difficulties. The pronunciation recordings were then labeled and stored securely, while questionnaire responses were coded numerically and compiled for analysis. Finally, all data were checked for completeness, including missing responses or unclear recordings, to ensure consistency and accuracy of the collected data.

Pronunciation Test

The pronunciation test was designed to assess second-semester students' ability to pronounce eight English diphthongs in American English, namely /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /oʊ/, /ɪr/, /ɛr/, and /ʊr/. The test aimed to measure students' pronunciation accuracy and identify which diphthong was the most difficult for students to pronounce. Since diphthongs require smooth articulatory movement from one vowel position to another, evaluating students' production of these sounds was considered important for understanding pronunciation difficulties experienced by EFL learners.

To obtain reliable data, the pronunciation test consisted of a sentence reading task containing 24 sentence items. Each diphthong was represented by three different target words embedded in meaningful sentences. The use of sentence-level context was intended to create a controlled yet relatively natural pronunciation environment, allowing students to produce the target sounds in connected speech rather than isolated word pronunciation. This approach was expected to provide more authentic evidence of students' pronunciation performance.

During the pronunciation test, students read the sentence items aloud individually. Their pronunciation was recorded using a digital audio recorder to ensure clarity, accuracy, and reliability during scoring. The recordings served as the main source of data for identifying pronunciation accuracy and possible diphthong mispronunciations. Students' pronunciation performance was evaluated using a scoring rubric in which each correctly pronounced diphthong received a score of 1, while incorrect pronunciation received a score of 0. Incorrect pronunciation referred to diphthongs that were reduced into monophthongs, substituted with other vowel sounds, or pronounced unclearly.

Questionnaire

In addition to the pronunciation test, a questionnaire was administered to collect data regarding students' perceptions of the factors contributing to their pronunciation difficulties. While the pronunciation test measured students' actual performance in producing English diphthongs, the questionnaire aimed to provide insight into how students perceived the causes of their pronunciation problems. Therefore, the questionnaire played an important role in answering the second research question related to contributing factors.

The questionnaire was designed in the form of close-ended statements using a four-point Likert scale consisting of Strongly Disagree, Disagree, Agree, and Strongly Agree. The use of a Likert scale was considered appropriate for quantitative descriptive research because it enabled the researcher to quantify subjective perceptions and analyze them numerically using descriptive statistical procedures such as frequencies, percentages, and mean scores. The questionnaire consisted of sixteen statements distributed across four categories of perceived difficulty.

The first category focused on students' difficulties in articulating English diphthongs, including controlling tongue movement, jaw position, mouth shape, and vowel glides during pronunciation. The second category examined difficulties related to teaching materials, particularly students' perceptions regarding the adequacy of explanations, examples, and pronunciation practice activities. The third category investigated difficulties related to learning media, including students' access to pronunciation videos, audio recordings, online dictionaries, and digital learning platforms. The final category focused

on the learning environment, including opportunities to practice English, teacher feedback, peer interaction, classroom atmosphere, and exposure to English outside the classroom. Before being administered, the questionnaire was reviewed to ensure clarity, relevance, and alignment with the research objectives. All participants were instructed to respond honestly based on their personal learning experiences, and the responses were later coded numerically for quantitative analysis.

Data Analysis

The data obtained from the pronunciation test and questionnaire were analyzed using descriptive statistical methods. The analysis aimed to answer the two research questions: (1) which English diphthong was the most difficult for second-semester students to pronounce, and (2) what factors contributed to students' difficulties in pronouncing English diphthongs. The pronunciation test data were analyzed by classifying students' responses into two categories, namely correct pronunciation and incorrect pronunciation. A correctly pronounced diphthong received a score of 1, while an incorrect pronunciation received a score of 0. Since each diphthong was represented by three sentence items, the total frequency of correct and incorrect pronunciations for each diphthong was calculated to determine students' pronunciation performance.

The analysis included frequencies, percentages, and mean scores to identify pronunciation patterns and determine the most difficult diphthong. The percentage of correct pronunciation was calculated by dividing the number of correct pronunciations by the total responses and multiplying the result by one hundred. The diphthong with the lowest percentage of correct pronunciation was identified as the most difficult diphthong for second-semester students. In addition, questionnaire responses were coded numerically using a four-point Likert scale ranging from Strongly Disagree to Strongly Agree. Frequencies, percentages, and mean scores were calculated for each questionnaire item and factor category, including articulatory difficulty, teaching materials, learning media, and learning environment. Higher mean scores indicated stronger perceived difficulty or stronger influence of a factor on students' pronunciation problems. The findings from both instruments were then interpreted comprehensively to provide a deeper understanding of students' diphthong pronunciation difficulties and the factors contributing to those difficulties.

FINDINGS AND DISCUSSION

The findings presented in this chapter are derived from a pronunciation test administered to the participants. The test included eight English diphthongs, namely /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /oʊ/, /ɪr/, /er/, and /ʊr/. Each diphthong was represented by several words embedded in reading tasks to ensure that students pronounced them in a meaningful context rather than in isolation. Their spoken responses were carefully assessed and categorized as either correct or incorrect based on standard American English pronunciation. The resulting scores were then analyzed quantitatively to determine the overall level of accuracy for each diphthong.

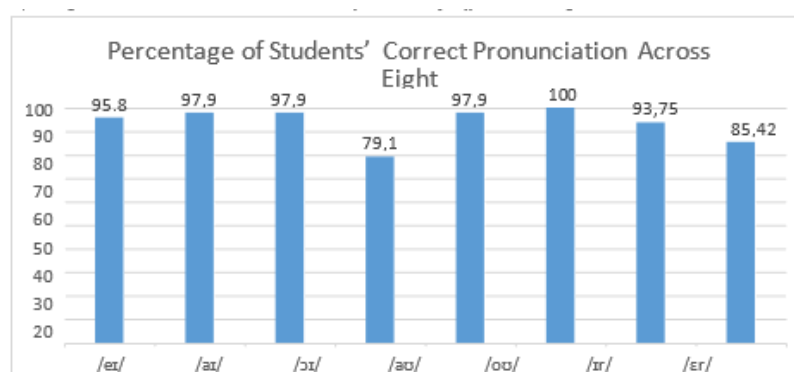


Figure 1. Percentage of Students' Correct Pronunciation Across Eight English Diphthongs

Figure 1 presents a visual representation of the percentage of students' correct pronunciation across the eight English diphthongs. The bar chart provides a clearer and more direct comparison of students' performance by illustrating the differences in accuracy levels for each diphthong. Through this visual representation, it can be observed that several diphthongs achieved higher percentages than others, while one diphthong recorded a noticeably lower percentage. This visual display supports the data presented in Table 1 and helps to highlight the variation in students' pronunciation performance across the tested sounds.

Table 1. Mean Scores of Valid Questionnaire Items

Item	Factor	Mean Score	Category
Q2	Articulatory Difficulty	2.95	High
Q5	Teaching Materials	2.68	Moderate
Q8	Teaching Materials	2.79	Moderate
Q12	Learning Media	2.90	High
Q15	Learning Environment	3.12	High

As shown in Table 1, the mean scores of the valid questionnaire items range from moderate to high levels, indicating that students generally agree that the identified factors contribute to their pronunciation difficulties. The variation in mean scores also shows that the degree of influence differs across factors, with some aspects being perceived as more challenging than others. These findings confirm that students' pronunciation problems are influenced by multiple interconnected factors rather than a single dominant issue.

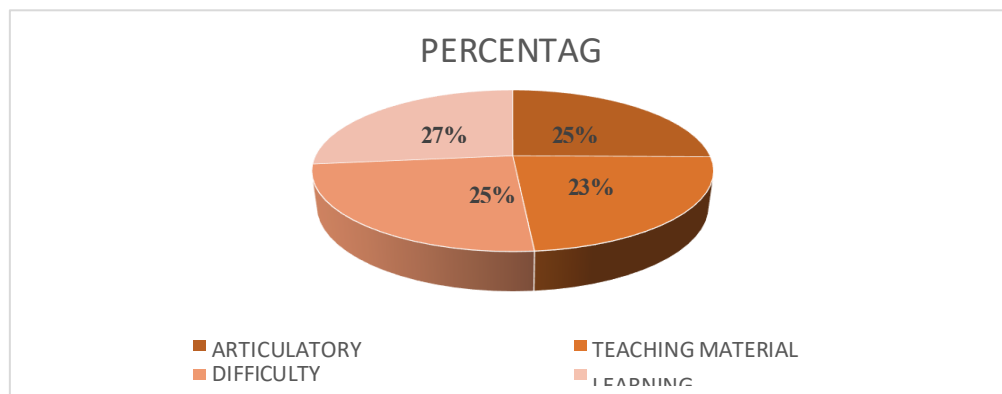


Figure 2. Distribution of Factors Affecting Students' Pronunciation Difficulties

The figure above illustrates the percentage distribution of the four main factors affecting students' difficulties in pronouncing diphthongs. Overall, the distribution shows that all factors contribute to pronunciation difficulties in a relatively balanced manner, although slight differences can still be observed in their respective levels of influence. This indicates that students' pronunciation problems are not caused by a single dominant factor, but rather by the combined effect of several interrelated aspects within the learning process.

Discussion

The discussion section presents a detailed interpretation of the findings obtained from both the pronunciation test and the questionnaire. It aims to explain the empirical results of the study in relation to the theoretical framework presented in Chapter II. Rather than merely describing the data, this section focuses on interpreting why the findings occur and how they can be understood through relevant linguistic and pedagogical theories. In this way, the discussion provides a deeper explanation of the factors influencing students' difficulties in pronouncing English diphthongs.

In general, the findings of this study indicate that students experience a range of

difficulties in pronouncing English diphthongs. These difficulties are reflected both in their actual pronunciation performance and in their perceptions of the learning process. The pronunciation test results demonstrate the students' real ability in producing diphthongs, while the questionnaire results reveal their subjective experiences regarding the factors that contribute to these difficulties. By combining these two sources of data, the study is able to provide a more comprehensive and balanced understanding of the issue, as it captures both performance-based and perception-based evidence.

Discussion of Pronunciation Test Results (Research Question 1)

The pronunciation test results provide a detailed overview of students' actual performance in pronouncing English diphthongs. The findings indicate that most students experience noticeable difficulties in producing diphthongs accurately and consistently. These difficulties are reflected in various types of pronunciation errors, such as incorrect vowel glides, substitution of diphthongs with monophthongs, and lack of smooth transition between two vowel sounds. This condition suggests that students have not yet fully mastered the articulatory requirements needed to produce English diphthongs in a natural and accurate way.

From a phonetic perspective, this finding can be explained by Roach, (2009) and Ladefoged & Johnson, (2015), who emphasize that diphthongs are complex vowel sounds that require a continuous glide from one vowel position to another within a single syllable. This process is not static but dynamic, involving precise and coordinated movement of the tongue, lips, and jaw. The results of the pronunciation test suggest that many students struggle with this coordination process, which leads to inaccurate or incomplete diphthong production. In many cases, students tend to hold one vowel position too long or fail to complete the glide properly, resulting in pronunciation that sounds closer to monophthongs rather than diphthongs.

Furthermore, these findings can be interpreted through Contrastive Analysis Theory proposed by Lado, (1957). This theory explains that second language learners often experience difficulties when there are significant differences between the sound systems of their first language and the target language. In the context of this study, Indonesian language does not contain diphthongs that function with the same level of phonetic complexity as English diphthongs. As a result, students tend to rely on their first language sound system when producing English words. This leads to sound transfer, where diphthongs are simplified into single vowel sounds or replaced with the closest equivalent sounds available in Indonesian phonology.

Discussion of Questionnaire Results (Research Question 2)

The results of the questionnaire provide a more detailed explanation regarding students' perceived difficulties in pronouncing English diphthongs. Unlike the pronunciation test, which focuses on students' actual performance, the questionnaire reflects students' subjective experiences and perceptions of the factors that influence their pronunciation learning. The findings show that students' pronunciation difficulties are influenced by four main factors, namely articulatory difficulty, teaching materials, learning media, and learning environment, with the learning environment identified as the most dominant factor.

The dominance of the learning environment (27%) indicates that external classroom conditions and social interaction play a crucial role in shaping students' pronunciation development. This finding can be explained through Krashen's Input Hypothesis, which emphasizes that language acquisition occurs effectively when learners are exposed to sufficient comprehensible input in meaningful contexts. In this study, the results suggest that students may not receive enough opportunities to actively use English in classroom interaction. The limited use of English reduces their exposure to correct pronunciation models and meaningful communicative practice.

Articulatory difficulty (25%) is also identified as a significant factor contributing to students' pronunciation problems. This finding is strongly supported by phonetic theory proposed by (Roach, 2009) and (Ladefoged & Johnson, 2015), who explain that diphthongs require a dynamic and smooth transition between two vowel positions within a single syllable. The questionnaire results indicate that students experience difficulties in controlling the movement of their tongue, lips, and jaw when producing diphthongs. In addition, they also struggle to differentiate diphthongs from similar monophthong sounds.

CONCLUSIONS

This study concluded that second-semester students of the English Education Study Program at Tadulako University experienced considerable difficulties in pronouncing English diphthongs accurately and consistently. Common errors included incorrect vowel glides, substitution of diphthongs with monophthongs, and incomplete transitions between vowel sounds. These findings indicate that many students have not yet mastered the articulatory processes required for accurate diphthong production due to the complexity of English diphthongs and the influence of their first language. Questionnaire results identified four contributing factors: articulatory difficulty, teaching materials, learning media, and learning environment. Among these, the learning environment emerged as the most influential factor, followed by articulatory difficulty and learning media, while teaching materials contributed the least. The findings suggest that pronunciation development is influenced by both linguistic and contextual factors. Therefore, improving students' mastery of English diphthongs requires systematic pronunciation instruction, increased exposure to accurate pronunciation models, effective learning resources, and supportive learning environments that encourage continuous speaking practice and pronunciation development.

ACKNOWLEDGEMENTS

The researcher would like to express sincere gratitude to Allah SWT for His blessings, strength, and guidance throughout the process of completing this article. Deep appreciation is also extended to the researcher's beloved family for their endless prayers, encouragement, and unwavering support during every stage of the writing process. The researcher would also like to thank the supervisor and co-supervisor for their valuable suggestions, guidance, and support that contributed greatly to the completion of this study. Sincere thanks are given to all respondents who willingly participated and contributed to the data collection process. Finally, heartfelt appreciation is addressed to friends for their encouragement, motivation, and meaningful support throughout the completion of this article.

REFERENCES

- Alimin, A. (2022). English diphthong pronunciation problems encountered by the fourth semester students of English Language Education Program at Qomaruddin University. *JET Adi Buana*, 7(1), 97-117. <https://doi.org/10.36456/JET.V7.N01.2022.5436>
- Aziez, F., & Sofiana, I. (2023). Deep in the diphthong problem: A study on Indonesian EFL learners' pronunciation development and the influence of individual differences. *Komposisi: Jurnal Pendidikan Bahasa, Sastra, dan Seni*, 24(2), 91. <https://doi.org/10.24036/KOMPOSISI.V24I2.122553>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems.
- Gilakjani, A. P., Ahmadi, S. M., & Ahmadi, M. (2011). Why is pronunciation so difficult to learn? *English Language Teaching*, 4(3). <https://doi.org/10.5539/ELT.V4N3P74>
- Inka, A., Lubis, P., Nadia, C., Rahmah, M., & Perangin-angin, A. B. (2024). Analysis of English diphthongs pronunciation errors of senior high school EFL students, 4, 6774–6781.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). Cengage Learning.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
- McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Pearson.
- Nafi'a, I., Salsabila, B., Munawaroh, B., El-Sulukiyyah, A. A., & Chuquin, E. G. P. (2025). The analysis of pronunciation challenges of EFL students in senior high school: Causes and solutions. *Indonesian Journal of Multidisciplinary Educational Research*, 3(1), 28–41. <https://doi.org/10.30762/IJOMER.V3I1.4514>
- Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge University Press.
- Suciati, T., Kurniawan, I., & Elfrida, E. (2024). Error analysis in pronouncing English words containing diphthong sounds made by EFL students. *Journal of English Education and Teaching*, 8(2), 414–428. <https://doi.org/10.33369/JEET.8.2.414-428>