

## STUDENTS' PERCEPTIONS OF THE USE OF SPEECH RECOGNITION TECHNOLOGY IN ENGLISH LEARNING

Ayunda Putri Amirah Siregar <sup>a\*)</sup>, Yani Lubis <sup>a)</sup>

<sup>ab)</sup> Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

<sup>\*)</sup>e-mail korespondensi: [ayunda0304212156@uinsu.ac.id](mailto:ayunda0304212156@uinsu.ac.id)

*Article history: received 01 June 2025; revised 12 June 2025; accepted 06 July 2025*

*DOI : <https://doi.org/10.33751/jmp.v13i2.12354>*

**Abstrak.** Kebutuhan global terhadap penguasaan bahasa Inggris terus meningkat secara signifikan, terutama dalam konteks komunikasi internasional, pendidikan, dan dunia kerja. Namun demikian, di Indonesia, banyak mahasiswa masih mengalami kesulitan dalam penguasaan pelafalan bahasa Inggris akibat terbatasnya interaksi dengan penutur asli serta minimnya akses terhadap media pembelajaran interaktif. Teknologi Pengenalan Ucapan (Speech Recognition Technology/SRT) hadir sebagai solusi inovatif yang memberikan umpan balik secara langsung serta memungkinkan pembelajaran mandiri yang lebih adaptif. Salah satu alat bantu SRT yang mulai dikenal adalah *Cambridge Learning Dictionary* (CLD). Penelitian ini bertujuan untuk menganalisis persepsi mahasiswa terhadap penggunaan CLD dalam pembelajaran bahasa Inggris di Universitas Islam Negeri Sumatera Utara. Penelitian ini menggunakan pendekatan kualitatif dengan metode fenomenologi, melibatkan sepuluh mahasiswi Program Studi Pendidikan Bahasa Inggris berusia 21–22 tahun. Teknik pengumpulan data dilakukan melalui observasi, wawancara semi-terstruktur, dan angket terbuka. Hasil penelitian menunjukkan bahwa mayoritas mahasiswa memiliki persepsi positif terhadap penggunaan CLD, baik dalam meningkatkan akurasi pelafalan, membangun kepercayaan diri, maupun mendorong kemandirian belajar. Selain itu, ditemukan tiga tema utama yaitu: (1) peningkatan kompetensi berbahasa melalui model pelafalan penutur asli dengan aksen British dan American; (2) kemandirian belajar yang tinggi melalui pemanfaatan CLD secara mandiri; dan (3) tantangan teknis yang mencakup keterbatasan konektivitas, ketiadaan umpan balik langsung, serta minimnya kosakata Islami dalam platform tersebut. Dengan demikian, CLD dinilai memiliki potensi yang besar dalam menunjang pembelajaran bahasa Inggris di lingkungan perguruan tinggi Islam, khususnya dalam hal pelafalan dan pengembangan kompetensi komunikasi. Agar pemanfaatannya optimal, diperlukan dukungan infrastruktur, pelatihan formal, serta integrasi konten yang kontekstual dengan budaya institusi pendidikan Islam.

**Kata Kunci:** Teknologi Pengenalan Ucapan, Kamus Cambridge, Pelafalan Bahasa Inggris, Perguruan Tinggi Islam

### STUDENTS' PERCEPTIONS OF THE USE OF SPEECH RECOGNITION TECHNOLOGY IN ENGLISH LEARNING

**Abstract.** The global demand for English language proficiency continues to grow significantly, particularly in the contexts of international communication, education, and the workforce. However, in Indonesia, many students still face difficulties in mastering English pronunciation due to limited interaction with native speakers and insufficient access to interactive learning tools. Speech Recognition Technology (SRT) has emerged as an innovative solution, offering real-time feedback and facilitating more adaptive, autonomous learning. One SRT tool gaining attention is the *Cambridge Learning Dictionary* (CLD). This study aims to analyze students' perceptions of using CLD in English language learning at Universitas Islam Negeri Sumatera Utara. Employing a qualitative phenomenological approach, the study involved ten female students aged 21–22 from the English Education Study Program. Data were collected through classroom observation, semi-structured interviews, and open-ended questionnaires. The findings indicate that most participants perceived the use of CLD positively, particularly in enhancing pronunciation accuracy, boosting speaking confidence, and fostering learner autonomy. Furthermore, three main themes emerged: (1) improved language competence through native-speaker pronunciation models in both British and American accents; (2) increased independent learning through self-directed use of CLD; and (3) technical challenges, including connectivity issues, the absence of real-time feedback, and the limited inclusion of Islamic vocabulary. Overall, CLD is considered to hold strong potential in supporting English learning in Islamic higher education settings, especially in terms of pronunciation and communicative skill development. To maximize its effectiveness, infrastructural support, formal training, and contextual content integration tailored to the cultural identity of Islamic educational institutions are recommended.

**Keywords:** Speech Recognition Technology, Cambridge Learning Dictionary, English Pronunciation, Islamic Higher Education

## I. INTRODUCTION

The global demand for English proficiency has increased significantly in recent years, driven by the language's role as a lingua franca in international communication, education, and business (Crystal, 2018). However, many non-native English speakers face challenges in mastering English pronunciation, which is often considered one of the most difficult aspects of language learning (Derwing & Munro, 2019). In response to these challenges, Speech Recognition Technology (SRT) has emerged as a promising tool, offering learners immediate feedback on their pronunciation accuracy (Liakin et al., 2018). Despite its potential, the adoption of SRT in language learning remains uneven across regions, with differing levels of acceptance and effectiveness (McCrocklin, 2019). This highlights the need for further research into how learners perceive and interact with such technologies in various educational contexts.

In the national context of Indonesia, English is a compulsory subject at both school and university levels, yet many students continue to struggle with pronunciation due to limited exposure to native speakers and insufficient practice opportunities (Suryani et al., 2020). While the national curriculum emphasizes communicative competence, traditional teaching methods often fail to address individual pronunciation difficulties (Wahyudi, 2021). The integration of technology, such as speech recognition applications, holds the potential to bridge this gap by providing personalized and interactive pronunciation practice. However, studies on the use of such tools in Indonesian higher education are still limited, especially within Islamic universities like Universitas Islam Negeri Sumatera Utara (UINSU), where students may encounter unique linguistic and cultural challenges (Nurhidayah, 2022). Moreover, in the context of Islamic higher education, the application of technology in English language teaching also presents its own challenges and opportunities. Students in Islamic universities have specific needs in mastering English particularly in speaking skills which requires adaptive and innovative learning approaches. The use of technology such as Speech Recognition Technology (SRT) is considered to enhance learning engagement and outcomes, especially when adapted to the cultural background and characteristics of the institution (Lubis, 2018). Similarly, in addressing individual learning challenges Adawiyah and Daulay emphasized that individual learning differences, particularly among slow learners, require adaptive teaching strategies that incorporate cognitive therapy approaches to enhance language competence and build learner confidence. (Adawiyah & Daulay 2022)

Previous studies have explored the role of SRT in language learning, emphasizing its potential to improve pronunciation and learner confidence. For example, (Liakin et al 2018) found that SRT-based applications enhanced pronunciation skills among French-speaking learners of English. (McCrocklin 2019) also noted positive perceptions of SRT among ESL learners in the U.S., citing its capacity for immediate and objective feedback. In Indonesia, (Suryani et al 2020) examined the use of SRT in secondary schools and found that students appreciated the interactive features of the technology. Previous studies have explored the role of SRT in language learning, emphasizing its potential to improve pronunciation and learner confidence. For example, (Liakin et al 2018) found that SRT-based applications enhanced pronunciation skills among French-speaking learners of English. (McCrocklin 2019) also noted positive perceptions of SRT among ESL learners in the U.S., citing its capacity for immediate and objective feedback. In Indonesia, (Suryani et al 2020) examined the use of SRT in secondary schools and found that students appreciated the interactive features of the technology. In addition, (Andayani and Fitriani 2021) conducted a study on students' perceptions of Google Speech Recognition in English language learning and found that most students considered the technology useful and easy to use for improving their pronunciation skills. This finding supports the growing recognition of speech-based technologies as effective tools in language education.

However, these studies largely focus on general SRT tools. There is a notable research gap regarding the use of specific speech recognition applications, such as the Cambridge Learning Dictionary, especially in the context of Islamic universities in Indonesia. This study aims to address this underexplored area by examining how learners respond to the integration of this specialized resource in pronunciation learning. Therefore, the main purpose of this study is to explore and analyze students' perceptions of the Cambridge Learning Dictionary as a speech recognition tool for improving English pronunciation, specifically within the Islamic university context of Universitas Islam Negeri Sumatera Utara. By focusing on this specific educational environment, the research seeks to provide insights into the potential role of SRT in supporting English language instruction at Islamic higher education institutions.

This study employs a phenomenological approach to explore students' lived experiences with the Cambridge Learning Dictionary. The approach focuses on understanding how students perceive and make meaning of their use of speech recognition tools, particularly in terms of improving pronunciation, building confidence, and overcoming technical or contextual challenges. The theoretical underpinnings of this study draw from Krashen's Input Hypothesis and Swain's Output Hypothesis, which emphasize the importance of comprehensible input and meaningful output in language acquisition (Krashen, 1982; Swain, 1985). Building on the theoretical foundations of linguistics, Daulay and Lubis (2018) in their comprehensive work on general linguistics highlight that the study of phonology and morphology provides essential foundations for understanding pronunciation accuracy and speech production mechanisms in second language acquisition. These theories support the use of SRT tools, which provide real-time feedback and opportunities for active language use. However, the effectiveness of such technologies greatly

depends on learners' engagement and perceptions (Suryani et al., 2020). Building on these insights, this study focuses on the Cambridge Learning Dictionary as a specialized tool that supports both pronunciation and lexical development.

Based on the rationale above, this study aims to investigate the implementation of speech recognition technology in English language learning by focusing on students at Universitas Islam Negeri Sumatera Utara. Specifically, the research explores students' perceptions of using the Cambridge Learning Dictionary as a speech recognition tool, analyzes how this technology influences their confidence in English learning, and identifies the challenges they encounter when utilizing such tools. By addressing these questions, the study seeks to provide insights into the integration of digital pronunciation aids in Islamic higher education and their potential to enhance language learning outcomes.

## II. RESEARCH METHODS

This study employed a qualitative research design with a phenomenological approach to explore students' lived experiences and perceptions in using the Cambridge Learning Dictionary as a speech recognition tool for English learning. Phenomenology, as introduced by Edmund Husserl in the early 20th century, emphasizes the exploration of individual consciousness and experiences, aiming to uncover the essence of a phenomenon as it is perceived by individuals themselves without interference from the researcher's assumptions (Husserl, 1931; Creswell, 2013). This philosophical approach, later expanded by Heidegger, Merleau-Ponty, and Sartre, has proven to be a valuable methodology in fields such as education and social sciences for its ability to provide deep, meaningful insights into human experiences (Moustakas, 1994). The qualitative nature of the study facilitated an in-depth exploration of students' perceptions, experiences, and the challenges they faced when using speech recognition technology in English learning.

The subjects of this study were 10 female students aged 21-22 years from the English Education Study Program at Universitas Islam Negeri Sumatera Utara (UINSU) who had prior experience using the *Cambridge Learning Dictionary* for pronunciation practice in English learning class. The selection of participants in this study was based on a purposive sampling approach as proposed by Patton (2015), which emphasizes the importance of selecting individuals who have direct involvement in the use of the Cambridge Learning Dictionary as a speech recognition tool. Other criteria include the participants' ability to articulate their experiences in depth and their willingness to participate actively and with full commitment throughout the research process. Their voluntary involvement also supports the authenticity and depth of the data obtained. Meanwhile, UINSU was chosen as the research location because this institution represents an Islamic higher education environment, which is likely to have its own dynamics in foreign language learning, including unique linguistic and cultural challenges.

Data were collected using three qualitative techniques: classroom observation, semi-structured interviews with open-ended questionnaires. The initial stage of data collection involved observing the participants during classroom sessions that incorporated speech recognition technology. This allowed the researchers to document students' natural interactions, behaviors, and engagement in real-time contexts, which helped in identifying spontaneous learning processes and contextual influences (Kawulich, 2005; Creswell, 2014). Following the observations, semi-structured interviews were conducted with selected participants. These interviews, each lasting approximately 30 minutes, enabled the researchers to explore in greater depth the students' experiences, perceptions, challenges, and suggestions regarding the use of speech recognition technology in their language learning activities (Kvale, 2007).

To supplement the insights gathered from observations and interviews, open-ended questionnaires were distributed to all participating students. These allowed students to articulate their thoughts freely in written form, offering additional perspectives and reflections that might not surface in spoken conversations. Despite being in questionnaire format, the data collected remained qualitative in nature, focusing on the participants' subjective interpretations and insights (Cohen, Manion, & Morrison, 2018). The integration of multiple data collection methods not only enriched the study but also enabled triangulation, enhancing the validity and credibility of the findings. Furthermore, previous research has underscored the value of observational data in educational settings. For example, McDonough and McDonough (1997) highlighted how observation can uncover behavioral patterns and barriers that may be overlooked during interviews, while Wang and Chen (2020) demonstrated the significance of real-time observation in understanding students' responses to feedback during speech recognition-assisted learning sessions.

## III. RESULT AND DISCUSSION

Beyond the direct pedagogical advantages, the integration of CLD also reflects a broader transformation in students' learning autonomy and digital engagement. The study uncovered that many participants voluntarily utilized the CLD outside of course requirements, particularly during exam preparation periods or when encountering unfamiliar academic texts. This independent use signifies an important shift from teacher-centered instruction to learner-centered practices, where students take active responsibility for their language development. Such behavior supports the growing emphasis on self-regulated learning in higher education, particularly within the context of foreign language acquisition, where personalized exposure and practice are critical to long-term success. The repeated and voluntary use of CLD as a pronunciation aid suggests that students not only view it as a

reliable resource, but also as a personal learning tool that enhances their ability to interact with English in authentic communicative settings, including debates, conferences, and interviews. This self-driven engagement contributes meaningfully to the development of lifelong learning habits that are essential in a rapidly globalizing academic and professional landscape.

Moreover, the findings illustrate that the CLD played a valuable role in bridging formal academic English with the distinctive linguistic needs of students at Islamic universities. While its core functions were praised, many students expressed a desire for a more culturally responsive version of the dictionary one that includes Islamic terms or theological vocabulary that is often used in academic discussions within their unique institutional environment. The absence of such vocabulary limited the tool's full potential as a contextually integrated learning aid. Nevertheless, the students' adaptive use of CLD to approximate pronunciation for Arabic-rooted English terms demonstrates both creativity and linguistic flexibility. This highlights the evolving nature of English language education in faith-based institutions, where English is not only a global lingua franca but also a medium for expressing religious and cultural identity. Thus, the application of tools like CLD must consider not only linguistic accuracy, but also the cultural and religious relevance of the learner population.

In addition, peer influence emerged as an important factor contributing to the popularity and sustained use of CLD. Several students reported discovering the tool through recommendations from classmates or upper-year peers who had previously used it for similar purposes. This informal diffusion of learning tools within student communities reinforces the idea that technology adoption in educational settings often occurs organically through interpersonal networks, rather than solely through formal curricular endorsement. The peer-driven spread of CLD usage created micro-learning communities within the student body, in which learners shared strategies, compared pronunciation attempts, and motivated each other to engage more deeply with pronunciation practice. This phenomenon underscores the social dimension of digital learning tools in higher education, suggesting that speech recognition technologies, when embraced by peer groups, can generate collaborative learning ecosystems that extend beyond the physical classroom and foster a culture of academic support and shared growth.

The findings of this study illuminate the complex dynamics surrounding students' engagement with the Cambridge Learning Dictionary (CLD) as a Speech Recognition Technology (SRT) tool within the unique context of Islamic higher education in Indonesia. The overwhelmingly positive perceptions reported by participants at Universitas Islam Negeri Sumatera Utara reflect a broader paradigm shift in language learning, where digital technologies are increasingly recognized as essential components of effective pronunciation instruction. These results fundamentally align with the theoretical foundations established by Krashen's Input Hypothesis and Swain's Output Hypothesis, which collectively emphasize the critical importance of comprehensible input and meaningful output in second language acquisition processes. The theoretical convergence becomes particularly evident when examining how the CLD functions as both an input provider and output facilitator, delivering native-speaker pronunciation models available in both British and American accents that represent the comprehensible input Krashen (1982) identified as fundamental to language acquisition, while simultaneously encouraging students to actively reproduce and practice these pronunciation patterns in support of Swain's (1985) assertion that meaningful output production is essential for developing communicative competence.

The unanimous agreement among participants regarding the CLD's effectiveness in improving pronunciation skills represents a significant finding that extends beyond mere user satisfaction and resonates strongly with previous research conducted by Liakin et al. (2018), who demonstrated that SRT-based applications enhanced pronunciation skills among French-speaking learners of English. This improvement is directly tied to the tool's capacity to reduce uncertainty and allow for self-paced practice, highlighting the psychological value of SRT tools in language learning. These findings corroborate Adawiyah and Daulay's (2021) research on communication competences, which demonstrated that technological interventions can significantly impact language learners' communicative abilities and reduce performance anxiety in academic contexts. Statements from participants demonstrated a transformative shift from hesitation to fluency, supporting the notion that technological tools can foster learner autonomy and reduce classroom-related performance anxiety. The consistency of positive outcomes across different linguistic backgrounds and educational contexts suggests that the fundamental mechanisms underlying SRT effectiveness transcend cultural and linguistic boundaries, while McCrocklin's (2019) findings regarding positive perceptions of SRT among ESL learners in the U.S., particularly citing its capacity for immediate and objective feedback, find parallel expression in the current study's results where students valued the CLD's clear pronunciation models and authentic accent variations. The confidence-building aspect revealed in this study deserves particular attention as it addresses one of the most persistent challenges in second language learning, with transformative narratives shared by participants describing shifts from hesitation to fluency and from anxiety to confidence that align closely with findings from Indonesian contexts. Suryani et al. (2020) previously examined the use of SRT in secondary schools and found that students appreciated the interactive features of the technology, while Andayani and Fitriani (2021) discovered that most students considered Google Speech Recognition technology useful and easy to use for improving their pronunciation skills, yet the current study extends these findings by demonstrating that confidence building occurs not merely as a byproduct of improved pronunciation accuracy, but as a direct result of reduced uncertainty and increased preparation opportunities that the CLD provides.

The emergence of learner autonomy as a significant theme in this study reflects broader trends in digital language learning, with many participants voluntarily engaging with the CLD outside of class requirements in a pattern that demonstrates what Crystal (2018) identified as the growing recognition of English as a global lingua franca, driving learners to seek additional resources for skill development. This autonomous engagement behavior supports Derwing and Munro's (2019) assertion that

pronunciation improvement requires sustained, individual practice opportunities that traditional classroom settings often cannot provide, while the peer influence networks discovered in this study, where students discover and adopt the tool through classmate recommendations, illustrate how digital literacy spreads organically within learning communities, creating supportive ecosystems for technology integration. When examining the comparative advantages of the CLD over traditional classroom methods, the findings reveal fundamental differences in feedback quality and accessibility, as conventional classroom instruction often provides limited individual correction due to time and group size constraints, while the CLD offers unlimited, private, and pressure-free practice opportunities that address what Wahyudi (2021) identified as the failure of traditional teaching methods to address individual pronunciation difficulties within Indonesia's communicative competence-focused curriculum.

However, the study also reveals important contextual considerations that reflect the unique challenges facing Islamic universities in Indonesia, with technical barriers identified by participants, including limited internet access and the absence of built-in speech recognition features, echoing broader infrastructure challenges that Nurhidayah (2022) highlighted in her examination of linguistic and cultural challenges within Islamic higher education contexts. The students' recommendation for including Islamic vocabulary and culturally relevant content reflects what Lubis (2018) described as the need for adaptive and innovative learning approaches that consider the cultural background and characteristics of Islamic institutions, while the absence of interactive feedback mechanisms, such as speech-to-audio comparison and visual articulation guides, represents a significant limitation that participants consistently identified. This gap highlights the distinction between passive pronunciation models and active feedback systems, suggesting that while the CLD excels in providing input, it falls short in offering the kind of immediate corrective feedback that characterizes more advanced SRT applications, a limitation that becomes particularly significant when considered alongside McCrocklin's (2019) emphasis on automatic speech recognition's potential for providing immediate and objective feedback, indicating that the CLD, while valuable, represents only one point on the spectrum of available SRT tools.

In the broader context of Indonesian English education, this study contributes to the growing body of evidence supporting technology integration while highlighting the importance of contextual adaptation, suggesting that while SRT tools like the CLD can effectively address pronunciation challenges that have historically limited Indonesian students' English proficiency, successful implementation requires attention to local infrastructure, cultural relevance, and institutional characteristics. The study ultimately demonstrates that the Cambridge Learning Dictionary serves as an effective bridge between traditional pronunciation instruction and more advanced SRT applications, offering an accessible, user-friendly introduction to technology-enhanced pronunciation learning that builds student confidence and competence despite not providing the full range of features that characterize cutting-edge speech recognition systems. The implications extend beyond individual tool evaluation to encompass broader questions about the future of pronunciation instruction in diverse educational contexts, suggesting that effective SRT integration requires careful consideration of user needs, institutional contexts, and cultural factors while maintaining focus on fundamental pedagogical principles, with the success of the CLD at UINSU demonstrating that thoughtful technology adoption can enhance learning outcomes even when tools are not perfectly aligned with local contexts, provided that institutions and learners remain engaged in ongoing adaptation and improvement processes.

#### IV. CONCLUSION

This study explored students' perceptions of using speech recognition technology in English learning at Universitas Islam Negeri Sumatera Utara, revealing predominantly positive student perceptions of speech recognition technology such as the Cambridge Learning Dictionary for English learning, characterized by enhanced pronunciation accuracy, increased speaking confidence, and voluntary autonomous engagement with the technology. The findings demonstrate that students perceive speech recognition technology such as the Cambridge Learning Dictionary as an effective tool for English learning that provides reliable native-speaker pronunciation models in both British and American accents, reduces speaking anxiety through private practice opportunities, and supports learner autonomy beyond classroom requirements, though technical limitations including internet connectivity issues, absence of direct speech recognition feedback, and insufficient cultural contextualization for Islamic university settings present ongoing challenges. The study concludes that speech recognition technology such as the Cambridge Learning Dictionary holds significant potential for English learning in Indonesian Islamic higher education contexts, provided that institutions address infrastructure limitations, incorporate culturally relevant content, and integrate formal technology training into their curricula to maximize the pedagogical benefits of these digital learning resources for English learning.

#### V. REFERENSI

- Adawiyah, R. & Daulay, S. H. (2021). The differences of bilingual and monolingual communication competences in adolescents. *English Franca: Academic Journal of English Language and Education*.
- Adawiyah, R. & Daulay, S. H. (2022). Using a cognitive therapy to enhance slow learners' competence: Teacher's strategy. *English Education: Jurnal Tadris Bahasa Inggris*.

- Andayani, A. & Fitriani, S. S. (2021). Students' perceptions on the use of Google speech recognition in English language learning. *Journal of Language and Language Teaching*, 9(2), 159–168.
- Cohen, L., Manion, L. & Morrison, K. (2018). *Research Methods in Education* (8th ed.). Routledge.
- Creswell, J. W. (2013). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (3rd ed.). SAGE Publications.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). SAGE Publications.
- Crystal, D. (2018). *English as a Global Language*. Cambridge University Press.
- Daulay, S. H. & Lubis, K. R. (2018). *Introduction to General Linguistics*. UIN Sumatera Utara Press.
- Denzin, N. K. & Lincoln, Y. S. (Eds.). (2011). *The SAGE Handbook of Qualitative Research* (4th ed.). SAGE Publications.
- Derwing, T. M. & Munro, M. J. (2019). *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. John Benjamins Publishing.
- Kawulich, B. B. (2005). Participant observation as a data collection method. *Forum: Qualitative Social Research*, 6(2).
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon Press.
- Kvale, S. (2007). *Doing Interviews*. SAGE Publications.
- Liakin, D., Cardoso, W. & Liakina, N. (2018). Learning L2 pronunciation with a mobile speech recognizer: A case study of a French-speaking learner of English. *ReCALL*, 30(2), 207–224.
- Lubis, Y. (2018). The effect of DRTA teaching strategy and learning style on students' achievement in reading comprehension at Faculty of Tarbiyah Science and Teacher Training State Islamic University of North Sumatra.
- McCrocklin, S. (2019). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, 82, 25–36.
- McDonough, J. & McDonough, S. (1997). *Research Methods for English Language Teachers*. Arnold.
- Nurhidayah, N. (2022). Challenges and strategies in teaching English pronunciation at Islamic universities in Indonesia. *Journal of English Language and Education*, 7(2), 89–101.
- Suryani, I., Nurhidayah, N. & Wahyudi, A. (2020). The use of speech recognition technology in Indonesian EFL classrooms: Challenges and opportunities. *Indonesian Journal of Applied Linguistics*, 10(1), 123–134.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 235–253). Newbury House.
- Wahyudi, A. (2021). Integrating technology in English language teaching: A case study of Indonesian schools. *Journal of Language Teaching and Research*, 12(3), 456–465.
- Wang, Y. & Chen, N. S. (2020). The effectiveness of automatic speech recognition technology for language learners: A meta-analysis. *Computer Assisted Language Learning*, 33(8), 849–879.