



The Use of the *Aussprachetraining* Web-Based Application on German Pronunciation: A Study of Student Performance and Perception

Talitha Ardelia Syifa Rabbani^{1*}

¹Universitas Indonesia, Indonesia, talitha.ardelia81@ui.ac.id

**Corresponding author:*

E-mail:

talitha.ardelia81@ui.ac.id

Abstract

This research examines the use of the *Aussprachetraining* web-based application developed by the Goethe-Institut in improving students' German pronunciation skills as well as explores their perceptions of this technology. Utilizing a mixed-methods sequential explanatory design, the study analyzed data from 39 first-year students of the German Studies Program through pre- and post-tests, activity logs, and questionnaires. The quantitative analysis demonstrated a 3.28% improvement in average pronunciation scores, which increased from 78.45 to 81.03, alongside high engagement levels where over 66% of participants achieved significant accuracy within the application-based exercises. Qualitatively, the students reported a positive experience, specifically highlighting the visually appealing interface (66.67%), its user-friendly navigation (71.79%), and the relevance of its content to course materials (69.23%). The most helpful feature identified was the immense value of native speaker pronunciation models. Despite these benefits, the study identified critical technical limitations, particularly regarding the accuracy of the speech recognition system and the quality of automated feedback, which the majority of respondents found suboptimal. This research concludes that, while the application proves to be a powerful tool for fostering self-directed learning and increasing learner confidence, technical refinements are necessary to maximize its effectiveness.

Keywords: *Computer-Assisted Pronunciation Training (CAPT), German as foreign language, pronunciation*

Introduction

In the contemporary globalized era, German continues to play a significant role in science, technology, engineering, and international commerce. Consequently, the demand for high levels of communicative competence among learners of German as a Foreign Language/*Deutsch als Fremdsprache* (DaF) continues to rise. A key component of this competence is pronunciation, which plays a crucial role in successful communication. Recent research emphasizes that intelligible pronunciation is essential not only for comprehension but also for social integration and interactional success (Levis, 2018; Saito, 2021). Without a sufficient level of phonetic accuracy, even learners with advanced grammatical knowledge may still experience frequent communication breakdowns.

For non-native speakers, German phonology poses particular challenges, including the mastery of umlauts (ä, ö, ü), the distinction between long and short vowels, and complex consonant clusters. Learners of DaF with Indonesian as their first language tend to

How to cite:

Rabbani, T., (2026). The Use of the *Aussprachetraining* Web-Based Application on German Pronunciation: A Study of Student Performance and Perception. *International Journal of Pedagogical Language, Literature, and Cultural Studies*. Nexus Publishing. ISSN: 3047-2202. Pages 17-26. doi: [10.63011/ip.v3i1.72](http://dx.doi.org/10.63011/ip.v3i1.72)

experience persistent difficulties in producing certain vowels (e, o, ö, ü), consonants (e.g., /f/, /l/), and consonant clusters (e.g., *sch*, *st*, *sp* in word-initial position), as well as the *ach-Laut* and *ich-Laut* (Schäfermeier, 2018). These challenges are often influenced by phonological transfer from the first language and limited exposure to authentic input.

Current studies in pronunciation pedagogy increasingly emphasize technology-assisted pronunciation learning, particularly through Computer-Assisted Pronunciation Training (CAPT) applications that support autonomous and flexible language learning (Levis, 2018; Saito & Plonsky, 2019). However, this goal is often insufficiently addressed in traditional classroom settings. Due to time constraints and curricular priorities, teachers frequently focus on grammar and vocabulary at the expense of systematic pronunciation training. Moreover, Dieling & Hirschfeld (2000) observed that large class sizes limit opportunities for individualized corrective feedback, which is crucial for preventing the fossilization of incorrect pronunciation patterns.

To address these limitations, Computer-Assisted Language Learning (CALL) has introduced innovative digital solutions that facilitate pronunciation development. Technological advancements in CALL have enabled the integration of multimedia elements such as audio, video, and interactive interfaces, which enhance learners' engagement and promote deeper cognitive processing of linguistic input (Bogach et al., 2021). Recent research also indicates that CALL environments help learners improve their pronunciation skills by providing opportunities for repeated exposure, individualized practice, and flexible learning contexts beyond traditional classroom settings (Iftikhar et al., 2025). Applications designed for German phonetics, such as *Aussprachetraining* (<https://aussprachetraining.goethe.de/>), utilize audio-visual modeling and interactive exercises to encourage autonomous and student-centered learning.

Despite the growing availability of digital tools, their pedagogical effectiveness is often assumed rather than empirically validated. Existing studies tend to focus either on measurable learning outcomes or on user perceptions but rarely integrate both dimensions. Research on technology acceptance in education highlights that perceived usefulness and perceived ease of use remain central predictors of technology adoption (Scherer et al., 2019). Therefore, combining objective performance data (e.g., improvements in phonetic accuracy) with subjective learner perceptions is essential for evaluating the sustainability and effectiveness of web-based pronunciation tools.

While the *Aussprachetraining* application offers a promising alternative to traditional instructional methods, there is still a lack of context-specific research examining its impact on DaF learners, particularly in the Indonesian higher education context. This study seeks to address this gap by investigating the extent to which the application improves learners' phonetic accuracy and how it is perceived in terms of usability and learning value. By integrating both cognitive (learning outcomes) and affective (learner perceptions) dimensions, this research aims to contribute to a more comprehensive understanding of how web-based technologies can be effectively integrated into contemporary German language pedagogy.

Computer-Assisted Pronunciation Training

Computer-Assisted Pronunciation Training (CAPT) has emerged as a specialized branch of CALL focusing on the development of learners' phonetic and phonological competence through technology-mediated instruction. CAPT systems typically provide learners with access to rich auditory input, opportunities for high-frequency and self-paced practice, and immediate, individualized feedback through technologies such as automatic speech recognition (ASR) and visual acoustic analysis (Amrate & Tsai, 2025; Rogerson-Revell, 2021).

These features allow learners to engage in repeated practice and receive corrective feedback that is often unavailable in traditional classroom settings.

Moreover, CAPT environments have been shown to facilitate both segmental (e.g., vowels and consonants) and suprasegmental (e.g., stress, rhythm, and intonation) aspects of pronunciation, although many existing systems still emphasize segmental accuracy (Amrate & Tsai, 2025). The integration of CAPT into language learning also promotes learner autonomy and reduces anxiety by creating a low-stakes environment for experimentation and self-correction. Empirical studies further suggest that CAPT can lead to measurable improvements in pronunciation accuracy and intelligibility, particularly when combined with pedagogically grounded instructional design and meaningful feedback mechanisms (Hsu, 2024; Rogerson-Revell, 2021). For example, studies comparing AI-based pronunciation tools with traditional instruction suggest that CAPT systems can produce comparable or even enhanced learning outcomes, particularly in contexts where access to native speakers is limited (Ondrušková, 2025).

Despite these advantages, several challenges remain. One of the most frequently reported issues is the accuracy of automated feedback generated by ASR systems. Research shows that CAPT feedback may be inconsistent or insufficiently detailed, particularly when dealing with non-native speech patterns (Amrate & Tsai, 2025).

Technology Acceptance Model

The Technology Acceptance Model (TAM) provides a widely recognized theoretical framework for explaining how users adopt and engage with digital technologies. The model is widely used in language education to explain why learners accept or reject digital tools. Across many contexts, it links beliefs about a tool to attitudes, intentions, and actual use. TAM focuses on perceived usefulness (PU), perceived ease of use (PEOU), attitude toward use, behavioral intention, and often actual use of the technology (Barz et al., 2024; Liu & Ma, 2024; Sprenger & Schwaninger, 2021).

In foreign language learning, TAM has been widely applied to examine learners' interaction with digital tools. Research shows that learners are more likely to engage with technology when they perceive it as beneficial for improving language skills and easy to use (Sulistiyo et al., 2022). In the context of pronunciation learning, perceived usefulness refers to learners' beliefs about the effectiveness of a tool in improving their pronunciation skills, while perceived ease of use relates to the usability and accessibility of the system. Research on online pronunciation instruction indicates that learners generally exhibit positive attitudes toward CAPT tools when they are user-friendly and aligned with learning objectives (Dalman, 2025).

Pronunciation Learning Outcomes

Pronunciation learning outcomes can be broadly categorized into cognitive and affective dimensions. From a cognitive perspective, pronunciation proficiency involves the accurate production of phonemes, appropriate use of stress and intonation, and overall speech intelligibility. Research emphasizes that intelligibility should be the primary goal of pronunciation instruction, as it directly affects communication effectiveness. Rather than striving for native-like accuracy, learners should aim to produce speech that is easily understood by listeners (Levis, 2018).

Technology-assisted pronunciation training has been shown to improve various aspects of pronunciation, including segmental accuracy and overall intelligibility. A review of technology-assisted pronunciation training studies reports consistent positive effects on

learners' pronunciation performance, highlighting the potential of digital tools to enhance language learning outcomes (Anolin, 2025).

In addition to cognitive gains, pronunciation learning also involves affective outcomes such as motivation, confidence, and learner satisfaction. CAPT environments can reduce anxiety by providing a safe and private space for practice, allowing learners to experiment with language without fear of judgment. For example, studies on ASR-supported pronunciation practice show that learners benefit from low-anxiety environments and individualized practice opportunities, which contribute to increased speaking confidence and engagement (Mukhoyyaroh & Tarwiyah, 2026). Furthermore, exposure to native-speaker models and interactive feedback has been shown to enhance learners' motivation and emotional engagement in pronunciation tasks, particularly when feedback is personalized and dialogic (Martin et al., 2022).

Nevertheless, technical limitations—particularly those related to ASR accuracy—can negatively impact both cognitive and affective outcomes. Research indicates that inaccurate recognition and feedback can lead to learner frustration and reduce trust in the system (McCrocklin & Levis, 2025). Similarly, limitations in CAPT systems, including false error detection, may result in demotivation and decreased learner confidence (Rogerson-Revell, 2021). These challenges highlight the need for continuous technological refinement to ensure that automated feedback is both reliable and pedagogically meaningful.

Method

This study employed a mixed-methods sequential explanatory design to investigate students' pronunciation performance and perceptions toward the *Aussprachetraining* web-based application. Quantitative data were collected through pronunciation pre- and post-tests, questionnaires, and activity logs involving 39 beginner-level students enrolled in a German language course. Qualitative data were obtained through open-ended questionnaire responses to provide deeper explanations of the quantitative findings. The pronunciation scores were assessed using predetermined pronunciation criteria. To improve the credibility of the assessment process, scoring consistency was carefully maintained throughout the evaluation. Questionnaire items were adapted from previous studies related to technology acceptance and pronunciation learning. The quantitative data were analyzed using descriptive statistics, while qualitative responses were analyzed thematically to identify recurring perceptions and learning experiences. The integration of both data types was conducted during interpretation to provide a more comprehensive understanding of students' learning experiences. More analysis, especially through collecting qualitative data, will make the broad view clearer, add to it, or explain it (Creswell, 2012).

Participants & Context

The participants in this study were 39 first-year students enrolled in the German Studies Program at a university in Indonesia. Thirty-three of them were female, and the remaining six were male. They were between 18 and 21 years old and shared the same first language, i.e., Indonesian. All participants were beginner-level learners (CEFR A1) who were taking both an introductory German language course and a German phonetics and phonology course. Prior to and during the intervention, participants received classroom-based phonetic instruction focusing on the articulatory features of German sounds, with guided pronunciation practice.

The study was conducted over ten weeks, during which students used the web-based application *Aussprachetraining*, developed by the Goethe-Institut. The application provides

structured pronunciation exercises, native-speaker audio models, and automatic speech recognition (ASR)-based feedback. There are 12 units (*Einheit*) and 36 lessons (*Lektion*), and in total of 1020 words to be exercised within the application. Participants were required to complete all assigned exercises and report their progress to the lecturer at the end of the intervention. The application served as a complement to, rather than a replacement for, classroom instruction. Tasks were completed flexibly, allowing students to access the materials anytime and anywhere within the study period.

Table 1. List of Units and Topics in the *Aussprachetraining* application

Units	Topic	Units	Topic
1	Word Stress (<i>Wortakzent</i>)	7	Consonants at the End of Words (<i>Konsonanten am Wortende</i>)
2	Long and Short Vowels (<i>Lange und kurze Vokale</i>)	8	Ich- and Ach-Sounds (<i>Ich- und Ach-Laut</i>)
3	E-Sounds (<i>E-Laute</i>)	9	Sch-Sounds (<i>Sch-Laut</i>)
4	Ö and Ü (<i>Ö und Ü</i>)	10	R-Sound (<i>R-Laut</i>)
5	Endings -e and -er (<i>Die Endungen -e und -er</i>)	11	Consonant Clusters (<i>Konsonantenhäufungen</i>)
6	Consonant Blends (<i>Konsonantenverbindungen</i>)	12	Assimilation (<i>Angleichung</i>)

Data Collection

Data were collected using three instruments: pronunciation tests (pre-test and post-test), application activity logs, and a questionnaire. To measure students' pronunciation improvement, a pre-test and a post-test were administered, consisting of word- and sentence-level tasks aligned with the course objectives. Student performance was evaluated using a rubric focusing on segmental accuracy (e.g., vowels and consonants) and suprasegmental features (e.g., stress and intonation).

Learning analytics data were obtained from the application, including frequency of use, task completion rates, and accuracy scores. These log data provide insights into learner engagement and behavioral patterns in digital learning environments.

A structured questionnaire was administered after the intervention to capture students' perceptions of the application. It included Likert-scale items and open-ended questions addressing usability, perceived usefulness, content relevance, and feedback quality.

Data Analysis

Quantitative data from the pre- and post-tests were analyzed using descriptive statistics to examine mean score differences and percentage improvement. Responses to Likert-scale items were analyzed to identify general trends in students' perceptions of the application, while activity log data provided insights into engagement levels and task performance.

Qualitative data from open-ended questionnaire responses were analyzed using thematic analysis, including coding, identification of recurring themes, and categorization into key aspects such as usability, learning experience, and technical challenges.

Finally, findings from both data strands were integrated during the interpretation phase, with qualitative results used to explain and contextualize the quantitative outcomes, providing a more comprehensive understanding of the effectiveness and user experience of the *Aussprachetraining* application.

Results

This section presents the study findings, organized into two subsections: quantitative and qualitative results. The quantitative results report pronunciation test performance, engagement and application usage, and students' perceptions of the application. The qualitative results provide deeper insights into students' experiences, highlighting both strengths and limitations of the *Aussprachetraining* application.

Pronunciation Test Performance

The comparison of pre- and post-test scores indicates a modest overall improvement in students' German pronunciation skills. The mean score increased from 78.45 in the pre-test to 81.03 in the post-test, reflecting a gain of 3.28%. This result suggests that engagement with the web-based *Aussprachetraining* application contributed positively to learners' pronunciation development.

Despite this overall improvement, individual performance varied considerably. Several participants demonstrated substantial gains, with increases exceeding 20%, while others showed minimal progress or slight declines in their scores. This variability indicates that the effectiveness of the application may be influenced by individual learner differences, such as prior experience, learning strategies, or level of engagement with the tasks.

Table 2. Comparison of Pre- and Post-test Scores (N=39)

Item	Pre-Test	Post-Test	Difference
Mean	78.45	81.03	2.57 (3.28 %)
Lowest Score	60.31	62.50	2.19 (3.63 %)
Highest Score	91.88	92.50	0.63 (0.68 %)

Engagement and Application Usage

Activity log data reveal consistently high levels of engagement throughout the intervention period. The majority of participants (87.18%) completed between 10 and 12 units, as well as 31 to 36 lessons within the application, indicating that most students fulfilled the expected workload.

In terms of practice frequency, over half of the participants (56.41%) engaged with the application one to two times per week, while 25.64% reported practicing three to four times weekly. A smaller proportion (10.26%) used the application more than four times per week, whereas only 7.69% reported less frequent use. These findings suggest that most learners maintained regular, though moderate, engagement with the application.

The duration of practice sessions further supports this pattern. The largest group of participants (43.59%) reported spending 20-30 minutes per session, followed by 25.64% who practiced for 10-20 minutes and 23.08% who exceeded 30 minutes per session. This indicates that learners generally engaged in sustained practice during each session.

Performance data derived from the application also demonstrate strong outcomes. A substantial proportion of participants (66.67%) correctly pronounced between 817 and 1020 words, reflecting a high level of task completion accuracy. This finding aligns with the observed improvement in test scores and suggests that repeated exposure and practice within the application supported pronunciation accuracy.

Students' Perception of the Application

Responses to Likert-scale questionnaire items indicate that students generally held positive perceptions of the application across multiple dimensions. In terms of interface and usability, the majority of participants agreed that the application was visually appealing (66.67%) and easy to navigate (71.79%). Furthermore, 64.10% of respondents reported that the application could be used independently without assistance, suggesting that the platform supports autonomous and self-directed learning.

Regarding content and instructional design, most students perceived the materials as appropriate for their proficiency level, with 66.67% indicating agreement. The exercises were also considered effective in helping learners distinguish German sounds, with a combined agreement rate of 92.31%. In addition, 69.23% of participants reported that the application content was relevant to their classroom learning. These findings suggest alignment between the application and course objectives.

Students also reported positive perceived learning outcomes. A majority agreed that the application contributed to improved pronunciation accuracy (66.67%) and increased awareness of pronunciation errors (58.97%). Confidence in pronunciation likewise improved, with 76.92% of participants indicating greater confidence when producing German words and sentences.

In terms of motivation and learner autonomy, 69.23% of respondents reported that the application motivated them to practice pronunciation independently. Notably, all participants expressed a sense of responsibility for their own learning progress (100% agreement), and a majority indicated willingness to use similar applications in the future.

However, perceptions of automated feedback were less favorable. More than half of the participants (53.85%) disagreed that the feedback provided by the application effectively helped them identify pronunciation errors. This suggests that, despite the application's strengths, limitations in feedback quality may reduce its overall effectiveness.

Perceived Benefits and Challenges

The most frequently identified beneficial feature was the availability of native speaker pronunciation models, reported by 41.03% of participants. Students emphasized that these models offered clear and reliable references, which facilitated accurate imitation and improved pronunciation awareness.

Other commonly mentioned benefits included opportunities for repeated practice (23.08%) and vocabulary support (15.38%). These aspects suggest that learners valued the structured and repetitive nature of the exercises, which enabled them to reinforce their pronunciation skills over time. Additional responses highlighted the clarity of instructions and the systematic organization of materials, although these were mentioned less frequently.

The most prominent challenge reported by participants was related to the speech recognition system. A large majority (79.49%) indicated difficulties with voice detection accuracy, suggesting that the system often failed to accurately capture or evaluate their

pronunciation. This issue directly corresponds with the negative perceptions of automated feedback observed in the quantitative data.

Other reported challenges included technical issues with the website (28.21%), limited accessibility of certain lessons (7.69%), and insufficient or unclear feedback (5.13%). These findings indicate that technical and system-related factors significantly affected the user experience.

Discussion

This study addressed two research questions: (1) to what extent the *Aussprachetraining* web-based application improves students' German pronunciation skills and (2) how students perceive its use. The findings indicate that the application contributes to measurable improvement in pronunciation performance and is generally positively perceived by learners, although limitations in automated feedback reduce its overall effectiveness.

Regarding the first research question, the increase in post-test scores suggests that the application supports pronunciation development through repeated exposure and structured practice. This finding aligns with previous CAPT research highlighting the importance of individualized and flexible pronunciation training (Amrate & Tsai, 2025; Rogerson-Revell, 2021). The presence of native-speaker pronunciation models also likely contributed to learners' phonological awareness and speech intelligibility, which are central goals of pronunciation instruction (Levis, 2018).

However, the improvement remained relatively modest, indicating that technology-assisted pronunciation learning alone may not guarantee substantial progress. This finding supports studies suggesting that CAPT applications are most effective when combined with teacher guidance and explicit corrective feedback (Anolin, 2025; Iftikhar et al., 2025). In this study, classroom instruction complemented the application by providing explanations and guided pronunciation practice that could not be fully delivered through automated systems.

In relation to the second research question, students generally perceived the application positively in terms of usability and learning value. From the perspective of the Technology Acceptance Model (TAM), these findings reflect strong perceived usefulness and perceived ease of use, both of which influence technology acceptance in educational settings (Scherer et al., 2019; Sulistiyo et al., 2022). Students particularly appreciated the accessibility of the platform and the opportunity to practice independently outside the classroom environment.

Nevertheless, dissatisfaction with the automated speech recognition (ASR) feedback system emerged as a major limitation. Several students reported that the feedback was unclear or insufficiently accurate, reducing its usefulness for pronunciation improvement. Similar limitations have been identified in previous studies on ASR-based CAPT systems, where inaccurate recognition negatively affects learner trust and learning effectiveness (Hsu, 2024; McCrocklin & Levis, 2025). These findings suggest that while CAPT applications can effectively facilitate autonomous practice, the pedagogical value of the technology largely depends on the quality and reliability of the feedback provided.

Overall, the findings demonstrate that the *Aussprachetraining* application has pedagogical potential as a supplementary pronunciation learning tool. Its integration into classroom instruction may help learners benefit from both technological flexibility and teacher-supported feedback, thereby creating a more balanced and effective pronunciation learning environment.

Conclusion

This study investigated the effectiveness of the *Aussprachetraining* web-based application in improving German pronunciation skills and explored students' perceptions of its use. The findings indicate that the application contributes to modest yet measurable improvements in pronunciation performance, primarily by increasing opportunities for structured, repetitive practice and enhancing learner engagement. Students generally perceived the application positively, particularly in terms of its usability, relevant content, and access to native-speaker models, which supported their learning experience and confidence. However, limitations related to the accuracy and clarity of automated speech recognition (ASR) feedback reduced its effectiveness.

These results suggest that while CAPT tools like *Aussprachetraining* can play a valuable role in fostering learner autonomy and supporting pronunciation development, their impact is maximized when used as complementary resources alongside teacher-led instruction. Future improvements in feedback quality and ASR technology are essential to enhance their pedagogical value and learning outcomes.

This study also contributes to the growing body of research on Computer-Assisted Pronunciation Training (CAPT) in German as a Foreign Language education, particularly within the Indonesian higher education context, where research on web-based pronunciation learning remains limited. However, this study is limited by its relatively small sample size and focus on a single institutional context. Therefore, future research is recommended to involve larger participant groups, apply inferential statistical analysis, and examine the long-term effectiveness of CAPT applications in pronunciation learning.

References

- Amrate, M., & Tsai, P. (2025). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 37(1), 22-42. <https://doi.org/10.1017/S0958344024000181>
- Anolin, C. (2025). Effectiveness of Technology Assisted Pronunciation Training Devices for ESOL Learners. *JELP: Journal of English Language and Pedagogy*, 4(2). <https://doi.org/10.58518/jelp.v4i2.3434>
- Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2024). Students' acceptance of e-learning: extending the technology acceptance model with self-regulated learning and affinity for technology. *Discover Education*, 3(1). <https://doi.org/10.1007/s44217-024-00195-7>
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th ed.)*. Pearson.
- Dalman, M. (2025). Second Language (L2) Learners' Perceptions of Online-Based Pronunciation Instruction. *Languages*, 10(4). <https://doi.org/10.3390/languages10040062>
- Dieling, H., & Hirschfeld, U. (2000). *Phonetik lehren und lernen: Fernstudieneinheit 21*. Langenscheidt.
- Hsu, H.W. (2024). An examination of Automatic Speech Recognition (ASR)-based Computer-assisted Pronunciation Training (CAPT) for less-proficient EFL students using the Technology Acceptance Model. *International Journal of Technology in Education (IJTE)*, 7(3), 456-473. <https://doi.org/10.46328/ijte.681>
- Iftikhar, S., Binti Romly, R., & Begum, S. (2025). Impact of Computer-Assisted Language Learning (CALL) on Pronunciation Proficiency Among ESL Learners. *Forum for Linguistic Studies*, 7(10), 1438-1453. <https://doi.org/10.30564/fls.v7i10.11349>
- Levis, J. M. (2018). *Intelligibility, Oral Communication, and the Teaching of Pronunciation*. Cambridge University Press. <https://doi.org/10.1017/9781108241564>

- Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2), 125-138. <https://doi.org/10.1080/17501229.2023.2240316>
- Martin, S., Alvarez, I. M., & Espasa, A. (2022). Video feedback and Foreign Language Anxiety in online pronunciation tasks. *International Journal of Educational Technology in Higher Education*, 19(1). <https://doi.org/10.1186/s41239-022-00324-y>
- McCrocklin, S., & Levis, J. (2025). Automatic speech recognition and pronunciation learning. *Language Teaching*. <https://doi.org/10.1017/S0261444825100852>
- Mukhoyyaroh, F., & Tarwiyah, S. (2026). ASR-Based Practice: Exploring EFL Learners' Experiences in Developing Pronunciation and Speaking Confidence. *Journal of Educational Sciences*, 10(2), 2755-2769. <https://doi.org/10.31258/jes.10.2.p.2755-2769>
- Ondrušková, D. (2025). Evaluating the comparative effectiveness of CAPT-based vs. native speaker-led pronunciation training. *Contemporary Educational Technology*, 17(4). <https://doi.org/10.30935/cedtech/17481>
- Rogerson-Revell, P. M. (2021). Computer-Assisted Pronunciation Training (CAPT): Current Issues and Future Directions. *RELC Journal*, 52(1), 189-205. <https://doi.org/10.1177/0033688220977406>
- Saito, K. (2021). What Characterizes Comprehensible and Native-like Pronunciation Among English-as-a-Second-Language Speakers? Meta-Analyses of Phonological, Rater, and Instructional Factors. *TESOL Quarterly*, 55(3), 866-900. <https://doi.org/10.1002/tesq.3027>
- Saito, K. and Plonsky, L. (2019), Effects of Second Language Pronunciation Teaching Revisited: A Proposed Measurement Framework and Meta-Analysis. *Language Learning*, 69, 652-708. <https://doi.org/10.1111/lang.12345S>
- Schäfermeier, I. (2018). *Aussprachetraining für Unterricht und Selbststudium (1. Auflage)*. 42 Signals GmbH.
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Sprenger, D. A., & Schwaninger, A. (2021). Technology acceptance of four digital learning technologies (classroom response system, classroom chat, e-lectures, and mobile virtual reality) after three months' usage. *International Journal of Educational Technology in Higher Education*, 18(1). <https://doi.org/10.1186/s41239-021-00243-4>
- Sulistiyo, U., al Arif, T. Z. Z., Sulistiyo, U., Handayani, R., Ubaidillah, M. F., & Wiryotinoyo, M. (2022). Determinants of Technology Acceptance Model (TAM) Towards ICT Use for English Language Learning. *Journal of Language and Education*, 8(2), 18-31. <https://doi.org/10.17323/jle.2022.12467>