

Comparative Study of AI-Based Speech Recognition Tools for L2 Pronunciation Improvement: A Proposed Research Framework

Ms. S. Janani Sri¹, Dr. V. Anuradha²

¹Ph.D Research Scholar, Department of English,, Vellalar College for Women (Autonomous), Thindal, Erode-12

²PG & Research Department of English, Vellalar College for Women (Autonomous), Thindal, Erode-12

Abstract

Speech recognition technologies enabled by Artificial Intelligence are modifying personalised learning of second language (L2) pronunciation through instantaneous feedback. Though there is an enormous usage, comparative evidence on instructional effectiveness found to be limited. This paper proposes a systematic research framework for assessing three popular AI-driven tools such as Google Speech-to-Text, ELSA Speak, and Duolingo. In terms of their effects on pronunciation accuracy and speaking fluency among L2 learners.

Using quasi-experimental design including pre-test and post-test measurements intra-group and inter-group differences can be evaluated. Statistical methods with paired sample t-test and one-way ANOVA can also be used for evaluation. The framework incorporates qualitative dimensions such as engagement of the learner, accessibility and the nature of corrective feedback, which provides thorough evaluation for performance outcome.

The desired outcome indicates that AI-assisted learning atmosphere can enable quantifiable gains in oral fluency and phonological accuracy. This study contributes to the advancement of Computer-Assisted Language Learning (CALL) and provides useful information for emerging teachers who are looking to incorporate AI technologies into pronunciation pedagogy by offering a reproducible comparative model. The approach also lays the groundwork for futures empirical validation across a wide range of learner groups.

Keywords: AI in Language Learning, Speech Recognition, L2 Pronunciation, Proposed Study, CALL, Fluency

1. Introduction

In second language acquisition, pronunciation plays a pivotal role in achieving communicative competence. Traditional way of teaching lacks in providing individualized feedback for the learners, which restrict students' ability to fix phonetic errors. AI-driven applications such as ELSA Speak and Duolingo provide real time learning experience and immediate feedback. On the other hand, systems like Google Speech-to-Text provide precise speech-to-text conversion for pronunciation analysis.

Though these tools are increasingly being adopted, there is still a requirement for a systematic comparative

assessment of them. This paper introduces a research framework to address this gap.

2. Literature Review

Research has shown that AI is highly effective in pronunciation training and Computer Assisted Language Learning (CALL).

- Neri et al. (2008) demonstrated that speech recognition improves pronunciation accuracy through feedback.
- Liakin et al. (2015) emphasized the benefits of mobile-assisted language learning tools.
- McCrocklin (2016) highlighted increased learner autonomy using AI-based tools.

Despite these findings, there is a notable gap in research regarding comparative frameworks that evaluate various AI tools.

3. Research Gap

Though earlier studies highlight the advantages of AI in second language acquisition, few investigations:

- Compare many AI-powered speech recognition programs at once.
- Evaluate accuracy and fluency of your pronunciation.
- Establish well-organised experimental frameworks for replication.

4. Research Objectives

- To formulate a comparative framework for assessing AI-based speech recognition tools.
- To examine their possible effects on the accuracy of L2 pronunciation.
- To determine how they contribute to increased speaking fluency.
- To find out the main elements affecting effectiveness.

5. Proposed Methodology

5.1 Research Design

Proposed a quasi-experimental design with pre-test and post-test evaluation.

5.2 Participants

60 undergraduate learners will be selected approximately and divided into three groups.

5.3 Tools for Analysis

- ❖ Google Speech-to-Text
- ❖ ELSA Speak
- ❖ Duolingo

5.4 Data Collection Plan

- For Preliminary evaluation Pre-test will be conducted.
- Structured intervention (4-6 weeks) using assigned tools.
- For improvement measurement Post-test evaluation will be conducted.

5.5 Proposed Data Analysis

- Mean and Standard deviation
- Paired t-test for pre-post comparison
- ANOVA for inter-group comparison

6. Expected Outcomes

The study is anticipated to demonstrate, based on the body of existing literature:

- AI technology has improved pronunciation significantly.
- Increased accuracy at the phoneme level using technologies that provide comprehensive feedback (like ELSA Speak).
- Enhanced fluency through continuous practice.
- Effectiveness varies according to tool design.

7. Conceptual Framework

The study is based on the interaction between:

- AI-based feedback mechanisms
- Learner engagement
- Pronunciation accuracy
- Speaking fluency

The above-mentioned variables altogether influence L2 pronunciation improvement.

8. Discussion

The suggested structure highlights how AI can be used to provide individualized pronunciation instruction. Google Speech-to-Text facilitates speech analysis whereas Duolingo provide gamified learning opportunities. Future empirical research and classroom integration techniques are intended to be guided by this study.

9. Implications

It is expected to have several implications for the findings of the proposed study.

- Teachers are inculcating AI tools into pronunciation teaching.
- It supports the curriculum designers in adopting technology-enhanced learning models.
- Enabling learners to engage in self-directed and personalized practice.

10. Limitations

- At present, the study is conceptual and lacks actual data.
- The suggested sample size might make generalizability more difficult.
- There are just three widely used applications for tool selection.

11. Future Research Directions

This research can be expanded further in future studies by:

- Performing extensive empirical validation.
- Incorporating diverse learner populations.
- Comparing other AI-based tools.
- Integrating sophisticated analytical tools such as machine learning techniques.

12. Conclusion

This research proposes a thorough approach for assessing AI-based speech recognition tools in L2 pronu-

unciation learning. It advances AI-assisted language instruction by filling up current research gaps and providing a methodical framework for next empirical investigations.

References

1. Fouz-González, Jonás. (2020). Using apps for pronunciation training: An empirical evaluation of mobile pronunciation tools. *Computer Assisted Language Learning*, 33(4), 1–27.
2. Gottardi, William, Almeida, J. F., & Tumolo, C. H. S. (2022). Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement: Reflections on their affordances. *Texto Livre*, 15, 1–20. <https://doi.org/10.35699/1983-3652.2022.36736>.
3. Cámara-Arenas, Enrique, Tejedor-García, C., Tomas-Vázquez, C. J., & Escudero-Mancebo, D. (2023). Automatic pronunciation assessment vs. automatic speech recognition: A study of conflicting conditions for L2-English. *Language Learning & Technology*, 27(1), 1–19. <https://doi.org/10.64152/10125/73512>.
4. Dizon, Gilbert. (2021). Affordances and constraints of intelligent personal assistants for second-language learning. *RELC Journal*, 54(3), 1–16. <https://doi.org/10.1177/00336882211020548>.
5. McCrocklin, Shannon. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, 57, 25–42.
6. Liakin, Denis, Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, 32(1), 1–25.