

**THE ROLE OF AUTOMATIC SPEECH RECOGNITION IN
BRIDGING THE PRONUNCIATION GAP FOR ENGLISH
LANGUAGE LEARNERS**

THESIS

Submitted in Partial Fulfillment of the Requirement for Gaining the
Bachelor Degree in English Education Department



By :

NABELA ADELIA KUSUMA NINGSIH

2103046046

**ENGLISH LANGUAGE EDUCATION DEPARTMENT
FACULTY OF EDUCATION AND TEACHER TRAINING
WALISONGO STATE ISLAMIC UNIVERSITY
SEMARANG**

2025

THESIS STATEMENT

THESIS STATEMENT

I am the student with the following identity:

Name : Nabela Adelia Kusuma Ningsih
Student Number : 2103046046
Department : English Language Education

Certify that the thesis entitled:

THE ROLE OF AUTOMATIC SPEECH RECOGNITION IN BRIDGING THE PRONUNCIATION GAP FOR ENGLISH LANGUAGE LEARNERS

is purely my own work. I am responsible for the content of this thesis. The opinions and findings of other authors in this thesis are quoted or cited based on ethical standards.

Semarang, 24 February 2025

The Researcher,



Nabela Adelia Kusuma Ningsih
NIM. 2103046046

THESIS APPROVAL



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI WALISONGO SEMARANG
FAKULTAS ILMU TARBIYAH DAN KEGURUAN
Jl. Prof. Dr. Hamka (Kampus II) Ngaliyan Telp. 024-7601295 Fax. 024-7615387 Semarang 50185

THESIS APPROVAL

To:

The Dean of Education and Teacher Training Faculty
Walisongo State Islamic University Semarang

Assalamu'alaikum, wr.wb.

After correcting it to whatever extent necessary, we statue that the final project belongs to the student as below:

Name of Student : Nabela Adelia Kusuma Ningsih
Student Number : 2103046046
Department : English Language Education
Title : **The Role Of Automatic Speech Recognition In Bridging The
Pronunciation Gap For English Language Learners**

I state that the thesis is ready to be submitted to the Education and Teacher Training Faculty of Walisongo State Islamic University to be examined at the Munaqosyah session.

Wassalamu'alaikum Wr. Wb.

Semarang, 24 February 2025

Advisor,

Dr. Muhammad Nafi Annury, M.Pd.
NIP. 197807192005011007

RATIFICATION



**KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI WALISONGO SEMARANG
FAKULTAS ILMU TARBIYAH DAN KEGURUAN**

Jl. Prof. Dr. Hamka (Kampus II) Ngaliyan Telp. 024-7601295 Fax. 024-7615387 Semarang 50185

RATIFICATION

Thesis with the following identity:

Title : The Role Of Automatic Speech Recognition In Bridging The
Pronunciation Gap For English Language Learners
Name of Student : Nabela Adelia Kusuma Ningsih
Student Number : 2103046046
Department : English Education

Had been ratified by the board of examiners of Education and Teacher Training Faculty of
Walisongo State Islamic University and can be received as one of any requirements for gaining
a Bachelor's Degree in English Education.

Semarang, 14 April 2025

THE BOARD OF EXAMINERS

Chairperson,

Awwalia Fitrotin Izza, M.Pd
NIP. 199303012020122005

Secretary,

Dr. Muhammad Nafi Annury, M.Pd
NIP.197807192005011007

Examiner I,

Daviq Rizal, M.Pd
NIP. 197710252007011015

Examiner II,

Agus Mutohar, M.A., Ph.D
NIP.198408012019031007



Dr. Muhammad Nafi Annury, M.Pd.
NIP. 197807192005011007

ADVISOR NOTE



**KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI WALISONGO SEMARANG
FAKULTAS ILMU TARBIYAH DAN KEGURUAN**

Jl. Prof. Dr. Hamka (Kampus II) Ngaliyan Telp. 024-7601295 Fax. 024-7615387 Semarang 50185

ADVISOR NOTE

To:

The Dean of Education and Teacher Training Faculty
Walisongo State Islamic University Semarang

Assalamu'alaikum, wr.wb.

I inform you that I have given guidance, briefing, and correction to whatever extent necessary for the following thesis:

Title : **The Role of Automatic Speech Recognition in Bridging the Pronunciation Gap for English Language Learners**

Name of Student : Nabela Adelia Kusuma Ningsih

Student Number : 2103046046

Department : English Language Education

I state that the thesis is ready to be submitted to the Education and Teacher Training Faculty of Walisongo State Islamic University to be examined at the Munaqosyah session.

Wassalamu'alaikum Wr. Wb.

Semarang, 24 February 2025

Advisor,

Dr. Muhammad Nafi Annur, M.Pd.
NIP. 197807192005011007

ABSTRACT

Title : The Role of Automatic Speech Recognition in
Bridging the Pronunciation Gap for English
Language Learners
Name : Nabela Adelia Kusuma Ningsih
Student Number : 2103046046

This study explores the role of Automatic Speech Recognition (ASR) technology in bridging the pronunciation gap for English language learners, particularly third-semester students of the English Education Department at UIN Walisongo Semarang. In addition, this study analyzes students' learning experiences in using the ASR technology in ELSA Speak application, including the benefits and challenges they encounter. Using a qualitative case study approach, data were collected through questionnaires and interviews with 9 participants who actively use the ELSA Speak application, and the data were analyzed using thematic analysis. The findings indicate that ASR technology in ELSA Speak provides immediate feedback, allowing learners to identify and correct pronunciation errors efficiently. Moreover, the application enhances learner autonomy and motivation by offering flexibility in practice time and location. The study also highlights that ASR technology improves pronunciation awareness, particularly in distinguishing commonly mispronounced phonemes. However, challenges such as reliance on internet connectivity, limited access to premium features, and occasional inaccuracies in voice recognition were identified. Despite these limitations, participants reported increased confidence and engagement in pronunciation learning. The results suggest that ASR technology in ELSA Speak serves as an effective supplementary tool for pronunciation learning. However, for optimal results, it should be integrated with traditional speaking practice and instructor feedback. These findings contribute to understanding the impact of ASR technology in language learning and provide insights for educators on leveraging digital tools to support pronunciation development.

Keywords: *Automatic Speech Recognition, Pronunciation, ELSA Speak, Learners Experience, English Language Learners*

MOTTO

يَرْفَعُ اللَّهُ الَّذِينَ ءَامَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ ۚ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ

“Allah will raise those who have believed among you and those who were given knowledge, by degrees. And Allah is fully aware of what you do.” (Quran 58:11)

حَسْبُنَا اللَّهُ وَنِعْمَ الْوَكِيلُ

“Allah Subhanahu wa ta’ala alone is sufficient for us, And He is the best disposer of affairs for us.” (Quran 3:173)

“Do not fear mistakes. You will know failure. Continue to reach out.”
– Benjamin Zander.

“Technology will not replace great teachers, but technology in the hands of great teachers can be transformational.”
– George Couros

DEDICATION

All praise is due to Allah, the Most Gracious and the Most Merciful. Without His guidance, mercy, and endless compassion, this journey would not have been possible. This thesis is not just an academic work but also a testament to the struggles, sacrifices, and love of those who have supported me through every challenge and achievement. Therefore, with profound gratitude, I dedicate this to:

1. My first hero, my beloved father, Sih Waluyo.

A man of few words but immeasurable sacrifices. Perhaps he does not often express love through words, but every drop of sweat, every tiring workday, and every silent prayer is proof of his undying love and devotion. My father taught me the meaning of perseverance, how to stand firm even when life's storms hit, and never to give up despite the weight of the world. May Allah always bless my father with health, long life, and boundless happiness.

2. My paradise on Earth, my beloved mother, Suliyah.

No words can adequately describe the depth of love and the magnitude of sacrifices my mother has given. She is a woman who never tires, never complains, and never stops giving. My mother's hands may have aged from hard work, but to me, those hands are the most beautiful—hands that have nurtured, strengthened, and guided me. Her prayers are the invisible shield that always protects me. I pray that Allah grants her long life, happiness, and endless blessings.

3. My little brother, Muhammad Zheini Al Kahfi.

You are not just my sibling, but also a small light in my life, the source of my happiness, and the greatest reminder that love is unconditional. Watching you grow into an amazing person gives me the strength to keep moving forward, work hard, and set a good example for you. I hope one day you will look at me with pride, just as I am proud to have you as my brother. Never stop chasing your dreams, because I will always be here to support you.

4. My beloved grandmother, Sukatmi.

The warmth of your affection, the wisdom in your words, and the gentleness of your prayers are a quiet strength that always guards me. You have witnessed many generations grow, and in every story you tell, there is a hidden lesson to learn. I am forever grateful for the love and prayers you have bestowed upon my life. May Allah grant you peace, health, and happiness without end.

5. My dearest friend in all things, Anis Mafazatin Ummah, Risqi Fitrianiingsih & Tri Aprilia Azmi

Thank you for everything. Thank you for accepting me as your friend, for making me feel welcomed and valued. I am truly grateful to have you by my side—from the days of being new students until now. Thank you for being my discussion partner, my safe space to share stories, and a friend who listens

without judgment. Thank you for always supporting me, for being someone who understands, and for walking alongside me with the same vision and mission. I cherish our journey together, and I hope life continues to bless you with happiness, success, and all the goodness you deserve.

6. My amazing friends and companions: Syella, Nabilla, Azka, Sindy, Arin, and other beloved friends.

Thank you to everyone who has walked with me, supported me, and strengthened me through every step of this journey. Without you, I wouldn't have made it this far. May this togetherness always remain, and may success always be with us all.

7. I would also like to extend my gratitude to all those who cannot be mentioned individually, for their invaluable help and support, both morally and materially, in ensuring the smooth progress and success of this thesis.
8. Lastly, I would like to thank myself, Nabela Adelia Kusuma Ningsih.

I want to thank myself for believing in myself when the world seemed uncertain. I want to thank myself for overcoming every challenge with resilience and courage. I want to thank myself for rising after every fall, for continuing to walk even when the path seemed dark, and for always striving to give, even in exhaustion. I also want to thank the pain that has

always accompanied me, the silent teacher that has never stopped shaping me, and the invisible strength that constantly reminds me to grow. Thank you for every lesson hidden behind wounds, for every fall that taught me how to stand stronger, and for the ways in which I have been forged into a more resilient soul. Without you, I would never have truly known how strong I am. I am proud of myself. I am grateful for this journey. And I promise to continue to fight, grow, and become the best version of myself.

May this journey serve as a foundation for greater things, and may every challenge I have faced become a valuable lesson that shapes me into a wiser, stronger, and better person.

ACKNOWLEDGEMENT

Bismillahirrahmanirrahim

Alhamdulillah, all praise is due to Allah SWT, the Most Gracious and the Most Merciful. With His endless blessings, guidance, and mercy, I am able to complete this final project successfully. This journey has been filled with challenges, struggles, and moments of doubt, but by His will, I was given the strength to persevere. Every difficulty, every sacrifice, and every step in this journey has been a valuable part of the learning process and self-growth.

May blessings and peace be upon the Prophet Muhammad SAW, who has enlightened our lives with knowledge, faith, and noble character. Through His teachings, I have learned about patience, sincerity, and the importance of seeking knowledge as a lifelong journey. The completion of this thesis is not only the result of my own efforts but also thanks to the prayers, guidance, and support from many individuals who have always been by my side. Therefore, with deep gratitude, I would like to express my sincere thanks to:

1. Prof. Dr. Fatah Syukur, M.Ag., Dean of the Faculty of Tarbiyah and Teacher Training at Walisongo State Islamic University, for his leadership and dedication to the academic world.
2. Dra. Nuna Mustikawati Dewi, M.Pd., Head of the English Education Study Program and my academic advisor, for her guidance and support from the beginning of my studies until

the completion of this thesis.

3. Lulut Widyaningrum, M.Pd., Secretary of the English Education Study Program, for her dedication and assistance throughout my academic journey.
4. Dr. Muhammad Nafi Annury, M.Pd., my thesis advisor, who with great patience provided direction, constructive feedback, and invaluable support in completing this research.
5. Dra. Hj. Ma'rifatul Fadhillah, M.Ed., and Awwalia Fitrotin Izza, M.Pd., my academic advisors, for their advice and guidance throughout my studies.
6. All lecturers and staff of the English Education Study Program at Walisongo State Islamic University Semarang, for their knowledge, guidance, and support during my education.
7. All the participants of the research from the third-semester students of the English Education Study Program at Walisongo State Islamic University Semarang, who graciously took the time to participate in this research. Your contributions are greatly appreciated.
8. My dear friends from "PBI A 2021," for the teamwork, friendship, and support throughout this academic journey. The memories we created together will always remain with me.

This final project is proof of the power of prayers, perseverance, and the support from those who have always believed in me. May Allah SWT continually bless everyone who has been a part of this journey,

and may this work contribute to the world of education and knowledge.

Semarang, 24 February 2025

The Researcher,

A handwritten signature in black ink, appearing to read 'Nabela', with a stylized flourish extending from the end.

Nabela Adelia Kusuma N.

TABLE OF CONTENTS

THESIS STATEMENT	ii
THESIS APPROVAL	iii
RATIFICATION.....	iv
ADVISOR NOTE	v
ABSTRACT	vi
MOTTO	vii
DEDICATION.....	viii
ACKNOWLEDGEMENT	xii
TABLE OF CONTENTS.....	xv
LIST OF TABLE.....	xii
LIST OF APPENDICES.....	xiv
CHAPTER I	1
INTRODUCTION.....	1
A. Background Study	1
B. Research Questions	6
C. Research Objective.....	6
D. Limitation of the Study	7
E. Significance of the Study	7
CHAPTER II	9
REVIEW OF RELATED LITERATURE	9
A. Previous Research	9
B. Literature Review	14
C. Conceptual Framework	44
CHAPTER III.....	48

RESEARCH METHOD	48
A. Research Design.....	48
B. Research Settings	49
C. Research Participant	52
D. Research Focus.....	53
E. Data Collecting Technique.....	53
F. Data Analysis Technique	56
CHAPTER IV.....	58
FINDING AND DISCUSSION	58
A. Finding	59
4.1 Learners experience with the use of ASR technology in the ELSA Speak Application	59
4.2 The Role of ASR Technology in Supporting English Pronunciation Improvement	73
B. Discussion	86
CHAPTER V	94
CONCLUSION AND SUGGESTION.....	94
A. Conclusion.....	94
B. Suggestions	95
REFERENCES.....	98
APPENDICES	1
CURRICULUM VITAE	25

LIST OF TABLE

Table 3. 1 Instrument Guideline.....	51
Table 4. 1 Questionnaire Result of motivation.....	59
Table 4. 2 Questionnaire Result of Benefit	65
Table 4. 3 Questionnaire Result of Pronunciation Errors.....	74
Table 4. 4 Questionnaire Result of Improve Pronunciation	79

LIST OF PICTURE

Picture 2. 1 ELSA Speak application	34
Picture 2. 2 Dictionary	35
Picture 2. 3 Progress.....	36
Picture 2. 4 Coach-Daily Training Lessons.....	37
Picture 2. 5 Speech Analyzer	37
Picture 2. 6 Study Sets	38
Picture 2. 7 Course Finder.....	39
Picture 2. 8 ELSA AI	40

LIST OF APPENDICES

Appendix 1 Learners Questionnaire.....	1
Appendix 2 Interview Questions.....	3
Appendix 3 Result of Questionnaire	4
Appendix 4 Result of Interviews.....	7
Appendix 5 Interview Documentation	22
Appendix 6 ASR Technology in ELSA Speak	23
Appendix 7 Advisor Appoinment Letter.....	24

CHAPTER I

INTRODUCTION

This chapter consists of the background of the study, research questions, objectives of the study, limitations of the study, and significance of the study. This research aims to explain the role of Automatic Speech Recognition technology in bridging the pronunciation gap of English learners, especially for 3rd-semester students of the English Department at Walisongo State Islamic University of Semarang.

A. Background Study

Language is a communication system in which humans convey thoughts, feelings, and information through symbols, both in written and spoken form. (Kustyarini, 2021) states that language functions as a means to express every aspect of a person, both in the form of ideas and knowledge. Therefore, language plays an important role in human life as it enables social interaction as well as the exchange of knowledge and understanding between individuals and groups.

In learning a language, especially a foreign language, speaking proficiency becomes a major challenge, with one of the main challenges being achieving accurate pronunciation. According to (Sudarmaji & Yusuf, 2021), the first thing native speakers notice in conversation is pronunciation. They can judge someone as not proficient in English simply based on their poor pronunciation.

Pronunciation plays an important role in communication as it significantly affects the intelligibility of speech. Clear and correct pronunciation helps learners ensure that the message conveyed can be correctly understood by the listener, thus preventing miscommunication or confusion that can hinder effective interaction. Therefore, to prevent misunderstandings, learners need to develop pronunciation skills that can support successful communication (Samad & Ismail, 2020) .

Therefore, it is important for language learners to pay attention to not only a general understanding of the language, but also specialised aspects such as accurate pronunciation. This shows that good pronunciation not only affects comprehension, but can also strengthen confidence when interacting with native speakers, especially in foreign language learning such as English. The ability to pronounce words and sentences correctly not only affects comprehension, but also increases confidence when speaking with native English speakers (Bodnar et al., 2017). Accurate pronunciation allows for better comprehension by English speakers and fosters greater confidence in communicating. This skill is crucial as it provides advantages in a variety of situations, both in everyday conversations and in professional contexts. Ultimately, honing effective speaking skills can open doors to diverse opportunities and enrich one's communication experience.

Despite its crucial role in communication, pronunciation often poses a significant challenge for non-native English learners and those teaching native speakers for several reasons. One of the main reasons is that pronunciation is frequently overlooked in language teaching. In this context, (G. Brown & Yule, 1984; Derwing & Rossiter, 2002) highlight that pronunciation is often neglected in language instruction, even though it is essential for learners to speak accurately and clearly.

According to studies by (Couper, 2017; Elliott, 1995; Gilakjani & Ahmadi, 2011), improving pronunciation is often considered an insignificant goal in the classroom because it is time-consuming. In the long run, neglecting pronunciation can lead to persistent difficulties in speaking, ultimately hindering learners' progress in achieving language fluency. Therefore, it is crucial for language educators to give adequate attention to pronunciation, ensuring that learners have the opportunity to develop their pronunciation skills.

In the context of language learning, the challenge of achieving accurate pronunciation has become a primary focus. However, in this era of ever-evolving technology, innovative solutions have emerged to address this obstacle. Technology, particularly Automatic Speech Recognition (ASR), has played a crucial role in improving pronunciation skills for language learners. With the ability to provide precise and customized feedback, ASR

technology assists learners in identifying and correcting errors in their pronunciation. According to (Levis & Suvorov, 2014), ASR is a technology that translates and transcribes user speech. Some applications of ASR technology provide various features such as providing a voice user interface (e.g., voice dialing), performing actions based on voice instructions (searching for information on the Internet) or voice text processing (dictating the text in a word). One of the most widely recognized ASR applications for English pronunciation improvement is ELSA Speak, which utilizes AI-driven feedback to analyze and correct learners' pronunciation.

English Learning Speech Assistant, or ELSA Speak can be a very helpful tool for learning English, especially for pronunciation, which is very important for developing speaking skills (Akhmad & Munawir, 2022). Therefore, the use of Automatic Speech Recognition technology, especially through applications such as ELSA Speak, is expected to be an effective solution in improving English speaking proficiency for learners.

Based on previous research, several studies have indicated that the use of technology in language learning, particularly focusing on pronunciation aspects, has a positive impact on learners' abilities. For instance, a study by (Samad & Ismail, 2020) titled "The Application of ELSA Speak as a Supporting Medium in Improving Students' Pronunciation Skills" demonstrated the effectiveness of the ELSA Speak application in enhancing pronunciation skills

among first-semester students at STKIP Muhammadiyah Enrekang. Their findings highlights the positive effects of using ELSA Speak as a supporting tool to improve pronunciation. Similarly, (Bashori et al., 2022) in their study "'Look, I can speak correctly': learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology" explored the use of automatic speech recognition (ASR) technology on websites such as I Love Indonesia (ILI) and NovoLearning (NOVO) to enhance vocabulary and pronunciation skills among high school students in Indonesia. Their study found that ASR technology on these websites successfully contributed to improving students' vocabulary and pronunciation skills.

Although previous studies have demonstrated the effectiveness of technology in pronunciation learning, research on the specific role of ASR in addressing pronunciation difficulties remains scarce. Most existing studies focus on technology-enhanced learning in general, without a detailed exploration of how ASR technology directly mitigates pronunciation challenges. Furthermore, while some studies have investigated ASR-based websites, there is a lack of research on mobile-based ASR applications, such as ELSA Speak, which offers real-time, AI-driven feedback for pronunciation improvement. Given this gap, the present study seeks to examine the role of ASR technology within the ELSA Speak application in bridging pronunciation gaps and

enhancing learners' pronunciation skills. To address this research gap, this study aims to investigate the role of ASR in bridging pronunciation gaps for English language learners, with a particular focus on the ASR feature within the ELSA Speak application. This research will analyze how ELSA Speak helps learners identify and correct pronunciation errors and will further explore learners' experiences in using ASR technology for pronunciation improvement. By examining the effectiveness of this mobile-based ASR tool, this study seeks to provide deeper insights into how ASR can support pronunciation learning in a more interactive and personalized manner.

B. Research Questions

Based on the research background, the researcher seeks answers to the following questions:

- A. How do learners' experience the use of Automatic Speech Recognition technology in the ELSA Speak application for learning English pronunciation?
- B. How is the role of Automatic Speech Recognition technology in the ELSA Speak application play in supporting the improvement of English pronunciation?

C. Research Objective

- 1. To describe the experiences of learners using the ASR technology in ELSA Speak for learning English pronunciation.

2. To investigate the role of ASR technology in providing accurate and specific feedback to support the improvement of English pronunciation among learners.

D. Limitation of the Study

This study is limited to nine third-semester English learners at UIN Walisongo, Semarang, who have used the ELSA Speak application as a tool to improve their pronunciation. The study explores the role of ASR technology in ELSA Speak in supporting pronunciation learning, as well as learners' experiences in using the application, focusing on their motivation, the benefits gained, and the challenges encountered during the learning process.

E. Significance of the Study

The significance of this study is that it addresses as follow :

1. Lecturers:

The researcher hopes that lecturers will guide learners more effectively in learning and improving English pronunciation, either by utilizing ASR technology in the ELSA Speak application or other similar tools.

2. Learners:

Learners benefit from ASR technology in ELSA Speak as they receive direct feedback on how to improve their pronunciation, thereby assisting them in learning more effectively.

3. The researcher:

This study contributes to understanding the role of technology in language learning, guiding further research, and potentially influencing the development of new educational technologies.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter encompasses an exploration of relevant literature types and previous research that supports the current study's data analysis and helps address the research questions.

A. Previous Research

This chapter presents previous research relevant to the study on the role of Automatic Speech Recognition in bridging the pronunciation gap for English language learners. This review provides a foundation for understanding how ASR-based applications, such as ELSA Speak contribute to pronunciation improvement and highlights the gaps that this study aims to address.

The first previous research was conducted by (Kholis, 2021), entitled “ELSA Speak App : Automatic Speech Recognition (ASR) for Supplementing English Pronunciation Skills.” This study aimed to examine how the ELSA Speak application helps students recognize, pronounce, and correct their English pronunciation using ASR technology. The study employed Classroom Action Research (CAR) with a mixed-methods approach, involving 18 second-semester students at Universitas Nahdlatul Ulama Yogyakarta. The findings showed that ELSA Speak significantly improved pronunciation accuracy, and the instant feedback feature played a crucial role in enhancing students' pronunciation skills. However, this study did not explore the long-term impact of ASR on

pronunciation retention and its integration into a comprehensive language curriculum.

The second previous research was conducted by (Samad & Ismail, 2020b), entitled “ELSA Speak Application as a supporting Media in Enhancing Students’ Pronunciation Skill.” This study aimed to to verify the effectiveness of using the ELSA Speak app in improving students' pronunciation skills of first-semester students at STKIP Muhammadiyah Enrekang. The study used a pre-experimental quantitative design with a purposive sampling method. The results showed a significant improvement in pronunciation skills, with the pre-test average score of 1.96 increasing to 5.79 in the post-test, confirmed by a t-test score of 6.28. However, this study lacked a control group for comparison, making it difficult to evaluate the effectiveness of ASR-based learning compared to traditional methods. Additionally, the study did not explore students’ experiences and perceptions of ASR technology.

The third previous research was conducted by (Dai & Wu, 2023), entitled “ Mobile-assisted Pronunciation Learning with Feedback from Peers and/or Automatic Speech Recognition: A Mixed-Methods Study.” This study focused on evaluating the effectiveness of peer feedback versus ASR feedback in pronunciation learning. Using Skill Acquisition Theory (DeKeyser, 2017) and Sociocultural Theory (Lantolf, 2012), the study applied a mixed-methods approach with 84 second-year university students in

China. The results showed that while ASR feedback was beneficial, peer feedback was more effective in improving intonation and fluency, whereas ASR was particularly useful for segmental pronunciation practice. However, the study only examined controlled pronunciation practice and did not assess its impact on spontaneous speech improvement.

The fourth previous research was conducted by (Bashori et al., 2022), entitled "Look, I can speak correctly": Learning Vocabulary and Pronunciation through Websites Equipped with Automatic Speech Recognition Technology". This study investigated the effects of two ASR-integrated applications (I Love Indonesia & NovoLearning) in pronunciation and vocabulary learning. The study applied Pronunciation and Comprehensibility Theory (Munro & Derwing, 2015) and Technology-Enhanced Language Learning (TELL) in a quasi-experimental mixed-methods study with 232 vocational high school students in Indonesia. The results indicated that ASR-based learning significantly improved pronunciation and vocabulary acquisition, with instant corrective feedback as a key advantage. However, the study had gender imbalance (222 males, 10 females) and lacked a longitudinal analysis to determine the long-term impact of ASR technology.

The fifth previous research was conducted by (Grasella Simanjuntak et al., 2023), entitled " Diagnosing Students' Errors in English Pronunciation by Using ELSA Speak Application." This

study aimed to analyze the pronunciation abilities of third-year English Department students at Universitas HKBP Nommensen Medan, focusing on identifying pronunciation errors using the ELSA Speak application. The study applied a qualitative descriptive method, involving 15 students, and employed pronunciation tests, documentation, and interviews as data collection techniques. The results showed that the average pronunciation score was 58.02, classified as "almost correct," with the most frequent errors occurring in vowel sounds (especially "i") and consonants (particularly "k"). However, the study was limited by its small sample size (15 students), lack of comparison with other pronunciation learning methods, and focus only on vowel and consonant errors without considering intonation and rhythm.

The sixth previous research was conducted by (Evers & Chen, 2022) entitled "Effects of an Automatic Speech Recognition System with Peer Feedback on Pronunciation Instruction for Adults." This study aimed to investigate the effectiveness of an Automatic Speech Recognition (ASR) system in improving pronunciation instruction for adult English learners, comparing peer feedback-based learning with individual practice. The study applied a quasi-experimental research design, involving 64 working adults in Taiwan, and employed pronunciation tests (reading aloud & spontaneous speech), satisfaction questionnaires, and ease-of-use assessments as data collection techniques. The results showed that

both groups experienced pronunciation improvements; however, the peer feedback group significantly outperformed the individual practice group in comprehensibility but not in accentedness. Despite the effectiveness of ASR-based learning, the study had limitations, including ASR's inability to provide explicit phonetic corrections, lack of improvement in accentedness, and limited generalizability to younger learners.

Previous research has similarities with this study in examining the role of ASR technology in improving English pronunciation. However, this study offers several new contributions that distinguish it from previous research. First, while earlier studies focused on ASR effectiveness in pronunciation training, this study specifically investigates learner experiences using ASR technology and their perceptions of its usability and effectiveness. Second, this research applies a qualitative case study approach, differing from prior experimental and quasi-experimental methods, which predominantly measured quantitative performance outcomes.

Additionally, this study not only assesses the effectiveness of ASR in pronunciation learning but also explores how ASR provides specific and accurate feedback, an aspect that has been less explored in previous research. While earlier studies have examined the impact of ASR applications like ELSA Speak and NovoLearning, they have not extensively analyzed how learners engage with the technology and perceive its usability in pronunciation practice. By addressing

these gaps, this study contributes a deeper understanding of ASR technology in language learning, particularly in bridging pronunciation gaps among English language learners. The findings will provide valuable insights for educators, developers, and learners in optimizing ASR applications for pronunciation training and improving language acquisition experiences.

B. Literature Review

1. Pronunciation Learning

According to (Richards & Schmidt, 2010), Pronunciation is the way a certain sound or sounds are produced. Unlike articulation, which refers to the actual production of speech sounds in the mouth, pronunciation stresses more the way sounds are perceived by the hearer. In the process of learning English, pronunciation is an essential component of language competence and effective communication. Pronouncing words clearly and accurately helps speakers understand each other better, thereby reducing the chances of misunderstanding. Pronunciation in English involves understanding phonetics and phonology, with a focus on how words are spoken and produced as part of a sound system. This skill plays a crucial role in ensuring that the message is clearly understood by the listener (Kholis, 2021).

Pronunciation refers to the way a word is spoken, the way a word is pronounced, and the way a person speaks the words

of a language (Hornby & Turnbull, 2010). As stated by (Offerman & Olson, 2016), clear and accurate pronunciation is essential for effective communication in a foreign language. While fluency or grammar may have certain limitations, mispronouncing words can increase the likelihood that the listener will have difficulty understanding the message being conveyed by the speaker (Crowther et al., 2015). Additionally, (Celce-Murcia et al., 1996) emphasize that pronunciation is related to the production of sounds that convey meaning. Therefore, it is essential to articulate words clearly to ensure that the listener can accurately understand the message and respond appropriately.

Based on the definitions and explanations provided, it can be concluded that pronunciation is a crucial aspect of language competence that involves the accurate production and perception of speech sounds. It requires mastery of phonetics and phonology to ensure words are articulated clearly and correctly, which is essential for effective communication and reducing misunderstandings. Clear and accurate pronunciation helps speakers understand each other better and is particularly important in foreign language learning, where mispronunciations can significantly hinder comprehension.

Pronunciation is often neglected in language teaching, even though it plays a crucial role in ensuring effective

communication. Pronunciation errors can lead to misunderstandings, even when learners have a good command of grammar and vocabulary. Unfortunately, many teachers lack the theoretical knowledge or practical skills to teach pronunciation, which is why this topic is often omitted from the curriculum. Additionally, pronunciation teaching is often reactive, where pronunciation issues that arise in class are immediately corrected without comprehensive planning. To address this, (Broughton et al., 2003) emphasizes that pronunciation teaching should be conducted in the context of authentic language use so that learners can effectively understand and speak. Good pronunciation supports effective speaking skills, as stated by (Kelly, 2000), that improper phoneme production can cause misunderstandings, even if learners have good grammar and vocabulary. Therefore, teachers play an important role in providing clear, accurate, and firm examples of pronunciation, while also using engaging teaching techniques to make the learning process more effective.

However, the challenges in teaching pronunciation cannot be overlooked. One of the main obstacles is learners' difficulty in distinguishing words that sound similar and their limited knowledge of correct pronunciation. For example, there are many words in English with similar spelling but different

pronunciation, such as “tree” and “three”, “though” and “through”, and “heal” and “heel”. To address this, Fraenkel (1984) suggests two approaches, Receptive Stage, where learners learn to distinguish sounds and language patterns through listening. And Productive Stage, where learners practice and produce the sounds they have learned (Al Husna et al., 2021).

Furthermore, (Kelly, 2000) underscores the importance of an active approach in teaching pronunciation. Teachers should not only correct learners' pronunciation errors but also actively teach the correct pronunciation techniques. (Harmer, 2001) adds that the main goal of pronunciation teaching is to help learners pronounce words accurately so that they are easily understood, rather than merely imitating a particular accent. Thus, pronunciation teaching not only improves learners' speaking ability but also creates communication that is more comfortable and effective for listeners.

Pronunciation refers to a person's ability to articulate English sounds accurately, encompassing elements such as stress, intonation, and other related speech (Annury, 2013). It also describes the manner in which a word or a language is articulated. Previously, it might have been stated that the objective should consistently be native-like pronunciation, although it was recognized that only a few would achieve it.

However, nowadays, most individuals believe that this is an unsuitable objective for the majority of learners. The vast majority of English language learners have practical aims and do not gain any particular advantage from attaining native-like pronunciation. However, there are certain learners who might aim to achieve a native-like accent due to their profession demanding them to interact on an equal footing with native speakers either in an English-speaking nation or overseas. So, the goal of learning pronunciation is to enable clear communication in English, minimizing misunderstandings, and serving as a measure of proficiency, even for non-native speakers (Zakiyyah et al., 2022).

In numerous nations, English serves a specific function as the medium of communication among individuals who speak diverse native languages. These second-language English speakers may have a limited audience, as they primarily interact with other non-native speakers and, as a result, native-like pronunciation is entirely unsuitable. Nonetheless, it should be acknowledged that when communicating with native speakers, differences in pronunciation could potentially cause communication breakdowns. Although achieving native-like pronunciation may be a target for specific learners, and while we should not discourage learners from aiming for ambitious

goals, for most learners, a much more realistic objective is to be easily understandable (Kenworthy, 1987).

According to (Kenworthy, 1987), the native language plays a crucial role in learning English pronunciation, as evidenced by the fact that a foreign accent often retains certain sound features from the learner's first language. These features can be noticeable enough for both casual listeners and trained experts to identify a person's origin. Even a few distinctive traits in pronunciation can reveal the influence of the learner's native language on their spoken English. Several factors influence pronunciation learning. Native language significantly affects English pronunciation. Differences in the sound system, rhythm, and intonation between the native language and English determine the difficulty level in learning pronunciation. However, despite these challenges, many people from various linguistic backgrounds can achieve pronunciation close to native speakers.

Age is also considered an influential factor in acquiring native-like pronunciation. A study by Oyama showed that the younger someone is when they start learning a language, the better their pronunciation. Other studies found that native-like pronunciation is rare after the age of 15. However, research by Snow and Hoefnagel-Höhle found that in the short term, older learners showed an advantage in imitating foreign sounds. Over

time, children showed more progress, while adults' progress tended to plateau. The differences in study results could be due to different assessment methods and other factors such as individual ability, motivation, and opportunities for language practice. Some studies support the theory of a sensitive period in language learning, but there is no conclusive evidence linking age to the ability to pronounce a new language.

Another factor is the amount of exposure to English. While greater exposure can enhance pronunciation, it is not the sole determining factor. Living in an English-speaking country does not always guarantee improvement, especially if one primarily communicates in another language. Conversely, individuals who live in non-English-speaking countries but actively use English in daily life may also have good pronunciation. What matters more is how individuals take advantage of opportunities to hear and use English.

Additionally, phonetic ability plays a role in pronunciation learning. Some people have a natural ability to mimic and distinguish foreign sounds, often referred to as phonetic coding ability or auditory discrimination ability. While everyone has this basic ability, practice still plays a crucial role. Studies show that learners with high phonetic ability benefit from repetitive pronunciation exercises, while those who struggle to distinguish sounds do not gain much from this method.

Therefore, teaching should incorporate various approaches to meet the needs of each learner.

Attitude and identity also influence pronunciation acquisition. Individuals react differently to accents-some adapt easily, while others maintain their original accents. People tend to adjust their speech to show solidarity in conversations. Factors such as age, length of stay, and the intention to settle or return to one's birthplace influence accent change. In foreign language learning, a positive attitude towards native speakers and their culture can lead to pronunciation closer to a native accent. Integrative motivation, which is the desire to integrate into the target language community, plays an essential role in developing a more accurate accent.

Lastly, motivation and concern for pronunciation vary among learners. Some learners are more concerned about their pronunciation than others. This concern is often expressed through requests for correction, worry about making mistakes, or even reluctance to speak if they cannot say something perfectly. Concern for pronunciation reflects achievement motivation. If a person doesn't see the value in something, they are less likely to be motivated to do well. Some learners may also be unaware that their way of speaking is causing difficulty or misunderstanding for the listener.

Given its comprehensive nature, pronunciation is a crucial skill in teaching the English language. It is not only a component of the English curriculum but also affects learners' speaking performance. Additionally, Harmer (2007) noted that teaching pronunciation not only helps learners become aware of various sounds and sound features but also greatly improves their speaking skills. Nonetheless, it is a difficult task because learners face many challenges when learning pronunciation (Rismawati et al., n.d.).

Many learners may encounter difficulties with pronunciation, particularly when it comes to pronouncing diphthong sounds. Despite receiving instruction in class, they often struggle to master proper pronunciation. There are several reasons why learners commonly make pronunciation errors. According to (Yiing, 2011), the primary challenge stems from the influence of the learners' mother tongue. When attempting to pronounce a foreign language, learners often encounter difficulties as their native language plays a significant role in the learning process. Consequently, learners may make mistakes and experience confusion, particularly when pronouncing English words, including diphthongs.

Sembiring and Ginting (2016) state that challenges in English pronunciation are also influenced by one's environment. Learners who are exposed to an English speaking

environment will have their pronunciation skills affected. Continuous exposure to the English environment can lead to the development of habitual pronunciation patterns (Saadah & Ardi, 2020). Similarly, (Suryatiningsih, 2015), assert that another reason for pronunciation difficulties is the unfamiliarity of words. Learners, not only in academic contexts but also in daily life, tend to use familiar and common words. Therefore, when encountering unfamiliar words, their pronunciation may be affected.

(Yousif & Ameen, 2018) also identify a cause of pronunciation issues among learners, which is the inefficiency of the phonetic teaching process. In the realm of education, the teaching process holds significance as it serves as the primary medium for learners to learn and effectively practice English pronunciation.

2. The role of Automatic Speech Recognition in pronunciation learning

Automatic speech recognition (ASR) has been an active area of research for more than five decades. It has consistently been regarded as a crucial link in enhancing both human-to-human and human-to-machine communication. Historically, however, speech did not emerge as a prominent mode of communication between humans and machines. This was partly due to the inadequacy of technology at that time to meet

the usability standards for most real-world users and conditions. Additionally, alternative communication methods such as keyboards and mice often surpassed speech in terms of communication efficiency, limitations, and accuracy. In recent years, speech technology has begun to revolutionize our lifestyles and work environments, emerging as one of the primary means for individuals to interact with various devices (Yu & Deng, 2015).

Automatic speech recognition (ASR) is a technology which converts voice into text transcriptions and is one of the core techniques in man-to-machine communications. In recent years, several applications have extensively used ASR-related speech technologies for information access and speech-to-speech translation services (Kidawara et al., 2020). Meanwhile (Levis & Suvorov, 2014) explain that Automatic Speech Recognition (ASR) is a technological tool designed to interpret and transcribe spoken language. ASR technology has uses in several domains, including facilitating voice-based user interfaces, executing commands via voice instructions, and converting speech into text, such as for dictation purposes in word processing applications.

Automatic speech recognition (ASR) is the procedure wherein a spoken language signal is transformed into a series of words utilizing an algorithm executed as software on a

computer (Keshet & Bengio, 2009). Furthermore, the application of Automatic Speech Recognition (ASR) technology has shown significant promise in addressing these pronunciation challenges among language learners. Let's delve deeper into how ASR systems provide valuable feedback and facilitate interactive learning experiences.

The use of Automatic Speech Recognition (ASR) technology has proven to be very beneficial in improving the pronunciation skills of language learners. ASR systems serve as a useful tool to transcribe spoken language into text and provide instant pronunciation feedback. This feature makes learning dynamic and interactive by allowing learners to compare spoken words and phrases with models from native speakers. The instantaneous feedback from ASR technology helps learners recognise and correct pronunciation errors quickly, leading to more efficient skill development (Kidawara et al., 2020).

ASR technology's role in pronunciation learning aligns closely with Skill Acquisition Theory (DeKeyser, 2020), which posits that language learning occurs in stages, each requiring different levels of cognitive effort. According to this theory, learners progress through three stages in acquiring language skills. First, in the declarative knowledge stage, learners acquire an understanding of pronunciation rules. Next, in the

procedural knowledge stage, learners practice with conscious effort, refining their pronunciation. Finally, in the automatic knowledge stage, fluent and accurate pronunciation is achieved without conscious effort.

DeKeyser's theory evolves from earlier cognitive learning theories, particularly the distinction between declarative and procedural knowledge. It draws inspiration from Anderson's ACT-R Theory (1982), which emphasizes skill learning in cognitive psychology. DeKeyser then adapted this theory to the realm of second language learning, emphasizing that language can be learned like other skills through repetitive practice. Further research shows that a combination of practice and feedback is crucial in the transition from declarative knowledge to automaticity, particularly in speaking a foreign language. This theory highlights the shift from learning that is initially conscious and effortful to performance that becomes more automatic and unconscious over time.

The connection between Skill Acquisition Theory and ASR technology in pronunciation learning lies in the stages outlined by DeKeyser. ASR technology aids in each stage by providing feedback that supports learners' progress. Initially, learners receive explicit feedback on pronunciation rules (declarative knowledge), then engage in practice with feedback for improvement (procedural knowledge). Eventually, learners

refine their pronunciation skills to the point where they produce them automatically (automatic knowledge). Thus, the ELSA Speak application, which utilizes ASR technology, can accelerate this process by providing real-time, iterative feedback.

ASR technology, powered by advanced algorithms, evaluates pronunciation by considering phonetic features, intonation patterns, and accent variations. This ensures a comprehensive assessment of learners' speech. ASR supports learners through targeted feedback, continuous practice, and customization based on individual needs, facilitating a smoother transition from procedural to automatic knowledge. As explained by (Huang & Deng, 2010) automatic speech recognition (ASR) is the process and technology used to translate speech signals into sequences of words or corresponding linguistic entities, using computer software algorithms. By teaching algorithms to identify word variants, ASR attempts to address variability in word pronunciation.

ASR is widely used in open and commercial systems, and studies have shown that ASR improves learners' mastery of speech acts, the accuracy of spoken grammatical structures, and pronunciation quality (Chiu et al., 2007; de Vries et al., 2015; Neri et al., 2008; van Doremalen et al., 2016). ASR technology, with applications that range from voice user interfaces to

speech-to-text processing, is reported to be fun, useful and less complex for learners, offering an effective tool for pronunciation practice and error detection (Levis & Suvorov, 2014; Liakin et al., 2017; McCrocklin, 2019; Mroz, 2018). According to (Wallace, 2016), ASR dictation software serves as an excellent resource for practicing pronunciation while also aiding in identifying common errors.

In summary, Automatic Speech Recognition (ASR) technology has evolved significantly over the past five decades, overcoming early challenges and becoming a pivotal tool in both human-machine interactions and language learning. Its ability to provide instant feedback and personalized learning experiences makes it an invaluable resource for improving pronunciation and facilitating more effective communication. As ASR technology continues to advance, its applications will undoubtedly expand, further enhancing its utility and impact across various domains.

3. Learner's Experience in using ASR Technology

Learner experience is an important topic in educational studies because it reflects how individuals experience, process, and reflect on the learning process. A focus on learner experience provides insights into individual needs, motivations, and preferences, which are crucial in designing effective learning methods and tools (Kolb, 1984). In the

context of learning English, technology-based applications like *Elsa Speak* offer a new perspective on learner experiences, making them more personalized and focused on individual needs.

One important aspect of learner experience is motivation. Motivation is a complex element involving various models and theories (Dörnyei, 2001L). Self-Determination Theory (SDT) by (Ryan & Deci, 2017) focuses on motivation and the psychological needs that influence it in learning environments. According to Self-Determination Theory, motivation in learning is influenced by three psychological needs. The first is Autonomy, which refers to the ability to control one's learning experience. *ELSA Speak* provides autonomy as learners can practice pronunciation independently at their own pace. The second is Competence, which involves the feeling of mastery and effectiveness. ASR technology in *ELSA Speak* helps learners feel competent by providing instant feedback and highlighting areas for improvement. The third is Relatedness, which is a sense of connection with others. Although ASR learning is mostly self-guided, some learners might feel more engaged if there are social features or teacher interactions in the learning process.

The SDT was developed by Deci and Ryan in the 1980s and is rooted in cognitive and humanistic psychology. The

theory evolved from earlier motivation theories that focused more on external rewards and punishments and proposed that intrinsic motivation arises when learners feel autonomous, competent, and connected to others. SDT has been widely used in education, sports, and healthcare to understand motivation and improve learning outcomes. In the context of the ELSA Speak application, SDT provides a perspective to understand how learners interact with ASR technology. The design of this application supports autonomy by allowing learners to practice pronunciation independently. It also supports competence by offering real-time feedback that helps learners recognize their strengths and weaknesses, thereby promoting mastery. While relatedness may be less direct in an ASR-based application, interactive features such as teacher feedback or group activities can further increase learner engagement by addressing this need. Understanding how ELSA Speak supports these psychological needs can help explain why learners are motivated to use the application and how their experiences with it contribute to improving their pronunciation.

Along with the rapid development of technology in education, the use of technology in learning plays an increasingly significant role. Elsa Speak, as an AI-based application, offers an interactive and personalized English pronunciation learning experience by providing pronunciation

analysis, tailored exercises, and real-time feedback. This feedback is particularly valuable as it helps learners identify their mistakes and correct them immediately after completing tasks or exercises. Such features not only enhance learners' awareness of their errors but also encourage independent improvement. By offering instant evaluation, the application allows learners to track their progress effectively, making the learning process more engaging and interactive. This technology is considered capable of providing exercises relevant to individual needs, thereby increasing learners' motivation and involvement in the learning process (Jaelani & Sutari, 2021).

The learner experience with this application highlights several key aspects, such as increased phonetic awareness. Through *Elsa Speak*, learners become more sensitive to how they pronounce words, thanks to the fast and concrete feedback provided by the system. This encourages learners to actively improve their pronunciation. Moreover, this technology transforms how learners perceive and evaluate their pronunciation, as feedback is obtained from an automated system rather than solely from teachers or peers. However, this experience can vary depending on how effectively learners respond to feedback and interact with the technology. For some learners, this technology can enhance their confidence and

motivation to learn. Conversely, some may feel frustrated if the system is not entirely accurate in assessing their pronunciation (Mroz, 2018b).

Thus, although technology like *Elsa Speak* has great potential to support more effective and personalized learning, the learner experience largely depends on their interaction with the application and the feedback they receive during the learning process.

4. ELSA Speak application

ELSA, short for English Language Speech Assistant, is an Android application available for free download on the AppStore or Google Play (Samad & Ismail, 2020). English Learning Speech Assistant (ELSA) was developed by Vu Van in 2015, headquartered in San Francisco, USA. Leveraging Artificial Intelligence (AI) and speech recognition, this app aids in enhancing and refining English pronunciation (Anggraini, 2022). According to (Luu et al., 2021), ELSA Speak utilizes speech recognition technology to assist users in improving their English pronunciation, detecting over 95% of pronunciation errors. Tran (2019) mentions that users can receive feedback to correct their pronunciation errors, with over 1300 lessons and 70 relevant topics available for practice, catering from words to phrases (M. Saragih et al., 2023). Another feature is an interactive dictionary, aiding users in learning how to

pronounce searched words or phrases (Anggraini, 2022). The aim of this application is to boost users' confidence and fluency in English by aiding them in overcoming pronunciation difficulties. The ultimate objective is to empower non-native speakers to communicate effectively and confidently in various situations, be it daily conversations, professional, or academic environments (Sholekhah & Fakhurriana, 2023).

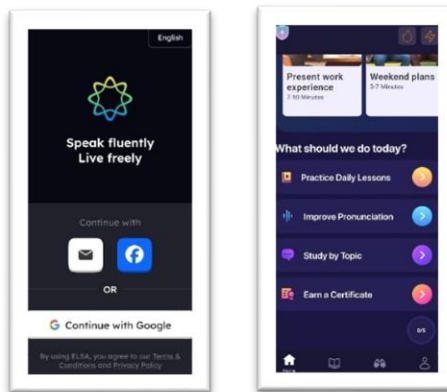
With the world transitioning into an industrial revolution, teaching media must align with modern learning methods. Learners can directly utilize the microphone icon in ELSA Speak to practice speaking as if they were listening to audio. This app benefits all learners greatly due to its user-friendly interface and assistance in learning English pronunciation. Additionally, it incorporates engaging elements to capture learners' interest in learning (Grasella Simanjuntak et al., 2023). ELSA Speak assists in practicing and honing English speaking skills by correcting incorrect pronunciation and explaining the correct pronunciation. Practice sessions can be conducted anytime, anywhere, aiding in improving English speaking abilities (Akhmad & Munawir, 2022).

As mentioned by (Zakiyyah et al., 2022), ELSA Speak offers several advantages. One of them is animated diagrams, which provide visual aids showing the correct positions for each vowel and consonant sound, helping users visualize

pronunciation. Another advantage is immediate feedback where the app uses speech recognition software to provide instant corrections, allowing users to repeat words for improvement. Additionally, accessibility is a key feature, as the app can be easily accessed on laptops or android phones, enabling users to practice pronunciation anytime and anywhere (Saragih et al., 2023).

Picture 2. 1

ELSA Speak application

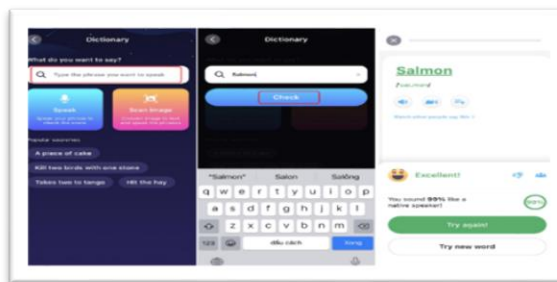




ELSA Speak offers several key features to support pronunciation learning. One of them is the dictionary, which provides the pronunciation of any word along with the option to connect to Youglish.com, a platform featuring YouTube clips emphasizing specific sounds, words, or phrases.

Picture 2. 2

Dictionary

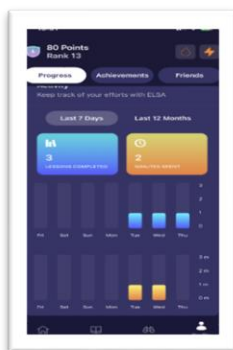


The progress section includes Elsa's pronunciation score and evaluation. The score summarizes progress across three

categories: Word Sound, Word Stress, and Conversation. Word Sound provides feedback on individual phonemes, while Word Stress evaluates syllable emphasis. The English pronunciation score compiles scores from Elsa activities, and the Scoring feature allows users to record sentences and receive feedback on segmental issues like consonant clusters, aspiration sounds, and schwa. Elsa presents correct score percentages and offers detailed reports.

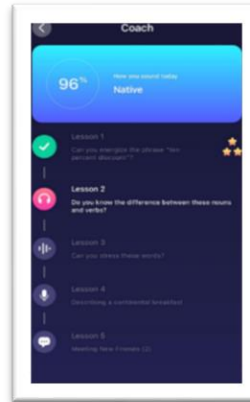
Picture 2. 3

Progress



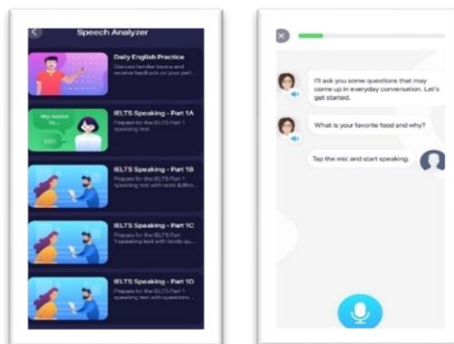
Another feature is coach-daily training lessons. This feature offers practice lessons recommended by Elsa based on the user's progress. Users can enhance their speaking pronunciation skills through these recommended practices. Additionally, users can select from available lessons, such as "Real-life Communication," which includes various conversations for pronunciation practice.

Picture 2. 4
Coach-Daily Training Lessons



The speech analyzer enables users to engage in conversational practice by responding to questions provided by Elsa.

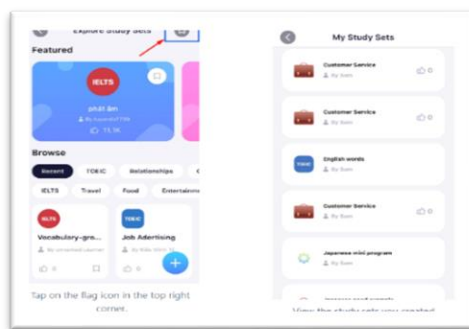
Picture 2. 5
Speech Analyzer



Meanwhile, the study sets feature allow users to compile or organize their learning materials into provided columns by simply inputting the name and category of the material. For instance, under the food category, users can list vocabulary related to desserts, bread, meat, and more.

Picture 2. 6

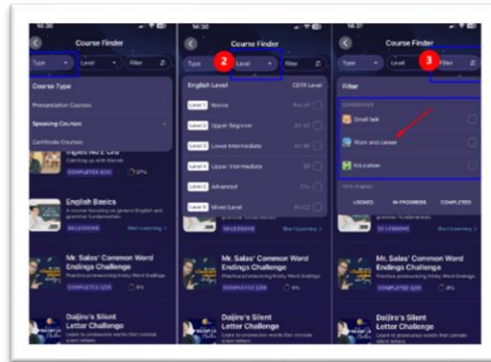
Study Sets



The course finder feature allows users to select topics to study according to their requirements, such as expressing gratitude, giving opinions, self-introduction, or describing something. They can then practice and receive feedback from Elsa. Users can view their scores and improve their pronunciation based on the provided feedback.

Picture 2. 7

Course Finder



Lastly, ELSA AI, supported by exclusive voice recognition technology and the world's largest accent data collection, presents a revolutionary new capability: ELSAAI Tutor. This AI Talk Tutor allows users to engage in real-life conversations that flow freely, spontaneously, and intriguing about a variety of topics or scenarios. Users will then receive personalized and comprehensive feedback about their English language skills in various dimensions.

Picture 2. 8

ELSA AI



ELSA Speak is a language learning application that utilizes Automatic Speech Recognition (ASR) technology to help users improve their English pronunciation skills. The ASR technology in ELSA Speak works by detecting and analyzing the user's voice as they pronounce words or sentences in English. The application then provides immediate feedback on pronunciation accuracy, offers suggestions for improvement, and shows how to pronounce words or phrases more accurately (Levis & Suvorov, 2014).

For instance, when users pronounce specific words, ELSA Speak employs ASR to compare their speech with a standard pronunciation model derived from thousands of native speaker recordings. Through this process, the application provides an

objective and detailed evaluation of pronunciation strengths and weaknesses. Additionally, it personalizes exercises to address individual learning needs, ensuring more effective pronunciation training (McCrocklin, 2019; Mroz, 2018). This ASR-based approach distinguishes ELSA Speak from traditional pronunciation learning methods, which often rely on human feedback that may be subjective or inconsistent. By integrating ASR into the learning process, ELSA Speak allows users to practice independently, refine their pronunciation skills at any time and place, and receive real-time, data-driven feedback. Research suggests that ASR-based pronunciation training enhances learners' pronunciation accuracy, fluency, and overall speaking confidence compared to traditional methods (Chiu et al., 2007; Neri et al., 2008; van Doremalen et al., 2016). Thus, the relationship between ELSA Speak and ASR technology is crucial in improving the effectiveness of English language learning, particularly in pronunciation. ASR's ability to provide instant feedback, objective assessment, and personalized practice makes applications like ELSA Speak an invaluable tool for English learners striving for greater pronunciation proficiency (Wallace, 2016).

The ELSA Speak Application offers a variety of benefits and drawbacks to its users, impacting their learning experience and interaction with the platform. Understanding these

advantages and disadvantages can help users make informed decisions regarding their usage and expectations. There are several advantages offered by Elsa Speak Application for its users. One of them is AI-Powered Feedback, where the application uses advanced AI, including ASR, to evaluate pronunciation and offer real-time feedback on the user's accuracy, helping them quickly identify areas for improvement. Another benefit is flexibility, as learners can choose from a variety of topics and lessons tailored to their interests and proficiency levels, making the learning experience more personalized and engaging.

The application also provides pronunciation practice, allowing users can listen to correct pronunciation models and compare them with their own recordings. This promotes active learning and better pronunciation retention. Additionally, the app offers detailed vvaluation, providing specific feedback on both correct and incorrect pronunciations, helping users understand their mistakes and make necessary corrections. From these advantages, the researcher concludes that the ELSA Speak Application provides a more interactive and personalised learning experience for users. With advanced AI technology, users can receive immediate feedback on the accuracy of their pronunciation, while flexibility in choosing topics and classes as well as categorisation by proficiency level

helps to improve learning efficiency. In addition, the ability to listen to correct pronunciation and record oneself for comparison, along with the evaluation provided by ELSA, helps users correct their mistakes quickly.

However, ELSA Speak also has some disadvantages. One of the main limitations is device dependency, as users must have a smartphone or tablet capable of downloading and installing the app, which may limit access for some learners. Another drawback is paid content, where some topics or lessons within the app may be restricted to paid users, creating barriers for learners who cannot afford premium content. The application also relies on internet access, requiring a stable internet connection to function, which may be inconvenient for users in areas with unreliable internet access. Additionally, there are accent limitations, as ASR technology's effectiveness may be restricted when dealing with specific accents, possibly reducing the accuracy of feedback for learners who do not have a standard accent (Aswaty & Indari, 2022).

From these disadvantages, the researcher concludes that while the ELSA Speak Application has advantages in English language learning, there are some constraints to consider. Reliance on smartphones to download the app limits access, while topic restrictions by paid content and dependence on internet connection can also be barriers. Lastly, the app's

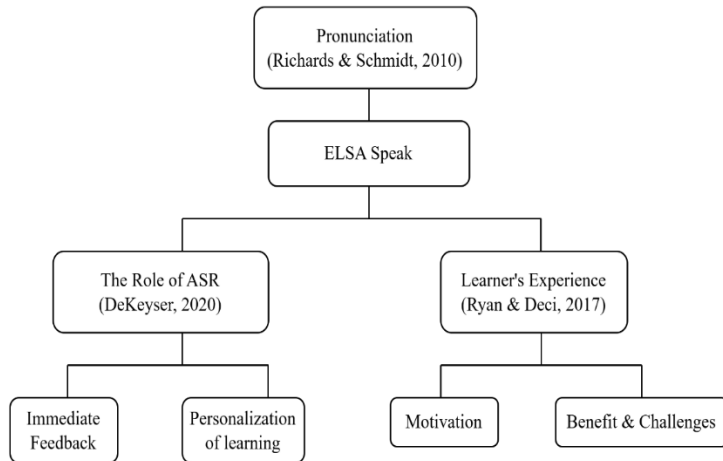
possibly limited effectiveness against certain accents suggests that some users may not get the optimal benefit from using this app.

C. Conceptual Framework

A conceptual framework is a network, or “plane”, of interconnected concepts that collectively offer a thorough understanding of a phenomenon or phenomena. The concepts that make up a conceptual framework support each other, clarify their respective phenomena, and establish a framework-specific philosophy (Jabareen, 2009). The conceptual framework provides numerous advantages for research. For example, it helps the researcher in defining and developing their perspective on the phenomenon being studied (Grant & Osanloo, 2014). The conceptual framework for this research was presented in the diagram on the following page.

Figure 2. 1

Conceptual Framework



This conceptual framework explains the relationship between various elements that support the research on the use of ELSA Speak in improving English pronunciation. The primary focus of the research is pronunciation, as defined by (Richards & Schmidt, 2010), which is the desired outcome of the learning process. ELSA Speak, as a learning tool, serves to bridge theory and practice by utilizing Automatic Speech Recognition (ASR) technology as its main feature, providing immediate feedback to users. Based on the illustrated conceptual framework, this study aims to explore the role of ASR technology in helping to enhance the English pronunciation of learners and to investigate their experiences in using ASR technology through the ELSA Speak app.

The use of ELSA Speak as a medium is based on its ASR feature that supports the learning process by providing immediate feedback on phonological accuracy. The app allows learners to receive real-time assessments and suggestions for improvement, which play a critical role in mastering pronunciation.

The role of ASR, according to (DeKeyser, 2020), involves two important aspects: immediate feedback and personalization of learning. Immediate feedback enables users to recognize their errors promptly, which is essential in the initial stages of skill acquisition (the declarative stage). Additionally, the personalization of learning tailors the exercises based on individual needs and abilities, accelerating the transition to more automated skill stages.

On the other hand, the learner's experience plays a significant role in the learning process, as described by (Ryan & Deci, 2017) through their self-determination motivation theory. This experience includes aspects such as motivation, perceived benefits, and challenges faced. Users' motivation is enhanced through engaging exercises and the personalized features offered by ELSA Speak, while the benefits include increased confidence and improved understanding of pronunciation. Challenges, such as technical issues or ASR limitations, are also explored to provide a holistic view of the learning experience.

Thus, this framework connects the role of ASR technology in providing technical learning support with the subjective

experiences of users. Together, these factors contribute to improving English pronunciation skills, establishing ELSA Speak as an effective tool in bridging the gap between pronunciation teaching theories and real-world practice.

CHAPTER III

RESEARCH METHOD

This chapter explains the methodology of the study. The descriptions include research design, The setting and the time of the research, population and sample of the research, data collecting technique, data analysis technique, and instrument of research.

A. Research Design

This research employs a qualitative case study design to investigate the role of the ELSA Speak application in bridging the pronunciation gap among English language Learners. According to (Creswell, 2018), qualitative research is a method for exploring and understanding the meanings individuals or groups attribute to social or humanitarian problems. This approach is particularly useful in educational research as it allows for an in-depth understanding of complex issues such as language acquisition and the challenges faced by learners.

Hancock and Algozzine stated that case study research focuses on specific phenomena, events, situations, programs, or activities, providing an in-depth and detailed description of the subject being studied. By investigating specific cases, researchers can explore various aspects and variables affecting the subject comprehensively. This approach is beneficial for understanding the nuanced ways in which technological tools like the ELSA Speak

application impact language learning processes and outcomes (E. E. Saragih et al., 2021). This method allows for a thorough exploration of learners' experiences with the application as well as the unique challenges and improvements they encounter.

In the context of this research, the case study method will be used to understand how English language learners utilize ASR technology in the ELSA Speak application and its impact on improving their pronunciation skills. This approach aims to provide rich and nuanced insights into learners' experiences with the application within the context of language learning environments. The qualitative research conducted is expected to make a significant contribution to the field of language education, particularly regarding the use of technology to improve pronunciation skills. Furthermore, the findings from this study can serve as a basis for developing more effective and efficient learning methods in English language education, ultimately promoting better outcomes for learners in their pronunciation development.

B. Research Settings

a. Place

This study was conducted on third semester students of English Education Department of UIN Walisongo Semarang in the academic year 2023/2024, which is located at Jalan Walisongo No. 3-5 Tambakaji, Ngaliyan, Semarang, Central Java. The researcher chose this location with the consideration

that the English Education Department at UIN Walisongo has significant potential to be researched. In addition, its proximity to the researcher, who is also a student in the same department, facilitates the data collection process as well as streamlining time and costs.

b. Time

This research was conducted in November – Desember 2024. This research aims to understand the role of Automatic Speech Recognition (ASR) technology in the ELSA Speak application in enhancing English language learners' pronunciation skills. To achieve this goal, the researcher will follow a structured series of steps, beginning with the development of a questionnaire designed to gather information about learners' experiences with the ELSA Speak application. This questionnaire is intended to identify various aspects of learners' experiences, including usability, benefits, and the challenges they encounter when using the ASR feature.

After distributing the questionnaire, an analysis of the responses will be conducted to select a diverse group of participants for interviews. Through interviews, the researcher will explore learners' experiences in more detail, allowing for a better understanding of how ASR technology supports pronunciation learning. The findings from this research are expected to contribute significantly to the understanding of

technology use in language education, as well as to inform better teaching practices. The table below outlines the complete schedule of research activities. This scheduling is designed to ensure that the entire research process can proceed in a regular and organized manner, facilitating effective and efficient data collection.

Table 3. 1
Instrument Guideline

No. Step	Activity	Description
1.	Questionnaire Development	Create and test the questionnaire to collect data on learners' experiences with the ELSA Speak application and ASR technology.
2.	Questionnaire Distribution	Administer the completed questionnaire to learners who have used ELSA Speak to identify potential interview participants.
3.	Response Analysis	Analyze the questionnaire results to select a diverse group of participants for interviews based on their experiences with the application.
4.	Conduct Interviews	Schedule and conduct interviews with selected participants to explore their experiences and perceptions of ASR technology in ELSA Speak.

5.	Interview Data Analysis	Transcribe and analyze the interview data to identify key themes related to the role of ASR and learners' experiences.
6.	Findings Report	Integrate the findings from the questionnaires and interviews into a comprehensive research report.

C. Research Participant

Subject of this research is the third-semester English Education students at UIN Walisongo Semarang. The researcher will use purposive sampling. Purposive sampling is a non-random sampling method in which participants are chosen according to particular characteristics or criteria that align with the research question or topic, making them especially pertinent to the study (Sugiyono, 2020). The research sample consists of 9 third-semester students from the English Education Department at UIN Walisongo Semarang who have used the application. The participants were selected from those who completed the questionnaire based on predetermined criteri. Subsequently, the participants who met the criteria will be interviewed to provide relevant answers to the research questions. The criteria are as follows, They must be students in the third semester, enrolled in the English Education Department at UIN Walisongo Semarang, and actively using the ELSA Speak application

Selecting participants with direct experience is particularly relevant because this study aims to explore the role of Automatic Speech Recognition (ASR) technology in supporting the improvement of English pronunciation. Purposive sampling allows the researcher to target a specific group of individuals who meet these criteria, which is especially useful when it is challenging to obtain a random or representative sample. However, it is important to acknowledge that purposive sampling can introduce bias and may not fully reflect the perspectives of a broader population.

D. Research Focus

In this research, the researcher solely focuses on the Automatic Speech Recognition (ASR) technology in the ELSA Speak application, accessible through <https://id.elsaspeak.com/>. The study aims to explore the role of ASR and learners' experiences with utilizing the ASR feature in the ELSA Speak app and how it aids in improving pronunciation gaps.

E. Data Collecting Technique

In this research, the primary data collection instrument that will be used is Questionnaire and Interviews. The data explored through these interviews will focus on the opinions, experiences, and statements of the participants, who are third-semester students from the English Education Department at UIN Walisongo Semarang.

1. Questionnaire

According to Brown (2007), a "questionnaire" is defined as a written instrument that asks respondents to answer a series of statements either by writing their responses or by choosing from a list of pre-prepared answers. The questionnaire in this study was carefully prepared and self-piloted to ensure that the instrument effectively collects the necessary information regarding learners' perceptions (Dörnyei & Dewaele, 2022). The questionnaire used in this study consists of closed-ended items designed to gather information about learners' experiences with the ELSA Speak application as a medium to support English pronunciation learning. The researcher prepared eleven questionnaire items and distributed them to all third-semester English Education Department students at UIN Walisongo Semarang who had previously used the ELSA Speak application. From the distributed questionnaires, nine participants completed the survey. Data were collected through a survey that utilized a five-point Likert scale, with response options ranging from Strongly Disagree to Strongly Agree.

To facilitate understanding, the questionnaire questions were presented to participants in both English and Indonesian, reducing the risk of misunderstanding regarding the intended meaning of the questions. This approach aims to enhance the clarity of the questions and ensure accurate responses from the

participants. The questionnaire were adapted from (Jaelani & Sutari, 2021) to improve their clarity and ensure that learners could easily comprehend them. Furthermore, the questionnaire was presented in the form of a Google Form, allowing for easy distribution and completion by the participants. A copy of the questionnaire can be found in the Appendix 1.

2. Interview

According to Kvale & Brinkmann, interviews are structured conversations centered around a theme of shared interest, characterized as a dialogue where two individuals exchange perspectives. Unlike regular conversations, interviews have a defined format and objective. The interview method used in this research is a semi-structured interview. A semi-structured interview is an interview where the interviewer determines the topics and key questions but provides flexibility for participants to develop their answers (Kvale & Brinkmann, 2015).

After completing the questionnaire, the nine participants proceeded to the interview stage to provide a more in-depth understanding of their experiences with the ASR technology in the ELSA Speak application. The interviews are conducted orally through face-to-face meetings with each individual and online meeting with google meet. Prior to the interviews, the researcher prepares an interview instrument called the interview guide. The questions asked cover facts, data comprehension, concepts,

arguments, perceptions, or participants' views related to the research focus. The researcher utilizes interviews to obtain information relevant to the research questions. The researcher records, takes notes, listens carefully, and then transcribes the information provided by the informants. The interview questions were adapted from (Mroz, 2018) to enhance their clarity and ensure that learners could easily understand them. A copy of the interview questions can be found in the Appendix 2.

F. Data Analysis Technique

After collecting data from the questionnaire and interview, both data needs to be analyzed. The questionnaire data were organized into a table and summarized using frequencies and percentages. These results were directly obtained from the google form responses. The researcher the analyzed and interpreted the data presented in the table. Data collected from semi-structured interviews were analyzed using thematic analysis. Thematic analysis is a method for analyzing qualitative data that involves searching across a data set to identify, analyze, and report repeated patterns (Kiger & Varpio, 2020). This method is particularly suitable for understanding experiences, thoughts, or behaviors within a data set.

(Braun & Clarke, 2021) outlined the steps of thematic analysis into 6 phases. The first phase is familiarizing with the data, where the researcher reads and rereads the questionnaire data and transcribes the interview results to gain a deep understanding of the

data. The second phase is generating initial codes, where the researcher determines codes based on the interesting points from the interviews. In the third phase, searching for themes, the analysis is expanded from collected code-shaped points into formed themes. Visual representations can greatly assist in grouping different codes into themes. The fourth phase is reviewing themes, where the researcher checks back if the themes align with the predetermined codes. Then, in the fifth phase, defining and naming themes, the researcher refines the themes by filtering and identifying each theme. Finally, the sixth phase is producing the report, where the researcher presents a complex narrative about the data in a convincing manner, demonstrating the benefits and validity of the analysis.

By following these steps, the researcher can gain a deeper understanding of the collected data, identify emerging patterns, and explore the meaning behind these patterns. These steps help the researcher organize, analyze, and report research findings systematically and structurally. Consequently, the researcher can present research findings comprehensively and convincingly, thereby enhancing the reliability and trustworthiness of the analysis results.

CHAPTER IV

FINDING AND DISCUSSION

This chapter presents the results of the research on the role of Automatic Speech Recognition (ASR) technology in bridging the pronunciation gap for English language learners. The study focuses on third-semester students from the English Education program at UIN Walisongo Semarang, who were selected as participants. A total of 9 students who use the ELSA Speak app to learn English pronunciation were involved in this study.

In this chapter, the data obtained from the questionnaire and interviews will be presented and analyzed. The questionnaire was used as supporting data to explore the participants' learning experiences with ASR technology in the ELSA Speak app, specifically in supporting the improvement of their English pronunciation skills. Meanwhile, interviews were conducted with the 9 selected participants to gain a deeper understanding of their views on the effectiveness of the app and how ASR technology contributes to improving their English pronunciation.

The questionnaire data were collected through Google Form, while the interview data were obtained through google platform and face-to-face interviews. The researcher scheduled the meetings by sending messages to the participants via WhatsApp. The results from the questionnaire and interviews will be discussed in detail in this

chapter to provide a clearer understanding of how the ELSA Speak app affects the learners' English pronunciation and the role of ASR technology in helping bridge the gap in their pronunciation skills.

A. Finding

4.1 Learners experience with the use of ASR technology in the ELSA Speak Application

4.1.1 Motivation for using ASR technology in ELSA Speak

In this section, the data obtained from the questionnaire provides an overview of learners motivation of using Automatic Speech Recognition (ASR) technology in the ELSA Speak application. Based on the processed results of the questionnaire, the analysis of the two statements asked to the participants will be described below.

Table 4. 1

Questionnaire Result of motivation

No	Statement	Alternative answer					Total
		SA	A	N	D	SD	
1.	ELSA Speak is easy to use for practicing English pronunciation	4 44,4%	4 44,4%	1 11,1%	0 0%	0 0%	9 100%
4.	The pronunciation exercises in ELSA Speak are interesting and motivate	1 11,1%	5 55,6%	3 33,3%	0 0%	0 0%	9 100%

Based on the table above, the results for statement 1 (S1) show that the majority of participants, 44.4%, strongly

agreed (SA), and another 44.4% agreed (A), indicating that they found the ELSA Speak application easy to use for effectively practicing English pronunciation. Meanwhile, only 11.1% chose a neutral (N) response, suggesting that while the application was generally considered easy to use, some participants might still face certain challenges in navigation or available features. No participants selected disagree (D) or strongly disagree (SD), which indicates that ELSA Speak was generally well received by participants as a pronunciation learning tool.

Additionally, most participants also stated that the pronunciation exercises in ELSA Speak were engaging and motivated them to continue learning. The results for statement 4 (S4) show that the majority of participants, 55.6%, agreed (A), indicating that they found the pronunciation exercises in ELSA Speak engaging and motivating them to continue practicing. A total of 33.3% chose neutral (N), suggesting that while they were not entirely stimulated by the exercises, they were also not disturbed by them. Meanwhile, 11.1% strongly agreed (SA), indicating that for a small portion of participants, the pronunciation exercises in the application were highly engaging and significantly boosted their motivation to learn. No participants selected disagree (D) or strongly disagree

(SD), showing that the majority felt that using ELSA Speak positively impacted their motivation to learn pronunciation.

Overall, the results of the questionnaire presented in the table above indicate that most participants have a positive perception of using ASR technology in ELSA Speak. They found the application easy to use and believed that the pronunciation exercises provided were engaging and could enhance their motivation to continue learning. These findings suggest that ELSA Speak has great potential in supporting English pronunciation learning, particularly in increasing learners' engagement and motivation in pronunciation practice.

The questionnaire results are also reinforced by interview findings, which provide deeper insights into users' experiences with the ELSA Speak application. Through interviews, participants were able to share their views directly and describe their experiences in utilizing various application features and the benefits they perceived while using ELSA Speak. The following statements from the interview provide a more detailed picture of how users experienced ASR technology in this application and its impact on their pronunciation learning, as expressed below:

P3

.....With this, ELSA Speak gives me control over my learning process. I can practice on my own without having to wait for corrections from others, making me more independent in improving my pronunciation. With instant feedback, I feel more confident to try speaking English in real-life situations.

P3 highlighted that the application gave them full control over their learning process, allowing them to practice without waiting for corrections from others. This made them more independent in improving their pronunciation and feel more confident in attempting to speak in real-life situations. This experience demonstrates that ease of access and flexibility in learning are crucial factors that encourage participants to use ELSA Speak consistently.

P8

..... it gives me the freedom to practice according to my needs, making me feel more independent in learning. I can also repeat the exercises as many times as I want until I feel confident in my pronunciation. As a result, I feel

more confident when speaking English because my pronunciation skills have improved.

Meanwhile, P8 emphasized that the instant feedback provided by ASR helped enhance their pronunciation skills and confidence in speaking English. With a feature that allows unlimited practice repetitions as needed, they felt they had the freedom to regulate their own learning pace. This motivated them to practice pronunciation without feeling pressured by time constraints or rigid learning methods.

P5

..... With this feedback, I feel more motivated to keep practicing and more confident in my pronunciation skills.

On the other hand, P5 provided a more specific perspective on how ASR in ELSA Speak helped distinguish the pronunciation of similar words (minimal pairs). With direct feedback from the application, they felt more confident in pronouncing words correctly, which ultimately increased their motivation to continue practicing pronunciation.

Based on the findings from the questionnaire and interviews, it can be concluded that participants' motivation to use ELSA Speak is strongly related to their experience in

controlling their learning process, increasing confidence, and receiving support through the interactive features of the application. The ease of use and flexibility in pronunciation practice encouraged participants to use it consistently. Additionally, the presence of instant feedback helped them understand and correct their mistakes, making them feel they were making real progress in their pronunciation skills.

However, some participants remained neutral regarding the application's effectiveness in enhancing their motivation. This indicates that, in some cases, the experience of learning pronunciation through ELSA Speak is still influenced by other factors, such as their usual learning methods or individual preferences in receiving feedback. Therefore, the use of ASR in ELSA Speak may be more effective when combined with other learning strategies that better suit each user's needs.

For example, P6 stated that the instant feedback feature greatly helped them understand pronunciation accuracy. They said:

P6

..... For example, when I struggled to pronounce the word "comfortable," the app demonstrated the correct sound and what needed improvement.

Meanwhile, P5 highlighted how the application helps distinguish between similar-sounding words, which further motivates them to practice. They stated:

P5

.....For example, in distinguishing the pronunciation of "Stationery" and "Stationary." With this feedback, I feel more motivated to keep practicing and more confident in my English pronunciation skills.

4.1.2 Benefits of Using ASR Technology in ELSA Speak

The results of the questionnaire show several benefits perceived by users when using ASR technology in the ELSA Speak application.

Table 4. 2

Questionnaire Result of Benefit

No	Statement	Alternative answer					Total
		SA	A	N	D	SD	
3.	I can use ELSA Speak anytime and anywhere to improve my pronunciation	2	7	0	0	0	9
		22,2%	77,8%	0%	0%	0%	100%
6.	ELSA Speak helps me understand the correct way to pronounce English	2	6	1	0	0	9
		22,2%	66,7%	11,1%	0%	0%	100%
10.	Learning pronunciation with ELSA Speak deepens my understanding of English pronunciation	2	5	2	0	0	9
		22,2%	55,6%	22,2%	0%	0%	100%

Based on the table above, Statement 3 (S3) indicates that the majority of participants (77.8%) agreed (A) with this statement, showing that they feel flexible using ELSA Speak anytime and anywhere to improve their pronunciation. This highlights that the application provides users with the freedom of time and location flexibility, which significantly supports the process of learning pronunciation. Only a small portion (22.2%) strongly agreed (SA) with this statement, indicating that they feel greatly assisted by the easy accessibility of the application. None of the participants selected disagree (D) or strongly disagree (SD), which demonstrates that ELSA Speak is highly accepted by users in terms of convenience and accessibility.

Statement 6 (S6) shows that the majority of participants (66.7%) agreed (A) that the ELSA Speak application helps them understand the correct way to pronounce English words. This indicates that the application is considered effective in providing guidance and exercises that help users grasp proper pronunciation. Additionally, 22.2% of participants strongly agreed (SA), signifying that they find the application very useful for understanding the correct pronunciation of words. A small portion (11.1%) gave a neutral response (N), which might indicate that

although they find the application useful, its impact on their understanding is not very significant or immediately felt.

The results of the questionnaire for Statement 10 (S10) show that the majority of participants (55.6%) agreed (A) that learning pronunciation using ELSA Speak helps them deepen their understanding of English pronunciation. This suggests that the application successfully provides a comprehensive learning experience for improving their pronunciation. However, 22.2% of participants strongly agreed (SA), indicating that they feel significantly helped in deepening their understanding of pronunciation. On the other hand, another 22.2% gave a neutral response (N), which might suggest that while they see some benefits, their level of understanding has not improved significantly, or they still need more time to fully experience its impact.

Overall, the results of these statements indicate that the majority of participants feel that ELSA Speak is very helpful in improving and understanding English pronunciation. The main advantages of this application lie in its flexibility for use anytime and anywhere (S3) and its ability to assist users in understanding the correct way to pronounce words (S6). However, there are some participants who provided neutral responses (N), indicating that the benefits may not yet be fully realized by all users, or they need more time and

practice to see clearer results. The presence of neutral responses shows that not all participants have the same experience, and there may be other factors influencing their perception of this application.

The results of the questionnaire above are also reinforced by statements from interview findings, as expressed in the following statements:

P7

..... I can practice English pronunciation anytime and anywhere. this app highlights specific sounds that I need to improve, such as the 'th' sound, and provides guidance on how to correct it.

P1

..... instant feedback, directed practice, and consistency. For instance, I have trouble distinguishing between the pronunciation of 'Sheep' and 'Ship', and this app repeatedly points out my mistakes and guides me toward the correct pronunciation.

The statements from participants P7 and P1 emphasize that one of the main advantages of using this application is the flexibility it offers for practice. Users can practice anytime and anywhere, which greatly supports English

language learning outside the classroom. The questionnaire results, showing that most participants feel the application enables them to practice at any time, reflect the convenience and ease provided by this application. With easy access, learners can practice their pronunciation without being constrained by time or location, making it a significant added value in the learning process.

P4

The instant feedback allows me to correct my pronunciation mistakes immediately. Additionally, it offers personalized learning, focusing on the areas I find most challenging.

P5

The ASR technology in the ELSA Speak app provides immediate feedback by showing where the mistakes are and offering precise corrections for the pronunciation.

The statements from participants P1, P4, and P5 highlight the importance of immediate feedback in the process of learning pronunciation. One of the key advantages of ASR (Automatic Speech Recognition) technology in the ELSA Speak application is its ability to provide instant feedback on pronunciation errors. For example, the difficulty in distinguishing the pronunciation

of the words “Sheep” and “Ship,” as mentioned by P1, demonstrates how the application helps users recognize their mistakes and provides accurate corrections. The technology's ability to pinpoint specific errors (such as the "th" sound or similar vowel pronunciations) facilitates a better understanding of how to correctly pronounce words, aligning with questionnaire findings that participants reported significant improvements in their pronunciation comprehension.

P4's statement also emphasizes that the application offers personalized learning. Each user can focus on areas they find challenging, based on their specific pronunciation difficulties. This supports the questionnaire results, which show that most participants feel the application helps them better understand how to pronounce English words. Personalized learning allows learners to address their weaknesses more effectively and according to their needs, improving the quality of the learning process.

As explained by P1, the use of this application provides learners with the opportunity for repeated practice, enabling them to consistently correct their mistakes. Pronunciation challenges are repeatedly addressed with feedback, allowing learners to continuously improve. Consistency in practice and receiving accurate feedback are

key to language learning, helping learners enhance their pronunciation skills in the long term.

Overall, the results of the questionnaire and interviews indicate that ASR technology in the ELSA Speak application plays a significant role in helping learners improve their English pronunciation. Features such as instant feedback, specific corrections, flexible accessibility, and personalized learning greatly support the learning process. This demonstrates that the use of ASR technology can be an effective tool to help learners overcome pronunciation gaps and improve their English pronunciation skills independently.

4.1.3 Challenges in Using ASR Technology on ELSA Speak

In addition to its benefits, users also identified several limitations and challenges in using ASR technology in the ELSA Speak application. The interview findings reveal that while ELSA Speak offers numerous advantages, there are a number of limitations and challenges faced by users. One of the obstacles is the need for a stable internet connection. As expressed by one participant,

P5

This app is difficult to access without using mobile data.

This becomes a barrier for users in areas with limited internet access or those facing data quota constraints, reducing the flexibility of using the application. Additionally, some features and materials in the application are only available to subscribed users, as noted by one participant:

P7

The downside is that some topics can only be accessed if the user subscribes. Therefore, those who haven't subscribed can only access limited content.

This indicates that limited access to premium content can hinder the learning experience, especially for users who lack the resources to subscribe. Another challenge faced is the inaccuracy of ASR technology in recognizing users' voices. One participant described this situation by saying,

P2

Sometimes, this technology does not recognize my voice well, especially when I am nervous or speak too quickly. There was a time when the app said I pronounced a word incorrectly, even though I felt I had pronounced it correctly.

This indicates that factors such as emotions or speaking speed can affect the application's ability to provide accurate

feedback, potentially leading to frustration for users. These challenges highlight that, despite the significant benefits of ASR technology in ELSA Speak for learning pronunciation, technical aspects such as the need for internet access, limited access to premium content, and inaccuracies in voice recognition require attention. Solutions to address these limitations, such as offering offline content options, expanding free access, or improving voice recognition algorithms, could make the application more inclusive and effective for all users.

4.2 The Role of ASR Technology in Supporting English Pronunciation Improvement

4.2.1 Identification of Pronunciation Errors and Corrections

This section reviews the role of Automatic Speech Recognition (ASR) technology in the ELSA Speak application in helping users identify pronunciation errors and providing useful corrections. Based on the questionnaire results, the majority of participants gave positive feedback on the application's ability to provide accurate feedback and assist in the pronunciation learning process.

Table 4. 3*Questionnaire Result of Pronunciation Errors*

No	Statement	Alternative answer					Total
		SA	A	N	D	SD	
2.	ELSA Speak provides clear feedback and helps me identify my pronunciation	2	6	1	0	0	9
		22,2%	66,7%	11,1%	0%	0%	100%
5.	The feedback from ELSA Speak helps me improve my pronunciation effectively.	3	5	1	0	0	9
		33,3%	55,6%	11,1%	0%	0%	100%

Based on the data obtained, in statement 2 (S2), the majority of participants stated that the application provides clear and effective feedback. As many as 22.2% of participants strongly agree (SA) that ELSA Speak provides useful feedback, while 66.7% agree (A) with the statement. This shows that almost all participants (88.9%) feel significant benefits from the feedback provided by the application. Users feel that the app is able to indicate the location of pronunciation errors in an easily understandable way. Conversely, only 11.1% of participants were neutral (N), with none disagreeing or strongly disagreeing.

In statement 5 (S5), 33.3% of participants strongly agree (SA) and 55.6% agree (A). This means that the majority of participants (88.9%) acknowledge that the feedback provided by the app is effective in improving their

pronunciation skills. The feedback provided by the app not only indicates mistakes but also provides targeted correction guidance. Only 11.1% of participants were neutral (N), with no participants providing negative feedback.

These results highlight the ability of ELSA Speak as an innovative and effective language learning tool, especially in identifying pronunciation errors. The ASR technology in this app is designed to recognize sounds or pronunciations that do not conform to standard English. By specifically highlighting these errors, users can identify areas that need improvement. Additionally, the correction feature provides direct suggestions on how to pronounce specific words or sounds correctly.

Overall, the data from the questionnaire indicates that ASR technology in ELSA Speak plays an important role in improving English pronunciation skills. The error identification and correction features provided by this app offer significant benefits to learners, allowing them to better recognize their weaknesses and correct them effectively. The combination of instant feedback, targeted guidance, and personalized learning experiences makes ELSA Speak a highly supportive tool in the process of mastering English pronunciation.

The questionnaire results are further reinforced by the in-depth insights from the interviews, which provide a deeper understanding of how ASR technology in ELSA Speak supports learners in overcoming their pronunciation challenges. One of the interview participants stated,

P4

..... This app can detect small mistakes that are often overlooked and provide the correct guidance to fix them.

This statement highlights the significant advantages of ASR technology, particularly its ability to detect small pronunciation errors that might go unnoticed in conventional learning environments. The app's ability to identify these nuances allows learners to focus on areas that require improvement, thereby systematically enhancing their pronunciation. Another interview participant elaborated further by sharing their personal experience:

P8

.....This technology allows me to immediately identify the areas that need improvement, making it easier to understand the correct pronunciation.

This statement emphasizes the instant feedback provided by ELSA Speak. Unlike traditional methods, which may take time to provide feedback or rely on external

assessments, ASR technology offers learners immediate insights into their mistakes. The quick feedback allows learners to correct their pronunciation errors right away, while their memory of how they pronounced the word is still fresh. This helps them eliminate incorrect pronunciation habits and replace them with more accurate ones.

P9

The ASR technology in ELSA Speak acts as a virtual coach that provides instant feedback. This is very helpful in identifying small mistakes that I might not have noticed before, such as intonation or word stress.

This statement provides an opinion about the use of ASR technology in the ELSA Speak app, which functions as a virtual coach providing instant feedback after participants complete their exercises. This feedback is highly beneficial for identifying and correcting small mistakes that users often overlook, such as errors in intonation or word stress. It shows that ASR technology in ELSA Speak plays a crucial role in improving English speaking skills, particularly in more subtle aspects that learners may not detect on their own.

The combination of these three responses provides a comprehensive view of how learners perceive the role of

ASR technology in their pronunciation practice. Data from both the questionnaire and interviews align in highlighting two main strengths of ELSA Speak: accurate error detection and timely corrective feedback. Together, these features help learners understand proper pronunciation patterns more clearly, enabling them to make consistent improvements over time.

Moreover, these findings suggest that ASR technology functions as a personalized learning tool, replicating some of the benefits of one-on-one teaching. By analyzing pronunciation mistakes individually, the app tailors its feedback to the user's specific needs, fostering a sense of confidence and motivation among learners. This aligns with the primary goal of integrating ASR in language learning: to bridge the gaps in traditional methods by harnessing technology's potential to offer personalized, immediate, and practical solutions to pronunciation difficulties.

Thus, the interview responses not only support the findings from the questionnaire but also highlight the transformative role of ASR technology in enhancing the pronunciation learning experience, making it an essential tool for English learners.

4.2.2 Improvement of Pronunciation Skills

This section examines the responses from participants indicating the extent to which they felt the benefits of the ELSA Speak app in improving their pronunciation skills.

Table 4. 4

Questionnaire Result of Improve Pronunciation

No	Statement	Alternative answer					Total
		SA	A	N	D	SD	
7.	I feel that my pronunciation has not improved by using ELSA Speak	0	0	1	7	1	9
		0%	0%	11,1%	77,8%	11,1%	100%
8.	ELSA Speak helps me be more confident in speaking English with correct pronunciation	1	5	3	0	0	9
		11,1%	55,6%	33,3%	0%	0%	100%
9.	Using ELSA Speak makes me more active in practicing pronunciation outside of class	0	4	5	0	0	9
		0%	44,4%	55,6%	0%	0%	100%
11.	My experience with ELSA Speak helps me apply pronunciation skills in daily conversations	0	6	3	0	0	9
		0%	66,7%	33,3%	0%	0%	100%

The questionnaire results presented above show that the majority of participants experienced an improvement in their pronunciation skills after using the ELSA Speak app. In statement 7 (S7), most participant (77.8%) disagreed (D) with the statement, indicating that they felt an improvement

in their pronunciation after using the app. Even 11.1% of participants strongly disagreed (SD), suggesting that they were confident the app had made a clear improvement in their pronunciation. Only 11.1% of participants were neutral (N), indicating uncertainty about whether any improvement had been made.

In statement 8 (S8), the majority of participants (55.6%) agreed (A) with the statement, meaning that the app helped improve their confidence, although not as strongly as those who strongly agreed (SA). Meanwhile, 11.1% strongly agreed (SA) with the statement, indicating that they experienced a significant increase in confidence in speaking English with correct pronunciation after using the app. However, 33.3% of participants were neutral (N), which shows that not everyone felt a significant change in their confidence.

On the other hand, in statement 9 (S9), 44.4% of participants agreed (A) with the statement 11 (S11), meaning the majority felt that the ELSA Speak app encouraged them to practice English pronunciation more actively outside of class. They felt motivated to continue practicing after formal learning sessions. However, more than half of the participants (55.6%) chose neutral (N), indicating that while the app was beneficial, not everyone

felt motivated to spend additional time practicing pronunciation outside of lessons.

Furthermore, in statement 11 (S11), 66.7% of participants agreed (A) and 33.3% were neutral (N). No participants strongly agreed, disagreed, or strongly disagreed. This result suggests that most participants felt the app was helpful in allowing them to apply their pronunciation skills in everyday conversations.

Overall, these questionnaire results indicate that ELSA Speak plays a significant role in improving users' pronunciation skills. The app not only helps correct pronunciation but also boosts users' confidence in speaking English. Additionally, some participants feel more motivated to practice pronunciation outside of class and are able to apply their skills in daily conversations. Participants who were neutral about some statements suggest that the learning experience through the app may still require a more personalized approach to meet individual needs.

The findings regarding pronunciation improvement are further supported by interviews, which provide deeper insights into the user experience with ELSA Speak. One participant stated:

P3

I think so. I believe my pronunciation has improved, especially in individual sounds and stress patterns. This app points out small mistakes that I might not have noticed myself. It has been very helpful in stress and intonation. The app breaks down sentences or words and highlights the parts I need to focus on.....

The statement from P3 above mentions that the application helps improve individual sound pronunciation and word stress patterns. According to P3, the app is able to show small mistakes that may go unnoticed, and the word stress and intonation features are especially helpful. With the feature that highlights the syllables that need emphasis, users can practice more effectively and correct their pronunciation through several attempts.

P5

..... I can identify my mistakes in English pronunciation. The corrections provided make me more careful in my pronunciation, as a mispronounced word can change its meaning.

The pronunciation mistakes they make. The corrections provided by the app make them more cautious in their pronunciation, as they realize that mispronouncing

words can change the meaning of a sentence. This indicates that the participants have become more aware of the importance of correct pronunciation in English, and the app raises their awareness about the impact of incorrect pronunciation.

P8

Yes, I feel that my pronunciation has improved after using the ELSA Speak application. It has also changed my perspective on the importance of paying attention to correct, precise, and clear pronunciation.....

Meanwhile, participant P8 stated that using the ELSA Speak app had significantly improved their pronunciation. The app also changed their perspective on the importance of clear and accurate pronunciation. Previously, they did not pay much attention to pronunciation, but after using the app, they realized how crucial good pronunciation is for effective communication. This experience made the participant feel more confident speaking English in front of others, showing the positive impact of the app on the user's confidence in communication.

Overall, the interview results reinforce the findings from the questionnaire, showing that the ASR technology in the ELSA Speak app not only helps users identify

pronunciation mistakes but also provides useful corrections and raises users' awareness of the importance of correct pronunciation. With the feedback provided, users feel more prepared and confident in using English, both in personal and professional contexts.

4.2.3 Contribution of ASR to the Effectiveness of English Language Learning

Based on the interview results, participant provided positive views regarding the contribution of ASR technology in the ELSA Speak app to the effectiveness of English language learning. Participant P9 stated that the ASR technology in the ELSA Speak app is very effective and easy to use, providing a personalized and efficient learning experience. Participant P9 said:

P9

.....it is highly effective and easy to use. The ASR technology provides a personalized and efficient learning experience, especially for those who want to improve their pronunciation independently.

This statement indicates that for users who want to improve their pronunciation independently, ELSA Speak with ASR technology provides a personalized experience tailored to their individual needs. This technology offers

instant feedback that allows users to practice efficiently without external assistance. This makes the app particularly useful for learners who wish to improve their pronunciation skills independently and in a more flexible way.

Meanwhile, participant P3 provides a more cautious view regarding the use of the app. They acknowledge the benefits of ASR technology but also emphasize the importance of speaking in real-life conversational contexts for further development. Participant P3 stated:

P3

.....Use this app as a learning tool, not something to rely on completely. This app is great for practicing and improving your pronunciation, but real conversations are where you truly develop. This app is very useful for building confidence before stepping into the real world.

This statement provides an understanding that while ELSA Speak is effective for pronunciation practice, users also need to engage in real-life conversations to truly hone their speaking abilities. The app is more effective as a supplementary tool that helps users prepare and build confidence before speaking in real-life situations. This participant highlights the importance of balancing self-study

using ASR technology with direct speaking experiences in social contexts.

Overall, both interviews indicate that ASR technology in ELSA Speak plays a significant role in improving users' pronunciation skills. While the app is highly effective as a self-learning tool, integrating speaking practice in real-life conversations remains essential for further development of English speaking skills.

B. Discussion

The first findings of this study indicate that the majority of participants were motivated to use the Automatic Speech Recognition (ASR) technology in the ELSA Speak application due to its ability to provide immediate feedback and allow independent practice. According to Self-Determination Theory (Ryan & Deci, 2017), motivation in learning is enhanced when learners experience autonomy, competence, and relatedness. In this study, participants reported that ASR technology encouraged them to practice pronunciation independently, without having to rely on direct teacher supervision. One participant mentioned that the instant feedback feature gave them more confidence to speak in real-life situations because they could recognize and correct their mistakes independently. This suggests that ASR technology in ELSA Speak fosters learner autonomy, empowering students to take control of their pronunciation learning process. This aligns

with research by (Senowarsito & Ardini, 2023), which found that ASR-based applications can enhance learners' autonomy in improving their pronunciation skills.

Furthermore, another participant highlighted that the ability to repeat exercises as many times as needed made them feel more comfortable experimenting with different sounds, ultimately increasing their confidence and competence in pronunciation. This finding is supported by (Darsih et al., 2021), who emphasized that repeated exposure to pronunciation exercises with instant feedback enhances learners' confidence in speaking English. Additionally, participants in this study noted that the structured practice recommendations provided by the application helped them stay motivated in their pronunciation learning journey. Similarly, (Bashori et al., 2022), in their research on ASR-based applications (I Love Indonesia & NovoLearning), concluded that ASR technology significantly improves vocabulary acquisition and pronunciation. They emphasized that instant corrective feedback is a key advantage of ASR, which aligns with the results of this study—highlighting how feedback in ELSA Speak motivates users to continue practicing and boosts their confidence in speaking English.

One of the primary benefits of ASR in ELSA Speak is its ability to enhance pronunciation awareness through real-time feedback. Participants reported that they could quickly identify

their pronunciation errors and make self-corrections, which accelerated their learning process. This finding aligns with research by (Kidawara et al., 2020), which suggests that ASR offers a dynamic and interactive learning experience by enabling learners to compare their speech with native speaker models. The immediate feedback provided by the application allows learners to adjust their pronunciation in real-time, reinforcing their learning and promoting continuous improvement.

Another significant benefit of ASR technology in ELSA Speak is its flexibility, allowing learners to practice pronunciation anytime and anywhere. Participants reported that the mobile accessibility of the application enabled them to integrate pronunciation practice into their daily routines, making learning more convenient and efficient. This finding aligns with (Aswaty & Indari, 2022), who emphasized that the flexibility of ASR-based applications allows learners to practice at their own pace without being restricted by time or location constraints. The ability to access pronunciation exercises on-demand enhances learning engagement and ensures that students can reinforce their skills whenever they have available time, ultimately leading to more consistent and effective language acquisition.

Despite its benefits, this study also highlights several challenges associated with using ASR in ELSA Speak. One major challenge reported by participants is the reliance on a stable

internet connection. Many learners found that interruptions in connectivity disrupted their learning sessions, making it difficult to maintain consistent practice. This finding is consistent with (Bashori et al., 2022), who noted that technical limitations, such as the requirement for a stable internet connection, can hinder the effectiveness of ASR-based learning. Without a reliable internet connection, learners may struggle to access feedback and practice materials, which can negatively impact their learning progress.

Another challenge faced by learners is the limited access to premium features in the application. Some participants expressed frustration over the restricted content available in the free version, which limited their ability to explore more advanced pronunciation exercises. This challenge aligns with the findings of (Samad & Ismail, 2020), which indicate that while mobile ASR applications enhance pronunciation skills, restricted access to comprehensive content remains a barrier for some learners. The inability to utilize advanced features may result in slower progress and limit the effectiveness of ASR technology in supporting pronunciation learning.

Moreover, participants also reported issues with ASR's accuracy in recognizing pronunciation, particularly when speaking quickly or with different accents. While (McCrocklin, 2019) suggested that ASR technology can provide accurate feedback in diverse accent and environmental contexts, findings from this

study indicate that recognition accuracy varies among users. These inconsistencies may stem from the software's sensitivity to accent variations and differences in voice processing methods used by different ASR applications. (Dai & Wu, 2023) further highlight that although ASR is effective in detecting pronunciation errors, its ability to accurately recognize diverse accents remains a challenge, which may affect learners' confidence in its reliability.

Learners' experiences indicate that ASR in ELSA Speak is beneficial in increasing pronunciation awareness and learning motivation, although challenges such as reliance on an internet connection and accuracy in recognizing accents still exist. Moreover, access to premium features and voice recognition accuracy are also factors that can influence the effectiveness of learning. With improvements in technical aspects and accessibility, ASR has the potential to become a more optimal tool for pronunciation learning in English.

The second findings of this study explain that the use of ASR technology in ELSA Speak plays a crucial role in enhancing participants' pronunciation skills, particularly in error identification and providing accurate corrections. This is supported by Skill Acquisition Theory (DeKeyser, 2020), which asserts that technology-based learning like ASR helps learners transition from declarative to procedural stages and eventually to automatization in acquiring language skills. Participants in this study reported

feeling more confident in speaking English after using this application, which aligns with the Skill Acquisition Theory, stating that repetitive practice with accurate feedback can improve skill automatization. (Dai & Wu, 2023) found that while ASR is beneficial for pronunciation improvement, peer feedback is more effective in enhancing intonation and fluency. This does not fully align with the findings of this study, in which ASR in ELSA Speak is still considered an effective tool for improving phonetic awareness and providing instant feedback on pronunciation errors.

Additionally, this study found that ASR helps learners understand commonly mispronounced phonemes, such as the difference between /i:/ and /ɪ/, which was also confirmed in research by (Chiu et al., 2007). Thus, ASR in ELSA Speak can be considered an effective tool in helping learners overcome pronunciation difficulties commonly encountered in English language learning. However, some participants mentioned that although ASR provides accurate feedback, they still struggle to understand the corrections without additional guidance from a teacher. This indicates that ASR should be used as a supplementary tool in pronunciation learning rather than the sole learning method.

ASR technology in ELSA Speak plays a vital role in improving pronunciation skills by providing instant feedback and helping learners independently recognize and identify errors. Although effective in enhancing phonetic awareness and

pronunciation accuracy, ASR still has limitations in aspects such as intonation and speaking fluency, which tend to develop more effectively through peer interaction or teacher guidance. Therefore, ASR should be combined with other learning methods, such as direct speaking practice and instructor support, to achieve more optimal results (García et al., 2020).

Overall, the findings of this study confirm that ASR technology in ELSA Speak plays a crucial role in supporting the improvement of English pronunciation skills, particularly through instant feedback and accurate error correction. However, to optimize its effectiveness, its use should be combined with other learning methods, such as direct speaking practice and instructor guidance. While ASR is effective in enhancing phonetic awareness and pronunciation accuracy, its limitations in recognizing accents, intonation, and speaking fluency indicate that this technology cannot fully replace human interaction in language learning. Furthermore, this study provides insights into how ASR can be integrated into pronunciation teaching, as well as how learners' experiences in using this technology contribute to their increased autonomy and motivation in learning. Therefore, further development of ASR technology to be more adaptive to accent variations and improvements in the accessibility of learning features in applications like ELSA Speak could further enhance the effectiveness of pronunciation learning in the future. Additionally,

future research on combining ASR with traditional learning methods and its impact on overall speaking skill development could be a valuable direction for further exploration in this field.

CHAPTER V

CONCLUSION AND SUGGESTION

In this chapter, the researcher presents the conclusion and suggestions based on the findings and discussion in the previous chapter.

A. Conclusion

This study explores the role of Automatic Speech Recognition (ASR) in the ELSA Speak application in facilitating English pronunciation learning. Through a comprehensive investigation involving questionnaires and interviews, this research also uncovers key insights into learners' experiences, benefits, and challenges when using ASR technology, as well as its contribution to supporting pronunciation learning.

1. Regarding learners' experiences, most participants found that ELSA Speak is easy to use, engaging, and increases motivation in pronunciation learning. The instant feedback feature helps learners identify and correct their pronunciation errors independently, reinforcing autonomous learning. The flexibility of using the application enables learners to practice pronunciation at their own pace making it more accessible and personalized. However, challenges persist, such as dependence on a stable internet connection, limited access to premium features, and ASR's difficulty in recognizing certain accents,

which sometimes hinder learners from fully optimizing their learning experience.

2. ASR in ELSA Speak plays a significant role in pronunciation learning by providing real-time error detection, improving phonetic awareness, enhancing confidence, and fostering learning motivation. Participants reported that regular practice with ASR helped them refine their pronunciation accuracy and recognize commonly mispronounced phonemes. However, ASR alone is insufficient for developing fluency and intonation, which require interactive learning methods, such as direct speaking practice with a tutor or peer collaboration. Thus, the integration of ASR with traditional pronunciation training methods can lead to more comprehensive and effective learning outcomes.

B. Suggestions

Although this study has limitations in terms of time and the number of participants, the researcher strives to provide useful insights into the experiences of English department learners who have studied pronunciation skills using ASR technology. The researcher hopes that these findings will offer valuable insights for lecturers, students, and future research. Below are some recommendations from the researcher:

1. For the lecturers

Lecturers can integrate ASR technology into pronunciation learning to create a more interactive and

technology-based learning experience. With ASR, learners can receive instant feedback on pronunciation errors, making the learning process more effective. Additionally, lecturers are expected to guide learners in optimizing the use of ASR, balancing independent practice with instructor feedback. Combining ASR with traditional teaching methods, such as direct speaking practice and group discussions, is also recommended to ensure that learners not only focus on phonetic accuracy but also on suprasegmental aspects such as intonation and rhythm in speaking.

2. For the learners

Learners are encouraged to use ASR as a self-learning tool to improve pronunciation but should complement it with other methods, such as group discussions or tutor guidance. By using ASR regularly, learners can enhance pronunciation accuracy and boost their confidence in speaking English. Moreover, they should critically evaluate ASR feedback and compare it with other sources, such as online dictionaries or consultation with instructors, to ensure that pronunciation learning is not solely dependent on a single technology.

3. For future researchers

Future research is recommended to explore the integration of ASR with other learning methods, such as technology-based classrooms or peer feedback approaches, to

gain deeper insights into its effectiveness in education. Additionally, longitudinal studies on the impact of ASR on speech fluency, intonation, and learners' confidence are necessary to examine its long-term benefits. Further studies should also involve a more diverse sample, covering various proficiency levels and cultural backgrounds, to evaluate the effectiveness of ASR in different English learning contexts.

REFERENCES

- Akhmad, N. W., & Munawir, A. (2022a). Improving the Students' Pronunciation Ability by Using Elsa Speak App. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 10(1), 846–857. <https://doi.org/10.24256/ideas.v10i1.2868>
- Akhmad, N. W., & Munawir, A. (2022b). Improving the Students' pronunciation ability by using Elsa Speak app. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 10(1), 846–857. <https://doi.org/10.24256/ideas.v10i1.2868>
- Al Husna, F., Perdana, I., & Dayu, A. T. (2021). Teaching Pronunciation : An exploratory case study. *Proceeding Studium Generale*, 181–182.
- Anggraini. (2022). Improving students' pronunciation skill using Elsa Speak application. *Journal of English Language and Pedagogy*, 5(1), 135–141.
- Annury, M. N. (2013). The impact of structural competence towards speaking competence of the fourth semester students of English Department. *REGISTER JOURNAL*, 6, 213.
- Aswaty, P., & Indari, A. (2022). The effect of using ELSA (English Language Speech Assistant) Speak application on students' speaking ability for the eleventh grade of MAS Darul Al Muhajirin in the academic year 2021/2022. *Jurnal Serunai Ilmu Pendidikan*, 8(1).
- Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022a). 'Look, I can speak correctly': learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*.
- Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022b). 'Look, I can speak correctly': learning vocabulary and pronunciation

through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*.

- Bodnar, S., Cucchiarini, C., Penning de Vries, B., Strik, H., & van Hout, R. (2017). Learner affect in computerised L2 oral grammar practice with corrective feedback. *Computer Assisted Language Learning*, 30(3–4), 223–246.
- Braun, V., & Clarke, V. (2021). *Thematic Analysis : A Practical Guide* (1st ed). SAGE Publications.
- Broughton, G., Brumfit, C., Flavell, R., Hill, P., & Pincas, A. (2003). *Teaching English as a Foreign Language, Second Edition* (2nd ed.). Taylor & Francis e-Library.
- Brown, G., & Yule, G. (1984). *Teaching the Spoken Language*. Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation : A reference for teachers of English to speakers of other languages*. Cambridge University Press.
- Chiu, T. L., Liou, H. C., & Yeh, Y. (2007a). A study of web-based oral activities enhanced by Automatic Speech Recognition for EFL college learning. *Computer Assisted Language Learning*, 20(3), 209–233. <https://doi.org/10.1080/09588220701489374>
- Chiu, T. L., Liou, H. C., & Yeh, Y. (2007b). A study of web-based oral activities enhanced by Automatic Speech Recognition for EFL college learning. *Computer Assisted Language Learning*, 20(3), 209–233. <https://doi.org/10.1080/09588220701489374>
- Couper, G. (2017). Teacher cognition of pronunciation teaching: Teachers' concerns and issues. *TESOL Quarterly*, 51(4), 820–843. <https://doi.org/10.1002/tesq.354>
- Creswell, J. W. (2018). *Research Design : Qualitative, quantitative, and mixed approaches* (5th ed). SAGE Publications Ltd.

- Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a speaking task affect second language comprehensibility? *Modern Language Journal*, 99(1), 80–95. <https://doi.org/10.1111/modl.12185>
- Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: a mixed-methods study. *Computer Assisted Language Learning*, 36(5–6), 861–884. <https://doi.org/10.1080/09588221.2021.1952272>
- Darsih, E., Wihadi, M., & Hanggara, A. (2021, March 8). Using ELSA app in speaking classes: students' voices. *UNiSET 2020*. <https://doi.org/10.4108/eai.12-12-2020.2304993>
- de Vries, B. P., Cucchiarini, C., Bodnar, S., Strik, H., & van Hout, R. (2015). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning*, 28(6), 550–576. <https://doi.org/10.1080/09588221.2014.889713>
- DeKeyser, R. (2020). *Theories in second language acquisition* (B. VanPatten, G. D. Keating, & S. Wulff, Eds.; 3rd ed.). Routledge.
- Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30, 155–166.
- Dörnyei, Z. (2001). *Motivational Strategies in the Language Classroom*. Cambridge University Press.
- Dörnyei, Z., & Dewaele, J. M. (2022). *Questionnaires in second language research: Construction, administration, and processing* (3rd ed.). Routledge.
- Elliott, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in spanish pronunciation. *The Modern Language Journal*, 79(4), 530–542.
- Evers, K., & Chen, S. (2022). Effects of an automatic speech recognition system with peer feedback on pronunciation

instruction for adults. *Computer Assisted Language Learning*, 35(8), 1869–1889. <https://doi.org/10.1080/09588221.2020.1839504>

García, C., Nickolai, D., & Jones, L. (2020). Traditional versus asr-based pronunciation instruction: An empirical study. *CALICO Journal*, 37(3), 213–232. <https://doi.org/10.1558/cj.40379>

Gilakjani, A., & Ahmadi, M. R. (2011). Why is pronunciation so difficult to learn? *English Language Teaching*, 4(3), 74–81. <https://doi.org/10.5539/elt.v4n3p74>

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your “house.” *Administrative Issues Journal Education Practice and Research*, 4(2). <https://doi.org/10.5929/2014.4.2.9>

Grasella Simanjuntak, A., Tuahman Sipayung, K., Tampubolon, S., & Manik, E. (2023). Diagnosing students errors in English pronunciation by using Elsa Speak application. *Journal on Teacher Education*, 4, 762–771.

Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Longman Group Ltd.

Hornby, A. S., & Turnbull, J. (2010). *Oxford advanced learner's dictionary* (8th ed.). Oxford University Press.

Huang, X., & Deng, L. (2010). An overview of modern speech recognition. In N. Indurkha & F. J. Damerau (Eds.), *Handbook of Natural Language Processing* (2nd ed., pp. 339–363). CRC Press.

Jabareen, Y. (2009). Building a conceptual framework: philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4).

- Jaelani, A., & Sutari, D. R. (2021). Students' perception of using Duolingo application as a media in learning vocabulary. *Bogor English Student And Teacher (BEST) Conference*, 40–47.
- Kelly, G. (2000). *How to teach pronunciation* (J. Harmer, Ed.; 1st ed.). Pearson Education Limited.
- Kenworthy, J. (1987). *Teaching English pronunciation* (1st ed.). Longman Group.
- Keshet, Joseph., & Bengio, Samy. (2009). *Automatic speech and speaker recognition : Large margin and kernel method*. J. Wiley & Sons.
- Kholis, A. (2021). Elsa Speak App: Automatic speech recognition (ASR) for supplementing English pronunciation skills. *Pedagogy : Journal of English Language Teaching*, 9(1), 01. <https://doi.org/10.32332/joelt.v9i1.2723>
- Kidawara, Y., Sumita, E., & Kawai, H. (2020). *Speech to speech translation*. Springer Nature Singapore Pte Ltd. <https://doi.org/https://doi.org/10.1007/978-981-15-0595-9>
- Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, 42(8), 846–854. <https://doi.org/10.1080/0142159X.2020.1755030>
- Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*.
- Kustyarini, K. (2021). The role of language in supporting the effectiveness of solution focused therapeutic counseling. *European Journal of Literature, Language and Linguistics Studies*, 5(1), 40–45. <https://doi.org/10.46827/ejll.v5i1.255>
- Kvale, S., & Brinkmann, S. (2015). *Interviews : Learning the craft of qualitative research interviewing* (3rd ed). SAGE Publications Ltd.

- Levis, J., & Suvorov, R. (2014a). Automatic speech recognition. In *The Encyclopedia of Applied Linguistics*. Wiley. <https://doi.org/10.1002/9781405198431.wbeal0066>
- Levis, J., & Suvorov, R. (2014b). Automatic speech recognition. In *The Encyclopedia of Applied Linguistics*. Wiley. <https://doi.org/10.1002/9781405198431.wbeal0066>
- Liakin, D., Cardoso, W., & Liakina, N. (2017). Mobilizing instruction in a second-language context: Learners' perceptions of two speech technologies. *Languages*, 2(3). <https://doi.org/10.3390/languages2030011>
- Luu, L. P. T., Nguyen, T. N. Q., Vo, N. T. T., & Nguyen, M. T. H. (2021). The need of applying English learning apps to help Van Lang University students improve their spoken English performance. *AsiaCALL Online Journal*, 12(2), 72–86. <https://asiacall.info/acoj>
- McCrocklin, S. (2019a). Learners' feedback regarding ASR-based dictation practice for pronunciation learning. *CALICO Journal*, 36(2), 119–137. <https://doi.org/10.1558/cj.34738>
- McCrocklin, S. (2019b). Learners' feedback regarding ASR-based dictation practice for pronunciation learning. *CALICO Journal*, 36(2), 119–137. <https://doi.org/10.1558/cj.34738>
- Mroz, A. (2018a). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals*, 51(3), 617–637. <https://doi.org/10.1111/flan.12348>
- Mroz, A. (2018b). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals*, 51(3), 617–637. <https://doi.org/10.1111/flan.12348>
- Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for

foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393–408.
<https://doi.org/10.1080/09588220802447651>

Offerman, H. M., & Olson, D. J. (2016). Visual feedback and second language segmental production: The generalizability of pronunciation gains. *System*, 59, 45–60.
<https://doi.org/10.1016/j.system.2016.03.003>

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Education Limited.

Rismawati, D., Suryana, Y., & Agustiana, V. (n.d.). *The effectiveness of ELSA Speaking application in improving English pronunciation*. 2580–7528.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness* (K. W. Brown, J. D. Creswell, & R. M. Ryan, Eds.). The Guilford Press.

Saadah, F., & Ardi, H. (2020). The analysis of students' pronunciation error on English diphthong made by fifth semester of English Language Education Program Universitas Negeri Padang. *Journal of English Language*, 9(1).

Samad, I. S., & Ismail, I. (2020a). ELSA Speak Application as a Supporting Media in Enhancing Students' Pronunciation Skill. *MAJESTY JOURNAL*, 2(2), 1–7.
<https://doi.org/10.33487/majesty.v2i2.510>

Samad, I. S., & Ismail, I. (2020b). ELSA Speak application as a supporting media in enhancing students' pronunciation skill. *MAJESTY JOURNAL*, 2(2), 1–7.
<https://doi.org/10.33487/majesty.v2i2.510>

Saragih, E. E., Tabrani, N. P., & Muthmainnah, N. (2021). The use of digital feedback on ELSA Speak in learning pronunciation for

seventh grade of junior high school. *Journal of English Education, Linguistics and Literature*, 08(1), 136–138.

Saragih, M., Sipayung, K., & Manurung, L. W. (2023). The effect of using Cake and ELSA Speak application on students' pronunciation skills at the eighth grade of SMP Negeri 25 Medan. *Journey: Journal of English Language and Pedagogy*, 6(3), 634–647. <https://doi.org/10.33503/journey.v6i3.3728>

Senowarsito, & Ardini, S. N. (2023). The use of artificial intelligence to promote autonomous pronunciation learning indonesian. *Journal of English Language Teaching and Applied Linguistics*, 8(2), 135–144. <https://doi.org/http://dx.doi.org/10.21093/ijeltal.v8i2.1452>

Sholekhah, M. F., & Fakhurriana, R. (2023). The use of ELSA Speak as a mobile-assisted language learning (MALL) towards EFL students Pronunciation. *JELITA: Journal of Education, Language Innovation, and Applied Linguistics*, 2(2), 93–100. <https://doi.org/10.37058/jelita.v2i2.7596>

Sudarmaji, I., & Yusuf, D. (2021). The effect of Minecraft video game on students' English vocabulary mastery. *JETAL : Journal of English Teaching & Applied Linguistics*, 3(1), 30–38.

Sugiyono. (2020). *Metode Penelitian Kuantitatif, Kualitatif dan R & D*.

Suryatiningsih, N. (2015). A Study on the Students' Ability in Pronouncing Diphthongs at STKIP PGRI Pasuruan. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 3(2).

van Doremalen, J., Boves, L., Colpaert, J., Cucchiaroni, C., & Strik, H. (2016). Evaluating automatic speech recognition-based language learning systems: a case study. *Computer Assisted Language Learning*, 29(4). <https://doi.org/10.1080/09588221.2016.1167090>

- Wallace, L. (2016a). Using Google Web Speech as a springboard for identifying personal pronunciation problems. In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds.), ;*Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference* (Vol. 7, pp. 180–181). Ames, IA: Iowa State University.
- Wallace, L. (2016b). Using Google Web Speech as a springboard for identifying personal pronunciation problems. In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds.), ;*Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference* (Vol. 7, pp. 180–181). Ames, IA: Iowa State University.
- Yiing, I. K. C. (2011). *An analysis of pronunciation errors in English of six UTAR Chinese studies undergraduates*. Tunku Abdul Rahman.
- Yousif, M. M., & Ameen, A. A. (2018). Analysing the improper pronunciation of diphthongs by Iraqi EFL learners. *JURNAL ARBITRER*, 5(1), 17–22. <https://doi.org/10.25077/ar.5.1.17-22.2018>
- Yu, D., & Deng, L. (2015). *Signals and communication technology Automatic Speech Recognition - A deep learning approach*. Springer- Verlag London. https://doi.org/10.1007/978-1-4471-5779-3_1
- Zakiyyah, F., Setyaji, A., & Ardini, S. N. (2022). The analysis of pronunciation application based on the concept of artificial intelligence. *Undergraduate Conference on Language, Literature, and Culture (UNCLLE)*, 2(1). <http://publikasi.dinus.ac.id/index.php/uncle>

APPENDICES

A. Appendix 1 Learners Questionnaire

Adapted from (Jaelani & Sutari, 2021)

English Learner Experience Questionnaire on Pronunciation Improvement Using Elsa Speak Application

Name :

Age :

Semester :

Instructions for Completion

1. Read all the statements carefully and thoroughly.
2. Choose one criterion that best represents your opinion on one of the score criteria.
3. Explanation of score criteria :
SD : Strongly Disagree
D : Disagree
N : Neutral
A : Agree
SA : Strongly Agree

NO.	Statement	SD	D	N	A	SA
1.	ELSA Speak is easy to use for practicing English pronunciation					
2.	ELSA Speak provides clear feedback and helps me identify my pronunciation					
3.	I can use ELSA Speak anytime and anywhere to improve my pronunciation					
4.	The pronunciation exercises in ELSA Speak are interesting and motivate					
5.	The feedback from ELSA Speak helps me improve my pronunciation effectively					
6.	ELSA Speak helps me understand the correct way to pronounce English					

7.	I feel that my pronunciation has not improved by using ELSA Speak					
8.	ELSA Speak helps me be more confident in speaking English with correct pronunciation					
9.	Using ELSA Speak makes me more active in practicing pronunciation outside of class					
10.	Learning pronunciation with ELSA Speak deepens my understanding of English pronunciation Using ELSA Speak makes me more active in practicing pronunciation outside of class					
11.	My experience with ELSA Speak helps me apply pronunciation skills in conversations					

B. Appendix 2 Interview Questions

The researcher adapted the interview protocol from (Mroz, 2018), focusing on students' experiences, perceptions, and suggestions. In the first part, they were asked to recount their experiences using the ASR feature in the ELSA Speak app. In the second part, they described the advantages and disadvantages of the feedback configurations. In the final part, they offered suggestions for improving pronunciation through automatic speech recognition. The interviews were conducted in Indonesian and lasted about 30 minutes

- a. How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
- b. What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
- c. What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
- d. Do you feel that your English pronunciation has improved or become much better after using this technology?
- e. How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
- f. Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?

C. Appendix 3 Result of Questionnaire

No	Statement	Alternative answer					Total
		SA	A	N	D	SD	
1.	ELSA Speak is easy to use for practicing English pronunciation	4	4	1	0	0	9
		44,4%	44,4%	11,1%	0%	0%	100%
2.	ELSA Speak provides clear feedback and helps me identify my pronunciation	2	6	1	0	0	9
		22,2%	66,7%	11,1%	0%	0%	100%
3.	I can use ELSA Speak anytime and anywhere to improve my pronunciation	2	7	0	0	0	9
		22,2%	77,8%	0%	0%	0%	100%
4.	The pronunciation exercises in ELSA Speak are interesting and motivate	1	5	3	0	0	9
		11,1%	55,6%	33,3%	0%	0%	100%
5.	The feedback from ELSA Speak helps me improve my pronunciation effectively	3	5	1	0	0	9
		33,3%	55,6%	11,1%	0%	0%	100%
6.	ELSA Speak helps me	2	6	1	0	0	9
		22,2%	66,7%	11,1%	0%	0%	100%

	understand the correct way to pronounce English						
7.	I feel that my pronunciation has not improved by using ELSA Speak	0	0	1	7	1	9
		0%	0%	11,1%	77,8 %	11,1 %	100%
8.	ELSA Speak helps me be more confident in speaking English with correct pronunciation	1	5	3	0	0	9
		11,1%	55,6%	33,3%	0%	0%	100%
9.	Learning pronunciation with ELSA Speak deepens my understanding of English pronunciation Using ELSA Speak makes me more active in practicing pronunciation outside of class	0	4	5	0	0	9
		0%	44,4%	55,6%	0%	0%	100%
10.	Learning pronunciation with ELSA Speak deepens my	2	5	2	0	0	9
		22,2%	55,6%	22,2%	0%	0%	100%

	understanding of English pronunciation Using ELSA Speak makes me more active in practicing pronunciation outside of class						
11.	My experience with ELSA Speak helps me apply pronunciation skills in conversations	0	6	3	0	0	9
		0%	66,7%	33,3%	0%	0%	100%

D. Appendix 4 Result of Interviews

Theme:

- **Immediate Feedback (IF):** ELSA Speak provides instant feedback to help users correct pronunciation errors immediately.
- **Competence Development (CD):** Users enhance pronunciation, stress patterns, and intonation through guided practice.
- **Sensitivity Issues (SI) :** The app may misinterpret accents or background noise, leading to incorrect feedback.
- **Skill Refinement (SR) :** Repeated practice improves pronunciation accuracy, fluency, and natural speech.
- **Autonomous Motivation (AM):** Users stay engaged by managing their own learning pace and progress.
- own learning pace and progress.

Participant	Speaker	Interview Transcription
P1	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it? .
	P1-IF	I feel highly motivated to learn English pronunciation using ELSA Speak. This application is very engaging; it feels like having a personal tutor who listens carefully and provides immediate feedback. With real-time feedback, I can instantly recognize my mistakes and correct them, which makes me feel more competent and confident when speaking. Additionally, its ease of use and interactivity make English pronunciation practice enjoyable and motivating.
	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?

	P1-SR	The advantages of using technology include instant feedback, directed practice, and consistency. For example I have trouble distinguishing between the pronunciation of 'Sheep' and 'Ship', and this app repeatedly points out my mistakes and guides me toward the correct pronunciation.
P1	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P1-SI	The challenges of using this technology in the application include sensitivity issues and limited context. For example, in one session, I was marked incorrect for pronouncing the word “water” with a British accent, even though the expected pronunciation was in an American accent.
P1	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P1-CD	Yes, my pronunciation has improved. This application has made me more aware of stress patterns, intonation, and sound articulation. Now, I can pronounce words like “word” and “rural” more clearly. By analyzing my speech, the application provides detailed phonetic comparisons and customized exercises for specific sounds.
P1	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P1-IF	The ASR technology in ELSA Speak has helped me overcome issues such as inconsistent vowel sounds and incorrect stress placement. This technology serves as a guide, providing step-by-step corrections and positive reinforcement.

P1	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P1-AM	I would recommend this as a learning tool for interactive feedback. Its personalized learning path and user-friendly interface make it a powerful aid for improving English pronunciation. However, users should complement it with other methods, such as conversation practice, to overcome its limitations.
P2	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P2-IF,AM	I feel that the ASR technology in the ELSA Speak application is very helpful. The application listens to my pronunciation and provides immediate feedback, allowing me to identify my mistakes and work on improving them. finally, I feel more confident in my pronunciation.
P2	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P2-SR	One major advantage is that I can practice speaking anytime. For example, when I use it to practice difficult words, I receive instant feedback on how well I pronounce them.
P2	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P2-SI	Sometimes, this technology does not recognize my voice well, especially when I am nervous or

		speak too quickly. There was a time when the app said I pronounced a word incorrectly, even though I felt I had pronounced it correctly.
P2	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P2-CD	Yes, I believe my pronunciation has improved after using ELSA Speak. I think this application is very helpful as it accurately identifies my mistakes. It also motivates me to practice more.
P2	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P2-IF	This ASR technology helps me identify weaknesses in my pronunciation. It feels like having a personal tutor who provides instant corrections.
P2	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P2-AM	Yes, I would recommend using ELSA Speak. It is a useful tool for learning pronunciation easily and enjoyably.
P3	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P3-IF,AM	Honestly, I think this app is pretty cool because it can carefully listen to my pronunciation and provide feedback. With this, ELSA Speak gives me control over my learning process. I can practice on my own without having to wait for corrections from others, making me more independent in improving my pronunciation. With instant feedback, I feel more confident to try speaking English in real-life situations.

P3	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P3-SR	The advantage I get from this application is the instant feedback. I don't have to wait for someone to correct me, making practice more efficient. Once, when I practiced words like "rural" and "squirrel," they felt quite tricky.
P3	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P3-SI	The drawback is that the application sometimes struggles to recognize context accurately. For example, in noisy environments or when speaking with a regional accent, the feedback may be inaccurate. I once practiced in a café, and the app kept indicating my pronunciation was incorrect when, in reality, it was just picking up background noise.
P3	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P3-CD	Yes, I think so. I believe my pronunciation has improved, especially in individual sounds and stress patterns. This app points out small mistakes that I might not have noticed myself. It has been very helpful in stress and intonation. The app breaks down sentences or words and highlights the parts I need to focus on. For example, the app marks syllables that should be stressed, and after several attempts, I can pronounce them correctly. It feels like having someone constantly reminding me to speak more slowly and clearly.

P3	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P3-IF	ELSA Speak feels like having a coach available at all times. ASR technology highlights areas for improvement without making me feel judged. For me, this app is very helpful in pronouncing difficult sounds like "t" or distinguishing between words like "sheep" and "ship."
P3	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P3-AM	Yes, I recommend it, but with a note. Use this application as a learning tool, not something to rely on completely. This app is great for practicing and improving your pronunciation, but real conversations are where you truly grow. It is very useful for building confidence before stepping into the real world.
P4	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P4-IF,AM	My experience using the ELSA Speak application, particularly its Automatic Speech Recognition (ASR) feature, has been very positive. The app evaluates my pronunciation in real-time and provides specific feedback on sounds, word stress, and intonation. This feature is highly beneficial in quickly identifying and correcting my mistakes.
P4	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?

	P4-SR	Instant feedback allows me to correct pronunciation mistakes immediately. Additionally, it offers personalized learning by focusing on the most challenging areas for me.
P4	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P4-SI	The challenge I faced while using this app is its limited understanding of context. The app focuses on pronunciation but does not provide much help with rhythm or conversational fluency.
P4	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P4-CD	I feel an improvement in my pronunciation when using ELSA Speak, especially in reducing errors in vowel and consonant articulation. For example, repeated practice on a word is very helpful.
P4	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P4-IF	ASR technology is very helpful in overcoming pronunciation challenges. The app can detect small mistakes that are often overlooked and provide correct guidance to fix them.
P4	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P4-AM	I highly recommend using the ELSA Speak app for anyone learning English pronunciation. Its accessibility, affordability, and detailed feedback make it a valuable tool. However, it should still be combined with human feedback to further enhance the learning experience.

P5	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P5-AM	The application makes it easier for me to improve my English pronunciation. For example, in distinguishing the pronunciation of "Stationery" and "Stationary." With this feedback, I feel more motivated to keep practicing and more confident in my English pronunciation skills.
P5	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P5-SR	The ASR technology in the ELSA Speak application provides instant feedback by identifying errors and offering precise corrections in pronunciation.
P5	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P5-SI	The application is difficult to access without mobile data.
P5	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P5-CD	Yes, it is very helpful for me because, through this application, I can identify my mistakes in English pronunciation. The corrections provided make me more careful in my pronunciation, as mispronouncing a word can change its meaning.
P5	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?

	P5-IF	ASR technology provides effective feedback for improving word pronunciation and makes the English learning process more efficient and structured.
P5	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P5-AM	I highly recommend this app for beginners learning English, especially for pronunciation skills. This app helps learners practice the pronunciation of words first before applying them in conversation to avoid misunderstandings.
P6	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P6-IF	In my opinion, using ASR technology in the ELSA Speak application is very interesting and effective. It allows me to practice English pronunciation anytime and receive instant feedback. Overall, my experience has been positive, although sometimes the app misunderstands my speech, especially with certain accents or unclear audio.
P6	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P6-IF,SR	Provides instant and detailed feedback on pronunciation accuracy. For example, when I struggled to pronounce the word "comfortable," the app demonstrated the correct sound and what needed improvement.
P6	R	What disadvantages have you encountered in using this technology? Could you provide an

		example of a situation that occurred while using it in the ELSA Speak application?
	P6-SI	Less sensitive to context or accents. For example, when I pronounce "can" with a British accent, the app sometimes interprets it as "can't," leading to inaccurate and frustrating feedback.
P6	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P6-CD	Yes, I feel that my pronunciation has improved. This app helps me recognize phonetics and correct mistakes in my speech. Now, I can pronounce phenomenon and architecture, which used to be difficult for me.
P6	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P6-IF	This technology plays a crucial role in identifying and correcting mistakes that I wouldn't notice on my own. I compare it to a mirror for my pronunciation.
P6	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P6-AM	Yes, I recommend it because this app is a practical, affordable, and effective tool for supporting English pronunciation learning. However, I also suggest that users combine it with other features or methods.
P7	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P7-IF,AM	I feel that using Automatic Speech Recognition (ASR) technology in the ELSA Speak app is very helpful and engaging. The app provides instant

		feedback on my pronunciation, allowing me to understand the mistakes I make. It is also easy to use and motivates me to keep improving my pronunciation skills. Moreover, the app tracks my pronunciation progress over time, which boosts my confidence.
P7	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P7-SR	One of the advantages is that I can practice English pronunciation anytime and anywhere. For example, when preparing for a speaking test, this app highlights specific sounds that I need to improve, such as the 'th' sound, and provides guidance on how to correct it.
P7	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P7-SI	The downside is that some topics can only be accessed if the user subscribes. Therefore, those who haven't subscribed can only access limited content.
P7	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P7-CD	Yes, my English pronunciation has improved after using this app. It helps me by breaking words into syllables and providing step-by-step guidance on how to pronounce them. It works by detecting mistakes and offering suggestions for correction, making it easier for me to improve gradually.
P7	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?

	P7-SR,IF	In my opinion, the role of ASR technology in ELSA Speak is crucial in overcoming pronunciation challenges. It identifies my weaknesses and helps me address them one by one, building my confidence in speaking English.
P7	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P7-IF,AM	Yes, I would recommend ASR technology in the ELSA Speak app. It is very practical for improving pronunciation, especially for learners who do not have access to native speakers.
P8	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P8-IF,AM	In my opinion, the use of ASR technology in the ELSA Speak app is very helpful in learning English, especially in pronunciation skills. The app provides instant feedback, which is crucial for users to improve their language abilities. My personal experience with this app has been very positive, as it gives me the freedom to practice according to my needs, making me feel more independent in learning. I can also repeat exercises as many times as I want until I feel confident in my pronunciation. As a result, I feel more confident when speaking in English because my pronunciation skills have improved.
P8	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P8-SR	The benefit I have experienced is an improvement in my pronunciation accuracy. For example, when I practice certain phrases, the application provides

		feedback on how to correct my mistakes. This helps me immediately understand what went wrong and how to fix it.
P8	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P8-SI	The disadvantage I experience when using this application is that, often in noisy environments, the ASR technology struggles to detect voices clearly, resulting in inaccurate feedback.
P8	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P8-CD	Yes, I feel that my pronunciation has improved after using the ELSA Speak application. It has also changed my perspective on the importance of paying attention to correct, precise, and clear pronunciation. Good pronunciation plays a crucial role in effective communication, making me feel more confident to speak English in front of others.
P8	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P8-IF	ASR technology is very helpful in overcoming pronunciation difficulties. This technology allows me to immediately identify the areas that need improvement, making it easier to understand the correct pronunciation.
P8	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P8-AM	I would recommend the ELSA Speak application. The reason is that this app offers an interactive

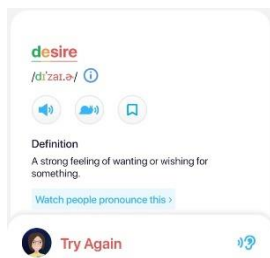
		and effective learning method, provides useful feedback, and helps users learn faster and become more confident in speaking English.
P9	R	How do you feel about using automatic speech recognition technology in the ELSA Speak application for your English learning? Could you describe your experience using it?
	P9-IF,AM	I feel greatly assisted by the automatic speech recognition technology in ELSA Speak. My experience shows that this application can provide highly specific feedback on my pronunciation. With this feature, I feel more confident in my pronunciation skills because I know I have been well-trained.
P9	R	What advantages have you found in using automatic speech recognition technology to study English? Can you provide an example of a situation that occurred while using this technology in the ELSA Speak application?
	P9-SR	The main advantage of this application is its ability to analyze pronunciation details down to the phoneme level, allowing me to correct minor mistakes.
P9	R	What disadvantages have you encountered in using this technology? Could you provide an example of a situation that occurred while using it in the ELSA Speak application?
	P9-SI	The drawback is that sometimes the application struggles to understand certain accents, making its analysis results not entirely accurate.
P9	R	Do you feel that your English pronunciation has improved or become much better after using this technology?
	P9-CD,SR	Yes, my pronunciation has improved significantly. This application helps me focus on words I often mispronounce, provides accuracy

		scores, and offers specialized training to correct those areas.
P9	R	How do you perceive the role of ASR technology in ELSA Speak in helping you overcome obstacles in English pronunciation?
	P9-IF	The ASR technology in ELSA Speak acts as a virtual coach that provides instant feedback. This is very helpful in identifying small mistakes that I might not have noticed before, such as intonation or word stress.
P9	R	Would you recommend the use of ASR technology in the ELSA Speak application as a tool to aid in our English pronunciation learning? Why or why not?
	P9-AM	I would definitely recommend this application because it is highly effective and easy to use. ASR technology provides a personalized and efficient learning experience, especially for those who want to improve their pronunciation independently.

E. Appendix 5 Interview Documentation

Participant 1 	Participant 2 	Participant 3 
Participant 4 	Participant 5 	Participant 6 
Participant 7 	Participant 8 	Participant 9 

F. Appendix 6 ASR Technology in ELSA Speak

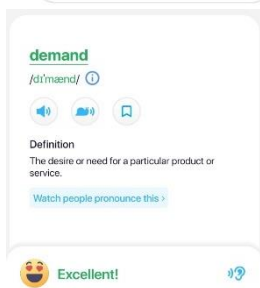


You sound 23% like a native speaker. 23%

Tap on each word for detailed feedback

Try Again

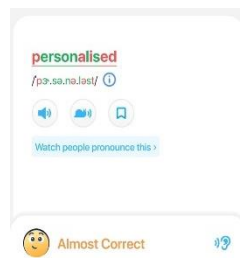
Try new word



You sound 96% like a native speaker. 96%

Try Again

Try new word



You sound 60% like a native speaker. 60%

Tap on each word for detailed feedback

Try Again

Try new word



G. Appendix 7 Advisor Appointment Letter



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI WALISONGO
FAKULTAS ILMU TARBIYAH DAN KEGURUAN

Sekretaris: Jl. Prof. Dr. Hamka Kampus II Ngaliyan Semarang Telp. 7601295 Fax. 7615387 Semarang 50185

Semarang, 06 Juni 2024

Nomor : 1516/Un.10.3/J4/DA.04/06/2024
Lamp : -
Perihal : Penunjukan Pembimbing Skripsi

Kepada Yth.

Dr. Muhammad Nafi Annury, M.Pd.

Assalamu 'alaikum Wr. Wb.

Berdasarkan hasil pembahasan judul penelitian di Jurusan Pendidikan Bahasa Inggris (PBI), maka Fakultas Ilmu Tarbiyah dan Keguruan menyetujui judul skripsi mahasiswa:

Nama : Nabela Adelia Kusuma Ningsih
NIM : 2103046046
Judul : The Role of Automatic Speech Recognition Technology in
Bridging The Pronunciation Gap for English Language Learner

Dan menunjuk saudara Dr. Muhammad Nafi Annury, M.Pd. sebagai pembimbing.

Demikian penunjukan pembimbing skripsi ini, dan atas kerjasamanya diucapkan terima kasih.

Wassalamu 'alaikum Wr. Wb.



Dr. Dekan

Jurusan Pendidikan Bahasa Inggris

Dr. Nuna Mustikawati Dewi, M.Pd.
NIP. 196506141992032001

Tembusan:

1. Dosen Pembimbing
2. Mahasiswa yang bersangkutan
3. Jurusan Pendidikan Bahasa Inggris

CURRICULUM VITAE

A. Personal Data

Name : Nabela Adelia Kusuma Ningsih
Place and Date of Birth : Trenggalek, 07 December 2003
Student Number : 2103046046
Major : English Education Department
Address : Prigi, Watulimo, Trenggalek
Phone : 081230960801
E-mail : nabelaakningsih73@gmail.com

B. Educational Background

1. RA MIN Prigi
2. MIN 1 Trenggalek
3. MTsN 4 Trenggalek
4. MAN 1 Trenggalek

Semarang, 24 February 2025
The Researcher



Nabela Adelia Kusuma N.
NIM.2103046046