

**Examining the Impact of AI-Assisted Language Learning Tools on Second Language
Production**

Dissertation Manuscript

Submitted to National University

School of Education

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF EDUCATION

by

NEGUIN SARRAF CALDWELL

San Diego, California

February 2025

Abstract

Second-language learners frequently encounter challenges in speaking proficiency, including pronunciation, grammar, vocabulary, cognitive fluency, anxiety, and cultural nuances. Some students at the Defense Language Institute (DLI) face challenges with confidence and motivation when speaking the target language, which can hinder their development of speaking skills and overall language proficiency. While traditional methods might fall short, emerging technologies like AI-assisted language learning tools offer promising solutions to encourage spontaneous language use. The purpose of this qualitative case study was to explore the impact of AI-assisted language learning tools on DLI students' motivation and confidence in speaking the target language. It drew on theories such as flow theory, mindset Theory, and the Communicative Language Teaching approach. The study surveyed students at the Language Center (LC) School of Language, located in the southeastern region of the United States, during the 19-week language course iteration. The participants were purposefully sampled, consisting of eight DLI linguists who were also language class students. The research method involved surveys and observation sessions where participants individually interacted with the TalkPal AI tool for speaking the Farsi language. The data analysis utilized an inductive and thematic data analysis method without any analysis software. Surveys and observation methods revealed that AI tools can reduce communication apprehension and boost learner confidence, especially for anxious learners. Future studies could investigate the effectiveness of these tools in various settings, utilize advancements in artificial intelligence, and assess their broader influence on language learning to incorporate them into the educational framework better. However, these tools must evolve to facilitate increasingly natural conversations, thus enhancing learner engagement and

motivation. Therefore, realistic interaction and tool performance limitations currently suggest that AI tools are best used to supplement traditional methods, not replacements.

Keywords: Language Learners, Speaking Anxiety, Confidence, Motivation, AI-assisted Language Learning Tools, TalkPal AI

Acknowledgment

I would like to express my sincere gratitude to all those who have contributed to completing this dissertation. First, I would like to thank my chair, Dr. Patricia Davis-Kennedy, for her invaluable guidance and support and for providing me with the resources I needed to delve deeper into my research. Her insights and encouragement have been instrumental in shaping this work. I would also like to extend my sincere thanks to my subject matter expert, Dr. Elaine Lux-Koman, for her knowledge and expertise in the field. Her guidance and support have significantly enriched this research. I am also indebted to my academic reader, Dr. Joanna Vance, for her careful review and constructive feedback. Her expertise has helped strengthen the quality of this work. I would also like to thank the Learning Center's stakeholders for their constant support and assistance in completing this work. Lastly, I would like to express my love and gratitude to my husband, David A. Caldwell. His support, patience, and understanding have been my anchor throughout this process. His love and belief in me have been my constant source of strength. This dissertation would not have been possible without all of you.

Table of Contents

| | |
|---|----|
| Section 1: Foundation | 1 |
| Statement of the Problem | 4 |
| Purpose of the Study | 6 |
| Research Questions | 7 |
| Conceptual/Theoretical Framework..... | 7 |
| Definitions of Key Terms | 10 |
| Review of the Literature | 13 |
| Theoretical Frameworks for Overcoming Speaking Anxiety and Promoting Language Learners' Confidence | 16 |
| The Impact of AI & VR Tools on Learning Outcomes | 20 |
| Critical Analysis and Limitations and Gaps in the Literature..... | 24 |
| Ethical Assurances | 27 |
| Summary | 28 |
| Section 2: Methodology and Design..... | 32 |
| Design and Method | 34 |
| Population and Sample | 39 |
| Data Collection and Analysis..... | 43 |
| Assumptions..... | 47 |
| Limitations | 47 |
| Delimitations..... | 50 |
| Summary | 51 |
| Section 3: Findings, Implications, and Recommendations | 54 |
| Findings..... | 55 |
| Trustworthiness of the Data | 57 |
| Credibility | 58 |
| Confirmability..... | 59 |
| Dependability | 60 |
| Transferability..... | 61 |
| Comfort and Usage | 67 |
| Motivation and Engagement..... | 68 |
| Features and Functionalities | 69 |
| Confidence and Anxiety | 70 |
| Strengths and Weaknesses | 71 |
| Comparison with Other Tools..... | 72 |
| Technology Skills & Experience | 72 |
| Achieving Goals..... | 73 |
| Evaluation of the Outcomes..... | 80 |
| Implications and Recommendations for Practice | 82 |
| Recommendations for Future Research | 83 |

Conclusions..... 85

References..... 88

Appendix A IRB Approval Letter..... 98

Appendix B Consent Form 99

Appendix C Observation List 101

Appendix D Survey/Questionnaire Form 106

Section 1: Foundation

Second language learners often face challenges when speaking the target language. Foreign language speaking is a constant struggle that second language learners face until they become fluent. This lack of proficiency can negatively impact the learner's academic success and career prospects in today's increasingly interconnected world, where strong oral communication skills are highly valued (El Majidi et al., 2021).

The students of second language learning in the Defense Language Institute (DLI) Extension Program (EP) are DLI's client partners. Their positions demand advanced language abilities to accomplish certain job objectives, including serving as civilian cryptologic linguists or supporting intelligence activities. Therefore, learners in the Defense Language Institute extension program must have advanced language abilities and a final course grade of at least 77% to produce accurate transcriptions and translations. Most learners are highly educated and have a background in the target language, enrolling in the Extension Program to preserve and advance their language skills. However, the speaking portion of learning a second language presents an ongoing challenge for DLI language learners. As Attia et al. (2021) described, speaking is one of the most challenging components of language learning in a Department of Defense (DoD) environment, and many DoD linguists need help acquiring this skill.

Communicating in a foreign language is a challenge (Al-Wossabli, 2016). One of the biggest challenges for second language learners is mastering the pronunciation of the target language. It may be challenging for second language (L2) learners to generate sounds like native speakers because of differences in sounds, intonation, and stress patterns, particularly regarding the study materials and practical aspects. Tonglo (2022) emphasized the importance of learning

through pronunciation for students to achieve maximum results. Formal learning is ineffective, so students should practice through direct practice in various activities, both inside and outside school, including recreational activities. However, according to Tonglo et al. (2022), students perceive learning pronunciation as challenging and time-consuming, arguing that it should be taught earlier in school to improve skills and be more effective. According to research on speaking effectiveness conducted at Jazan University in Saudi Arabia, EFL students showed a significant lack of competence in pronunciation skills (Al-Wossabli, 2016). In addition to finding that it is challenging to produce sounds, intonation, and stress patterns in the target language, second language learners frequently struggle with learning grammar rules, a large variety of terms, and idiomatic expressions. Kartal (2019) examined the linguistic difficulties experienced by students learning Turkish as a foreign language in French universities. The study indicated the first category of challenges related to suffixes, grammar, and syntax, with rates close to 40%. The second group encompasses pronunciation and consonance, verb tenses, and words, which were identified as challenges by approximately 20% of the students involved.

Second language learners often need extra time to mentally process and translate their thoughts from their native to the target language. As a result, they may struggle to find the right words or phrases quickly while speaking. This delay causes interruptions or gaps in their speech and can impede their ability to participate in fast-paced conversations or respond spontaneously. When someone speaks a second language, their processing, speed, and flexibility are called cognitive fluency (Sun, 2023). According to the research conducted by Sun (2023), one area of study on L2 speaking is the effect of cognitive processes on the production or delivery of L2 speech. Processing the language demands more concentration from language learners because their language knowledge has yet to be automated (Al-Wossabli, 2016). This delay in processing

and producing speech might cause a lack of confidence, discomfort, and anxiety when speaking in front of classmates or native-speaking professors. Mir Muhammad Ahmedani et al. (2022) suggested that learners frequently avoid speaking in second languages out of fear of making blunders. When the teachers question them, they are unsure of themselves and feel tense. Anxiety perhaps has a significant impact on second language speaking abilities. According to Teimouri et al. (2019), anxiety has been repeatedly shown to be negatively connected with L2 achievement and performance. The findings at Sindh University also revealed internal and external issues like anxiety, low motivation, and a negative attitude toward the English language, which led to a reluctance to speak the L2 (Mir Muhammad Ahmedani et al., 2022). Motivation is a critical component of successful language learning, as it is the expectation of reward from both external and internal sources. While extrinsic incentives come from the outside and influence learners physically, intrinsic motivations influence learners within. For instance, when a student gets good grades and receives a present from the instructor, they are motivated physically. If a student achieves good test results and receives compliments and acknowledgment from a teacher, this is referred to as having intrinsic motivation (Mir Muhammad Ahmedani et al., 2022).

Cultural differences also add to the linguistic challenges, as language is deeply intertwined with culture. Second language learners may encounter difficulties understanding cultural references, idioms, humor, or social norms embedded in the target language. These differences can lead to misunderstandings or inappropriate language use. Sun (2023) demonstrated that sociocultural factors, including attitudes toward the target language and culture, influenced advanced L2 Chinese speech performance. Consistent practice, exposure to authentic materials, and cultural immersion help overcome these difficulties. However,

traditional approaches and untrained teachers are obstacles, as is the atmosphere for practice, and sometimes, the syllabus needs to meet the student's needs or be motivating enough.

Leaver and Atwell (2002) investigated the traits of students with native-like proficiency. They questioned language learners about what had helped them and what had not throughout their language learning careers. The vast majority (82%) acknowledged the significance of instrumental motivation (Leaver & Atwell, 2002). The motives that make language learning smoother and more enjoyable might be categorized as technology. Students were often motivated to learn a skill to help them complete a particular task. Technologies such as Artificial Intelligence (AI) or Virtual Reality (VR) can directly substitute for functional improvement and allow for significant task redesign.

Statement of the Problem

The problem is that students in the Extension Program (EP) of the DLI lack enough confidence and motivation to speak the target language, which hinders their ability to develop speaking skills (Attia et al., 2021). Much research has been done to determine whether cutting-edge technologies like AI, VR, or Augmented Reality (AR) could aid in the development of language production. However, the speed at which these tools evolved prevented proper documentation of their impact on language learning and production. Therefore, their effectiveness and efficiency on language learners and how students perceived the utility of these tools in language learning remained unexplored to some extent.

Linguists from the Department of Defense (DoD), including both military and civilian personnel, are reluctant to speak the target language during language improvement classes offered by the National Cryptologic School (NCS) and DLI (Attia et al., 2021). According to Mir

Muhammad Ahmedani et al. (2022), this reluctance could have been due to linguistic, cognitive, psychological, and sociocultural factors, teaching methods, and material problems.

Speaking hours are mainly used to practice oral proficiency interview (OPI) modules and focus on specific topics to pass the OPI exam. However, fear of making mistakes and criticism could lead to a lack of confidence and motivation to speak the language (Habbiurrahim, 2020). Stakeholders called for investigating this phenomenon (Attia et al., 2021). The challenge for language centers was to find ways to motivate students to speak the target language during classroom activities (Attia et al., 2021).

The conventional methods are more teacher-centered than student-centered, don't motivate students to convey their ideas, and do not reflect the natural flow of speaking in everyday interactions (Lockiewicz, 2019). While the positive effects of initiatives like one-on-one speaking sessions, peer-to-peer practice, and conversation groups were evident, my analysis suggested they might not have been enough to fully address the complex issue of DLI students' lack of confidence and motivation in speaking the L2. As a result, it was crucial to explore a supplementary method such as AI-powered tools or Immersive VR, whose ability to simulate real-life scenarios and provide immediate feedback presented an opportunity to foster confidence and encourage spontaneous language production in a safe and engaging environment (Salido, 2023). One potential solution is immersive VR, which has been used to treat people with a phobia, social anxiety, or cognitive or social skills (Bisso et al., 2020). Exposing students to different public virtual places and avatars might decrease their apprehension about speaking in public.

As a result, there is a need for further studies in this field. According to Attia et al. (2021), future studies in the DoD could replicate or expand on this study to explore the role of

VR as a motivating factor for verbal language production. Huang et al. (2023) suggested that the potential of AI to enhance teaching and learning outcomes and apply it to language courses should be the subject of future research. Finally, Harris (2022) reinforced the idea that VR is a promising tool for language learning, but further study is needed to maximize its potential.

Purpose of the Study

The purpose of this qualitative case study was to explore the impact of AI-assisted learning tools on DLI students' motivation and confidence in speaking the target language. DLI Extension Program (EP) linguists faced challenges speaking L2 due to limited speaking practice opportunities and speech discomfort. This hindered their language proficiency and ability to communicate effectively in the target language. Current methods, which focused on reading and listening for the Defense Language Proficiency Test (DLPT), did not adequately address these challenges (Attia et al., 2021).

Students need to be adequately motivated to speak or write, despite the necessity of utilizing the target language in class (Attia et al., 2021). Meanwhile, barriers such as speech anxiety lessened students' intrinsic incentive to use or produce the target language (Sun & Teng, 2021). On the other hand, with an increasing collection of novel tools to improve L2 speaking, it became critical to consider how these tools currently impacted language learners; in other words, how students reacted after exposure to newer technologies. To address this gap, research should investigate the potential benefits and drawbacks of the most recent technologies in language learning and production. This should involve examining student experiences with these technologies to gain insight into their perceptions, challenges, and successes. For data collection, an open-ended questionnaire was used, along with class observation notes. The participants were purposefully sampled, consisting of eight DLI linguists who were also language class students.

The research was conducted at the Learning Center (LC) School of Language, located in the southeastern region of the United States. The LC is a continuing education (CE) school that offers 19-week courses and refresher courses to DoD linguists and civilians after graduation from the Defense Language Institute (DLI). The data analysis utilized an inductive and thematic data analysis method. This approach involved analyzing the data to see what it was saying (analysis) and seeking to understand what it meant (interpretation) (Bloomberg & Volpe, 2019). At this stage in the project, the L2 students' confidence and motivation were generally defined as a willingness to communicate in a language with belief in their language skills, allowing them to express themselves fluently, engage in conversations, and overcome any hesitations or anxieties related to speaking the second language (Alam et al., 2021).

Research Questions

The following research questions guided this study:

RQ1: How do DLI students perceive the impact of AI-supported language learning tools on their motivation to engage in language production activities?

RQ2: How do DLI students describe their experience with AI-supported language learning tools in terms of lowering their anxiety and enhancing confidence when speaking the target language?

Conceptual/Theoretical Framework

This qualitative study explores the interplay between motivation, reducing anxiety, and second-language communication, with a particular focus on how AI-assisted language learning can support these areas. The framework integrates these elements to provide a comprehensive understanding of the factors influencing language production.

The initial theoretical framework that enabled me to comprehend the factors influencing language students' motivation to speak in the target language was communicative language teaching (CLT). CLT is an approach that strongly emphasizes the value of meaningful communication in language learning. It encourages using genuine, real-world tasks and interactions to help students improve their speaking skills.

Students should practice communicating with their teacher, engaging with their classmates, and participating in scenarios that reflect situations they might encounter in their future careers (Simion, 2022). The idea originated from Noam Chomsky's theory, which emphasized the distinction between competence and performance and was later expanded by Michael Halliday and Dell Hymes to include the concept of communicative competence (Simion, 2022). CLT encourages learners to actively engage in communicative activities, such as role plays, collaborative group discussions, and project-based tasks, focusing on conveying meaning, expressing their thoughts and opinions, and achieving successful communication (Nggawu & Thao, 2023). This approach can foster the development of fluency and pragmatic competence in speaking a second language by providing opportunities to practice in realistic contexts.

However, CLT is not a single theoretical framework but a pedagogical approach aimed at enhancing learners' communicative competence through the use of authentic language in meaningful contexts (Savignon, 2018; Martini & Aizan, 2013; Taridi et al., 2024). As noted by Celce-Murcia (2011, as cited in Taridi et al., 2024), the ultimate goal of language instruction is to enable effective communication. For these reasons, CLT provided a suitable framework for exploring language learners' speaking abilities in the context of the Defense Language Institute's Extension Program.

Flow Theory

Flow theory emphasizes optimal experience and engagement in language learning activities. Flow theory is a psychological concept proposed by psychologist Mihaly Csikszentmihalyi (1990) that explores the state of optimal experience that occurs when individuals are fully engaged in an activity. Enhanced self-confidence and task performance have been linked to flow, a positive experience state characterized by focus and involvement in complex yet manageable tasks (Csikszentmihalyi, 2008; Payant & Zuniga, 2022). Achieving flow requires balancing an individual's abilities and the task's difficulty. When challenges are too great, they lead to worry and anxiety, disrupting the flow state (Csikszentmihalyi, 2014). Flow emerges when challenges and skills are in perfect equilibrium. This one-to-one ratio between challenge and ability creates ideal conditions for the flow experience (Csikszentmihalyi, 2014). One of the most notable aspects of flow is that being in a state of flow is the most fulfilling experience. In flow, individuals feel both happiness and focus simultaneously (Csikszentmihalyi, 2014). By selecting speaking tasks that are neither easy nor difficult, learners can immerse themselves in the learning process and experience a sense of enjoyment and accomplishment. This aligns with research suggesting that positive learning experiences can reduce anxiety in language learners (Csikszentmihalyi, 2014). This theory offers insights into how learners' engagement and positive experiences might alleviate the fear of speaking a second language.

Mindset Theory

Rather than solely focusing on individual practice, mindset theory addresses learners' beliefs about their abilities and intelligence. Mindset theory is a motivational theory developed by psychologist Carol Dweck that points out that our beliefs about intelligence and our ability to learn can influence how we handle challenges, respond to feedback, and set our goals (Wolcott et

al., 2021). This psychological framework explores individuals' beliefs about their abilities and intelligence and how these beliefs shape their behavior and approach to challenges (Dweck, 2006). It proposes that people tend to hold one of two mindsets: a fixed mindset or a growth mindset. Those with a growth mindset tend to be more resilient, have a stronger sense of self-efficacy, and demonstrate higher levels of achievement and success (Dweck, 2006). For example, students' mindset regarding their language proficiency and ability to speak the target language can significantly impact their willingness to engage in speaking activities. If students hold a fixed mindset and believe that their language skills are fixed and cannot be substantially improved, they may be more hesitant to speak, fearing mistakes or judgment. A growth mindset pertains to the conviction that one's language proficiency can be enhanced through diligent effort, suitable strategies, and hard work. Conversely, a fixed language mindset suggests that language aptitude remains unchangeable or fixed (Haoting, 2023; Lou & Noels, 2016). Applying mindset theory to language learning helps identify strategies to foster a growth-oriented mindset among students, enhancing their motivation, persistence, and performance in speaking the target language.

The foundational theories presented in the previous paragraphs provided a conceptual lens through which to examine the research problem of students' reluctance to speak the target language in the Defense Language Institute's Extension Program. This section establishes the theoretical underpinnings that guide the investigation into the impact of AI-assisted language learning tools on language learners' lowering anxiety and ability to speak the target language.

Definitions of Key Terms

AI-Supported Language Learning Tools

AI-supported language learning tools offer personalized, adaptive, and interactive learning experiences, analyze speech patterns, enhance language acquisition, and help learners acquire and master new languages more effectively. As Divekar (2020) described, AI agents can simulate social interactions with humans to teach languages, allowing students to experience real-world interactions and practice authentic conversation in an immersive virtual environment.

Anxiety

Anxiety refers to feelings of worry, fear, unease, or apprehension. It is a psychological and physiological response to stressful or threatening situations, which can manifest in physical, cognitive, and behavioral symptoms. The distressing anxiety induced when speaking a second language negatively impacts the learners' spoken performance (Rahman & Tomy, 2024).

Confidence

Confidence is a state of self-assurance and belief in one's abilities, qualities, and judgment, which can lead to personal growth, motivation, resilience, and the willingness to take risks. According to Ruth (2023), confidence is the belief that someone or something can be relied upon to do a task, and it develops over time due to experience. Self-confidence is the belief in one's capacity to deal with challenges in life.

Defense Language Institute (DLI)

DLI provides world-class, culturally based foreign language education, training, evaluation, and degrees for the Department of Defense (DoD) to create and maintain multilingual military and civilian, from apprentice to master, to protect American national security (DLIFLC, n.d.).

Department of Defense Linguists

Department of Defense linguists are bilingual or multilingual learners who focus on learning a foreign language to increase the country's security. Because of this, the terms "student" and "linguist" are used interchangeably throughout this research study (Attia et al., 2021). In this study, the term "linguist" is used to designate the participants who have received foreign language analysis training from the Department of Defense (DoD), not individuals with formal training as linguists in the academic field. These individuals are referred to as linguists while also being students of the Extension Program (EP) offered by our Learning Center (LC).

Extension Program (EP)

The EP offers post-basic second language instruction. To ensure that DoD employees are mission-ready, EP oversees a network of permanent post-basic language training to advance students' language proficiency, increase intercultural competence, and encourage a lifetime of language learning (DLIFLC, n.d.).

Immersive VLEs

Immersive VLEs are digital learning environments using virtual reality (VR) or augmented reality (AR) technologies to create a simulated environment replicating real-world scenarios, enabling learners to acquire knowledge and skills that are more engaging and experiential. The virtual environment is generally defined as computer-generated 3-dimensional environments like virtual reality (VR), augmented reality (AR), mixed reality (MR), virtual habitats displayed through desktops, and some 3D videos and video games (Divekar, 2020).

National Cryptologic School (NCS)

NCS is a Cryptologic Training School established in 1965 that offers language, information assurance, cryptology courses, and leadership. This school operates under the Government Employee Training Act and the Department of Defense (DoD) (National

Cryptologic University, n.d.). Both military and civilian employees of the U.S. government can receive training in cryptology, leadership, business, language, and cryptology (National Cryptologic University, n.d.).

Review of the Literature

Acquiring speaking skills in a second language is crucial for effective communication and language proficiency. With technological advancements, AI-assisted language learning tools and Virtual Reality (VR) have emerged as potential avenues to enhance second language speaking abilities. VR can provide a contextually rich environment for learners to practice language skills in assimilated real-world scenarios. Virtual reality settings have been used in language learning due to their enhanced immersion and higher levels of engagement. The results of a meta-analysis of VR between 2007 and 2019 revealed that students who used VR performed better in identifying structures and describing functional meanings (Ma et al., 2023). Ma et al. (2023) also noted that VR had a moderately positive effect on students' emotions, skills, and cognitive levels. Ayotunde et al. (2023) stated that the emergence of AI technology benefited students and teachers in learning English as a foreign language and mainly enhanced the processes of speaking skills, along with writing, reading, and listening skills. Ayotunde et al. (2023) also highlighted that an engaging and personalized learning experience positively affected students' involvement and success in learning foreign languages.

Thus, within the frameworks of flow theory, mindset theory, and CLT, this literature review explored the effects of AI-assisted language learning tools on second language speaking skills. The scope of this literature review is to investigate the impact of AI-assisted language learning tools, specifically on second-language speaking. The goal of this study was to understand how digital methods could enhance language production proficiency and to examine

students' perceptions of AI's impact on their confidence in speaking. This was accomplished by conducting a thorough qualitative analysis.

The review began with an overview of theoretical frameworks such as flow theory, mindset theory, and CLT approach for overcoming speaking anxiety and promoting language learners' confidence. After providing a brief history of virtual reality and artificial intelligence, the discussion concentrated on the function of AI-language-assisted tools in second-language speaking. The review of AI-assisted language learning tools was further refined regarding learning outcomes concerning the flow and mindset theory and CLT approach. As a result, the objectives included evaluating the effectiveness of AI tools in improving speaking skills, identifying key factors that contributed to their impact within the theoretical frameworks, and identifying any gaps or limitations in the existing research.

A Brief Restatement of the Practice-Based Study Problem

The students of second language learning in the Defense Language Institute extension program are DLI's client partners. Their positions demand advanced language abilities to accomplish certain job objectives. Additionally, they must earn a final course grade of at least 77%. In accordance with the established graduation criteria, students who have not demonstrated proficiency levels of ILR 2 or higher in listening and reading and ILR 1+ or higher in speaking are ineligible for graduation (Attia et al., 2021). In a working environment, they should routinely produce accurate transcriptions and translations at a level of language competency between 2+ and 3 (General Professional Proficiency). Most learners are highly educated and have a government and military background. They are already conversant in the target language. They enroll in the Extension Program to preserve and advance their language skills. Language learners who graduate from the Extension Program of Defense Language Institute are proficient in

reading and listening comprehension, but their written and verbal communication skills need some improvements (Attia et al., 2021).

Teachers, department chairs, and program managers express their worries about the absence of language production in the classroom at weekly and monthly leadership and staff meetings (Attia et al., 2021). As a result, the leadership of our learning center places a high priority on using a second language in the classrooms. Using the target language in the classroom is one of the teaching standards against which each teacher is evaluated (Attia et al., 2021). However, most of our Extension Program students concentrate more on reading and listening skills. This is particularly relevant in the DoD language learning context, where speaking, despite its crucial importance, receives less emphasis in traditional classroom settings (Binkley, 2017; Attia et al., 2021). Despite requiring students to use the target language in classrooms, they need more confidence speaking the L2. Understanding the DoD linguists' educational background helps clarify the causes of the problem and explains why speaking is often neglected in DLI's Extension Program's foreign language classes. One of the causes is the lack of confidence and potential discomfort and anxiety that result from speaking in front of their classmates and instructors who are native speakers. The other reason is that the Department of Defense solely evaluates listening and reading comprehension skills. Students must pass the Defense Language Proficiency Test (DLPT) to be paid and advance in their professional careers. Therefore, the incentive pay system doesn't persuade language learners to achieve and succeed in writing or speaking the target language (Belser, 2020; Attia et al., 2021). This could be one of the reasons that students pay less attention and need to be motivated enough to produce written and spoken language. Therefore, students are only encouraged to write in the target language when they are particularly interested in writing or take composition tests in 19-week courses. Although some

teachers can spark interest in speaking and writing in a second language by using their teaching techniques and skills to help students pass the composition or speaking tests in the 19-week courses, there are no other specific motivations or incentives to encourage students to produce spoken or written language.

Of course, there has been a notable shift in DLI's Extension Program's courses over the past few years. Unlike in the past, where most courses focused on refreshing skills, there has been a transition towards longer-term studies that span 19 weeks, and the graduates are awarded credits to complete their bachelor's degree after the end of these 19-week courses. Nevertheless, the credits toward a bachelor's degree are exclusively awarded to second language students who demonstrate proficiency by obtaining the required Defense Language Proficiency Test (DLPT) scores and completing the Oral Proficiency Interview (OPI) exam. The OPI exam is a standardized and global assessment of functional speaking proficiency in a second language. DLI students must obtain at least an ILR level 2 in this exam to obtain the required credits for the bachelor's degree. Therefore, the OPI exams can be considered the sole incentive to persuade DLI students to enhance their second language speaking abilities and overcome their reluctance and discomfort to some extent in this area of language.

Theoretical Frameworks for Overcoming Speaking Anxiety and Promoting Language

Learners' Confidence

Learning a language is a challenging topic that strongly emphasizes practical application (Atkinson, 2002). Learners must go through a prolonged and complex learning process, particularly when learning a new spoken language. Language acquisition is more of a social interaction activity in actual social language circumstances than a cognitive activity separated

from context and language use (Atkinson, 2002). Atkinson (2002) also highlighted that language communication and expression must be founded on a specific social environment since people use language to convey, construct, and express their thoughts, feelings, behaviors, identities, and other information. The context affects how a second language learner obtains meaningful data and understandable content. Language learning environments in classrooms built around text and audio materials fall short of providing students with a more realistic learning experience. Three-dimensional virtual reality technology offers a virtual community for students to apply a second language in context in a more diversified setting simulated in actual society.

According to Vygotsky (1978), language serves two purposes: while exterior speech is used to communicate with others, inner speech is used for mental reasoning. In the overall process of development, two distinct developmental pathways can be identified, each with a different origin: one involves basic processes that are biologically based, while the other involves higher psychological functions that stem from sociocultural influences. A child's behavior develops from the interaction of these two pathways (Vygotsky, 1978). These actions take place independently. A child uses words socially before age two when they do not have an internal language. However, when language and thought come together, social language is internalized and helps the child reason. As a result, a child's learning is influenced by their social surroundings (Vasileva & Balyasnikova, 2019). According to Vygotsky's (1978) sociocultural view of human learning, learning is a social process that develops in society or culture, just as the human intellect does (Vasileva & Balyasnikova, 2019). The central idea of Vygotsky's (1978) theoretical system is that social contact is crucial to cognition growth (Vasileva & Balyasnikova, 2019).

Flow theory posits that individuals experience a state of optimal engagement and immersion when faced with challenging yet manageable tasks. In language learning, students might overcome their fear and anxiety connected to speaking the target language by participating in tasks and activities that match their skill level and present an appropriate level of challenges (Payant & Zuniga, 2022). Exploring the impact of AI-assisted language learning tools and VR on second language speaking within the context of flow theory might facilitate a state of flow. The interactive and adaptive nature of AI tools, along with the immersive experiences offered by VR, contributes to heightened concentration, enjoyment, and intrinsic motivation, thereby enhancing learners' speaking performance. However, it is essential to recognize that achieving flow in language learning depends not solely on task design. The problem of students' reluctance to speak the target language is influenced by various factors, such as language proficiency, curriculum, untrained instructors, and the learning environment. Flow theory could be integrated with other theories and approaches to address these multifaceted aspects comprehensively and develop effective strategies to promote speaking a second language.

The mindset theory, developed by psychologist Dweck, offered a relevant framework for understanding students' reluctance to speak the target language. The theory explores individuals' beliefs about their abilities and intelligence and how these beliefs shape their behavior, motivation, and approach to challenges. In language learning, students' mindset regarding their language proficiency and ability to speak the target language can significantly impact their willingness to engage in speaking activities. If students hold a fixed mindset and believe that their language skills are fixed and cannot be substantially improved, they might be more hesitant to speak, fearing mistakes or judgment. This fixed mindset could hinder their motivation, resilience, and overall achievement in the language (Haoting, 2023; Lou & Noels, 2016).

On the other hand, cultivating a growth mindset among language learners could have a positive impact. A growth mindset involves the belief that abilities can be developed through effort and practice (Dweck, 2006). Students with a growth mindset are more likely to embrace challenges, persist in the face of setbacks, and actively seek opportunities for improvement (Dweck, 2006). They have higher levels of self-efficacy, viewing mistakes as opportunities for learning rather than failures. By fostering a growth-oriented mindset, students could become more open to speaking the target language, taking risks, and engaging in communicative activities.

The combination of flow theory, mindset theory, and CLT can be a compelling framework that explains students' motives for speaking a second language. Flow theory emphasizes engaging language activities that sustain learner motivation by immersing students in tasks that balance challenge and skill level (Csikszentmihalyi, 2014). Mindset theory complements this by fostering a growth mindset; students may be more likely to persevere through challenges and look at mistakes as a constructive part of the learning process (Eppler & Ironsmith, 2021). Building on this, adopting a growth mindset might reduce fear of mistakes and encourage risk-taking.

CLT emphasizes effective communication and active participation in real-life scenarios, encouraging learners to express themselves in the target language (Taridi et al., 2024). One essential component of effective communication within this approach is discourse competence, which focuses on the effective selection, arrangement, and organization of words, structures, and discourse elements to convey a clear and cohesive message (Celce-Murcia, 2007; Taridi et al., 2024). CLT fosters a dynamic and engaging learning environment that supports this

development, encouraging learner-centered activities and providing varied forms of linguistic feedback (Cloudia Ho, 2020; Taridi et al., 2024).

This combined framework explains how students can build confidence and develop practical communication skills by encouraging active engagement in language learning. By creating an environment that supports growth, resilience, and real-world application, educators can enable learners to navigate communication challenges effectively and achieve success in their language acquisition journey.

The Impact of AI & VR Tools on Learning Outcomes

AI-assisted language learning tools have undergone significant advancements over the years, with their history dating back to the late 20th century. The idea of AI began in the 1950s, when Alan Turing created his well-known Turing test to determine if machines could reason (Almelhes, 2023). Then, the first iterations of intelligent tutoring systems (ITS) focused on computer-based learning platforms that made an effort to adjust to the learner's needs (Almelhes, 2023; Self, 1998). Computer-assisted Language Learning (CALL), which today includes multimedia, virtual worlds, distant learning, and interactive whiteboards, is traced back to the 1960s (Almelhes, 2023). The field of intelligent computer-assisted language learning (ICALL), which blends CALL with AI, first appeared in the late 1970s (Almelhes, 2023). Despite its complexity and reliance on contemporary technology, ICALL benefits from deep learning and neural networks (Almelhes, 2023). The development of more advanced natural language processes (NLP) and computer vision systems began in the 1990s due to improvements in machine learning algorithms and computer power (Gold, 2023).

As technology progressed, the emergence of machine learning and natural language processing techniques in the 2000s led to further developments in AI-assisted language learning.

These tools began incorporating more advanced algorithms and linguistic models, enabling learners to receive personalized instruction tailored to their needs (Gold, 2023). Adaptive learning platforms developed the ability to analyze learners' performance and adapt the content and difficulty level accordingly, enhancing the effectiveness of language learning.

The Virtual Learning Environment (VLE) is generally defined as computer-generated 3-Dimensional environments like virtual reality (VR), augmented reality (AR), mixed reality (MR), virtual habitats displayed through desktops, and some 3D videos and video games (Divekar, 2020). AI-supported language learning refers to a mediating tool that employs AI technology in various ways to help the production and teaching of a second language, such as automatic writing evaluation systems, speech recognition software, AI robots such as chatbots, the intelligent or humanoid robots, interactive conversation simulations, virtual language tutors, personalized learning paths and AI-supported voice-based smartphone apps (Yang & Kyun, 2022). Language learning applications began leveraging VR to create immersive and interactive environments where learners could practice language skills in realistic scenarios, such as virtual conversations, cultural simulations, and virtual tours of foreign countries.

Studies investigating AI-assisted language learning tools have consistently reported positive effects on second-language speaking proficiency. These tools employed natural language processing algorithms, providing learners with interactive speaking practice and personalized feedback. Additionally, studies exploring the use of AI in language learning highlighted its immersive nature, promoting speaking confidence and enhancing learners' ability to engage in real-life scenarios. Almelhes (2023) suggested that NLP, AI, and chatbots can be used to create an intelligent self-learning environment for students learning Arabic as a second language. For example, he stated that an AI-based Arabic language and speech tutor could assist non-native

Arabic speakers by pointing out their pronunciation mistakes and keeping track of their progress. However, Almelhes (2023) believed that while the potential of AI in assisting language teaching and learning is substantial, and some scholars have initiated exploration in this area, there still exists a significant gap in our comprehension of the capabilities, advantages, and constraints of AI in aiding second language learners in achieving proficiency in pronunciation. One explanation could be that learning pronunciation takes a lot of time because the materials were perceived as complicated due to changes in sounds, making it challenging to practice correctly (Tonglo et al., 2022).

On the other hand, some students thought studying pronunciation was a sophisticated and challenging skill, and they were too old to learn it now (Tonglo et al., 2022). This is consistent with renowned developmental psychologist Kuhl's (2011) emphasis on the importance of early language exposure for optimal language learning. According to Kuhl (2011), even though adults have stronger cognitive abilities, newborns and young children are more adept at learning languages than adults. While anyone can learn a new language, it is widely agreed that after puberty, the proficiency level is likely lower than that of a native speaker. Although vocabulary acquisition appears less affected by age and remains good throughout life, mastering pronunciation and grammar to the level of a native speaker is unlikely to occur after puberty.

Although this assertion may be accurate, students might be instructed to rethink their perceptions that learning pronunciation happens only at a particular age. Everyone needs strong pronunciation to communicate effectively and advance their language skills. This is where adopting a growth mindset can alter students' perceptions of their abilities and limitations.

Another limitation of AI in helping second language learners is that students with varying degrees of proficiency may receive different learning outcomes when using AI (Huang et al.,

2023). To address this issue, personalized learning systems, one of the most significant AI techniques, enable learners with low language proficiency to progress at their own pace, maximizing their learning outcomes (Chen et al., 2022). Vocabulary is also crucial in second language learning (Hua, 2020). Accuracy cannot be disregarded when learning a new language because proper grammar and vocabulary are equally important (Simon, 2022). Speaking abilities are severely hampered by the absence of vocabulary words (Ahmedani et al., 2022). Even though some students can grasp the language, they need more skills to use it conversationally. Therefore, students need to master vocabulary in the second language. However, according to Hua (2020), traditional methods in L2 vocabulary learning focus solely on memorizing and analyzing the sound, spelling, and meaning of words without considering them in context. For example, traditional methods in China focus on reciting English words and analyzing sounds without cognitive factors, focusing on polysemous words, idiomatic expressions, and word connections to enhance students' metaphorical awareness and competence (Hua, 2020).

Different strategies, such as conceptual metaphor theory or the flipped learning environment, help students learn vocabulary more effectively and enhance their proficiency and learning autonomy. However, Lin and Hwang (2018) suggested that although the flipped learning environment is a sound instructional strategy for giving students more practice with the learning process, it may be necessary to offer suitable technology and scaffolding tools to help students organize their information to enhance their speaking abilities.

For instance, the mind mapping tool allows students to organize and present their vocabulary knowledge effectively. Thus, Lin and Mubarak's study established the mind map-guided AI chatbot strategy in an EFL flipped speaking classroom to facilitate students' speaking learning

performance. According to Lin and Mubarak (2021), students who used mind map-guided AI chatbot tools performed much better in learning and remembering new terminology.

On the other hand, Harris and Sun (2022) assessed the effect of interactivity design in a VR-based second language learning tool, a personalized context-aware system, to provide students with relevant vocabulary learning materials according to their locations, ability, learning time, and leisure time to engage students in learning in a contextualized fashion. Santo et al. (2016) found that AR-based vocabulary learning, which can easily connect vocabulary with the environment would lead to better retention of word learning. The system had the potential to suggest new words to be learned based on students' leisure time.

The results showed that the students who used the system with an elevated contextualized approach outperformed those who did not. Therefore, AI seems to have a great deal of potential for language learning and can address certain current concerns and problems in technology-enhanced language learning (Huang et al., 2023).

Critical Analysis, Limitations and Gaps in the Literature

Despite the promising findings, several studies have raised concerns about the limitations of AI-assisted tools, such as overreliance on technology and lack of human interaction. Similarly, while VR showed potential for improving speaking skills, access to VR equipment, limited content availability, or current low technical capabilities in conversational AI chatbots have limited their potential applicability and were identified as barriers. The challenge is the technological limitations of AI, particularly its ability to engage in meaningful dialogue, which has hindered its seamless integration into language learning (Weigand, 2019; Yang & Kyun, 2022). Furthermore, cultural and individual factors, learner motivation, and instructional design influenced the effectiveness of both AI and VR interventions.

While there is a growing body of literature on the use of immersive VR and AI for second language learning, the amount of research specifically focused on the impact of these technologies on second language learning still needs to be increased. Harris (2022) stated that more study in this field is required because there is still much to learn about second language learning in an immersive VR environment. Yang and Kyun (2022) also believed that further research should be done on successful communication and interaction in L2 through AI-supported collaborative language learning.

Although literature exists about the use of technology in the foreign language classroom, there needs to be more detailed studies of AI's potential advantages or drawbacks for language learning or teaching among second-language learners within a theoretical framework (Yang & Kyun, 2022). According to Yang and Kyun (2022), the effectiveness of AI-supported technology as a mediating tool for language acquisition has been somewhat proven. However, the language used for producing speech, communication, and collaborative design still needs improvement. Similarly, there is plenty of literature about augmented and virtual reality applications in different fields of study. For instance, in the medical field, a team of surgeons and individuals from around the world could observe an operation performed by a British surgeon in real time online via VR (Attia et al., 2021; Zaleski, 2016). However, using VR to enhance speaking in a foreign language needs to be explicitly addressed and discussed.

Moreover, there needs to be more literature to specifically investigate how AI can be utilized to enhance the language skills necessary for real-life language learning objectives (Almelhes, 2023). New research is required since the body of knowledge about second language acquisition in an immersive VR environment is still expanding (Harris, 2022). In unique contexts like the DLI, there needs to be more focus on how AI-assisted tools may improve language

production in a foreign language. More research is still required on how AI experiences can motivate DLI students/linguists to use their target language.

A significant gap exists in the literature regarding the impact of AI-assisted language learning technologies on learners' second language speaking abilities. The rapid advancement of AI-assisted language learning technologies necessitates further research in this area.

Understanding how to leverage these technologies effectively has significant potential benefits for language instruction, particularly within the context of national security.

AI-assisted language learning and VR tools can potentially improve the language learning experience. They can enable students to improve their language abilities in realistic, interactive environments that mimic real-life circumstances. For example, learners can practice their speaking abilities in a secure and controlled environment by using VR or conversational AI chatbots to immerse them in surroundings that replicate a real-life situation. This can aid students in getting over their fear and stress of speaking in a foreign tongue, which is a significant roadblock to language learning.

Furthermore, we need to study how language learners perceive their interactions with these digital environments and how this affects their confidence or motivation to speak the target language and their ability to cope with fear and anxiety. A qualitative case study method would help investigate the primary phenomenon of language learners' motivations and potential apprehensions about speaking a target language. Using AI-assisted language tools in language classrooms could be a portal for learning more about this subject.

This literature review examined and evaluated relevant literature, establishing a foundation for the proposed research methodology in addressing the problem statement. The search encompassed a wide range of sources, including journal articles and books available in

both print and electronic formats. Various databases, such as ERIC, National University Library (formerly NCU), Academic Search Premier, ProQuest Dissertations, and Google Scholar, were consulted due to the interdisciplinary nature of the study, covering multiple subject areas.

Throughout the review process, priority was given to primary sources over secondary sources.

Moreover, the quality and credibility of the studies and journals were taken into account.

Ethical Assurances

Research ethics rely heavily on informed consent, assent, confidentiality, and anonymity, especially when using human subjects (Bloomberg & Volpe, 2019). To address the ethical requirement of informed consent in this study, a clear and comprehensive document contains all the information about the study, including the purpose, procedures, risks, benefits, and compensation, if any. Informed consent is giving participants understandable information about a research project so they may make an informed decision (Hennessy & Doherty, 2022). Therefore, an informed consent document was given in a language understandable to the participants. Additionally, participants were provided ample time to review and ask questions about the informed consent document before signing it.

Research participants have a right to privacy; this is accomplished by offering guarantees of confidentiality and anonymity (Mertler, 2019). Researchers must ensure the confidentiality of research participants' data and personal information by storing data in secure locations, using coding, limiting access to authorized individuals, or using anonymization techniques (Bloomberg & Volpe, 2019). This study collected only the minimum amount of personal information necessary for the research and ensured that the data were stored securely and only accessible to authorized individuals. Besides, the data would not be shared unless a participant signed an explicit consent form.

Anonymity provides an additional layer of privacy protection and is frequently used in studies involving sensitive topics or vulnerable populations. Data submitted by participants should, at the least, be kept confidential and not shared with anybody (Mertler, 2019).

Anonymization techniques protect sensitive data for ethical purposes. These methods include data masking, pseudonymization, generalization, randomization, data swapping, data perturbation, synthetic data, etc. (Jin et al., 2023).

This study used data masking and pseudonymization techniques to protect the participants' identities. In the data masking technique, sensitive information such as names, student numbers, and addresses, are removed or partially obscured from the dataset, leaving only non-identifying details intact. To ensure that readers won't recognize individuals, pseudonyms or codes should be used instead of participants' real identities (Bloomberg & Volpe, 2019). Pseudonymization involves replacing identifiable data with fictitious names or codes, making it more challenging to directly link the data to a specific individual. These techniques allow the protection of the identities of participants while still retaining the data's utility for analysis.

Furthermore, this study determined a particular time to retain the data collected and ensured that the data were securely disposed of when they were no longer needed. It also provided participants with a debriefing that explained the study's purpose and allowed them to ask questions or express concerns about the research or their participation. Participants also had the option to withdraw from the study at any time and for any reason.

Summary

This study investigated the impact of AI-assisted language learning tools on second-language speaking skills through the lens of CLT, flow theory, and mindset theory. It sought to explore the impact of AI on language learners' motivation and confidence to speak the target

language. This study examined the theoretical frameworks, the historical context of AI in language learning, and the role of the recent developments in AI-language-assisted tools in second-language speaking.

Second-language learners face various challenges when speaking the target language, including pronunciation, grammar, and vocabulary difficulties (Ahmedani et al., 2022). DLI language students are often required to achieve high levels of proficiency in a second language for their work in the military or government. Therefore, proficiency in a second language can open up more job opportunities and make it easier to communicate with people from diverse backgrounds. Besides, finding ways to enhance DOD linguists' language learning experiences can significantly impact national security and diplomacy. DLI's language learners are already conversant in the target language but need to be sufficiently confident to speak the target language (Attia et al., 2021). One reason is that students often require more time to process and translate their thoughts, leading to cognitive fluency issues. Anxiety and cultural differences can further hinder their speaking abilities. Students naturally feel uncomfortable and anxious when speaking in front of their peers and instructors who are native speakers. Barriers such as speech anxiety lessen students' intrinsic incentive to use or produce the target language (Sun & Teng, 2021).

The problem in the Defense Language Institute (DLI) Extension Program (EP) is that students are reluctant to speak the target language, leading to language production and fluency challenges (Attia et al., 2021). The reluctance to speak the target language is attributed to linguistic, cognitive, psychological, and sociocultural factors, as well as teaching methods and material issues (Ahmedani et al., 2022). The short duration of refresher courses and long gaps between language courses also contribute to difficulties maintaining language fluency.

Therefore, the practice-based study problem revolves around the reluctance of DLI second language learners to speak the target language, with challenges stemming from the lack of confidence, discomfort, and anxiety in verbal interactions. As a result, through this qualitative case study, I aimed to investigate the effects of AI-assisted language learning technologies on second language learning, confidence, and self-assurance among DLI second language learners.

The theories jointly analyze the effects of virtual reality and AI-assisted language aids on learners' confidence and motivation to use the target language. For example, flow theory suggests that students can overcome fear and anxiety by engaging in challenging yet manageable language tasks (Csikszentmihalyi, 2008; Payant & Zuniga, 2022). At the same time, mindset theory explores the impact of students' beliefs about their language proficiency on their willingness to speak. A fixed mindset may hinder motivation and resilience, while a growth mindset can lead to embracing challenges and seeking improvement opportunities. Fostering a growth-oriented mindset can encourage language learners to be more open to speaking, taking risks, and engaging in communicative activities (Dweck, 2006). Communicative language teaching (CLT) promotes meaningful communication and authentic language use through role plays, discussions, and problem-solving tasks, exposing learners to real-life language situations instead of rote grammar and vocabulary memorization. CLT adopts a student-centered approach, empowering learners with autonomy in their learning process and encouraging active participation in communicative activities (Simion, 2022).

VR and AI-assisted language learning tools have a developmental history that can be traced from the 1950s to the present day. Integrating AI with CALL resulted in ICALL, and using NLP, AI, and chatbots created intelligent self-learning environments for language learners (Almelhes, 2023). The Virtual Learning Environment (VLE), including VR, AR, and mixed

reality, is explored to create immersive and interactive language learning experiences. Studies have shown positive effects on second-language speaking proficiency using AI-assisted language learning tools, which provide personalized instruction and feedback to learners. However, challenges remain, such as difficulties in pronunciation learning and varying learning outcomes among students with different language proficiency levels. To address these challenges, personalized learning systems allow learners to progress at their own pace. Vocabulary learning is also emphasized, focusing on contextualized approaches like mind map-guided AI chatbot strategies and AR-based vocabulary learning, which have been shown to improve word retention. There is potential in using AI and VR for language learning, but further research is needed to address the identified limitations and gaps.

Through this dissertation, the researcher intends to contribute to the growing knowledge of AI in language learning and provide insights into how this technology can effectively improve second-language speaking skills. By addressing the identified limitations and gaps in the literature, this research aimed to pave the way for more effective and engaging language learning experiences, particularly in unique contexts such as the Defense Language Institute. Ultimately, understanding the impact of AI on language learners' speaking abilities holds significant potential for advancing language instruction and fostering language proficiency in diverse educational settings.

Section 2: Methodology and Design

Researchers must consider how their decisions affect the study design, especially those related to methodology and potential biases (Bloomberg, 2019). Qualitative research is notably more challenging since it involves interacting with interpretations and biases. Positionality is a complex concept resulting from personal experiences, identities, and beliefs (Creswell & Poth, 2018). Human perspective plays a significant role in qualitative research, where qualitative researchers are perceived as the primary gathering, interpreting, and analysis tools. Therefore, transparency about biases and assumptions is crucial for ensuring rigor and trustworthiness in research. This qualitative research focused on the Defense Language Institute Extension Program (DLI-EP) students.

The DLI-EP serves students seeking to hone higher level language competency for jobs such as cryptologic linguists with prior language experience. However, speaking proficiency remains a notable challenge for these learners as they are influenced by linguistic, cognitive, psychological, and sociocultural aspects, teaching methods, and course scheduling. Second language learners in DLI-EP frequently struggle with motivation levels and linguistic components of the L2. Therefore, the problem is that students are reluctant to speak the target language in the DLI Extension Program. Attia et al. (2021) stated that the staff of DoD, comprising both military and civilian linguists, has historically been reluctant to speak the target language during language improvement classes conducted at the National Cryptologic School (NCS) and DLI. Given these gaps in language production, there was a need to investigate how the use of upcoming technologies like AI-assisted language learning tools could impact learners' language confidence and proficiency. Attia et al. (2021) recommended that similar research

needs to be done with the Department of Defense. Furthermore, Harris (2022) acknowledged the prospects of VR in language learning, and Huang et al. (2023) advocated for future research on AI's potential to enhance language education outcomes.

The purpose of this qualitative case study was to explore the impact of AI-assisted language learning tools on DLI students' motivation and confidence in speaking the target language. Specifically, it is intended to enhance post-basic level language speaking skills in the Extension Program (EP) and alleviate language-related discomfort and anxiety. Therefore, the researcher intends to answer the following questions: How do DLI students perceive the impact of AI-supported language learning tools on their motivation to engage in language production activities? How do DLI students describe their experience with AI-supported language learning tools in terms of enhancing their confidence and lowering their anxiety when speaking the target language?

Due to the localized nature of the issue at the LC, this study employed a qualitative methodology and a case study design. This centered on a particular group of students and their understanding of AI motivational roles. Quantitative research may not allow for capturing the qualitative nuances involved in students' perceptions and experiences, such as motivation and confidence. Additionally, it may need to provide adequate space for participants to express their thoughts and feelings in their own words, potentially hindering the exploration of multifaceted aspects of language learning with AI tools. Mixed methods may not be appropriate for this research due to its resource-intensive nature, including time, personnel, budget requirements, more participants, and potential challenges in integrating quantitative and qualitative data when the primary focus is qualitatively understanding student perceptions and experiences. Thus,

qualitative methodology seems appropriate for probing localized problems, such as the DLI students' motivation problems, to speak the target language within the Extension Program.

The research questions and data collection type determined the data analysis protocol. Therefore, a flexible method like thematic analysis was used to extract and formulate themes from qualitative data. By focusing on a specific student group and their experiences, I gained insights into how AI-assisted language learning tools can impact students' confidence, motivation, and anxiety related to language production.

Design and Method

The selection of a case study design is predicated on the contextual significance of a localized issue manifest within our institutional setting. Our learning center is a small, close-knit institution serving particular communities and pursuing educational goals. Therefore, a qualitative open-ended survey was conducted with students to understand how AI tools impact motivation for language production. This aligned with the case study approach, allowing for detailed, experiential insights essential for comprehensive analysis.

This study is primarily exploratory and descriptive, providing detailed descriptions of the research site, participants, and individual experiences within a specific group. It aimed to understand motivations for second language speaking through interactive AI experiences. Therefore, other qualitative research designs were not suitable for this research. For example, phenomenology, which focuses on individuals' subjective experiences, was unsuitable for this study, as it aimed to understand a single unit holistically and grasp the big picture. Additionally, this research was intended to build upon existing theories rather than create new ones, which made grounded theory less appropriate given the substantial existing research on AI for second language learning. Ethnography is particularly useful when studying a group's culture or social

context. Since this research didn't emphasize the cultural aspects of how DLI students interact with these technologies, there were better designs than ethnography. Narrative inquiry is valuable when a researcher wants to explore the stories and meaning-making processes behind students' experiences. My research focused on something other than the narrative aspect of their experiences; therefore, this design would not be appropriate. Action research is typically conducted to bring about practical change or improvement in a specific context or organization. It often involves problem identification, action planning, implementation, and reflection cycles to address real-world issues.

My research, on the other hand, focused on understanding and describing students' perceptions and experiences rather than driving immediate practical changes. Therefore, case studies excel in providing a rich, contextual understanding of a specific phenomenon within its real-world setting. This aligned with my research questions about DLI students' perceptions and experiences with AI-supported language learning tools, as it allowed me to explore a single group's experiences within their natural context. However, case studies can take various forms and be conducted in multiple ways. For example, descriptive case studies focus on providing a detailed account of a phenomenon, while explanatory case studies aim to explain causal relationships. These approaches would be for more advanced stages of research when a researcher has a clearer understanding of the phenomenon. Intrinsic case studies typically investigate a case of inherent interest, which may only apply if my specific case (DLI students) is inherently interesting for its own sake. Instrumental case studies use the case to explore broader issues, which wouldn't be aligned with this study's primary focus on understanding students' experiences. Collective case studies involve the study of multiple cases, which would be appropriate for a researcher to compare the experiences of DLI students from different contexts

or settings. However, since the study's goal is to gain a deep understanding of the experiences of a single group (DLI-EP students) with AI language learning tools, an exploratory case study would likely be the most suitable approach to start my investigation. An exploratory case study would be well-suited for investigating a relatively unexplored or complex phenomenon, which would be aligned with this research purpose and questions regarding students' perceptions and experiences with emerging technologies in language learning. An exploratory case study allows the researcher to delve deeply into the experiences and motivations of DLI students in a flexible and open-ended manner. It provides a foundation for understanding the phenomenon before diving into more specific research questions or hypotheses.

The data collection methods included in-depth surveys to gather detailed insights into individual experiences, opinions, and observations of students using these tools to understand their interactions and effects on language production. A benefit of the interview approach is that it can acquire rich, multifaceted data most appropriately and informally. However, I could not use the interview method because it is prohibited by our educational institution's policy, which forbade recording students' voices during interviews.

Therefore, I used the survey method to avoid recording interviewees' voices during the process. Although surveys might not be the best method for analyzing nuanced patterns of interaction, they can be adaptable to diverse research questions to collect various types of data, such as opinions, preferences, and behavior, which can be responded to anonymously. Anonymous responses can encourage more honest answers. Surveys also allow for a standardized set of questions to be administered to all respondents, ensuring consistency in data collection. This helps reduce bias and increases the reliability of the findings. According to Fink

(2013) and Fowler (2014), one benefit of the survey method is that it is generally discreet and straightforward to perform and manage.

The observation was also an apt method for this case study research. Observation is one of the key techniques used in qualitative research to gather data (Creswell & Poth, 2018). I observed students within a controlled classroom environment to understand their activities and behaviors related to the research topic, such as the impact of AI tools on their language production. A researcher can spot behavioral patterns and trends by closely observing what students do and how they use AI tools. This can aid in the researcher's comprehension of the general dynamics of the classroom as well as how students respond to and use the tools. It aligns with Mertler's (2019) description of observing, involving meticulous data collection through close attention to what occurs. Furthermore, observation aids in understanding participants' interactions and relationships. Researchers can gain valuable insights into their relationships and behavior by watching how participants interact with each other, the teacher, and their surroundings (Miles et al., 2020).

Document review was less appropriate for this case study research since it involved systematically analyzing textual or visual data, often for quantifying patterns or themes. While content analysis can generally be a valuable component of case study research, relying solely on this method may overlook the depth and context that case studies aim to capture. Furthermore, documents typically reflect the perspectives and interpretations of those who created them. They may not capture the viewpoints of all relevant stakeholders or provide insights into their experiences, beliefs, or motivations.

The research questions and data collection type determined this qualitative study's specific data analysis protocol. Thematic analysis is a flexible and widely used method well-

suited to various qualitative data sets, so I planned to use this approach in this study. The thematic analysis involves identifying and developing themes from the data; therefore, this method helped me answer my research questions comprehensively and informally.

Highlighting how crucial it is for researchers to consider the potential effects of all the decisions they make regarding the design of their studies raises the significance of positionality and reflexivity throughout the analytic research process. These decisions include the identification, justification, and limitations of all methodological choices and researchers' potential biases and assumptions (Bloomberg & Volpe, 2019). For researchers, a crucial aspect of qualitative research is critically challenging, as well as interacting with their interpretations and the biases that influence them (Bloomberg & Volpe, 2019).

It was essential to develop reflexivity, because I, as a qualitative researcher, positioned myself in my writings. As a result, transparency about my positionality in my research reports was essential. This allows readers to assess the potential influence of my positionality on the research process and findings. Denzin and Lincoln (2018) argued that researchers' positionality is a complex and multifaceted concept shaped by various factors, including their personal experiences, identities, and beliefs. They also argued that the researchers' positionality could impact the research process in several ways, including topic selection, participant selection, data collection techniques, and data interpretation. Thus, in contrast to quantitative researchers, who employ a variety of instruments and tests, qualitative researchers are frequently regarded as the primary means of collecting, interpreting, and evaluating data (Bloomberg, 2019). This reflects the unique role of human beings as both participants and researchers in qualitative research (Bloomberg, 2019). Hence, it highlights the crucial role of my positionality and reflexivity throughout the analytic research process. I had to be transparent about my biases and

assumptions and critically examine how these factors influenced my research design, data collection, analysis, and interpretation. By reflecting on my biases, I ensured the rigor and trustworthiness of my research.

Population and Sample

The population consisted of DLI linguists who were also students of the language class at the Learning Center, School of Language. The participants were purposefully sampled and composed of eight DLI linguists who are also language class students. The research question and data saturation level determine a qualitative case study's number of participants, indicating that the sample size is sufficient to answer the research questions (Guest et al., 2020).

Students were client partners of DLI whose jobs required enhanced language skills to meet their specific job objectives. The Defense Language Institute (DLI) Extension Program (EP) is designed for students who already have some proficiency in the target language and come to the program to maintain and improve their language abilities. Therefore, to complete the program, students must consistently produce accurate language proficiency Level 2+ to 3 (General Professional Proficiency) or above in transcriptions and translations in an operational environment. Learners in the Extension Program are typically well-educated and have military and other government experiences.

The research was conducted at the Learning Center (LC), School of Language, located in the southeastern region of the United States. The LC is a continuing education school that offers refresher courses to DoD linguists and civilians after graduation from the DLI. Purposeful sampling was used, and participants were selected based on their AI technology usage and motivation to speak a second language. The chosen population, DLI linguists, represented this group directly as they actively engaged in language learning within the DLI-EP program. DLI

linguists, both learners and practitioners of the language, offered an ideal real-world context for studying the effectiveness of AI tools within the DLI language learning environment.

The sample consisted of eight male and female participants aged 22 to 45, allowing for in-depth exploration of their experiences and perspectives. This diversity within the sample allowed for a range of perspectives and experiences, contributing to a richer and more comprehensive understanding of how AI tools impact language learners across different demographic characteristics. Studying their experiences provided valuable insights into how these challenges may have manifested in the context of AI-assisted language learning. This was consistent with the exploratory character of the study, which aimed to record a wide variety of perspectives. An exploratory case study is designed to provide a thorough understanding of a particular phenomenon in a real-world context.

I purposefully selected DLI linguists, who were also language class students, as the study population. This selection was directly relevant to the research focus, as these individuals actively engaged in language learning within the DLI-EP program. In qualitative research, purposeful sampling is applied to understand better the main topic under investigation and the research problem (Creswell & Poth, 2018). The question of what should be sampled, who should be tested, how to sample, and how many sites or persons to sample must all be resolved (Creswell & Poth, 2018). The research question and data saturation level determine a qualitative case study's number of participants, indicating that the sample size is sufficient to answer the research questions (Guest et al., 2020). The choice of eight participants allowed for an in-depth exploration of their experiences and perspectives. Case studies often have small sample sizes because of their distinctive features or particular setups (Schoch, 2020). This would be aligned

with the qualitative nature of the research, where the depth and richness of data are valued over a large sample size.

Additionally, many eligible case candidates might be required to be screened and turned into a smaller number of participants. For example, Yin (2018) advised asking knowledgeable people about the case candidates to avoid choosing a case that illustrates something other than what a researcher wants to study. He brings an example of the research team that screened 62 candidates and gathered information from 47 through structured interviews. Using a rigorous selection process, the team narrowed the pool to 22 qualified candidates, finally reduced to six participants.

Marshall and Rosman (2015) also highlighted the necessity of being flexible since sampling may alter throughout a study. They nonetheless emphasize the need for as much advanced planning as feasible for the sampling method (Creswell & Poth, 2018). Even if the sample size or sampling strategy changed while the study was being conducted, as a researcher, I wanted to ensure that I was gathering the data required to answer my research questions. I also wanted to use various strategies from multiple sources and compare the data to see if they reached saturation.

While data saturation is typically associated with qualitative methods, it may not directly apply to the survey component. However, in the context of the study's approach to observation and surveys, data saturation might be a minor concern for the survey component, as surveys often collect structured, predefined data. Triangulation also involves using multiple data collection methods and sources to corroborate the findings. This was good evidence of saturation if the same themes and concepts emerged from surveys and observations. Apart from employing multiple data collection methods to gather more diverse perspectives, another approach could be

document analysis, which involves examining relevant materials and documents or reports pertaining to the DLI-EP program. The consistency in findings between surveys and observations and alignment with the theoretical framework could also strengthen the case study for data saturation. The consistency with the theoretical framework also suggested that my findings were well-grounded in the existing body of knowledge on this topic.

The other evidence of data saturation is comparing and contrasting the data from diverse participants in further research, looking for similarities and differences. For example, most of my participants reported that AI-assisted language learning tools have made them feel more confident or motivated to engage in language production activities. If the same themes and concepts are expressed by participants of other studies' findings, regardless of their background or experience level, then it implies that my results were not exclusive to my particular sample and could apply to other students. Therefore, the topics and ideas that arise from my data would be aligned with the conclusions of earlier research on applying virtual reality and artificial intelligence in language learning aids. The heterogeneous sample's age, gender, degree of language skill, and other relevant factors make the results generalizable to various learners. In addition to providing evidence of data saturation in my study, my transparent reporting data analysis procedure, iterative data gathering, and analysis process helped me better formulate my research objectives. However, I was careful to take into account deviant or exceptional cases that could refute or offer an alternative viewpoint on the themes that have been identified. Demonstrating that I explored both confirming and disconfirming evidence could strengthen my evidence for data saturation and, therefore, support the validity and trustworthiness of my findings.

Data Collection and Analysis

For collecting data, I used a combination of data collection methods to provide a comprehensive understanding of student's perceptions of the role of AI tools in motivating language production. A researcher might select particular data collection methods or combination techniques to acquire the necessary data in the most appropriate and helpful way (Bloomberg & Volpe, 2019). The practical analysis started at the design phase, when I defined participants, chose methods, and formulated research questions. Then, data collection and methods start to create meaning by developing generative themes (Richards & Hemphill, 2018).

However, my recruiting participants and data collection began after the approval of the Institutional Review Board (IRB) (National University, n.d.). The approval of the IRB from the institution differs, but the first was applying to my school's IRB. The application included a study summary with the purpose, methods, subjects, possible risks, and benefits. It also contained copies of all the research material, such as surveys, interview guides, or consent forms.

There are a variety of methods of recruiting participants in a research study. The best choice varies from one study to another, as well as among different target populations. Having selected a participant recruitment strategy, I started recruiting participants. This involved reaching out to prospective participants directly; however, it should be made sure that all participants give informed consent before the commencement of the study. I made sure I clarified the study's objective and the benefits and drawbacks of participation and thanked the participants for the time they spent and participated.

Techniques for collecting data mediated by technology and ethically managing data are crucial. The emergence of social media and new technologies, including data accessible to the public, have made data security imperative. Researchers must consider privacy, administration,

and data storage ethically. They should consider possible breaches and compromises when organizing and carrying out studies, as well as who can access data and why. Therefore, the researcher should start the study with the appropriate safety measures and ensure the research proposal, IRB application materials, and dissertation text address them (Bloomberg & Volpe, 2019). Reporting the study ethically and adequately portraying research participants is essential (Bloomberg & Volpe, 2019).

I surveyed the students in this study to explore their perceptions of the role of AI in motivating language production. This method provided detailed insights into individuals' experiences, opinions, and attitudes toward using these tools. As the study's researcher, I observed students using AI to get insights into how they interacted with these tools and how the technologies affected their language production.

The study surveyed students at the Language Center (LC) School of Language, located in the southeastern region of the United States, during the 19-week language course iteration. Both language students and linguists received a survey through Microsoft Teams. The questionnaire encompassed questions about their previous and current experience with using AI for language practice, opinions regarding using AI in language practice, and perspectives on whether using such tools influenced their motivation and confidence and lowered emotional exhaustion. The survey also gathered demographic information, such as age, gender, and language proficiency.

In addition to the survey, I conducted observations, where I observed students utilizing AI-assisted language learning at the LC to learn whether the tools affected their language production. I took notes about how the students interacted with the tools, as well as their non-verbal responses along with verbal ones. This approach led to the detection of certain behaviors and usage patterns. As the researcher looks for details and records data, s/he focuses on the

interesting or important aspects. Patterns begin to appear in the data as observations are made over time (Mertler, 2019). Another observer assisted in collecting data on different aspects of a phenomenon to get a more comprehensive picture of what happened in the classroom.

I applied a qualitative method to analyze the survey and observation data. The students' experience and perspective themes were elicited using thematic analysis, a qualitative research approach. Based on this, this study offered insights into how AI tools could enhance second-language learners' production.

In qualitative research, coding is an almost universal procedure; it is a crucial component of the analytical procedure and the methods by which researchers dissect their data to create something new (Elliott, 2018). The first step in data analysis is developing a comprehensive analysis plan. I first described the project briefly. Then, I listed all qualitative data sources, acknowledged all theoretical or conceptual frameworks, and considered all the research questions. In that case, I developed a flexible analysis plan that broke down the tasks week by week. The weekly plan included all the steps for data analysis. As Creswell and Poth (2018) described, coding is a process that involves aggregating data and giving it meaning; it is a procedure that implies analysis while labeling concepts that represent data. I started by sorting my data based on their main informational categories. Next, I coded the data transcripts into generative categories. I used a thematic analysis technique to extract themes from data to formulate them. Thematic analysis is a structured data analysis process that entails identifying, analyzing, and deciphering patterns of meaning in data (Braun & Clark, 2006). The thematic analysis aims to gain in-depth knowledge of the data and extract essential themes and subthemes. I tried qualitative data analysis software like NVivo but preferred doing it manually. Software-based coding is faster, saves time, enhances data organization category formation, and handles

challenging parameters. However, critical thinking is crucial for qualitative analysis, and one must rely on their knowledge and thinking process. Relying on my knowledge and thought process allowed me to engage more deeply with the data. Therefore, I decided to code manually to ensure a more nuanced and critical analysis.

The objective is to find patterns that frequently recur in transcripts or to make a note of any instances of deviation (Richards & Hemphill, 2018). After coding the different participant's transcripts, I created a codebook. As Richards and Hemphill (2018) described, researchers create a preliminary codebook after reviewing the notes after the interviews are finished. The codebook often includes first- and second-order themes, explanations for each subject, and room-to-code transcript quotations. The codes need to be reviewed and edited if required. Researchers should review the segments given a code label and think about which would help develop topics in their study (Creswell & Poth, 2018). The codebook must also be modified and adjusted (Richards & Hemphill, 2018). The updated codebook is created to direct further analysis.

The researcher writes a report detailing the main themes and subthemes arising out of the data. It also comprises the researchers' interpretations of the findings and recommendations for further studies. Therefore, at this point, as Braun and Clark (2006) noted, it is necessary to acknowledge our theoretical stances and values regarding qualitative research. We do not hold to a naive realist interpretation of qualitative research in which the researcher can "give voice" to the subjects (Braun & Clark, 2006; Fine, 2002).

The last step is selective coding, in which the researcher creates propositions or hypotheses that connect the categories or puts together a plot that explains how the categories relate to one another. When theoretical saturation occurs, or when no new concepts can be inferred from the data and the theory has been thoroughly tested using the data gathered, analysis

is said to be finished (Bloomberg & Volpe, 2019). Researchers must also address their plans for handling potential ethical issues, such as policy relevance and dissemination, how they plan to communicate their findings, limitations, delimitations, and strategies for addressing these issues (Bloomberg & Volpe, 2019).

Assumptions

The assumption was that AI tools can enhance language confidence and proficiency by providing immersive, interactive, personalized learning experiences. It also assumed that DLI-EP students may hesitate to speak the target language due to fear of making mistakes. The research also suggested that student motivation, confidence, and anxiety level were crucial factors in language learning production and speaking skills. The study also assumed that all DLI-EP students had the necessary skills and were willing to use the tools and that the DLI-EP program supported their use. These implicit assumptions were not explicitly stated but were implied by the study design and methodology. On the other hand, these assumptions may have only partially been accurate.

I selected a qualitative case study design for exploring these presumptions in the context of the DLI-EP program, as it enabled detailed data collection from a limited number of participants. I critically reflected and carefully assessed these presumptions during the research process. The study aimed to ascertain the validity and accuracy of these assumptions within the context of the DLI-EP program.

Limitations

The first limitation was the small sample size. The study involved approximately eight participants. While this may allow for in-depth exploration of their experiences, the small sample size limits the generalizability of the findings. The participants' experiences and perspectives

may not represent the broader DLI student population. To mitigate this limitation, a researcher can consider expanding the sample type to include a more diverse group of DLI students who study other languages. Therefore, the research participants could represent diverse languages, proficiency levels, and demographics. A researcher can also consider conducting a multi-site study involving DLI programs in other languages and locations. This would provide a broader perspective on the impact of AI tools across various contexts to enhance the generalizability of the findings.

The other limitation of this study was limited data collection methods. This study primarily relied on surveys and observations, which may not have captured the full range of factors influencing motivation, confidence, and anxiety. Complementing these methods with interviews or focus groups could have provided a more comprehensive understanding of participants' experiences. Recording the participants' voices is not allowed in our learning center, so not being able to implement an interview or focus group in this study was the first and fundamental limitation. I employed the dual observation method to gather more data consistent with the ratings and yield more substantial findings.

Another limitation was the possible Hawthorne effect during the observation. Participants may react or behave differently if they know they are being observed, which could affect the accuracy of the study's observations. To lessen the Hawthorne effect, a researcher should ensure the observation procedure is discreet and minimize the participants' awareness of being watched. This might involve discreet observation techniques or ensuring the participants are accustomed to being observed in the DLI program.

Another limitation of this research was the technological limitation of language programs and the need for Farsi/Persian language programs in any AI-assisted language learning tools or

AI voice chatbots. For instance, there were a few AI voice chatbots that could have vocal conversations in Arabic, but at the moment, there were no AI voice chatbots available in public that could converse in Farsi. To overcome this technological limitation, a researcher can actively encourage participation from a diverse range of DLI students who speak different languages to be able to utilize various and broader AI tools. However, with the passage of time and a prolonged search, I ultimately found an AI-supported language learning tool named TalkPal AI, featuring a Persian language chatbot to employ for Farsi language learners.

General technical limitations also applied to the experimentation with this research on participants, not only for Persian/Farsi students but also in other languages. The fact that these AI voice chatbots were still under development was one of their main problems. Due to technological limitations, artificial intelligence agents cannot recreate human-like language. One of the challenges in developing a conversational AI is that it needs to understand and generate human speech in real-time (Davis et al., 2023). According to Davis et al. (2023), conversational AI agents currently in use need to improve more in terms of parsing and understanding human speech, particularly speech with foreign accents, as well as preserving the general conversational flow. This technical limitation impacts the interaction between language learners and these AI agents. This implies that future studies with advanced humanlike conversational AI that exceeds current technical limitations might produce different results. As a result, conversational AI agents used for second language learning in the future may differ significantly from the agents used in the studies described in this dissertation (Davis et al., 2023).

Little can be done to address the lack of progress in AI's ability to communicate linguistically like a human other than waiting for newer versions or constantly checking the most sophisticated AI devices. Researchers can also collaborate with other researchers with AI and

language learning expertise. By working together, they can share their knowledge and resources and be informed of the most advanced technology in AI or develop more sophisticated and practical solutions. The other way to mitigate technical limitations is to thoroughly evaluate the applied tools and methods. It is essential to thoroughly assess the current AI-assisted language learning tools and procedures to ensure they are effective and reliable. A researcher can accomplish this by conducting user studies and contrasting their tools and techniques with other inventive researchers.

Delimitations

This study was delimited in terms of population, time frames, language skill focus, and research designs. This indicates that only DLI-EP students who want to work as cryptologic linguists are the study's subject because the study focuses only on the challenges these students face in speaking proficiency. The research was conducted during the Persian/Farsi intermediate and advanced 19-week course. Furthermore, the focus was on speaking proficiency, other than language skills such as reading, writing, and listening. The study also confined its investigation to using AI tools to investigate the impact of the most recent technological tools on student motivation, confidence, and anxiety related to language production. A qualitative exploratory case study design was employed in this study to focus on a single unit, the DLI Extension Program, and to develop a comprehensive understanding of it. This design is also suitable for exploring new phenomena and generating insights. On the other hand, case studies are frequently used in educational intervention studies because they allow for gathering detailed information on participant experiences and implementation and identifying factors influencing success.

Summary

This research focused on a study targeting the Defense Language Institute Extension Program (DLI-EP) and students' language learning difficulties, particularly in speaking proficiency. This study focused on DLI-EP students aiming for cryptologic linguist jobs, addressing challenges in speaking proficiency due to linguistic, cognitive, psychological, sociocultural, and teaching methods, motivation, and linguistic components. I proposed investigating the use of AI tools to boost language confidence and proficiency, aiming to address the reluctance of students to speak the target language within the program.

The study employed a qualitative case study design to explore the impact of these tools on student motivation, confidence, and anxiety related to language production. It used qualitative methodology to capture subtle experiences and perceptions that quantitative research might overlook. However, a researcher must consider potential presumptions in qualitative research involving interpretations and biases. Positionality and reflexivity are crucial in qualitative research, requiring researchers to engage with their biases and interpretations critically. Transparency in positionality evaluation and acknowledging biases ensures rigor and trustworthiness. Furthermore, iterative data analysis and considering deviant cases strengthen evidence validity and reliability. This investigation excluded alternative qualitative research designs like phenomenology, grounded theory, ethnography, narrative inquiry, and action research due to their insufficient scope to describe student experiences with AI tools. The study justified the choice of a case study design due to its alignment with the localized issue within the specific institutional context. Various approaches were evaluated, such as intrinsic, instrumental, or collective case studies. However, an exploratory case study was chosen to comprehensively understand student experiences before addressing specific research questions.

For data collection, I focused on eight DLI linguists at the Learning Center, School of Language, aged 22-45, to explore how AI tools influence language production. The purposeful sampling ensured depth and validity, with multiple sources, triangulation, and comparison with existing literature confirming the findings' reliability. I applied purposeful sampling for language students as a sample type in this research study. I planned to recruit participants through in-person invitations in class, ensuring informed consent, clarifying the study's objectives, benefits, and drawbacks, and thanking participants for their time. However, before undergoing these processes, the study, which contained a study synopsis, research materials, and application, must first have been approved by the IRB. The study needed to be modified at the IRB's request for further adjustments, asking for additional information and safeguards to protect participants, or denying the application altogether.

Due to restrictions that prevented interviews that recorded students' voices, I planned to employ survey methods and observations. The research surveyed students at the Language Center School of Language in the southeastern United States during the 19-week language course iteration. The survey included questions about their experience, opinions, and perspectives on the impact of AI tools on motivation, confidence, and emotional exhaustion. I also conducted observations of students using AI-assisted language learning tools at the LC to understand how students interacted with these tools and produced language. I observed and noted interactions, non-verbal responses, and patterns with the assistance of another observer.

Coding involves aggregating data and labeling concepts. The written data were sorted into informational categories and then coded into generative categories. Qualitative data analysis software like NVivo was not used. NVivo involves software-based coding, which is faster, more efficient, and can handle complex parameters. However, thematic analysis was manually

conducted to extract themes and decipher patterns of meaning in data. The goal was to identify common themes and subthemes incorporated into a preliminary codebook. The codebook was then reviewed and edited, and the updated version was used for further analysis. The analysis was considered finished when theoretical saturation occurred.

The study had limitations, such as a small sample size, reliance on surveys and observations rather than interviews and focus groups, potential Hawthorne effect, and technical and technological limitations. To improve validity, I planned to include a diverse group of DLI students, apply dual observation methods, collaborate with AI experts, and verify the most recent version of the technology. I also employed established qualitative analysis techniques to minimize interpretation bias and ensure the reliability of the study's findings. To further understand the effects of AI tools on second language learning, future research should focus on advanced and more sophisticated conversational AI similar to human speech.

Section 3: Findings, Implications, and Recommendations

This qualitative narrative study explored the effects of AI-assisted language tools on second-language learners' speaking abilities. Specifically, it evaluated students' engagement and motivation when interacting with these tools. It examined how such interactions influence their confidence in speaking a second language compared to interactions with real people. Additionally, the study investigated whether AI tools could reduce anxiety associated with second-language speaking. The purpose of this qualitative case study was to explore the impact of AI-assisted learning tools on DLI students' motivation and confidence in speaking the target language.

While previous research had explored motivation in Virtual Reality (VR) environments for language learning, there was a notable gap in studies specifically addressing students' engagement and motivation in developing speaking skills through AI-assisted language learning tools. This research addressed this gap by conducting observations, surveys, and document reviews to gather firsthand insights into how AI tools impact language learners' motivation and confidence.

The problem is that students in the Extension Program (EP) of the Defense Language Institute (DLI) lack enough confidence and motivation to speak the target language, which may hinder their ability to develop speaking skills (Attia et al., 2021). Attia et al. (2021) noted that students hesitate to speak in class for fear of producing wrong utterances. The Learning Center (LC) 's challenge is finding a means to motivate students to speak the target language during classroom activities (Atia et al., 2021). To explore these issues, I recruited eight DLI language learners as participants, with data collection commencing only after obtaining approval from the

National University Institutional Review Board (NU IRB). A central requirement of the study was that participants had been studying Persian/Farsi for at least one year. The research methodology involved surveys and observation sessions where participants individually interacted with the TalkPal AI tool for speaking the Farsi language.

The study faced several limitations that impacted the interpretation of its findings. It relied solely on surveys and observations, excluding qualitative methods like interviews or focus groups due to logistical constraints. While allowing detailed insights, the small sample size of eight participants restricted generalizability to the broader DLI student population. Additionally, the potential influence of the Hawthorne effect must be considered. Participants might have altered their behavior simply due to awareness of being observed, potentially impacting the authenticity of the data.

Furthermore, technological limitations of the AI voice chatbot tools used in the study affected participants' engagement. Still under development, these tools struggled to maintain natural conversations, which may have influenced the learners' motivation and interaction with the tool. These factors collectively underscore the importance of contextualizing the findings within the study's inherent limitations.

Findings

The research methodology involved surveys and observation sessions where participants individually interacted with the TalkPal AI tool for speaking the Farsi language. All participants studied Persian/Farsi language for at least one year at the Defense Language Institute (DLI) and completed the DLPT and OPI tests. Based on the participants' DLPT and OPI scores, they had enrolled in either an intermediate or advanced-level language course. Table 1 contains demographic information for each deidentified participant.

Table 1*Participant Demographics*

| Participants | Gender | Age Group | Ethnicity | L2 Proficiency Level |
|---------------------|---------------|------------------|------------------|-----------------------------|
| Participant 1 (P1) | Male | 20-29 | White | Intermediate |
| Participant 2 (P2) | Female | 20-29 | White | Intermediate |
| Participant 3 (P3) | Male | 20-29 | White | Intermediate |
| Participant 4 (P4) | Male | 40-49 | Asian | Intermediate |
| Participant 5 (P5) | Male | 40-49 | White | Intermediate |
| Participant 6 (P6) | Male | 20-29 | White | Intermediate |
| Participant 7 (P7) | Male | 30-39 | Hispanic | Intermediate |
| Participant 8 (P8) | Male | 30-39 | White | Advanced |

All study respondents (n=8) met the inclusion criteria. The participants included (a) White males (n=5), (b) White females (n=1), (c) Hispanic males (n=1), and (d) Asian males (n=1). The participants ranged from 20–49 years, with a mean age of 30. The eight participants in the study were enrolled in intermediate and advanced Persian/Farsi language 19-week courses. They were personally invited during class and agreed to participate following a positive response to recruitment efforts. The study's purpose was clearly stated in a consent letter and verbally explained during class. All communication with potential participants occurred face-to-face or through direct messages via Microsoft Teams. After reading and signing the consent letter, each participant verbally confirmed their voluntary participation. Participants were briefed on the

study and instructed to interact with an AI tool by conversing in the target language. During these interactions, I observed and took notes to document participant behavior.

For the observation, I created a structured list that included various aspects of participants' performance, such as verbal output, cognitive processes, engagement levels, comfort levels, confidence, anxiety levels, and technological interaction approaches. In the observation session, I used the checklist to observe and record specific behaviors, facilitating my study on how participants produce the target language while interacting with an AI tool. I applied this list to mark observed behaviors for each participant I observed individually in a classroom.

To enhance the reliability and validity of the observational data, a second observer joined the observation room for some sessions. The second observer and I observed simultaneously and recorded detailed notes on participant behavior during the participants' interaction with the AI tool. To ensure data integrity, I combined both notes for the final analysis of each participant. The decision to include a second observer was driven by the need to minimize observer bias and ensure a more comprehensive and accurate account of each participant's interaction with the AI tool. By comparing and combining the second observer's notes with my own, I aimed to incorporate the triangulation to ensure that the findings did not result from one person's perspective. Therefore, this chapter includes an examination of the data's reliability, followed by the presentation of data analysis results categorized according to the research questions, an assessment of the outcomes, and a concluding overview.

Trustworthiness of the Data

The credibility and value of qualitative research hinges on its trustworthiness. Trustworthiness means that the academic and research community accepts your analysis and interpretation of what others have said and done in the field, supporting the findings' validity,

reliability, and applicability (Bloomberg & Volpe, 2019). To ensure this study's rigor, a multifaceted approach was employed to establish trustworthiness. This section delves into the essential components of trustworthiness: credibility, dependability, confirmability, and transferability, as outlined by Bloomberg and Volpe (2019). These validation strategies are paramount in assuring readers of the significance and value of the research findings. The subsequent discussion thoroughly explores each trustworthiness criterion, elucidating how these principles were applied to the data collection, analysis, and interpretation processes.

Credibility

Credibility involves assessing whether the researcher has accurately reflected and conveyed the participants' viewpoints and experiences (Bloomfield & Volpe, 2019). To enhance the credibility of my study, I employed triangulation along with several rigorous strategies throughout the data collection and analysis phases. For the instance of triangulation, I employed both surveys and observational methods to ensure that the findings accurately represent participants' interactions with the AI. By integrating these dual methods, I ensured that the observations and survey responses corroborated, providing a more comprehensive and reliable account of the participants' experiences. Another significant strategy was conducting member checks to validate the interview transcripts further. Although members checking in the traditional sense of reviewing interview transcripts was not applicable, I used a similar approach by seeking feedback from a subset of participants regarding the accuracy and clarity of survey items and observational notes. During the observation sessions, I sought verbal feedback from participants about their interactions with the AI tool. For instance, some students described being trapped in a repetitive loop with the AI tool and unable to exit the debate with AI due to misunderstandings caused by the AI's incorrect interpretation of their pronunciation. Another participant mentioned

that AI was trying to correct a misinterpreted word; again, perhaps because the participant could not pronounce the term correctly. Additionally, when I verified a Farsi word generated by the AI, the student confirmed that it was unfamiliar, and upon checking their screen, I noticed that the AI had generated a non-existent Farsi word. This feedback helped ensure that the data collected accurately reflected participants' experiences and perceptions.

Confirmability

According to Bloomfield and Volpe (2019), confirmability refers to the extent to which a researcher's findings and interpretations are directly derived from the data rather than influenced by the researcher's biases or subjectivity. While challenges exist in achieving this in survey and observation research due to the lack of direct interaction, several strategies can be employed to enhance confirmability. I actively engaged in reflexivity throughout the research process by acknowledging and reflecting on potential biases or preconceptions that could influence the study. By being transparent about these biases and their potential impact, I aimed to mitigate their effect on the findings and ensure that the conclusions were derived from the data rather than personal opinions. I also tried to provide a thick description of the research setting, participants, and data collection procedures, and, specifically, a rich explanation of the data analysis of each participant to facilitate understanding and evaluation. Data Triangulation also helped to ensure confirmability. I enhanced the study's confirmability by integrating survey results with observational data. This triangulation allowed for cross-validation of findings from different data sources, reinforcing the results' objectivity and reducing the potential for individual researcher bias to influence the conclusions.

Dependability

Dependability in qualitative research refers to the consistency and stability of the data over time and the extent to which the research process is documented and traceable (Bloomfield & Volpe, 2019). While survey and observation research might present challenges in demonstrating dependability compared to experimental designs, several strategies can be employed to enhance this aspect. The approach to ensuring dependability for my study was Data Triangulation. I combined the survey results with observational data. By integrating these two distinct data sources, I could cross-verify the consistency and accuracy of the information collected, providing a richer and more reliable understanding of the research questions. The second observer meticulously documented participant interactions and the AI tool in the examining room. This included detailed observations of how participants engaged with the technology while speaking the target language. In one instance, the second observer facilitated a visit from the DLI site director, who also noted the participants' interactions with the AI tool. This additional layer of observations provided further insights. It ensured a more comprehensive understanding of the participant's experiences and the effectiveness of the AI tool in facilitating target language-speaking practice.

To maintain dependability, I also employed consistent data collection methods. The surveys were administered using the same format and conditions for all participants to ensure uniformity. Observations were conducted using a structured protocol with predefined categories and criteria for recording participant behavior. For example, I developed detailed guidelines for observers, which contained a structured list that included various aspects of participants' performance, such as verbal output, cognitive processes, engagement levels, comfort levels, confidence, anxiety levels, and technological interaction approaches to ensure consistent data

recording. This consistency in data collection methods helped to produce reliable and comparable results.

Transferability

Qualitative research seeks to establish credibility, dependability, and confirmability and is concerned with the issue of transferability; that is, how and in what ways the findings of a particular study and the knowledge derived therefrom might apply or be useful in other similar contexts (Bloomfield & Volpe, 2019). Unlike quantitative research, which often seeks broad generalizability, qualitative research focuses on providing detailed and context-rich insights that can be transferred to similar situations or settings. Transferability emphasizes the importance of detailed contextual information, which helps readers make informed comparisons and draw relevant connections to their own situations (Bloomfield & Volpe, 2019).

Ensuring transferability involves providing sufficient contextual information and detailing the study's methodology to enable others to assess the relevance of the findings to their own contexts. To facilitate the transferability of the findings, I provided a detailed contextual description of the research context, including the participant characteristics, setting, and the specific conditions under which the data were collected. This detailed contextual information allows other researchers to determine how closely their own contexts align with the study environment and assess the potential applicability of the findings. I also meticulously recorded all research procedures, including questionnaire development, data collection, analysis, and interpretation. This documentation allows for scrutiny and replication of the study. To ensure the research findings could be traced to the data collected, I maintained detailed records of the study's methodology, including the survey design, observation protocols, and data analysis

procedures. This thorough documentation provided a transparent audit trail, allowing for the verification of the research process and the steps taken to derive conclusions.

Building on these trustworthiness criteria, the study sought to uncover how AI-assisted language tools affect learners' engagement, motivation, and confidence in their speaking abilities. Specifically, the research addressed two key questions:

RQ1: How do DLI students perceive the impact of AI-supported language learning tools on their motivation to engage in language production activities?

RQ2: How do DLI students describe their experience with AI-supported language learning tools in terms of lowering their anxiety and enhancing confidence when speaking the target language?

I manually extracted themes from the data to answer these questions and organized the findings using a color-coding system. The table below provides an overview of the questionnaire's themes and categories.

Table 2: Questionnaire's Themes

| Categories | Description | Question |
|--|--|--|
| 1. Comfort & Usage | Comfortable using AI tools while speaking with the AI agent | 1. How comfortable were you using AI-assisted language learning tools while speaking with the AI agent? |
| 2. Motivation & Engagement | Impact on motivation to speak the target language (enjoyable or engaging) | 2. How did the AI-supported language learning platform impact your motivation to speak the target language? Did it make it enjoyable or engaging? |
| 3. Features & Functionalities | Features to enhance motivation to speak L2 | 3. What features or functionalities within TalkPal AI do you think could further enhance your motivation to speak the L2? |
| 4. Confidence & Anxiety | Impact on confidence in speaking the target language | 4. How did the AI-supported language learning platform impact your confidence in speaking the target language? |
| 4. Confidence & Anxiety | Impact on anxiety reduction | 4. Did it help reduce your anxiety? Please explain. |
| 5. Strengths & Weaknesses | Strengths and weaknesses of AI platforms (e.g., TalkPal AI) | 5. What do you consider are the biggest strengths and weaknesses of AI platforms, for example, TalkPal AI? |
| 6. Comparison | Comparison with other AI-assisted language learning tools | 6. Have you used any other AI-assisted language learning tools? How does TalkPal AI compare to them? |
| 7. Technology Skills & Experience | Overall rating (1-10) of technology skills and TalkPal AI experience | 7. Overall, on a scale of 1-10, how would you rate skills using technology and your experience using TalkPal AI for language learning? |
| 8. Achieving Goals | Achievement of personal language learning goals | 8. Personal goals: Did TalkPal AI help you achieve your personal language learning goals? If so, how? |
| 8. Achieving Goals | Rating (1-10) of TalkPal AI's effectiveness in achieving goals | 8. How do you rate TalkPal AI on a scale of 1-10 in helping you achieve your language learning goals? |
| 9. Future Applications | Envisioning AI-assisted tools in language learning programs (e.g., DLI-EP) | 9. How do you envision AI-assisted tools being used more effectively in language learning programs like the DLI-EP? |

Table three presents key themes identified from one of the participant responses and coded data.

Table 3: Coded Participants' Answers

| Category | Keywords and Description | Codes |
|----------------------------|--|---------------------|
| Comfort & Usage | Extremely comfortable ; Overcome fear of mistakes; Flowing conversation | COMF |
| Engagement & Motivation | Increased motivation ; Enjoyment ; Self-practice | MOTV_ENJ |
| Features & Functionalities | Advanced debate feature; Difficulty in understanding some responses | ADV_DBT |
| Confidence & Anxiety | Reduces anxiety ; Alleviates stress; Increased confidence ; Coping with speaking anxiety faster | LES_ANX MOR_CONF |
| Strengths | Speaking partner ; Self-practice ; Utilizing AI everywhere; Real-time feedback; Corrects mistakes; Allows more speaking | SPK_PRC |
| Weaknesses | Misunderstands user input; Difficulty with subtle pronunciation; Hard to comprehend some responses | MSUND |
| Comparison | No experience with other AI tools | NO_EXPR |
| Technology | Rating technology skills (9) ; User-friendly /easy to navigate; Minor technical issues | FRND_9 |
| Achieving Goals | Increased practice speaking L2; Helped overcome speaking anxiety; Rating of TalkPal AI (10) | SPK_10 |
| Future Application | Useful for language learning schools; Must be implemented in DLI; Supplementing speaking time; Combination of old-fashioned & AI tool; Incredible for improving L2 speaking | SPLM |

After individually coding each participant's survey responses, I analyzed each in more depth. I then compared each participant's answers with those of the other participants. Finally, I combined the results for all eight participants, question by question, to complete the overall analysis of the survey responses. The table below presents the themes mentioned by each participant and the frequency with which each theme was discussed.

Table 4: Thematic Summary of Participants' Survey Responses

| Theme Code | Themes Description | Number of Participants | Total Mentions |
|---|---|-------------------------------|-----------------------|
| Category 1: Comfort & Usage | | | |
| 1/COMF | Comfort with AI-assisted tools | 8 | 13 |
| Category 2: Motivation & Engagement | | | |
| 2/MOTV | Motivation increase | 3 | 4 |
| 2/NOT_MOTV | Not motivated | 1 | 1 |
| 2/ENJ | Enjoyable | 5 | 9 |
| 2/ENG | Engaged in conversation | 3 | 5 |
| 2/SPK | Speak anytime | 1 | 1 |
| Category 3: Features & Functionalities | | | |
| 3/VART | Variety of features and functionalities | 2 | 4 |
| 3/DVRS | Diverse topics | 2 | 4 |
| 3/DEBT | Debate feature | 2 | 3 |
| 3/ROL | Roleplay feature | 2 | 2 |
| 3/TOPIC-ISSU | Topic issue | 2 | 2 |
| 3/LIM-SCEN | Limited scenarios | 1 | 1 |
| 3/WRD-RECG | Word recognition issue | 3 | 7 |
| 3/LOOP | Loop, circular conversation | 3 | 3 |
| Category 4: Confidence & Anxiety | | | |
| 4/CONF | Confidence Increase | 3 | 3 |
| 4/RED-ANX | Anxiety reduction | 6 | 13 |
| 4/TRNSC | Transcription errors | 1 | 1 |

| | | | |
|--|---|---|----|
| 4/JUDG | No Judgment | 4 | 4 |
| Category 5: Strength & Weaknesses | | | |
| 5/RL-CNVR | Realistic conversation situation | 2 | 3 |
| 5/RL-SCEN | Strength: Real-world scenarios | 2 | 3 |
| 5/VART | Strength: Variety of scenarios | 2 | 3 |
| 5/ACES | Strength: Anytime access | 2 | 2 |
| 5/SPK-PRCT | Strength: Speaking practice flexibility | 5 | 12 |
| 5/LMT-CNVR | Weakness: Limited conversation | 3 | 3 |
| 5/RL-RSPN | Weakness: No realistic responses | 1 | 1 |
| 5/UNDR | Weakness: Understanding nuances | 6 | 10 |
| 5/RPET | Weak: Repetition/Circular conversation | 3 | 4 |
| 5/CULTR | Weakness: Lack of cultural nuances | 1 | 1 |
| Category 6: Comparison | | | |
| 6/OTHER-AI | Utilizing other AI tools | 1 | 1 |
| 6/NO-EXP | No experience with other AI tool | 7 | 7 |
| 6/GPT | ChatGPT: More useful for learning | 1 | 1 |
| Category 7: Technology Skills & Experiences | | | |
| 7/USR-FRND | User-friendly tool | 2 | 2 |
| 7/TKPL-EXP | Experience using TalkPal AI | 0 | 0 |
| 7/RATE | Rated 8-9/10 | 4 | 4 |
| 7/RATE | Rated 7/10 | 2 | 2 |
| 7/RATE | Rated 4&5 /10 | 2 | 2 |
| 7/HLPFL | Helpful tool for language practice | 6 | 5 |

| Category 8: Achieving Goals | | | |
|---------------------------------------|--|---|---|
| 8/GOAL | Achieving personal goals | 6 | 7 |
| 8/RATE | Rated 6/10 | | |
| 8/NATV-INTR | Preferring native speaker interaction | 3 | 3 |
| 8/PRCT | Practice L2 without a native speaker | 4 | 3 |
| 8/LMT-IMPC | Limited impact on goals | 1 | 1 |
| 8/NTRL-EFCT | Neutral effect on L2 learning | 1 | 1 |
| 8/USR-FRST | User Frustration | 1 | 1 |
| Category 9: Future Application | | | |
| 9/USEFL | AI-assisted tools very useful | 3 | 4 |
| 9/DLI-IMP | Implementing in DLI-EP program | 3 | 4 |
| 9/COMPL | Complementary to address teacher shortages | 1 | 1 |
| 9/SMRT | Getting smarter | 1 | 2 |
| 9/TRANS | Better transcription programs | 1 | 1 |
| 9/LNG-LRN | AI Potential in language learning | 1 | 1 |

Comfort and Usage

Five out of eight participants reported a high level of comfort using AI for language learning, citing reduced anxiety and the absence of judgment as key factors. The significant benefits were the ability to practice without fear of making mistakes and to engage in conversation at their own pace. Participant 2 highlighted this advantage: "I feel extremely comfortable using this tool...I get very nervous...with this tool, though, I feel comfortable speaking freely and overcoming the fear of making mistakes..."

However, three out of eight participants expressed mixed comfort levels. They noted initial discomfort or a learning curve, suggesting that familiarity and time with the AI are necessary for optimal comfort. For example, Participant 6 indicated a period of adjustment, stating, "I'm pretty comfortable. It does take a few times to get used to how the AI works". Similarly, Participant 8 initially felt hesitant but became more comfortable after understanding how to use the tool: "Without much instruction, there was hesitation. But after realizing how to use the tool, the comfort level did increase". Therefore, a learning curve exists for some users, as not everyone achieves the same level of comfort when interacting with the AI agent immediately. While some individuals required an adjustment period, all participants ultimately experienced a degree of comfort using the TalkPal AI language assistance tool.

Motivation and Engagement

This analysis examined the AI language learning platform's impact on participants' enjoyment and engagement levels when speaking the target language. The platform enhanced the speaking experience for five out of eight participants. Features like diverse conversation topics, character interactions, and role-playing scenarios contributed to this. Participants also valued the control over conversation topics, which increased their motivation. Participant 5 said, "I feel that the AI...was motivational...offers a range of various historical and fictional personalities...control over the conversation...".

Additionally, the ability to practice at their own pace and without pressure from a native speaker made the experience more enjoyable. For instance, Participant 2 stated, "I found this tool to really help my motivation speaking...I think it is fun...allows for practice at home." The flexibility and reduced-pressure environment contributed to a more enjoyable learning experience.

However, limitations in AI understanding emerged as a barrier to motivation for three out of the eight participants. The inability to handle complex conversations or adapt responses led to frustration. The initial appeal of the platform's variety diminished for some as they encountered these limitations, as Participant 4 noted: "At first, it was good: the variety of topics... but once we started going through... the AI seemed to be very surface-level...". As they encountered these limitations, the initial appeal of the platform's variety diminished for some. For example, Participant 7 stated: "There are frustrations with the platform that make it difficult to have a conversation, such as it not being able to adapt to the conversation." Furthermore, repetitive responses and the need to adjust speaking style due to transcription errors detracted from the overall enjoyment.

Features and Functionalities

The platform offers features to make speaking practice more engaging, but AI shortcomings can hinder this over time. While the platform seems valuable for increasing motivation to practice, especially for those who dislike traditional exercises, some participants still prefer human interaction for language learning. Overall, by analysis, the platform can be a beneficial supplement, but not a replacement, for traditional methods.

Several participants highlighted specific features of TalkPal AI that they found motivating for practicing the target language (L2). Participant 3 emphasized the role-play feature as particularly beneficial, stating, "I think the most useful tool the AI has is the roleplay feature." Participant 5 valued the situational conversations and debates, remarking, "I especially enjoy the situational conversations...and I enjoy the debates as they do force a student to engage in realistic conversations."

However, there were also significant suggestions for improvement. Participant 1 identified a need for better lexical identification capabilities, noting, "It needs to improve its word recognition." Participant 4 recommended a more lenient approach to word recognition, similar to translation tools. Participants also pointed out limitations in the AI's ability to conduct natural conversations. Participant 7 expressed a desire for a "smarter" AI that could adapt to conversations rather than relying on pre-programmed responses. Participant 8 found the limited reach in the scenarios and suggested a wider range of responses. Participant 8 may have found the AI's capabilities insufficient for addressing a broad or diverse range of conversational contexts, which could limit its effectiveness in simulating real-world language use.

Confidence and Anxiety

Examining how the AI-supported language learning platform influenced participants' confidence in speaking the target language reveals several insights. Most participants reported a notable benefit: reduced anxiety and increased comfort associated with using the platform. Participants 1,3,5,7, all highlighted that the AI agent's nonjudgmental nature was a key factor in alleviating their speaking anxiety. Participant 1 noted, "It does help reduce anxiety because if I make a mistake, it is not a person I can embarrass myself in front of." Participant 2 said, "I think this tool is a great asset for those of us who struggle with speaking-related anxiety and stress." The AI's supportive environment encouraged experimentation with new phrases, as Participant 3 explained: "Talking to an AI makes me more likely to try new phrases or words." However, some participants had mixed experiences. Participant 4 found the platform only "helped a little bit" with managing stress, while Participant 8 described a "net neutral" impact on confidence due to issues with transcription accuracy. In total, six participants reported a reduction in their

L2 speaking anxiety due to various factors, such as the nonjudgmental environment and sufficient L2 practice, which helped alleviate their stress.

Strengths and Weaknesses

Participants appreciated the flexibility of practicing speaking anytime and anywhere. Participant 6 noted, "I can use it anywhere, anytime." This flexibility enhanced confidence and reduced anxiety. The platform's various features, including diverse languages and topics and the ability to converse with historical and fictional characters, were also praised. Participant 4 mentioned the "variety of languages and topics" as a strength. Additionally, the nonjudgmental nature of AI interactions created a safe space for practice, with Participant 5 appreciating not worrying about being judged by a human teacher.

Key limitations included the AI's inconsistent ability to understand spoken language, leading to misunderstandings due to pronunciation or vocabulary issues. Participant 5 highlighted the frustration of the AI misinterpreting neutral statements as offensive. Participants also noted the AI's struggle to hold natural conversations, with repetitive responses and a lack of cultural nuance. Participant 7 summarized this limitation: "The negative is that the AI is not very advanced, and the level of conversation you can have with it is limited." This AI shortcoming made certain features or modes frustrating, as pointed out by Participant 5: "Debate mode can often be frustrating for that reason." Additionally, Participant 8 felt the platform was not suitable for beginners due to its limitations.

Comparison with Other Tools

This analysis examined participants' experiences using TalkPal AI compared to other AI language learning platforms. Most participants (n=7), including Participants 1, 2, 3, 4, 5, 6 and 8, had not used other AI-assisted language learning tools. Therefore, they could not compare them with TalkPal AI. This limited the generalizability of any inferences about TalkPal AI relative to other products. One participant (n=1), Participant 7, had experience with ChatGPT and found it "much more useful for learning" but noted its lack of a listening function, which hindered spoken conversation practice: "ChatGPT is much more useful for learning but doesn't have a listening function, so speaking conversations are not possible at the moment." This suggests that while ChatGPT might be superior for learning, TalkPal AI may be better for facilitating spoken language practice.

Technology Skills & Experience

Overall, participants rated their experience with TalkPal AI for language learning at an average of 7.0 out of 10. This suggests that participants found TalkPal AI to be a helpful tool for language learning and practice. The table below is a breakdown of the participant ratings:

Table 5: Technology Skills and Experience Rating

| Participants | Rating |
|---------------------|---------------|
| Participant One | 8-9 |
| Participant Two | 9 |
| Participant Three | 7 |
| Participant Four | 5 |
| Participant Five | 8 |
| Participant Six | 7 |
| Participant Seven | 4 |
| Participant Eight | 8 |

The ratings ranged from 4 to 9, indicating that some participants found TalkPal AI more helpful than others. Two participants who rated TalkPal AI more highly mentioned that they found it a user-friendly and easy-to-navigate tool. They also appreciated the ability to practice speaking anytime and anywhere. However, as mentioned before, some participants, like Participant 7, noted that TalkPal AI had limited conversational abilities, and it was often too frustrating to be useful. Based on participants' responses and an average rating of seven out of ten, it can be concluded that six participants found the AI tool helpful for language practice, while two other participants did not find it useful.

Achieving Goals

Several participants reported that TalkPal AI helped them achieve their goals to some extent. Participant 1 indicated a positive impact, rating it a 7. Similarly, Participant 3 found it valuable for practicing without a native speaker and rated it a 7/10: "While not the best way to improve speaking, I'd rate it a 7/10". Participant 5 viewed TalkPal AI as a supplement to human interaction, rating it an eight compared to a 10 for speaking with a native speaker. These participants found the platform beneficial for practicing speaking and overcoming communication fear. Participant 2 highlighted this success, rating it a 10: "My personal language goal right now is to practice and to overcome speaking anxiety. I think TalkPal AI has absolutely helped me complete that...". Participant 4 rated it 4 out of 10, but he stated, "It gets me speaking." Six out of eight participants reported experiencing modest progress and benefits from using TalkPal AI.

However, not all participants felt TalkPal AI significantly impacted their goals. Participant 4 viewed it as a tool to "augment" their learning but not a primary resource, still preferring in-person interaction. Similarly, Participant 6 found it helpful for fluency but not a

replacement for speaking with a native speaker, rating it a 5: "5 – I still think speaking with a native speaker is the best way to practice speaking. This is just a great way to get more fluent without needing to have a physical person speaking to you at all times." Participant 8 found it had no impact on their goals, rating it a 5: "5, the tool didn't improve or decrease my language learning goals. It has zero to little effect on how I learn." These participants emphasized the value of human interaction for achieving their language learning goals. A review of the participant ratings is provided in the table below:

Table Six: TalkPal AI Effectiveness in Achieving Goals

| Participant | Rating |
|--------------------|---------------|
| Participant One | 7 |
| Participant Two | 10 |
| Participant Three | 7 |
| Participant Four | 0 |
| Participant Five | 8 |
| Participant Six | 5 |
| Participant Seven | 4 |
| Participant Eight | 5 |

Therefore, according to participant feedback, TalkPal AI's overall rating as a language learning platform for achieving language learning goals was 6 out of 10.

Observation results indicate that most participants show mixed motivation levels when interacting with AI-supported language learning tools. For Participant 1, for instance, while there were positive signs of engagement and effort to use complete sentences, the reliance on AI,

issues with AI understanding, and frustration with feedback mechanisms indicate potential areas where motivation could be improved. The AI-assisted language learning tools presented mixed results for Participant 4. He showed some engagement but struggled with motivation, confidence, and anxiety. However, Participant 2 showed strong motivation when using AI-supported language learning tools. The participant's use of complete sentences, correct grammar, self-correction, clear pronunciation, thoughtful cognitive processes, active engagement, and enjoyment of using AI tools all point to a high level of motivation to engage in language production activities. Participant 2's engagement and enjoyment of using AI tools further support the notion of enhanced confidence and lowered anxiety.

In contrast, Participant 4 displayed confusion and frustration, particularly during the role-play and debate. The AI tool did not significantly enhance his confidence or reduce his L2 speaking anxiety, as evidenced by his performance and feedback. The biggest problem was his hardship pronouncing certain Persian words, which caused technical issues with the app. He switched topics multiple times due to boredom and difficulty, indicating a lack of sustained engagement. Therefore, the participant expressed boredom and expected features that could enhance motivation.

Participant 4 expressed boredom through several noticeable behaviors. He frequently switched topics during the role plays, starting with "coffee shop" and then moving on to others, indicating a lack of interest in the activities. This constant shifting suggests he found the initial role-play unengaging and did not find the subsequent ones any more stimulating. The participant also used very short sentences and minimal language during these activities, which could signal disengagement and a lack of motivation to participate fully. Signs of frustration and discomfort were also evident, especially during the complicated role-play about renting a car. This

frustration likely stemmed from challenges completing the tasks due to general difficulty speaking and pronouncing certain Farsi words.

Technological issues with the app further disrupted his focus and enjoyment and added to his disengagement. Moreover, the debate on "Graffiti as art" was not well-suited to his language level. Participant 4 had the weakest speaking skills among the other seven students, making the debate topic particularly challenging. He could not achieve Level 2 or 2+ speaking during the debate, leading to frustration and discomfort, especially while performing in front of three observers. These factors, together with constant topic changes, short responses, technological problems, inappropriate activity choices, and hardships in speaking Farsi, reflected the participant's boredom and disengagement.

Participant 3 demonstrated a positive overall experience with the AI-assisted language learning tool. He was actively engaged, motivated, and receptive to feedback. Unlike Participant 4, Participant 3's pronunciation was very clear, and he effectively self-corrected mistakes. However, there were occasional minor grammatical errors and room for improvement in vocabulary usage, as the participant sometimes needed to repeat sentences due to the AI's initial misunderstanding. He also managed to initiate interactions with the AI without prompts or sometimes would initiate the prompt. Participant 3 cognitively exhibits strong skills in applying logical reasoning and demonstrating practical problem-solving abilities. His facial expressions and engagement with specific AI features showed his interest in the tool. However, he sometimes showed confusion or amusement when the AI struggled to understand or looped responses.

Regarding engagement, Participant 3 remained focused and involved, asking for clarification when needed and participating actively in turn-taking conversations. He displayed

no signs of distraction or frustration and seemed to enjoy using AI tools, a positive indicator of his engagement and motivation. This feature was also confirmed and described by the second observer's note, who wrote: "Participant 3 showed confidence and displayed a strong grasp of language, using the right vocabulary and grammar in his interactions with the AI."

The participant's motivation remained high despite the need to repeat some sentences for the AI to understand. In terms of confidence and anxiety, Participant 3 showed increased confidence and reduced stress in his speaking abilities. He sought reassurance from the AI and felt more comfortable when the conversation flowed smoothly. He also frequently checked with the instructor to ensure correct sentence structure and vocabulary, which indicates a proactive approach to language learning.

Regarding feedback and adaptation, Participant 3 responded positively to AI feedback, utilized it effectively, and adjusted his communication style based on the feedback provided. The participant also adeptly managed tools, followed instructions, and adjusted settings for better interaction despite occasional pauses and misunderstandings due to AI interaction delays. Overall, Participant 3 exhibited strong language production skills and effective engagement with AI tools.

Like Participant 3, Participant 8 demonstrated strong cognitive engagement, verbal proficiency, robust engagement with the AI tool, and effective use of it. He also showed high motivation, confidence, and compelling technological interaction. However, compared to Participant 3, Participant 8 experienced fewer issues with AI misunderstandings and had a slightly different approach to feedback and reassurance. He only used the feedback feature once but responded positively and adjusted his behavior based on it. Participant 8 also appeared to think before speaking, used logical reasoning, and demonstrated problem-solving skills.

However, Participant 8 reported a minor feeling of anxiety and discomfort, mainly due to being observed during the session. This suggested that while AI tools could significantly enhance confidence and reduce stress in many contexts, the presence of an external observer might still introduce a degree of discomfort. Despite this, his overall experience with the AI tool remained positive, highlighting the tool's ability to facilitate interactions and potentially lower language-related anxiety. The initial feelings of anxiety also highlight the importance of providing a supportive environment for learners.

Participant 7's motivation to engage with the AI tool varied based on the AI's responsiveness. He demonstrated increased motivation when the AI tool understood him and avoided producing repetitive responses. However, he felt trapped in a conversational loop when the AI's responses were circular. This indicates that while the AI tool can boost motivation when it functions effectively, limitations or repetitive patterns can negatively impact motivation and engagement. Participant 7 reported increased confidence in speaking, particularly compared to interactions with a native speaker. However, he noted that the TalkPal AI never gave him the feeling of interacting with a natural person, which could affect his overall engagement and perception of the tool's effectiveness. Despite this, he did not seek reassurance from the AI or express discomfort, suggesting that his confidence and anxiety levels were positively influenced by the AI's support, even if it fell short of providing an entirely realistic interaction experience. According to Participant 7, he could see increased motivation if the AI had fewer restrictions, allowing for a broader and more engaging range of conversation.

Participant 6 also reported increased confidence and reduced anxiety while using the AI tool, particularly at the beginning of the 19-week language course. The TalkPal AI tool was noted to help reduce anxiety and stress, suggesting that it positively impacted his comfort level

with speaking the target language and led to increased motivation in language production activities. However, like participants 3, 5, and 7, there were moments of expressed boredom and lack of interest due to repetitive circular responses from the AI. According to Participant 6, AI repeated the same error, although the participant corrected the tool.

Like Participants 6 and 7, Participant 5's motivation to engage in language production activities with the AI tool varied based on the effectiveness of the AI responses. He showed increased motivation when the AI tool provided accurate and relevant responses. Participant 5 also expressed moments of boredom and frustration when the AI tool produced incorrect or repetitive answers. The participant's motivation was positively influenced by engaging interactions but was diminished by issues such as unfamiliar or non-existent terms used by the AI tool. For example, the AI tool generated unfamiliar Farsi words, such as رحان (rahan) and تکویض (takviz), or provided incomprehensible information due to the use of non-existent terms. Therefore, these limitations, such as generating unfamiliar or incorrect terms, impacted Participant 5's engagement and motivation. Nevertheless, Participant 5 experienced reduced anxiety when using the AI tool, as it provided a safer space for practicing language skills without the fear of immediate judgment. The AI provided a controlled environment where Participant 5 felt less pressure than when interacting with native speakers.

Regarding feedback adaptation, participants 5, 6, and 7 exhibited varied reactions and behaviors toward this feature. Participant 7 used the feedback button minimally, possibly only once, and did not use the translation feature. Unlike Participant 7, Participant 6 paid attention to the feedback provided by the AI tool and used it every time. He incorporated the AI's feedback into his subsequent responses, indicating effective adaptation to feedback. He adjusted his communication style based on the AI's feedback, particularly in vocabulary, showing responsiveness to the AI's input.

Participant 5 did not use or pay attention to the feedback feature of the AI tool. During member checking, I asked Participants 5 and 7 why they had not utilized the feedback feature, and both responded that they were not fully aware of its existence. In my view, this lack of engagement with feedback tools highlights potential areas for improvement in making these features more noticeable and user-friendly.

Evaluation of the Outcomes

The findings revealed that the platform can significantly reduce communication fear and boost comfort for learners, particularly those who experience high levels of anxiety. The nonjudgmental environment encourages learners to experiment and practice without embarrassment. Nonetheless, a learning curve is evident for some users, and not everyone reaches the same degree of comfort with the AI agent. For instance, while the AI platform includes features designed to make speaking practice more engaging, limitations in AI comprehension can impede progress over time. The study also highlights those features mimicking real-world scenarios, such as role-playing and situational dialogues, that serve as strong motivators for second-language (L2) learners. Creative expression and opportunities for debate are also valued aspects. However, the AI's limitations in word recognition and inability to maintain natural conversations require further improvement to bolster learner motivation.

Ratings for TalkPal AI in technology skills and experience varied between 4 and 9, reflecting that individual experiences with the tool differed. Participants who rated TalkPal AI highly noted its user-friendliness and the convenience of practicing speaking at any time and

from any location. Overall, participants rated their technology skills and experience with TalkPal AI at an average of 7.0 out of 10, suggesting that while the tool is somewhat effective for language learning and speaking practice, its usefulness is limited.

In terms of achieving language learning goals, TalkPal AI offers a flexible and supportive platform for language learning with an average rating of 6 out of 10, but its impact on achieving specific goals depends on individual needs and preferences. Regarding comparisons between TalkPal AI and other language learning tools, the limited available data prevents definitive conclusions about how TalkPal AI measures up against other AI-assisted language learning platforms. Future research could benefit from a broader participant base and comparisons of various language learning tools.

In conclusion, the observation results underscored a range of participant experiences with AI-supported language learning tools, revealing both successes and challenges. The findings indicated that while some participants, such as Participant 2, exhibited high levels of motivation, confidence, and effective use of the AI tools, others, like Participant 4, faced notable difficulties. Issues such as AI misunderstandings, technical problems, and lacking engagement with feedback mechanisms affected motivation and language learning outcomes. Participant 3 and Participant 8 demonstrated strong cognitive engagement and adaptation to feedback, although Participant 8 experienced minor anxiety due to the observational setting. In contrast, Participants 5 and 6 showed varying responses to feedback, highlighting the need for improvements in making feedback tools more accessible and impactful.

Observations suggested that while AI tools have the potential to enhance language learning by improving confidence and reducing anxiety, their effectiveness can be limited by

technical issues and user engagement with feedback features. This highlights the importance of continuous improvement in AI technology to ensure it aligns with learners' evolving needs. Future developments should address these limitations, refine the AI's responsiveness, and create a more engaging and supportive learning environment to better meet diverse learner needs.

Nevertheless, the observations revealed that all participants demonstrated a notable capacity to proficiently manage the AI tool and handle its technical aspects effectively. Their ability to operate tools and interfaces competently, follow on-screen instructions accurately, and navigate technical difficulties with minimal assistance highlights their adaptability and familiarity with technology. This proficiency suggested that not only did the participants, whose ages ranged from 20 to 49, effectively handle the tool's technology without assistance, but it also highlighted the platform's user-friendliness. This proficiency is a testament to their comfort and ease with digital tools, reflecting the positive impact of their age and technological fluency.

Implications and Recommendations for Practice

The findings of this study suggest that AI-assisted tools like TalkPal AI can address critical challenges in language learning programs, particularly in settings such as DLI-EP. These tools provide learners with additional opportunities for speaking practice, which is often limited in traditional classroom environments due to time constraints or an insufficient number of native Farsi speakers. Furthermore, AI tools create a supportive, nonjudgmental environment, which can help reduce learner anxiety, particularly for beginners.

Educational practitioners can apply these findings by integrating AI tools into language programs as supplemental resources. For example, instructors can recommend AI-assisted tools for independent practice to complement classroom instruction. This approach aligns with Participant 2's observation that AI tools can provide valuable feedback outside class time,

compensating for the limited opportunities for teacher-student interactions. Additionally, as Participant 5 highlighted, these tools can offer an alternative practice platform for students who lack access to native speakers or experience heightened anxiety in real-life interactions. To maximize the benefits of these tools, educators should provide structured guidance on their use, ensuring students engage meaningfully with the technology. This recommendation is supported by Participant 3, who suggested that a more structured implementation would enhance the tool's effectiveness. Moreover, programs should remain adaptable as AI technology evolves, ensuring that new features, such as improved conversational capabilities, are leveraged to enrich the learning experience further.

Recommendations for Future Research

Future researchers can build upon this study by addressing its limitations and expanding its scope to gain a more comprehensive understanding of AI-assisted language learning tools.

One of the key limitations of this study was its small sample size of eight participants, all from a single DLI program. Future researchers can include a more extensive and diverse group of participants to improve the generalizability of findings. Incorporating students from various languages, proficiency levels, and demographics would provide a more representative view of AI tools' effectiveness. Conducting a multi-site study across different DLI programs and languages could also enhance the breadth of the research and provide insights into how AI tools perform in varying linguistic and cultural contexts.

This study relied primarily on surveys and observations, which limited the depth of insights into participants' experiences. Although the restriction on voice recording at the study site might persist and continue to pose a challenge, it may eventually be lifted. Future researchers

should consider complementing these methods with interviews or focus groups if this restriction is removed.

Future researchers should also explore discreet observation techniques or methods that allow participants to acclimate to observation over time. The Hawthorne effect can alter participants' behavior because they will change it when they are aware of being observed, which may have influenced the study's results. Therefore, future researchers can implement long-term embedded observations or indirect monitoring methods to reduce the probable Hawthorn effect and yield more natural participant behaviors.

Future research could explore implementing more sophisticated AI language tools to assess the impact of enhanced AI capabilities. Given the current limitations of AI tools in replicating human-like conversational abilities, investigating newer, more advanced systems should be a priority. For instance, researchers might examine the effects of emerging technologies, such as conversational AI agents with improved natural language understanding and real-time interaction capabilities.

Additionally, researchers could collaborate with AI developers to create or enhance tools for less commonly supported languages, like Farsi, thereby expanding accessibility to a broader range of learners. Beyond conversational abilities, future studies could investigate the impact of AI on other aspects of second language acquisition, such as reading, listening, writing, grammar, and pronunciation, to provide a more comprehensive evaluation of language learning at DLI. Although some of these areas may have already been explored at DLI, the rapid advancement of AI underscores the importance of incorporating cutting-edge tools into research. As AI continues to evolve, assessing whether these advancements lead to meaningful improvements in language learning programs like DLI-EP is crucial.

Repeating this research as AI technology evolves would help assess the consistency and applicability of findings over time. Future studies with next-generation AI tools may yield different results, offering fresh insights into how these technologies influence learner motivation, confidence, and anxiety. By addressing these limitations and adopting these strategies, future researchers might enhance the robustness and applicability of their findings, contributing to a deeper understanding of AI-assisted language learning tools and their potential in diverse language contexts.

Conclusions

This research provided valuable insights into the effectiveness of AI-supported language learning tools as adjuncts to traditional methods. The analysis revealed that these tools offer significant advantages, such as increased accessibility, diverse features, and a supportive environment. For example, TalkPal AI's role-playing and situational conversations have effectively motivated learners and provided a nonjudgmental space for practice, which has generally reduced stress and boosted confidence.

However, the study also identified notable limitations. While some participants experienced increased motivation and confidence, as well as an anxiety reduction, others encountered challenges such as repetitive responses, technical issues, and a lack of realistic interaction. These limitations highlight the need for improvements in areas such as word recognition, conversational adaptability, transcription accuracy, and feedback mechanisms.

The varying impacts of AI tools on motivation, confidence, and anxiety among participants underscored the importance of addressing these limitations. Improving the AI's responsiveness and feedback features could significantly enhance its effectiveness and user

engagement. The participants' adeptness with technology further suggests that AI tools have the potential to be effectively integrated into educational practices, leveraging their strengths while addressing areas of user concern. Additionally, despite variations in the participants' levels of engagement and motivation, all participants reported that interacting with the AI tool in the second language reduced their stress levels and enhanced their confidence compared to conversing with native speakers.

AI-assisted platforms demonstrate promise as valuable resources within language learning programs. Nevertheless, they should be seen as complementary to, rather than a replacement for, traditional methods that involve native-speaker interactions because speaking a language involves more than just grammar and vocabulary. Speaking a second language involves cultural understanding, social interaction, and the development of interpersonal skills. Native interactions provide authentic language exposure, cultural context, and immediate, nuanced feedback that AI struggles to match with its current limitations. On the other hand, the effectiveness of AI tools can vary greatly depending on individual learner characteristics, such as their mindset about their L2 speaking ability, age, attentiveness, learning style, and technology proficiency. Some learners may find AI tools highly beneficial, while others may prefer traditional methods.

Despite the AI language learning platform's effectiveness in increasing confidence for those who find speaking a second language with AI less intimidating, the preference for human interaction remains strong. Participants appreciated the personal touch and tailored feedback that humans can provide. However, they also acknowledged blending AI tools with traditional methods. Until the AI reaches its peak performance and optimum points, this balanced approach leverages AI technology's and human interaction's strengths, ultimately enhancing learning

outcomes. Thus, while the platform is a beneficial adjunct to traditional methods, it is still not a replacement.

References

- Alam, M. R., Jahan, S., Khan Milon, M. R., Ansarey, D., & Faruque Al Hadi, S. U. (2021). Accelerating learners' self-confidence level in second language acquisition: A qualitative study. *ICRRD Quality Index Research Journal*, 2(3). <https://doi.org/10.53272/icrrd.v2i3.5>
- Almelhes, S. A. (2023). A review of artificial intelligence adoption in second-language learning. *Theory and Practice in Language Studies*, 13(5), 1259–1269. <https://doi.org/10.17507/tpls.1305.21>
- Al-Wossabi, S. (2016). Speaking in the target language: Issues and considerations. *Journal of Language Teaching and Research*, 7(5), 886.
- Atkinson, D. (2002). Toward a sociocognitive approach to second language acquisition. *The Modern Language Journal*, 86(4), 525-545. <https://doi.org/10.1111/1540-4781.00159>
- Attia, G. O., Sairany, Z., & Treacy, J. (2021). *DOD Linguists' perceptions of the use of virtual learning environments on the motivation to speak a foreign language* (Doctoral dissertation, Augusta University).
- Ayotunde, O. O., Jamil, D. I., & Cavus, N. (2023). The impact of artificial intelligence in foreign language learning using learning management systems: *A systematic literature review*. *Information Technologies and Learning Tools*, 95(3), 215. <https://doi.org/10.33407/itlt.v95i3.5233>
- Beyond Boredom and Anxiety (2000, December 1). (Book Review) (Brief Article). *SciTech Book News*, 24, 3.
- Berkeley, U. (n.d.). *Human Research Protection Program*. UC Berkeley Human Research Protection Program. <https://cphs.berkeley.edu/>

- Bisso, E., Signorelli, M. S., Milazzo, M., Maglia, M., Polosa, R., Aguglia, E., & Caponnetto, P. (2020). Immersive virtual reality applications in schizophrenia spectrum therapy: A systematic review. *International Journal of Environmental Research and Public Health*, 17(17). <https://doi.org/10.3390/ijerph17176111>
- Bloomberg, L. D., & Volpe, M. (2019). *Completing your qualitative dissertation: A road map from beginning to end* (4th ed.). Sage.
- Bloomberg, L. (2019). *Role & responsibilities of the qualitative Researcher* [Webinar]. Northcentral University/Center for Teaching and Learning.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chen, Y.-L., Hsu, C.-C., Lin, C.-Y., & Hsu, H.-H. (2022). Robot-assisted language learning: Integrating artificial intelligence and virtual reality into English tour guide practice. *Education Sciences*, 12(7), 437. <https://doi.org/10.3390/educsci12070437>
- Creswell, J. W., and Poth, C. N. (2018). *Qualitative Inquiry & Research Design: Choosing among Five approaches*. SAGE Publication Inc.
- Csikszentmihalyi, M. (2014). *Applications of flow in human development and education: The Collected Works of Mihaly Csikszentmihalyi*. Springer Netherlands.
- Data Anonymization: Use cases and 6 common techniques*. Satori. (2022, November 22). <https://satoricyber.com/data-masking/data-anonymization-use-cases-and-6-common-techniques/>
- Davis, G. M., Padilla, A. M., Landay, J. A., & Schwartz, D. L. (2023). *Humanlike conversational artificial intelligence agents for foreign or Second language learning: User perceptions, expectations, and interactions with agents* (dissertation).

- Denzin, N.K. & Lincoln, Y.S. (2018) *The Sage handbook of qualitative research*. Los Angeles: Sage. Defense Language Institute Foreign Language Center (n.d.).
<https://www.dliflc.edu/>
- Divekar, R. R. (2020). *A.I. enabled foreign language immersion: Technology and method to acquire foreign languages with A.I. in immersive virtual worlds*. ProQuest Dissertations & Theses Global.
<https://go.openathens.net/redirector/ncu.edu?url=https://www.proquest.com/dissertations-theses/ai-enabled-foreign-language-immersion-technology/docview/2459192732/se-2>
- Du, X. (2022). *English-speaking students' learning experiences in a Chinese/English dual language immersion program: A multiple-case study [ProQuest Information & Learning]*. In Dissertation Abstracts International Section A: Humanities and Social Sciences (Vol. 83, Issue 10–A).
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Dweck, C. (n.d.). *The power of believing that you can improve*. Carol Dweck: The power of believing that you can improve, TED Talk.
- El Majidi, A., de Graaff, R., & Janssen, D. (2021). Debate as a pedagogical tool for developing speaking skills in Second language education. *Language Teaching Research*, 28(6), 2431–2452. <https://doi.org/10.1177/13621688211050619>
- Elliott, V. (2018). Thinking about the coding process in qualitative data analysis. *The Qualitative Report*, 23(11).
<https://link.gale.com/apps/doc/A568913418/AONE?u=pres1571&sid=ebsco&xid=2e155fac>

- Fink, A., (2013). *How to conduct surveys: A step-by-step guide*. Thousand Oaks, CA: Sage.
- Fowler, F. J. (2014). *Survey research methods (5th ed.)*. Thousand Oaks, CA: Sage.
- Gold, E. (2023, May 5). The history of artificial intelligence from the 1950s to today. freeCodeCamp.org. <https://www.freecodecamp.org/news/the-history-of-ai/>
- Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *PLoS ONE*, *15*(5), 1–17.
<https://doi.org/10.1371/journal.pone.0232076>
- Habiburrahim, H., Risdaneva, R., Putri, G., Dahliana, S., & Muluk, S. (2020). The effects of anxiety toward Acehnese students' English-speaking ability. *The Qualitative Report*.
<https://doi.org/10.46743/2160-3715/2020.3639>
- Hamed, I., Denisov, P., Li, C.-Y., Elmahdy, M., Abdennadher, S., & Vu, N. T. (2022). Investigations on speech recognition systems for low-resource dialectal Arabic–English code-switching speech. *Computer Speech & Language*, *72*, 101278.
<https://doi.org/10.1016/j.csl.2021.101278>
- Haoting Li. (2023). Perceived teacher-student relationship and growth mindset as predictors of student engagement in foreign student engagement in foreign language learning: the mediating role of foreign language enjoyment. *Frontiers in Psychology*, 1–14.
<https://doi.org/10.3389/fpsyg.2023.1177223>
- Harris, C. E. (2022). *Assessing the effect of interactivity on virtual reality second language learning*. Doctoral dissertation, Rowan University.
<https://go.openathens.net/redirector/ncu.edu?url=https://www.proquest.com/dissertations-theses/assessing-effect-interactivity-on-virtual-reality/docview/2670022856/se-2>

Harris, C., & Sun, B. (2022). Assessing the effect of interactivity design in V.R. based on the second

A language learning tool. *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. <https://doi.org/10.1109/ismar55827.2022.00015>

Hennessy, M., Dennehy, R., & Doherty, J. (2022). Outsourcing transcription: Extending ethical considerations in qualitative research. *Qualitative Health Research*.

https://doi.org/10.1177_10497323221101709

Hua, H. (2020). A study of English vocabulary learning in China—from the perspective of conceptual metaphor theory. *Journal of Language Teaching and Research*, *11*(3), 427-434. <https://doi.org/10.17507/jltr.1103.11>

Jin, F., Hua, W., Francia, M., Chao, P., Orłowska, M. E., & Zhou, X. (2023). A survey and experimental study on privacy-preserving trajectory data publishing. *IEEE Transactions on Knowledge and Data Engineering, Knowledge and Data Engineering, IEEE Transactions on, IEEE Trans. Knowl. Data Eng*, *35*(6), 5577–5596.

<https://doi.org/10.1109/TKDE.2022.3174204>

Kartal, E. (2019). The artificial neural network modeling of language learning challenges of French-speaking students learning Turkish as a foreign language: *The case of France. Kuram ve Uygulamada Eğitim Bilimleri/Educational Sciences: Theory & Practice*, *19*(1), 55–77. <https://doi.org/10.12738/estp.2019.1.0264>

Kuhl, P. K. (2011). Early language learning and literacy: Neuroscience implications for education. *Mind, brain, and education*, *5*(3), 128-142.

Lazar, V. (2022). *An Investigation of the Relationship between the Use of Modern Digital*

Technologies, Language Learning Strategies, and Development of Second Language Skills (Doctoral dissertation, The University of North Dakota).

- Leaver, B.L. & Atwell, S. (2002). Preliminary qualitative findings from a study of the processes leading to the advanced professional proficiency level (ILR 4). *Developing Professional-Level Language Proficiency*, 260–279.
<https://doi.org/10.1017/cbo9780511606465.014>.
- Lin, C. J., & Hwang, G. J. (2018). A learning analytics approach to investigating factors affecting EFL students' oral performance in a flipped classroom. *Journal of Educational Technology & Society*, 21(2), 205-219.
- Lin, C.-J., & Mubarak, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society*, 24 (4), 16–35.
- Lockiewicz, M. (2019). The role of motivation in learning foreign languages: Towards a global self-identity – a literature analysis. *Current Issues in Personality Psychology*, 7(2), 81–90.
<https://doi.org/10.5114/cipp.2019.84569>
- Marshall, C., & Rossman, G. B. (2015). *Designing qualitative research (6th ed.)*. Thousand Oaks, CA: Sage.
- Mertler, C. A. (2019). *Introduction to educational research (2nd ed.)*. SAGE.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative Data Analysis: A methods sourcebook (4th ed.)*. Thousand Oaks, CA: Sage.
- Mir Muhammad Ahmedani, Ammara Zafar, Sadia Munir, & Hira Komal. (2022). Investigating the problems in speaking English as a second language: A study of Sindh University

- undergraduate students. *Sindh Journal of Linguistics*, 1(1), 1–16.
<https://doi.org/10.58921/sindhjol.v1i1.6>
- National Cryptologic University (n. d.). American Council on Education.
<https://www.acenet.edu/National-Guide/Pages/Organization.aspx?oid=4a099b28-9016-e811-810f-5065f38bf0e1>
- National University. (n.d.). *Procedures, data collection, and data analysis*. Applied Doctoral Proposal. <https://ncuone.ncu.edu/d21/le/content/321954/Home>
- Nggawu, L.O., & Thao, N.T. (2023). The impact of communicative language teaching (CLT) approach on students' speaking ability in a public Indonesian University: Comparison between introverts and extrovert groups. *International Journal of Language Education*, 7(3). <https://doi.org/10.26858/ijole.v7i3.50617>.
- Payant, C., & Zuniga, M. (2022). Learners' flow experience during peer revision in a virtual writing course during the global pandemic. *System*, 105, 102715.
<https://doi.org/10.1016/j.system.2021.102715>
- Rahman, A., & Tomy, P. (2024). Intelligent personal assistant - an interlocutor to mollify foreign language speaking anxiety. *Interactive Learning Environments*, 32(8), 4569-4586.
<https://doi.org/10.1080/10494820.2023.2204324>
- Richards, K.A. & Hemphill, M.A. (2018). A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical Education*, 37(2), pp. 225–231.
<https://doi.org/10.1123/jtpe.2017-0084>.
- Ruth, M. (2023). Confidence. *Salem Press Encyclopedia*.
- Rieder-Bünemann, A. (1970, January 1). *Second language learning*. SpringerLink.
https://link.springer.com/referenceworkentry/10.1007/978-14419-1428-6_826

- Salido, V. (2023). Impact of AI-powered learning tools on student understanding and academic performance. *Cavite State University*. <https://doi.org/10.13140/RG.2.2.17259.31521>
- Santos, M. E. C., Taketomi, T., Yamamoto, G., Rodrigo, M., Mercedes, T., Sandor, C., & Kato, H. (2016). Augmented reality as multimedia: the case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning*, 11(1), 1-23.
- Shao, S., Alharir, S., Hariri, S., Satam, P., Shiri, S., & Mbarki, A. (2022). AI-based Arabic Language and Speech Tutor. *2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)*, 1-8.
- Schoch, K. (2020). Case study research. *Research design and methods: An applied guide for the scholar-practitioner*, 245-258.
- Simion, O. M. (2022). The effects of Communicative Language Teaching (CLT) on the post-communicative methodology. *Research and Science Today*, 23, 93-98.
- Snyder, C. R., Lopez, S. J., Edwards, L. M., & Marques, S. C. (2021). *The Oxford Handbook of Positive Psychology*, 3rd ed (C. R. Snyder, S. J. Lopez, L. M. Edwards, & S. C. Marques (Eds.)). Oxford University Press.
- Sun, P. P. (2023). The influence of cognitive, affective, and sociocultural individual differences in L2 Chinese speech performance: A multiple indicators multiple causes approach. *Journal of Psycholinguistic Research*, 52(1), 217–239.
<https://doi.org/10.1007/s10936-022-09847-x>
- Taridi, M., Risnita, Yaakob, M. F. M., & Khairani, M. (2024). An evaluative study for communicative language teaching (CLT) on online teaching and learning in higher education: Indonesian and Malaysian university context. *Education and Information Technologies*, 29(9), <https://doi.org/10611-10647>.

- Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition, 41*(2), 363–387.
- The communicative language approach in ESL education* (2022). University of Louisiana Monroe. <https://online.ulm.edu/degrees/education/med/curriculum-and-instruction/communicative-language-approach/>
- Tonglo, G. G. R., Sirande, N., Tulaktondok, L., Patanduk, S. T., & Hadi, Z. R. (2022). Investigating EFL students' perspective on learning English pronunciation at English study program of Uki Toraja. *International Journal of Early Childhood Special Education, 14*(5), 7090–7096. <https://doi.org/10.9756/INTJECSE/V14I5.886>
- Vasileva, O., & Balyasnikova, N. (2019). (Re) Introducing Vygotsky's thought: from historical overview to contemporary psychology. *Frontiers in psychology, 10*, 1515.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard University Press.
- Wolcott, M. D., McLaughlin, J. E., Hann, A., Miklavec, A., Beck Dallaghan, G. L., Rhoney, D. H., & Zomorodi, M. (2021). A review to characterize and map the growth mindset theory in health professions education. *Medical Education, 55*(4), 430–440. <https://doi.org/10.1111/medu.14381>
- Xinyi Huang, Di Zou, Gary Cheng, Xieling Chen, & Haoran Xie. (2023). Trends, research issues and applications of artificial intelligence in language education. *Journal of Educational Technology & Society, 26*(1), 112–131. [https://doi.org/10.30191/ETS.202301_26\(1\).0009](https://doi.org/10.30191/ETS.202301_26(1).0009)
- Yang, H. & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective.

Australasian Journal of Educational Technology, 38(5), 180–210.

<https://doi.org/10.14742/ajet.7492>.

Yin, R. K. (2018). *Case study research and applications: Design and methods*. SAGE.

Eppler, M., & Ironsmith, M. (2021). Dweck's Mindset Theory as Context for Service Learning.

Journal of Community Engagement and Higher Education, 13(3), 3.

Appendix A

IRB Approval Letter

IRB-FY23-24-847 - Modification: Modification (Verification of Exemption) do-not-reply@cayuse.com <do-not-reply@cayuse.com>

Mon 4/8/2024 11:19 AM

To: Neguin Sarraf Caldwell <N.SarrafCaldwell9049@o365.ncu.edu>; pdavis@ncu.edu <pdavis@ncu.edu>



9388 Lightwave Ave. San Diego, CA 92123
irb@nu.edu

Notice of Exemption

April 8, 2024

To: Neguin Sarraf Caldwell

Project Title: Examining the Impact of AI-Assisted Language Learning Tools on Second Language Production

NU IRB Number: IRB-FY23-24-847

Modification Determination: Exempt from further review 45 CFR 46.101

Status: Active - Research activities may begin as of April 8, 2024 Dear

Neguin Sarraf Caldwell:

The study referenced above has been reviewed by the National University IRB. The IRB has determined the proposed modification to your research is exempt from further review under 45 CFR 46.101, which means you will not need to renew your study and may begin your study effective immediately. However, if you find the need to change your study in any way, you will need to submit a modification to the IRB prior to implementing the changes. This will allow the IRB to determine whether or not the study still meets exemption criteria.

Please review your Post Approval Responsibilities here: [Approved Documents Guidelines](#) For

any questions regarding your protocol, please reach out to the IRB at irb@nu.edu.

Handwritten signature of Joseph M. Marron in black ink.

Dr. Joseph Marron, IRB Chair

Handwritten signature of Brianne Mongeon in black ink.

Dr. Brianne Mongeon, Director, HRPP & IRB

Handwritten signature of Jenessa Eberhardt in black ink.

Jenessa Eberhardt, Associate Director, HRPP & IRB

Appendix B

Consent Form

Introduction

My name is Neguin S. Caldwell. I am a doctoral student at the National University (former Northcentral University) conducting a qualitative case study. I also hold a role as an assistant professor at the Extension Program (EP) of the Defense Language Institute (DLI).

I'm asking you to take part in a research study to explore the impact of AI-assisted language learning tools on second language learners' motivation and confidence in producing L2. The study's name is "Examining the Impact of AI-Assisted Language Learning Tools on Second Language Production."

You may participate in this research if you meet all of the following criteria:

1. You are 18 years or older.
2. You have studied a language for at least one year at the Defense Language Institute (DLI).
3. You have studied one of the following languages or dialects: Persian/Farsi, Dari, Arabic, or Chinese.
4. You completed the DLPT and OPI tests.
5. You are an intermediate or advanced-level linguist based on your DLPT and OPI scores.

I hope to include eight language learners in this research.

Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What you will be asked to do: If you agree to be in this study, you will be asked to do the following activities:

While the researcher observes the activities and takes notes, each participant in this study:

1. interact with the AI-supported language learning platform online for 20 minutes.
2. then complete a survey on the Word file lasting about 30 minutes.
3. upload the completed survey to Microsoft Teams' Assignment.

Each activity is optional, and participants are free to choose whether to participate.

During these activities, you will be asked questions about the following:

- impact of AI-assisted language learning tools on confidence and lowering anxiety

- impact of AI-assisted language learning tools on motivation and features or functionalities that could enhance motivation to speak the target language.
- **comfort level when using AI-assisted language learning tools.**
- strengths and weaknesses of AI platforms for language learning
- comparison of TalkPal AI to other AI-assisted language learning tools
- and last, if you achieve your personal language learning goals using these tools

Risks

There are minimal foreseeable risks or discomforts associated with this research. You can still skip any question you do not wish to answer, skip any activity, or stop participation at any time.

Benefits

If you participate, there are no direct benefits to you. This research may increase the body of knowledge in the subject area of this research.

Confidentiality

I will keep the records of this study private and take reasonable measures to protect the security of all your personal information. In any report I make public, I will not include any information that will make it possible to identify you. Taking part is voluntary: Participation in this study is completely voluntary. You may quit at any time.

Sponsorship: This project is sponsored by the Department of Defense. Authorized DoD representatives may review de-identified research records as part of their oversight authority. If you have questions: Please ask any questions you have now. If you have questions later, please contact us at N.SarrafCaldwell9049@o365.ncu.edu or (980) 288-1137. If you have any questions about your rights as a volunteer in this research, contact the Institutional Review Board (IRB) via email at irb@nu.edu. For questions concerning the Sponsor's Human Research Protection Program review, contact the Sponsor Component Office for Human Research Protections (COHRP) via email at NSA-hrpp@uwe.nsa.gov. This study has been reviewed and approved by the COHRP.

Name:

Date:

Signature:

Appendix C

Observation List

| | |
|--|--|
| Class number: | Date of the visit: |
| Week of Instruction: | Time of the activity: |
| Tools used: TalkPal AI | Activity: Roleplay, debate, Pic description |
| Observer: | Second Observer: |
| Observation Objectives: Interaction | Focus Area(s): Speaking Persian/Farsi |
| Number of Participants: 8 | Participant's name: |

I used this checklist to observe and record specific behaviors to facilitate my research on how participants produce the target language while interacting with an AI tool. Below is a structured list that includes various aspects such as verbal output, physical reactions, engagement levels, and problem-solving approaches. I applied this list to tick off observed behaviors for each participant:

Participant Observation Checklist

1. Verbal Output

- Uses complete sentences.
- Uses fragments or incomplete sentences.
- Uses correct grammatical structures.
- Makes grammatical errors.
- Uses target language vocabulary accurately.
- Uses incorrect or inappropriate vocabulary.

- Self-corrects verbal mistakes.
- Repeats phrases or questions.
- Pronounces words clearly.
- Has pronunciation difficulties.

2. Cognitive Processes

- Appears to think before speaking.
- Reacts impulsively.
- Uses logical reasoning during discussions.
- Demonstrates problem-solving skills.
- Seems to guess or make random responses.
- Reflects on and revises responses.

3. Use of Tools and Resources

- Refers to notes or textbooks.
- Uses online resources (e.g., dictionaries, translators).
- Notes down important information.
- Relies heavily on AI for answers.

4. Interaction with AI Features

- Initiates interaction without prompts.
- Responds only after prompts.
- Maintains eye contact with the screen.
- Shows interest through facial expressions (e.g., smiling, nodding).
- Uses gestures or other non-verbal cues.
- Utilizes specific AI features effectively (e.g., voice recognition, feedback).

- Interacts with AI to enhance understanding (e.g., asking follow-up questions).

5. Engagement

- Appears distracted (e.g., looking away, fidgeting).
- Displays confusion or frustration.
- Asks for clarification or repetition.
- Engages in turn-taking conversations.
- Monologues or does not engage in dialogue.
- Appears engaged and enjoys using AI tools.

6. Comfort Level

- Appears comfortable using AI tools.
- Expresses discomfort or hesitation.
- Needs assistance to use AI tools.
- Independently navigates AI tools.

Comments on the user-friendliness of AI tools:

7. Motivation

- Shows increased motivation to practice speaking.
- Expresses boredom or lack of interest.
- Expects features that could enhance motivation.

Comments on how AI impacts motivation:

8. Confidence and Anxiety

- Displays increased confidence in speaking.
- Shows reduced anxiety levels.
- Expresses feelings of anxiety or discomfort.
- Seeks reassurance or validation from AI.

Comments on how AI impacts confidence:

9. Feedback and Adaptation

- Responds positively to AI feedback.
- Utilizes feedback feature for correctness.
- Shows resistance or indifference to AI feedback.
- Adjusts behavior or responses based on feedback.
- Ignores feedback and continues previous behavior.

10. Technological Interaction

- Manages tools and interfaces proficiently.
- Faces difficulties with digital interfaces.
- Follows on-screen instructions accurately.
- Needs assistance with technical aspects.
- Adjusts settings for a better interaction experience (e.g., volume, display).

Additional comments:

Interaction with AI Feedback:

Attention to feedback: Does the participant pay attention to the feedback provided by the AI tool? (High/Medium/Low)

Incorporation of feedback: Does the participant incorporate the AI's feedback into their subsequent responses? (Always/Sometimes/Never)

Adapting communication: Does the participant adjust their communication style (grammar/vocabulary/ complexity) based on the AI's feedback? (Yes/No)

How many times do you use the feedback button? Throughout use of the tool. At least ...times

How many times do you use translation? I used translation about

Appendix D

Survey/Questionnaire Form

Qualitative questionnaire items:

After using an AI-assisted language learning platform:

1. How comfortable were you using AI-assisted language learning tools while speaking with the AI agent?

Motivation and Engagement

2. How did the AI-supported language learning platform impact your motivation to speak the target language? Did it make it enjoyable or engaging?
3. What features or functionalities within TalkPal AI do you think could further enhance your motivation to speak the L2?

Confidence and Anxiety

4. How did the AI-supported language learning platform impact your confidence in speaking the target language? Did it help reduce your anxiety? Please explain.

General Experience and Feedback

5. What do you consider are the biggest strengths and weaknesses of AI platforms, for example, TalkPal AI?
6. Have you used any other AI-assisted language learning tools? How does TalkPal AI compare to them?
7. Overall, on a scale of 1-10, how would you rate skills using technology and your experience using TalkPal AI for language learning?

Impact on Learning

8. **Personal goals:** Did TalkPal AI help you achieve your personal language learning goals? If so, how? How do you rate it on a scale of 1-10?

9. **Future applications:** How do you envision AI-assisted tools being used more effectively in language learning programs like the DLI-EP?

Thank you for your time and willingness to complete this survey/questionnaire.