

CHAPTER I

INTRODUCTION

1.1 Background of The Study

Artificial Intelligence (AI) has emerged as one of the most transformative innovations in education, reshaping how learning is delivered, personalized, and accessed beyond the constraints of the traditional classroom (Marzuki et al., 2026). Unlike earlier waves of educational technology, which largely digitized existing materials without changing the underlying mode of instruction, AI-based tools are distinguished by their capacity to respond adaptively to individual learners in real time, adjusting content, feedback, and pacing to the user's specific needs rather than delivering a uniform experience to an entire class (Ni'mah et al., 2025). This adaptive quality has made AI a particularly attractive vehicle for self-regulated learning, in which students take primary responsibility for managing their own practice with minimal reliance on teacher supervision or fixed classroom schedules (Ni'mah et al., 2025). As AI becomes further embedded in mobile and digital learning environments, its potential is no longer confined to content delivery alone; it increasingly extends into skills that have historically been the hardest to teach and practice at scale, English speaking chief among them.

Among the various language skills, speaking stands to benefit most directly from this shift, precisely because its difficulty is not purely linguistic

but deeply psychological. Previous studies have established that affective barriers, particularly speaking anxiety, significantly restrict oral performance more than receptive skills (Horwitz et al., 1986). This is further supported by Marzuki et al. (2026), who demonstrated that transitioning to autonomous digital platforms effectively lowers this psychological filter, thereby facilitating smoother language production. University-level EFL students frequently identify speaking as the most demanding skill to develop (Tom et al., 2013, as cited in Omar, 2023), a difficulty rooted less in gaps of grammatical or lexical knowledge than in speaking anxiety and the fear of negative judgment from the surrounding environment (Ayiz & Tauchid, 2024; Omar, 2023). Rather than functioning as an individual trait some learners simply happen to carry, this anxiety is actively reinforced by the social dynamics inherent to speaking in front of peers and instructors (Omar, 2023). Its behavioral consequence is a pattern of passivity: learners disengage from communicative risk-taking, choosing silence or minimal participation over the possibility of visible failure, even while recognizing the long-term academic and professional value of being able to speak English confidently (Salih & Omar, 2024; Younas et al., 2025). It is precisely this gap, between students' awareness of speaking's importance and their reluctance to actually engage in it, that positions AI's capacity for adaptive, non-judgmental feedback as a potentially significant intervention, one whose actual effect on learners' psychological experience of speaking remains to be understood.

This anxiety is not simply an individual disposition; it is reinforced by the structural conditions of conventional EFL classrooms. Widely used methods such as audiolingualism rely heavily on repetition and controlled drills rather than opportunities for spontaneous, self-directed speech production, meaning that even when students are speaking in class, they are often reciting rather than communicating in the kind of low-stakes, exploratory way that might ease their anxiety over time (Lightbown & Spada, 2011; Richards & Rodgers, 2014). Large class sizes compound this problem, leaving little room for every learner to receive individual speaking opportunities within a single session (Ayiz & Tauchid, 2024), while the lack of updated technology and instructional materials further limits teachers' ability to implement more responsive strategies (Salih & Omar, 2024; Younas et al., 2025). These structural constraints have pushed both teachers and researchers to look beyond the conventional classroom for new approaches, particularly toward self-regulated and gamified learning models that shift responsibility for practice onto the learner and reduce the social pressure associated with performing in front of others (Marzuki et al., 2026; Salih & Omar, 2024).

This search for lower-pressure, learner-directed speaking practice finds a concrete technological manifestation in Mobile-Assisted Language Learning (MALL), a development within Computer-Assisted Language Learning that leverages the portability of smartphones and tablets to extend speaking practice beyond the physical classroom (Marzuki et al., 2026). What makes MALL particularly relevant to the anxiety problem discussed above is not simply its

convenience, but its capacity to relocate speaking practice into a private setting, removed from the peer evaluation and performance pressure that characterize conventional classrooms (Sari & Rahman, 2025). The integration of Artificial Intelligence into MALL has made this shift considerably more consequential. Rather than simply digitizing practice, AI-driven applications are able to simulate real-world spoken interaction, functioning as a responsive conversational partner that delivers adaptive, instant feedback (Ni'mah et al., 2025). This feedback loop allows such tools to mimic the corrective role typically played by a teacher, without reproducing the social risk that role carries in a classroom setting (Denistiani, 2025; Shakira et al., 2026). Applications such as ELSA Speak, Duolingo, and Stimuler have emerged as leading examples of this approach.

Among these, Stimuler AI stands out as an application purpose-built for continuous oral production. It guides users through structured speaking practice via an AI speech-coach persona, evaluating pronunciation, fluency, grammar, and intonation in real time (Stimuler – AI English Speaking Coach with Instant Feedback). Beyond general conversation practice, the application offers impromptu speaking tasks, in which users respond to random topics under time pressure, as well as full-length IELTS speaking simulations complete with instant band-score estimates. Rather than simply assigning a score, Stimuler identifies specific mispronounced words and sounds and generates targeted practice sessions to correct them, a level of granular, corrective feedback that distinguishes it from more general-purpose language-

learning applications. Its popularity reflects this positioning: the application has surpassed 8.5 million users and 5 million downloads on Google Play, holds a 4.7-star rating, and was recognized as Google Play's Best AI App of 2023 (Stimuler Team, 2023). This shift culminates in the advancement of Mobile-Assisted Language Learning (MALL), a framework that encompasses both autonomous practice and anxiety-reducing methods. The integration of Artificial Intelligence (AI) into MALL has reshaped how speaking practice is approached, making this particularly relevant in contemporary language education. AI-driven applications such as ELSA Speak, Duolingo, and specifically the Stimuler app, offer simulated real-world interactions and two-way dialogues that function as personalized speaking partners (Denistiani, 2025; Ni'mah et al., 2025; Shakira et al., 2026). The key advantage of these AI applications lies in their ability to provide adaptive, instant feedback and facilitate repeated, autonomous rehearsal outside the classroom.

Yet this aggregate popularity does not necessarily reflect a uniform ground-level reality. A preliminary review of user reviews on Stimuler's Google Play listing, conducted independently by the researcher prior to this study, suggested a more mixed picture: some users described genuine improvement in pronunciation and speaking confidence, while others reported technical frustrations, such as the AI call feature failing to register spoken input, responses being cut off mid-speech, and pronunciation feedback that did not account for regional accent variation. These observations are anecdotal and are not presented here as established findings; rather, they point to a possible

gap between the application's broad, developer-reported success and the variability of individual user experience, a gap this study is positioned to examine directly among a specific and, so far, understudied group of learners.

Several studies have examined how Stimuler specifically is received by EFL learners. In a quasi-experimental design, Denistiani (2025) found that vocational high school students who used Stimuler achieved a markedly larger gain in speaking post-test scores compared to those taught through conventional methods, pointing to the application's measurable impact on oral performance. Using a narrative thematic review, Shakira et al. (2026) similarly found that Stimuler contributes positively to learners' speaking development, particularly in pronunciation, fluency, confidence, and classroom participation. Most directly relevant to the present study, Sari and Rahman (2025) conducted a qualitative case study examining senior high school students' experiences and perceptions of using Stimuler in their speaking classes, finding that the application positively contributed to students' motivation, confidence, and fluency, while also surfacing practical challenges such as internet connectivity issues and difficulty navigating certain app features. Taken together, these studies converge on a favorable view of Stimuler's impact on EFL learners' speaking development. However, all three were conducted exclusively at the high school or vocational level, leaving the experience of university-level EFL students, whose speaking anxiety may be shaped by different academic and social pressures, entirely unexamined.

To substantiate this contextual gap with local empirical evidence, a preliminary observation survey was conducted by the researcher involving English Language Teaching Department (ELTD) students at the target university. The survey revealed that while all participants possessed adequate digital infrastructure and prior familiarity with general language learning platforms (such as Duolingo and Elsa Speak), their self-perceived speaking confidence in academic settings remained highly polarized, ranging from high self-efficacy to severe speaking apprehension. Furthermore, students expressed a strong willingness and curiosity to integrate AI into their oral practice, primarily motivated by the desire to find a low-pressure, non-judgmental environment that could supplement their formal classroom learning. These preliminary findings confirm that speaking anxiety and fluctuating self-confidence are present realities among local ELTD students, further justifying the need to investigate how this specific group of tertiary learners perceives and experiences a dedicated AI speaking coach like Stimuler AI in their autonomous learning journey.

Taken together, both these gaps and preliminary research conducted point to the same unresolved question from two different directions. On one hand, the tension between Stimuler's aggregate popularity and the inconsistency of individual user experience, as noted above, suggests that students' actual, lived perspective on the application may diverge from what download counts, star ratings, or developer claims would suggest. On the other hand, the existing body of research on Stimuler, favorable as it is, has yet to

reach university-level EFL learners, whose speaking anxiety and classroom pressures differ from those of the high school and vocational students studied so far. Both gaps converge on the same population: students in the English Language Teaching Department (ELTD), who face many of the psychological barriers to speaking discussed earlier in this chapter, yet whose perspective on an AI-based speaking tool such as Stimuler remains undocumented. This gap forms the basis for the present study, which seeks to explore how ELTD students perceive and experience using Stimuler AI as a speaking practice tool, particularly in relation to their speaking anxiety and confidence, as well as the challenges and benefits they encounter along the way.

1.2 Identification of The Problem

Based on the background of the study, the specific problems identified are:

1. The psychological experience of English Language Teaching Department students in using AI-based speaking tools such as Stimuler, particularly with regard to speaking anxiety and confidence, remains unexamined, as existing research on Stimuler has been confined to high school and vocational-level learners.
2. The perceived challenges and benefits of using Stimuler AI for speaking development among university-level EFL students have not yet been documented, despite the application's wide adoption and generally positive reception in aggregate.

1.3 Limitation of The Problem

This research is limited to exploring the perspective of six purposively selected sixth-semester English Language Teaching Department (ELTD) students at Universitas Maritim Raja Ali Haji on using Stimuler AI as a speaking practice tool, based on their experience across a structured seven-day trial period. As a qualitative case study, this research is bounded by its participants, its focus on a single application, and its seven-day timeframe, and does not aim to produce findings generalizable beyond this particular case. The scope of inquiry is limited to participants' subjective accounts of their speaking anxiety, confidence, and perceived challenges and benefits, gathered through pre-interviews, a daily learner logbook, and post-interviews, rather than to measuring or proving any actual change in their speaking fluency. Accordingly, this study does not employ any quantitative instrument or pre-/post-test design to assess speaking performance before and after using the application.

1.4 Problem Formulation

Based on the background of the research, the problems formulated in this research are as follows:

1. How do English Language Teaching Department students perceive their speaking anxiety and confidence in using Stimuler AI as a speaking practice tool?
2. What challenges and benefits do students perceive in using Stimuler AI as a speaking practice tool for their speaking development?

1.5 Research Objective

Based on the research questions formulated above, this study aims to:

- a) To explore how English Language Teaching Department students perceive their speaking anxiety and confidence in using Stimuler AI as a speaking practice tool.
- b) To identify the challenges and benefits perceived by students in using Stimuler AI as a speaking practice tool for their speaking development.

1.6 The Significances of the Study

1.6.1 Theoretically

This study is expected to contribute to the development of English as a Foreign Language (EFL) teaching, particularly regarding the use of technology in language learning. It also provides additional insight into how Artificial Intelligence (AI) is experienced and perceived by learners in supporting speaking activities. Furthermore, this study is expected to enrich existing research on students' perspectives toward AI-based speaking tools, particularly within the context of university-level EFL learners.

1.6.2 Practically

- a) For Students

This study is expected to give students the opportunity to reflect on their own speaking practice experience through exposure to an AI-based speaking tool, and to provide insight into how such tools might serve as an alternative medium for independent speaking practice.

b) For Teachers

This study can serve as a reference for teachers in understanding students' perspectives toward AI-based tools, which may inform how such tools are introduced or integrated into speaking instruction.

c) For Future Researchers

This study can be used as a reference for further research related to the use of AI in language learning, particularly in exploring students' experiences and perspectives toward AI-based speaking tools at the university level.

1.7 Definition of Key Terms

a) Stimuler AI

Stimuler AI refers to an AI speech coach persona application that guides users through structured conversational practice, evaluating their speech across pronunciation, fluency, grammar, and intonation in real time.

b) Speaking Anxiety

In the context of this study, speaking anxiety refers to the affective barriers, psychological discomfort, and emotional apprehension experienced by ELTD students when required to perform oral communication in English. In this research, speaking anxiety specifically encompasses two main dimensions: fear of negative social evaluation (e.g., scrutiny from peers or lecturers in high-stakes classroom settings) and situational nervousness triggered by spontaneous speech requirements or complex vocabulary demands.

c) Perspective

Perspective refers to an individual's subjective views, attitudes, and interpretations formed through personal experience. In this study, perspective is operationalized as students' expressed thoughts, feelings, and evaluations regarding their experience of using Stimuler AI as a speaking practice tool, including their perceived challenges, benefits, and emotional responses such as confidence or anxiety. When operationalized in the plural form (perspectives) within this research, it denotes the diverse range of distinct viewpoints and multifaceted experiences shared by the six ELTD students regarding their 7-day trial of using Stimuler AI.

d) Mobile Assisted Language Learning / MALL

Mobile-Assisted Language Learning (MALL) is operationally defined as the pedagogical practice of utilizing mobile smartphone applications to facilitate autonomous and flexible foreign language learning outside the traditional classroom setup.

e) EFL Students

EFL (English as a Foreign Language) students are learners who study English in a context where it is not used as the primary language of daily communication (Lightbown & Spada, 2011). These learners typically have limited exposure to authentic English interaction outside the classroom, which may affect their language development, particularly in productive skills such as speaking.