



## **CHAPTER I**

### **INTRODUCTION**

#### **A. Background of the Study**

Speaking is considered one of the most difficult skills for EFL learners to master, as they often struggle with several aspects at once, including limited vocabulary, grammatical inaccuracy, pronunciation problems, and lack of fluency and confidence when communicating in real time (Normawati, et al., 2023). Unlike reading or writing, speaking allows little time to plan, so learners often produce hesitations, silent pauses, or fillers such as “uh” and “um,” which signal planning and word search rather than only lack of knowledge (Clark, et al., 2002). Empirical evidence shows that lower-proficiency learners tend to produce more hesitations, repetitions, and filled pauses than higher-proficiency learners, suggesting that disfluency reflects heavier cognitive demands and less developed automaticity in L2 speech production (Albarqi & Tavakoli, 2023).

Research shows that teachers often employ a variety of interactive and student-centered strategies to improve students’ speaking proficiency, such as role-plays, group discussions, storytelling, task-based learning, and technology-assisted activities. While these methods help create engaging learning environments and promote speaking practice, their effectiveness can vary depending on contextual factors like class size, time constraints, and learners’ individual differences (Tarmin et al., 2024). These activities do increase practice opportunities, but they can be constrained by large class sizes and limited instructional time; when classes are large, many students receive little individual speaking time (Xuan, 2024). Technology has also been used to improve speaking skills. Some examples include pronunciation apps, speech recognition programs, and online speaking tools that give automatic feedback. These technologies allow students to practice individually and receive quick responses about their pronunciation or rhythm (Shahab et al., 2024). Recent meta-analytic findings show that automatic speech recognition (ASR) technology can effectively improve learners’ pronunciation, particularly in segmental features, and provides immediate feedback that supports individualized practice (Thi-Nhu et al., 2024). However, since ASR-based training typically emphasizes controlled pronunciation tasks rather than extended spontaneous speaking, its impact on overall speaking fluency remains limited in comparison to communicative interaction.

Although many approaches have been developed to improve students' speaking fluency, most still have certain weaknesses when applied in real classrooms. Communicative tasks like role-plays and discussions do help students speak more, but they often depend on peer participation and classroom atmosphere. In large classes, for instance, many students remain passive listeners rather than active speakers. Similarly, pronunciation or speech training apps can support independent practice, but their responses are usually limited to fixed expressions and do not mimic real, two-way conversations (Tampubolon et al., 2025). These tools often lack emotional engagement and spontaneity, which are two elements that make real communication effective and enjoyable (Tampubolon et al., 2025). Consequently, many traditional classroom and digital speaking activities still lack a truly interactive speaking partner. While students may repeat phrases or respond to pre-set prompts, these exercises seldom mirror real dialogue, where speakers must interpret tone, adapt meaning, and respond to unpredictable cues. Genuine communication involves unpredictability and emotional feedback, which motivate learners and contribute to greater fluency development over time.

This gap provides an opportunity to explore the potential of emerging technologies such as voice-enabled conversational agents in language learning, particularly ChatGPT Voice. Recent studies suggest that AI-driven speaking tools can expand students' access to authentic oral practice and create more flexible learning opportunities (Pratiwi et al., 2024; Safitri et al., 2025). Compared to earlier applications that relied mainly on text input, voice-based models allow for more natural, two-way spoken interaction and can be used conveniently beyond classroom settings (Weran & Sa'adah, 2025). ChatGPT Voice, in particular, represents one of the most advanced yet accessible tools available today, offering highly natural speech, a human-like tone, and a warm conversational flow that can encourage learners to speak more confidently. While initial findings are encouraging, there remains limited empirical evidence on how effectively such tools improve EFL learners' speaking fluency in measurable terms, and on whether the amount of practice is related to the degree of improvement. Therefore, this study aims to quantitatively examine the impact of ChatGPT Voice on EFL students' speaking fluency through a one-group pre-test/post-test design and statistical analysis, including an examination of the relationship between practice duration and fluency gain.

## **B. Problem Formulation**

The research question is as follows: Does the use of ChatGPT Voice Mode significantly improve the speaking fluency of EFL students as measured by pre-test and post-test scores?

## **C. Objectives and Significance of the Study**

This study aims to determine whether the use of ChatGPT Voice Mode significantly improves the speaking fluency of EFL students as measured by pre-test and post-test results. This study is also expected to provide both theoretical and practical contributions to the field of English language teaching.

### **1. Theoretical Benefit**

This study aims to contribute to the growing body of literature on technology-enhanced language learning by providing quantitative evidence regarding the effectiveness of AI-based speaking tools, particularly ChatGPT Voice Mode, in improving EFL students' speaking fluency. The findings are expected to strengthen theoretical understanding of how interactive AI technologies can influence oral language development.

### **2. Practical Benefit**

For teachers and institutions, this study may serve as an evaluation tool to determine the potential of ChatGPT Voice Mode as a supplementary method for enhancing speaking instruction. The results could guide educators in integrating AI-driven voice interaction into existing curricula to create more engaging and accessible learning environments. For students, this study offers an opportunity to engage in independent speaking practice beyond the classroom. Through interaction with ChatGPT Voice Mode, learners can practice speaking in a low-pressure, self-paced environment that encourages continuous improvement in fluency and confidence.

## **D. Scope and Limitation**

This study focuses specifically on examining the impact of using ChatGPT Voice Mode on the speaking fluency of EFL students enrolled in the fourth semester of the English Department at Madura University. The scope of the research is limited to the measurement of speaking fluency, which is operationalized through three quantitative indicators: speech rate, mean length of runs, and pause frequency. These indicators reflect the speed and continuity (breakdown fluency)

aspects of oral production and do not cover other dimensions of speaking proficiency such as accuracy, complexity, pronunciation quality, or interactive competence. Furthermore, the alignment of quantitative fluency measures with CEFR levels in this study is interpretive and indirect, as it is based on descriptor correspondence rather than standardized CEFR testing procedures; therefore, the assigned levels should not be interpreted as official proficiency classifications.

The study employs a pre-experimental one-group pre-test and post-test design. All participants practiced speaking using ChatGPT Voice Mode over a span of two weeks, with each participant's total practice duration recorded individually, as usage time varied across participants; no separate control group was used. The findings therefore apply only to similar learning conditions and participant characteristics, and should not be interpreted as evidence of causality in the absence of a control group.

Several limitations should be acknowledged. First, the study focuses solely on short-term improvement and does not explore long-term retention or sustained fluency development beyond the post-test. Second, the measurement of fluency relies on recorded speech samples in controlled test tasks, which may not fully represent students' fluency in natural, extended conversation. Third, technological factors, such as the quality of internet connection, device performance, or students' familiarity with AI-based tools, may influence the consistency of the treatment. Because this study uses a one-group design without a control group, the observed improvement in fluency cannot be fully attributed to ChatGPT Voice Mode alone, as other factors such as maturation or repeated testing may have contributed. Finally, the sample size is relatively small (11 participants, after three of the initial 15 were excluded for not completing the speaking practice sessions), which may limit the generalizability of the findings to broader EFL populations. Despite these limitations, the study is expected to provide valuable empirical insights into the role of AI-based voice interaction in enhancing speaking fluency within an EFL context.

## **E. Definition of Key Terms**

To avoid ambiguity, the following key terms used in this study are defined operationally:

### **1. ChatGPT Voice Mode**

ChatGPT Voice Mode refers to the voice-based feature of the ChatGPT system that enables real-time spoken interaction between the user and the AI. In this study, it is used as the primary treatment tool for speaking practice.

## **2. Speaking Fluency**

Speaking fluency refers to the smoothness, continuity, and speed of spoken language production. In this research, fluency is operationalized through three measurable indicators: speech rate, mean length of runs, and pause frequency. These indicators capture the speed and breakdown (continuity) aspects of speech (utterance fluency). In addition, the results are interpreted in relation to CEFR fluency descriptors to provide a proficiency-oriented perspective on learners' performance.

## **3. EFL Students**

EFL (English as a Foreign Language) students refer to learners who study English in a context where English is not the dominant language. The participants in this study are forth-semester English Department students at Madura University.

