CHAPTER II

REVIEW OF RELATED LIERATUE

The researcher provides an explanation of the review of relevant literature in this chapter. There are several topics to cover, including the literature review, previous studies, and the conceptual framework. This section presents studies on proficiency in speaking, speaking instruction techniques, and websites as learning resources in relation to the literature review.

2.1 Definition of Speaking

Speaking is one of needed proficiencies by students to evaluate how they communicate with other individuals (Kurniasih, 2016). The proficiencies which need to enhance includes reading, writing, listening, and speaking. These skills are pivotal for English as a Foreign Language (EFL) students navigating diverse linguistic environments. Speaking proficiency is instrumental in enabling EFL students to communicate clearly and confidently, thereby enhancing their engagement in conversations, presentations, and discussions. Mastery of core skills not only augments language fluency but also cultivates overall confidence and competence in English communication, thus equipping EFL students with the necessary tools to excel academically, professionally, and socially in English-speaking settings (Folse and Bologna, 2003).

According Harmer (2007), speaking is a capacity to communicate effectively and takes not just an understanding of linguistic elements but also simultaneous understanding of language and information. It includes processing language and information at the same time as well as a deeper comprehension of linguistic elements. This suggests that learning the language's mechanics is simply one aspect of fluent speaking; another is being able to understand and react to spoken language in authentic contexts. To communicate concepts clearly and interactively, people must, in essence, seamlessly combine their cognitive and linguistic expertise when speaking.

Palmer (2014) also outlined the fundamental skills necessary for achieving proficiency in English communication, presenting a multifaceted approach. These competencies cover pronunciation, fluency, vocabulary, grammar, and discourse management. Pronunciation proficiency involves the accurate enunciation and intonation crucial for clear understanding. Fluency is characterized by the smooth flow of speech, devoid of hesitations and with a natural rhythm. Vocabulary mastery spans the breadth and depth of words, allowing for precise and varied expression. Grammar proficiency ensures the correct structure of sentences, contributing to coherence and clarity in communication. Lastly, discourse management skills involve the strategic organization of thoughts and effective participation in conversation, fostering meaningful interaction.

Speaking is a crucial skill for EFL students as it enables them to communicate clearly and confidently in academic, professional, and social contexts. Mastery of speaking involves not only understanding grammar and vocabulary but also the ability to process information and language simultaneously. Effective speaking

a combination of cognitive and linguistic skills, including accurate pronunciation, fluency, vocabulary range, grammatical accuracy, and the ability to manage conversations meaningfully.

2.1.1 Types of Speaking

According to Brown (2004), further states that there are some basic types of speaking as in the following taxonomy:

1) Imitative, at one end of a continuum of types of speaking performance is the ability to simply parrot back (imitate) a word or phrase or possibly a sentence. The imitative helps students improve their speaking skills by imitating native speakers' pronunciation, intonation, and speech patterns. It involves repeated practice of model sentences or dialogues to build fluency and accuracy. To enhance its effectiveness, imitative often uses authentic materials, choral or individual repetition, and meaningful activities like role-plays, giving students structured opportunities to practice speaking in a systematic and engaging way. Overall, the imitative speaking aims to provide learners with structured and systematic practice opportunities to develop their speaking skills through imitation and repetition Behzadi & Fahimniya (2014).
2) Intensive, a second type of speaking frequently employed in an assessment context, is the production of short stretches of oral language designed to demonstrate competence in a narrow band of grammatical, phrasal, lexical, or phonological relationships. To achieve this, intensive speaking offers various structured tasks to help learners improve their speaking skills through focused practice. These include directed response tasks, read-aloud activities to enhance

pronunciation and fluency, sentence completion for grammar and vocabulary practice, picture-cued tasks to encourage descriptive speaking, and translation exercises to deepen understanding of language structure. Intensive creates an interactive learning environment that builds students' confidence, accuracy, and fluency in oral communication.

3) Responsive, responsive assessment tasks include interaction and test comprehension but at the somewhat limited level of very short conversations, standard greetings and small talk, simple requests and comment, and the like. The stimulus is almost always a spoken prompt (in order to preserve authenticity), with perhaps only one or two follow-up or retorts. Responsive speaking assessment focuses on assessing how well students communicate spontaneously, using criteria such as fluency, accuracy, coherence, and interaction skills. This type of speaking allows teachers to give immediate, personalized feedback and uses tools like rubrics or recordings to measure performance effectively and support students' speaking development.
4) Interactive, the difference between responsive and interactive speaking is in the length and complexity of the interaction, which sometimes includes multiple exchanges and multiple participants. Interaction can take the two forms of transactional language, which have the purpose of maintaining social relationship.
5) Extensive (monologue), extensive oral production tasks include speeches, oral presentations, and story-telling, during which the opportunity for oral interaction from listeners is either highly limited (perhaps to nonverbal responses) or ruled out altogether.

The speaking type assessed in this research is responsive, as reflected in the design of the assessment. Test-takers are prompted to produce short stretches of discourse (typically no more than a sentence) in direct response to prompts or questions, demonstrating their linguistic ability at a specified level through immediate, contextually appropriate replies. These tasks align with what Madsen (in Brown, 2003) terms 'limited-response tasks', Underhill (1987) labels 'mechanical tasks', or classroom pedagogy's 'controlled-response tasks' all characteristic of responsive speaking's focus on targeted, reactive language use.

2.1.2 Component of Speaking Skill

According to Harmer (2007), speaking is a complex skill because at least it is concerned with components of grammar, vocabulary, pronunciation, fluency, and comprehension. Speaking has some important components, there are:

1) Grammar, is defined as the system of rules that govern the structure of sentences in a language. It encompasses the way words are combined to form phrases, clauses, and sentences, as well as the rules that dictate how these elements interact with one another.
2) Vocabulary, is the knowledge of words, phrases, collocations (words that often go together), and expressions needed to express ideas, opinions, and information effectively, focusing on choosing the right words and mastering "chunks of language" (ready-to-use phrases) for fluency.
3) Pronunciation, is the ability to produce English phonemes, word stress, sentence stress, rhythm, and intonation in a comprehensible manner. The goal is clarity of

communication, not a perfect native accent. It also includes mastery of suprasegmental features (intonation, rhythm, stress) that are vital to meaning.

4) Fluency, is the ability to speak fluently, flowingly, and at a reasonable pace, without excessive pauses, repetitions, or self-corrections that interfere with communication. Fluency relates to automaticity in language production and communication confidence. It's not just about speaking quickly, but also about maintaining a consistent flow of speech.
5) Comprehension, is the ability to understand spoken language spoken by the interlocutor to respond appropriately and relevantly.

2.2 Technology in Language Learning

2.2.1 Definition of CALL (Computer-Assisted Language Learning)

Computer Assisted Language Learning (CALL) refers to the use of technology including software programs, online resources, and mobile apps—to facilitate language learning and teaching. As defined by Gruba (in Lafford, 2005), it encompasses “the search for and study of applications of the computer in language teaching and learning”. Typically supplementing traditional instruction, CALL allows customization to accommodate individual learning styles.

CALL fundamentally functions as a supplementary tool that enriches language pedagogy. It does not replace the teacher’s role but serves as a technology-based facilitator providing: (a) flexible learning environments, (b) access to authentic materials, (c) instant feedback, and (d) personalized practice. Crucially, as Chapelle (2001) emphasizes, CALL’s effectiveness hinges on three factors: pedagogical design, curriculum integration, and students content technology. In

contemporary contexts, CALL has evolved into Web-Based (WBLL) and Mobile-Assisted Language Learning (MALL). Web Speechace a cloud-based platform using AI speech recognition exemplifies this evolution, positioning itself as a WBLL tool for pronunciation training in EFL settings like Indonesia.

2.2.2 Automatic Speech Recognition (ASR) in Language Learning

Automated Speech Recognition (ASR) converts audio streams of speech into written text. ASR is still imperfect but improving rapidly in terms of its accuracy in recognising spoken discourse and transcribing it into written text. ASR is based on big data-searching language corpora and finding matching patterns in data in order to convert the audio into written text. However, it does not analyse the audio semantically. The ASR output cannot assess meaning or coherence – it is not the same as Natural Language Processing which parses and analyses language. It merely transcribes speech and turns spoken language into written language using complex statical and language analysis model.

Recent advancements in ASR technology, particularly through deep learning and neural network models, have significantly enhanced its ability to recognize diverse accents and noisy environments. For instance, end-to-end systems like Speechace and Google's WaveNet have reduced error rates by learning directly from raw audio data, making ASR more accessible for EFL students with varying pronunciation backgrounds. Moreover, the integration of ASR into mobile language apps demonstrates its scalability for personalized, on-the-go practice. However, these improvements do not eliminate core pedagogical challenges; designers must still prioritize learner-centric features such as adjustable difficulty levels and

contextual feedback mechanisms to bridge the gap between transcription accuracy and language acquisition.

ASR serves as a powerful tool for converting speech to text and providing immediate phonetic feedback, thereby supporting pronunciation and fluency development in language learning. Its underlying technology relies on pattern recognition from large corpora, though it lacks semantic understanding. While ASR like Web Speechace offer valuable practice opportunities, their effectiveness is moderated by challenges such as accent bias and the need for complementary human instruction. Thus, the successful integration of ASR in EFL contexts requires both technological refinement and pedagogical scaffolding to address students' specific needs, particularly in under-resourced settings like rural Indonesia.

2.3 Web Speechace

Web Speechace is an AI-powered, web-based language learning platform designed to enhance learners' English-speaking skills through automatic speech recognition (ASR) and real-time feedback. It evaluates learners' oral production across several aspects, including pronunciation, fluency, vocabulary usage, and grammatical accuracy (Ningsih, 2024). According to Yuniar (2023), Speechace functions as a self-paced speaking assessment tool, which allows students to practice speaking anytime and receive instant corrective feedback without waiting for a teacher evaluation. The aligns with Computer-Assisted Language Learning (CALL) principles, which emphasize technology as a medium to facilitate independent and interactive language learning.

In the context of English as a Foreign Language (EFL) education, Web Speechace serves as a valuable resource for enhancing speaking skills among students. Its design promotes autonomous learning by providing students with the flexibility to engage in practice sessions at their own pace, which is particularly beneficial in large classroom settings where individual feedback from teachers may be limited. The platform's instant feedback mechanism allows learners to identify and correct pronunciation errors immediately, fostering a more effective learning experience. Additionally, the integration of various speaking tasks, such as role-plays and dialogues, aligns with the communicative language teaching approach, encouraging students to apply their skills in realistic contexts.

Web Speechace represents a significant advancement in language learning technology, combining ASR capabilities with pedagogical strategies that support independent practice and immediate feedback. Its alignment with CALL principles enhances its effectiveness as a tool for developing speaking skills, particularly in EFL contexts where traditional methods may fall short. For students at SMAN 1 Srengat, the platform not only provides a means to improve their English proficiency but also empowers them to take charge of their learning journey. As this study evaluates the efficacy of Web Speechace, it will be essential to consider both its technological strengths and the pedagogical frameworks that support its integration into the classroom.

2.3.1 Features of Web Speechace

The core foundation of Web Speechace is the integration of cutting-edge speech recognition technology. This technology allows the system to analyze the user's speech in real time, compare it to a native speaker model, and then identify

areas for improvement. This approach significantly helps learners hone their pronunciation accuracy, both at the level of individual sounds and suprasegmental aspects such as intonation and word stress patterns.

One of Web Speechace most prominent features is its provision of instant and comprehensive feedback. After the user pronounces a word or sentence, the system immediately displays a visualization indicating pronunciation accuracy, often with color indicators or graphics highlighting correctly or incorrectly pronounced sections. Additionally, learners can listen to their own voice recordings and compare them directly with the audio recordings provided by native speakers. According to experts in language pedagogy, this kind of rapid and specific feedback is vital because it allows learners to self-correct efficiently and strengthens muscle memory associated with speech production.

In addition to detailed feedback, Web Speechace also provides diverse and structured practice modules. These platforms typically offer a variety of activities, from practicing pronunciation of isolated words or phrases to practicing complex sentences, paragraphs, or even simulated dialogues. Progress tracking features allow learners to monitor their progress over time, identify common error patterns, and focus on areas that need the most attention. Some modules are even designed to build learners' confidence in specific communication contexts, such as presentations or interviews. This flexibility allows learners to adjust the pace and focus of the exercises to suit their individual needs.

2.3.2 Advantages and Disadvantages Using Web Speechace in Class

Web Speechace offers significant pedagogical advantages for English-speaking instruction in classroom settings. According to Liakin et al. (2017), its capacity to provide immediate, individualized pronunciation feedback reduces the feedback gap that often occurs in large classes, allowing students to self-correct efficiently. This instantaneity fosters learner autonomy and encourages frequent practice—factors directly linked to pronunciation improvement (Chen et al., 2020). Additionally, teachers can leverage its analytics dashboard to identify class-wide weaknesses and adjust instruction accordingly (Hwang & Fu, 2019). The platform's gamified elements, such as progress badges and leaderboards, further enhance student motivation by transforming rote practice into engaging challenges, aligning with Kurikulum Merdeka's emphasis on student-centered active learning (Kemdikbud, 2022).

Despite these strengths, Web Speechace faces several limitations when implemented in EFL classrooms. ASR's inherent accent recognition bias can undermine pronunciation scoring for learners with strong regional accents, such as Javanese-accented English, leading to frustration and misdiagnosis (Tatman, 2017). This technological constraint is compounded by infrastructure challenges; schools like SMAN 1 Srengat may experience unstable internet connectivity, disrupting practice sessions despite Speechace offline caching feature (Kusuma & Dewi, 2023). Pedagogically, over-reliance on the platform may overshadow communicative practice, as students focus on scoring well rather than contextual language use (Blake, 2016). Teachers must also dedicate time to develop strategies

for integrating Speechace into lesson plans a hurdle for educators with limited CALL training (Darmawan et al., 2024).

Web Speechace presents a potent yet imperfect solution for speaking skill development. Its advantages personalized feedback, actionable insights for instructors, and student engagement mechanisms make it a compelling tool for augmenting limited human resources in Indonesian high schools. Conversely, its disadvantages, particularly technological bias and infrastructural dependencies, necessitate teacher mediation and localized adaptation. Ultimately, its efficacy is optimized not as a standalone solution, but as a scaffolded component within a balanced Communicative Language Teaching framework, where teacher-guided interaction contextualizes machine-driven practice (Hubbard, 2009).

2.4 Efficacy of Using Web Speechace for Speaking Skill

Research demonstrates that Web Speechace significantly improves pronunciation accuracy among EFL students. In a 12-week study by Chen et al. (2023), Indonesian high school students using Speechace for daily 15-minute drills showed a 27% reduction in vowel/consonant errors compared to the control group. The platform's granular feedback identifying misarticulations like /ɪ/ vs. /iː/ ("ship-sheep") enables learners to recalibrate tongue positioning and airflow in real-time (Ningsih, 2024). This targeted remediation is particularly effective for addressing L1-interference patterns, such as Javanese speakers' tendencies to devoice final consonants (e.g., "dog" → /dɒk/). Quantitative metrics from Speechace scoring algorithms correlate strongly with human raters in phonemic accuracy (r = .78, Li, 2022), validating its reliability for foundational skill-building.

Beyond pronunciation, Speechace enhances oral fluency by reducing hesitation frequency and improving speech rhythm. Yuniar (2023) documented a 41% increase in words-per-minute and 33% fewer pauses among XI-grade students practicing with Speechace dialogue simulations. The tool's prosody visualization mapping intonation contours onto pitch graphs helps learners internalize English stress-timing patterns, counteracting syllable-timed L1 habits (Rachmawati et al., 2024). Crucially, Speechace reduces speaking anxiety by providing a judgment-free practice environment; 89% of low-confidence learners in Gilakjani's (2024) study reported increased willingness to participate in classroom discussions after consistent ASR practice. Gamified elements like progress badges further sustain motivation (Hwang & Fu, 2024).

Despite these gains, Web Speechace exhibits limitations in developing pragmatic and strategic competence. While excelling in bottom-up skills (phonemes, fluency), it lacks contextual awareness to evaluate discourse coherence, register appropriateness, or nonverbal communication (Blake, 2016). Kusuma & Dewi (2024) found no significant improvement in students' ability to structure persuasive arguments using Speechace alone—skills requiring teacher-guided instruction. Additionally, infrastructural challenges affect efficacy: unstable internet in Eastern Java caused task disruptions for 34% of learners (Rachmawati, 2023), while accent biases led to inconsistent scoring for Javanese-accented interrogatives (e.g., rising intonation misread as uncertainty).

In synthesis, Web Speechace proves highly efficacious for developing mechanical speaking sub-skills. Its strength lies in democratizing access to high frequency, personalized feedback critical in large Indonesian classrooms, where

teacher attention is scarce. However, its impact on higher-order speaking proficiency remains limited without human mediation. Thus, Speechace achieves optimal efficacy not as a standalone solution, but as an integrated component within blended-learning frameworks. Future studies should explore hybrid models pairing Speechace with teacher-led communicative tasks to bridge bottom-up and top-down speaking competencies.

2.5 Previous Studies

The following earlier studies on the use of websites to teach speaking has been uncovered by the researcher and supports this research as follows:

Table 2. 1 *Previous studies*

Writer's Research Title	Other Research Titles	Year	Researcher Names	Similarities	Differences
The Effectiveness of Web Speechace to Improve Eleventh Grade Students' Speaking Skill at SMAN 1 Srengat 2025	The Effectiveness of Speech-Ace Website on Students' Pronunciation	2024	Anggytha Mutiara	The research sampling technique	The research design, the research approach, the participants of research, and the data analysis.
	The Impact of Speechace on EFL Learners' Pronunciation Competency	2025	Nguyen Ngoc Vu. et al.	The research approach, the independent variable.	The research sampling technique, the research design, the dependent variable, the participant, and the research data analysis
	Improving Junior High School Students' English-Speaking Skills Using the Speechace Application.	2019	Arry Pauji H. et al.	The dependent and independent variable, and the data analysis	The sampling technique, The research approach, the research design and the participants.
	Speechace Facilitate Students' Pronunciation Accuracy in the English Class	2023	Siska Yuniar	The dependent variable, the sampling technique; purposive sampling.	The research approach, the research design, the research objectives, participants, the independent variable, and the data analysis
	Utilizing Speechace to Enhance Speaking Skills among English as a Second Language Pre-University Students	2024	Mazliyana Zainuddin. et al.	The research media	The research objectives, participants, the research design, the independent variable, and the data analysis.

A previous thesis entitled “The Effectiveness of Speech-Ace Website on Students’ Pronunciation” was carried out by Mutiara (Mutiara et al. 2024). This research was categorized qualitative research which utilized the design of case study, in which the samples were taken from two groups (control class and experimental class). The data were taken by observation and interview. The research took participants from Tidar University, who were in third semester. The researcher carried out a pretest before training either the students in the control class or experimental class. The trainings were implemented for experimental class. At the end of days, the researcher gave a post-test to the experimental class. The result showed that there was more improvement towards the experimental class after treatment with Speechace than the result of their pre-test.

Another case relating to speaking problem was also was found by Nguyen. et al. (Nguyen et al. 2025) entitled “The Impact of Speechace on EFL Learners’ Pronunciation Competency”. The research was utilized a mixed-method experimental design, and the research involved an experimental group using Speechace and a control group relying on traditional methods. Quantitative data were collected through pretests, post-tests, and questionnaires. Sampling was conducted randomly by selecting two EIC4 classes, namely EIC4-0500 and EIC4-0200. Both classes were selected as the whole group in the research, where class EIC4-0500 (n=37) participated as the experimental group and class EIC4-0200 (n=37) as the control group. The results of the pre-test and post-test given after the treatment showed differences between the two groups. In the control group, the average score increased from 30.93 to 51.8, while in the experimental group it increased more significantly from 31.47 to 58.03.

The next research was under the title “Improving Junior High School Students’ English-Speaking Skills Using the Speechace Application” conducted by Arry Pauji (2024). This research employed a quasi-experimental design involving two distinct groups: an experimental class that utilized Speechace in instructional sessions, and a control class receiving conventional teaching methods. Participants were selected through random sampling, with 25 seventh-grade students from class 7.1 assigned to the experimental group and 25 from class 7.2 to the control group. Analysis of pre-test assessments revealed that the experimental group had scores ranging from 56 to 81 (M:69.08, SD:7.000), while the control group scored between 55 and 84 (M:68.24, SD:7.742). Following the intervention, post-test results showed significant improvement in the experimental group (min:80, max:92, M:85.20, SD:3.594) compared to more modest gains in the control group (min: 67, max: 89, m: 77.80, sd: 5.212).

On the other hand, Yuniar (Yuniar 2023) also conducted a set of experimental research entitled “Speechace to Facilitate Students’ Pronunciation Accuracy in the English Class”. The research was categorized exploratory research with exploratory design, and the sample was 31 students of science. The researcher collected the data by doing survey in the form of questionnaires, conducting interview, and focus group discussion and analyzed those using transcription, coding, and constructing the ideas, categorizing the topics, drawing the themes, and drawing summaries to find out if there how Speechace impacted on their speaking and pronunciation. In conclusion, there were positive impacts towards the students’ speaking proficiency. The conclusion was said that the use of Speechace was effective enough to enhance students’ speaking ability, and pronunciation accuracy.

Lastly, there was a matter regarding to have a communication. It was the research conducted by Mazliyana Z. et al (Adillah 2022) entitled “Utilizing Speechace to Enhance Speaking Skills among English as a Second Language Pre-University Students”. After providing mixed-method research, the use of Speechace by students and the advantages they believe it offers at a Malaysian public university. Speechace is a useful tool for helping ESL students get better at speaking. The application of Speechace has been shown to aid students in both quantitative and qualitative research. Speechace enables students to imitate accurate pronunciation and facilitates conversation, providing a stress-free setting for language development.

The current research and previous research have both similarities and divergences. There are a number of similarities between this research and previous research in the realm of language learning technologies. The Space website served as a platform for the delivery of instructional materials in both this research and the previous research. Additionally, the participant groups in both cases consisted of high school students, indicating that adolescent language learning was the main focus. Both studies use an experimental study research method, namely a one-group pretest-post-test technique, demonstrating an effective methodological alignment. This research makes it possible to compare the speaking skills of the students before and after the intervention. Lastly, while the precise focus and resources may change, the overall objective of both the earlier and present studies is to help students become more proficient speakers.

The following are the differences between previous research and current research. First, there is a basic methodological difference: this research uses

quantitative research, whereas previous studies used qualitative research. Second, the research has a different objective. While earlier studies focused mostly on pronunciation, this research attempts to enhance speaking abilities in general, covering a wider range of linguistic capabilities than just pronunciation accuracy. Therefore, even though the goal of both studies was to improve students' speaking skill, the scope of this research and its use of the Web Speechace. A clear change in strategy and research objectives.

2.6 Thinking Framework

The need to master English-speaking skills has long been recognized as critical (Baker and Westrup, 2003). In today's digital era, this demand necessitates further developments in teaching methods. The context of the Independent Curriculum provides opportunities for teachers to be creative in learning, including by integrating innovative technology or media to improve students' communication skills. The use of technology-based learning media such as Web Speechace is based on the view that technology can provide authentic and engaging language learning experiences (S. E. . Smaldino et al. 2019). Based on a preliminary study at SMAN 1 Srengat involving two eleventh-grade classes (XI-C1 as experimental class and XI-E2 as try-out class for instrument validation the researcher identified an urgent need for innovative media in speaking instruction. This aligns with Smaldino, et al (2019), who emphasize that media serves as a supportive tool for English language learners by offering authentic and interactive language practice.

Relating to previous studies, the use of Speechace have primarily focused on isolated aspects of language learning, such as pronunciation and vocabulary enhancement. However, there is a gap in research regarding its effectiveness in

using web Speechace for students' speaking skills, which includes fluency, grammar, vocabulary, and pronunciation. Therefore, this study aims to investigate whether Web Speechace can be effective for students' speaking skills, not just specific components, and to evaluate its potential in increasing students' engagement and ease in speaking English accurately and fluently.

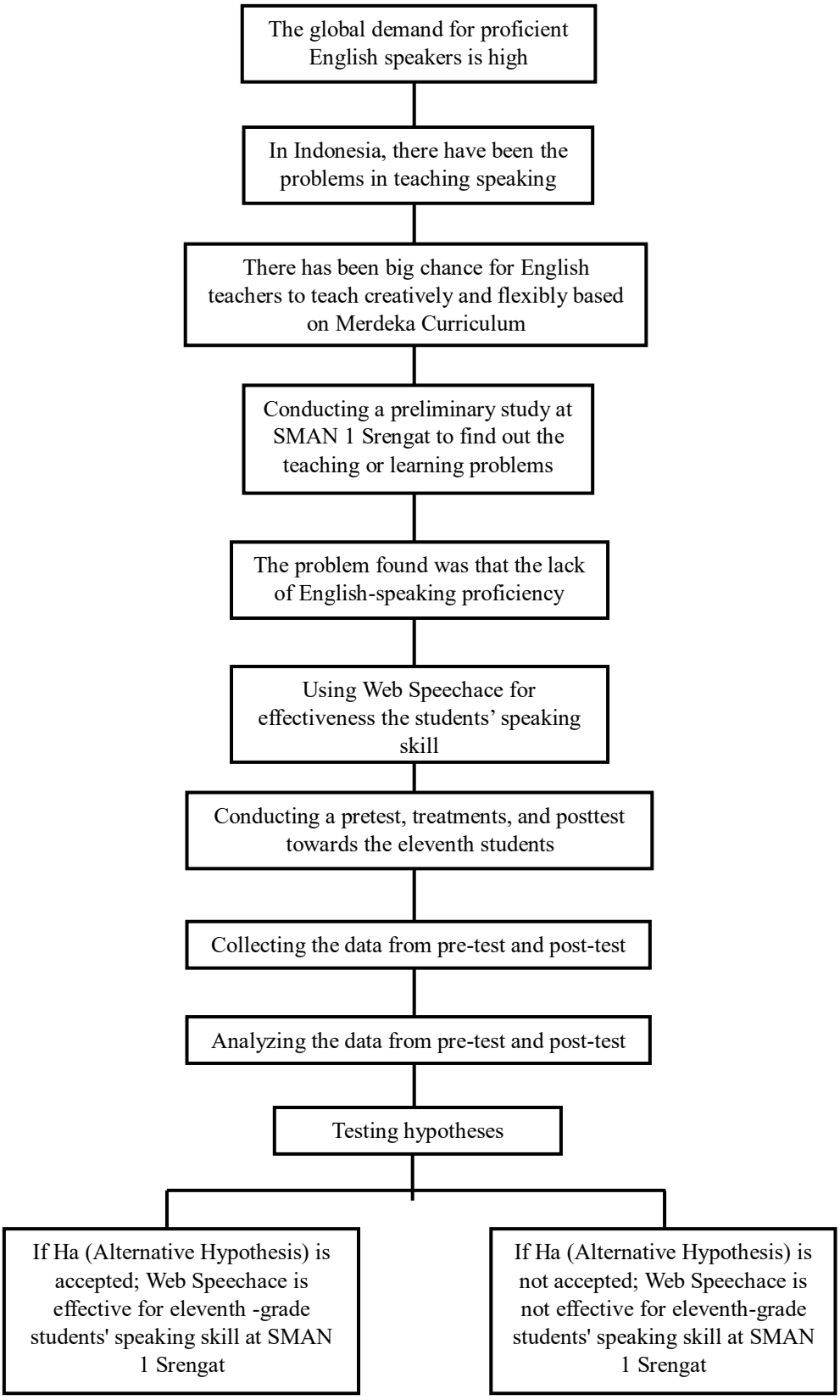

***Figure 2. a** Thinking Framework*