

CHAPTER I

INTRODUCTION

This chapter presents 1) research background, 2) problem identification, 3) problem limitation, 4) problem formulation, 5) research objectives, 6) research significance, 7) development specification, 8) development assumption, and 9) definition of key terms.

1.1 Research Background

English for Specific Purposes (ESP) is an English subject that prepares learners to use English in professional settings aligned with their field of employment. Anthony (2018) and Basturkmen (2021) agree that English for Specific Purposes (ESP) teaches English to target students according to their specific occupational needs, focusing on skills, genres, and related language expressions. Therefore, the emphasis is on practical English skills rather than theories, since professional human capital plays a critical role in operating positions in the industry while implementing English mastery.

In Bali, the tourism industry makes the largest contribution to the economy as an international destination. Data from the Central Statistics Board of the Bali Government (Badan Pusat Statistik, 2023) showed that 6.33 million tourists visited Bali in 2024, a significant 20% increase over 2023, when only 5.23 million visitors spent their holidays on the island. This circumstance demands more qualified human resources in the hospitality and accommodation sectors, where employees are required to have strong English proficiency in both written and spoken forms.

To perform their responsibilities, they must also possess good competencies,

including appropriate attitudes, product knowledge, and practical skills. As stated by Ratminingsih (2021), competence is the ability that involves skill, knowledge, and attitude, which must be possessed by every worker to carry out their duties effectively. These three integrated aspects are essential to achieving competency across several areas of expertise (Fadlia et al., 2022). Therefore, communication skills, especially in English, become crucial for hospitality workers to interact effectively with international visitors and colleagues.

To meet the demand for professional English communication, English for Specific Purposes (ESP) has been integrated into tertiary and vocational education (Ibrahim, 2010). ESP emphasises the practical and professional use of English (Simanjuntak & Ratmanida, 2024), ensuring that institutions offering vocational education prioritise industry-relevant English skills to meet both national and global employment standards. Through mastering ESP, human resources can maintain their value and competitiveness, supporting the continuous growth of the hospitality industry. ESP subjects in tourism vocational programs vary by professional role, such as English for Travel Operation, Front Office, Culinary, Telephone and Correspondence, Event and MICE, and Food and Beverage Service. Integrating ESP into the curriculum strengthens graduates' adaptability and competitiveness in dynamic professional environments by enabling them to function effectively in target workplace situations (Ratminingsih et al., 2018).

In Bali, many Vocational Education Training (VET) institutions focus on the hospitality industry. The increasing popularity of vocational education is due to students' desire for faster access to employment opportunities. Learners are trained through competency-based units in specialities such as front office, housekeeping, food and beverage service, bartending, and culinary arts, following a curriculum

aligned with the SKKNI (Indonesian National Working Qualification Standard) issued under Ministry of Manpower Regulation Number 125 in 2024. Jolly Roger International (JRI) is one of these VET institutions in Bali that offer tourism training under this standard. One of the general subjects taught is English for Specific Purposes (ESP), aimed at preparing future hoteliers during off-the-job training before they implement their skills during on-the-job training.

This circumstance requires instructors to adapt real-world professional contexts into classroom materials. As Widodo (2016) emphasises, needs analysis is essential in providing ESP learning materials tailored to industry requirements. The perspective of the tourism industry is important since it provides insights into what practical skills are essential for learners (Nova & Lastari, 2024). This connection bridges the gap between classroom learning and workplace expectations, so learners will not be surprised when entering professional environments.

An initial observation conducted in JRI's English for Food and Beverage class revealed that the module used as the main learning source lacks learning activities that limit student participation. Furthermore, the available English module for ESP was not fully appropriate for meeting students' needs (Yuliantini et al., 2024) as the module used by instructors mainly contained conversational examples in restaurant and kitchen settings and lacked a systematic structure aligned with the syllabus. It provided a limited vocabulary, a few language expressions, and monotonous exercises. The vocational learners need to develop both productive and receptive skills (Konstantinidou et al., 2023; Resmayani & Putra, 2022). However, the current module did not include sufficient reading or listening instructions, meaning learners were forced to produce output in speaking and writing without adequate input through reading and listening practice.

Focusing on listening practices, initial interviews with instructors showed difficulties in finding appropriate listening audio materials relevant to hospitality and cruise contexts. The available audio on the internet or from other sources was not aligned with current industry conditions, especially given technological changes such as digital payment systems replacing cheques. This mismatch between available materials and real-world practices limits learners' exposure to authentic communication. In addition, familiarisation with some English accents was found to be crucial for the learners. This is due to the need to forecast understanding of spoken English across various accents in teamwork communication. Therefore, developing updated, context-specific listening materials is essential to enhance learners' readiness for actual hospitality or cruise ship operations or for working overseas. In response, material personalisation using Generative Artificial Intelligence (GenAI) offers a promising approach for creating customised and relevant audio resources (Karimi et al., 2024; Mraković et al., 2025; Natalia, 2024). Integrating technology into education is also seen as inevitable (Santosa et al., 2022; Sari et al., 2023).

Generative AI (GenAI) is a type of artificial intelligence that can autonomously create realistic digital content, including text, images, audio, and video (Sakirin & Kusuma, 2023). Lv (2023) defines GenAI as an advanced AI system capable of generating new content and producing highly customised creative outputs. Similarly, Takale et al. (2024) explain that GenAI simulates human creativity through advanced algorithms and deep learning techniques, making it transformative in creating realistic and innovative content.

In educational contexts, AI integration is increasingly utilised for its significant advantages, although it also presents challenges. Santosa et al. (2023)

emphasise that to achieve optimal learning benefits, teachers must be selective and effective in their use of AI technology. AI expands students' opportunities to access learning materials and to practice independently and remotely (Trinh, 2024). Nguyen & Nguyen (2025) highlight that AI serves as a key instructional tool supporting teachers in providing effective media for learning, while Yeh (2024) states that AI integration can revolutionise language education by enabling innovative pedagogical approaches tailored to diverse learners' needs. Huang et al. (2023) add that AI effectively supports the teaching of four language skills: listening, speaking, reading, and writing, as well as language elements such as vocabulary, pronunciation, and grammar.

Recent research trends have shown growing attention to the use of AI in English language education. Liando et al. (2025) found that AI offers positive impacts from both students' and teachers' perspectives. Rahmi et al. (2024) discovered that the AI writing assistant ParagraphAI improves grammar accuracy and coherence, though it still requires refinement in content depth. Kusuma et al. (2024) verified that using ChatGPT as a GenAI tool provides comprehensive and relevant information sources despite some limitations in accuracy. Trinh (2024) reported that ChatGPT creates an interactive and personalised learning environment that enhances students' vocabulary acquisition. Similarly, Liao et al. (2023) and Kostikova et al. (2023) affirmed that ChatGPT is applicable for teaching all four language skills and useful in developing professional English courses.

Specifically focusing on AI-generated audio, Fitria (2023) explored the use of NaturalReader, an AI-powered voice tool, for teaching listening in TOEFL classes and found it effective in improving listening delivery. Chuan et al. (2025) examined MetaListen and found that it provides context-relevant listening

materials, helping teachers develop comprehensive listening-practice tools. They concluded that MetaListen enhances students' listening skills and learning attitudes. Ly et al. (2025) further found that AI-generated audio transcriptions (AIGATs) assist students by reducing listening complexity, thereby improving comprehension and fostering positive perceptions of AI-assisted learning.

Based on the reviewed empirical evidence, it is clear that although AI in education has been widely explored, there remains a lack of research on the use of Generative AI in developing listening audio. In particular, studies focusing on integrating Generative AI to create listening materials for English for Specific Purposes (ESP) remain limited. Therefore, this research aims to develop an English for Food and Beverage Service module that incorporates Generative AI. The developed module integrates AI-generated audio for listening activities to support their language development. Given the promising potential of Generative AI to enrich ESP learning materials, this study is entitled “Developing an English for Food and Beverage Service Module with AI-Generated Audio for Vocational Students.”

1.2 Problem Identification

From the background above, there are a number of problems that can be identified as it is presented in the following points.

- 1.2.1 The module of English for food & beverage service is less structured, and need enhancement comparing to the syllabus content.
- 1.2.2 The module used in teaching and learning English for food & beverage service was found lack of listening audio and exercise.
- 1.2.3 The students need to be familiar with some English accents for employment preparation.

- 1.2.4 There is an imbalance between activities between receptive skills (reading, and listening) and productive skills (speaking and writing).

1.3 Problem Limitation

This study is limited in developing an English for Food and Beverage Service module with AI-generated audio for vocational students of a hotel and cruise line college (vocational training institution) particularly in Food & Beverage Service Major at Jolly Roger International, in academic year 2025/2026 located in Tabanan, Bali. The research specifies the scope into Food and Beverage major since many opportunities offered from man power agent and most of the students conduct their internships period abroad and apply positions in food and beverage service. In addition, this study is conducted by following the Research and Development (R&D) instructional design recommended by Branch (2009) which covers Analyze, Design, Develop, Implement, and Evaluate. Meanwhile, the AI tools used are Fliki.ai for generating listening audio.

1.4 Problem Formulation

Based on the above background, the research problems are formulated as follows:

- 1.4.1 What are the students' needs and gaps in existing English for Food & Beverage Service module that is essential to be developed?
- 1.4.2 How is the content of the English for Food & Beverage Service module designed and developed with AI-Generated audio?
- 1.4.3 How is the quality of the developed English for Food & Beverage Service module with AI-generated audio?

1.4.4 How is the perception of the students toward the developed Food and Beverage Service Module with AI-Generated audio?

1.5 Research Objectives

Based on the research question, the study was conducted for several objectives as follows:

1.5.1 To identify the learners' needs and gaps in existing English for Food & Beverage Service module.

1.5.2 To describe the content of the English for Food & Beverage Service module designed and developed with AI-generated audio.

1.5.3 To investigate the quality of the developed English for Food & Beverage Service module with AI-generated audio.

1.5.4 To investigate the perception of the students toward the developed Food and Beverage Service Module with AI-Generated Audio.

1.6 Research Significance

There are two significances aimed in this study, theoretically and practically. The significance can be mentioned as follows:

1.6.1 Theoretical Significance

The result of this study is expected to be a reference for further research regarding to the field of the implementation of AI-generated audio in developing learning material especially English for Food and Beverage Service for vocational school students. Furthermore, this study is aligned with multimodal learning theory which underlines that learning will be more active by utilization of multiple mode of representation such as text, audio, and visual through incorporating in the developed module.

1.6.2 Practical Significance

1) For English Teachers

This study is expected to give a contribution of knowledge for English teachers especially English for specific purpose in providing a contextual guiding book for learning material during the English language teaching process. The resulted AI-assisted module is believed to support the source of learning material besides open up teachers' perspective to actively utilize the current technology of AI in teaching English for the vocational students.

2) For Students

This study is expected to assist students in studying English with engaging module as it is equipped with AI-integration improving their English language proficiency and other skills that is necessary required in 21st century learning. In addition, since the module has online access audio, students are able to independently study and practice their listening skill which can promote autonomous learning outside the classroom, therefore, students will have more time to improve their English proficiency through practical English to communicate with international guests in hospitality context.

3) For Future Researchers

This study is expected to be empirical study for future researchers and enrich their references for those who are encouraged to explore the integration of AI technology in preparing English material for teaching English for Specific Purpose which is very beneficial and create innovative, interactive, and engaging learning situation. Besides, the result of this study is also possible for future researchers to investigate the effectiveness of AI-generated audio toward students learning achievement.

1.7 Development Specification

In this study, the development specification of English module for food & beverage service includes the characteristics and criteria of product development. The product was designed and developed in the form of module which is specified for vocational students particularly food & beverage major students in term of English subject. The scope of topics developed in this module follows the syllabus as well as the need resulted from the need analysis conducted before the design and development. Therefore, it is designed and developed for the purpose of English learning material both for students to learn and for instructors as instructional learning material. The product is designed altogether with AI integration especially AI-generated audio to support the novelty of the product.

In the context of product criterion, every chapter of the module was developed into words corner, relevant language expressions, language use, example of practical dialogue, various exercise, as well as equipped with listening audio composed from “app.fliki.ai”. In addition to the listening audio, it can be accessed on google drive through quick response (QR) code or by provided link as it is available online. This allows students and users in general to be able to learn and practice remotely and independently via their mobile phone, tablet, laptop, or personal computer.

1.8 Development Assumption

Based on the development specification above, the following assumptions can be inferred.

1.8.1. Vocational students will have better learning engagement during the use of English module for food & beverage service which is designed and developed with AI-generated audio.

1.8.2. The designed and developed English module for food & beverage service will have positive perception by the students on its use in English language teaching.

1.8.3. The instructors perceive positive view on the use of English module for food & beverage service which is designed and developed with AI-generated audio as instructional material.

1.9 Definition of Key Terms

The definition of the research key terms includes conceptual definition and practical definition as follows.

1.9.1 Conceptual Definition

a. Need Analysis

Need analysis is defined as a basic process of curriculum practices to design syllabus, material development, and instruction design to tailor what should be taught and learnt (Widodo, 2016). In a line with the above argument, Ratminingsih et al. (2021) also note that that need analysis refers to the process of collecting data about learners' need in the purpose of a consideration to design learning material, teaching activities, evaluation strategies, and designing test.

Lasekan et al. (2023) also clarify need analysis is the process of language and communications assessment required by the learners to suit their career, company, and cultural background. They obviously outline that by integrating three analysis types including present situation analysis (PTA), learning situation analysis (LTA),

and target situation analysis (TSA), it is more effective to align the need of the learners. As also empowered by Al-Malki et al. (2022) it is pivotal to carry to meet the specific purpose of learning.

Furthermore, Sudana et al. (2024) also elaborate that need analysis has a clear significance not only promote the relevancy of teaching material but also optimises the practical usage of the designed course. Lastly, Yuliantini et al. (2024) also argue that need analysis is significantly crucial to carry out in the purpose of design appropriate material in English for Occupational Purposes or ESP according to students' requisite. Thus, need analysis is one of the steps that is necessary to be conducted in the purpose of design curriculum, syllabus, or learning instruction which is believed to align the learning needs.

b. Material Development

Material development of learning refers to anything prepared by teachers, students, or writers that functions to provide information and experience of language to promote language learning (Tomlinson, 2011). There are six criteria suggested by Hutchinson & Waters (1987) in designing and developing materials such as comprises a particular goal; provide opportunities to practice the language; include specific instruction; the content is fascinating and able to engage the learners attention; authentic; and relevant for the learners' level.

Howard & Major (2004) also elaborate material development should be developed by considering these six factors involving (1) material should be contextualised, (2) stimulate interaction and be generative in the language form, (3) encourage learners to develop skill and strategies, (4) should allow for a focus on form and function, (5) should be integrated across language skills, and (6) should be authentic material.

Furthermore, to cope the learners different attitudes toward the module, Tomlison (2011) concern to: (1) equip with different types of text; (2) equip with different types of activities; (3) equip with providing optional extras for the more positive and motivated learners; (4) consider activities which ask the learners to discuss their perspective about the course and the materials; and (5) be aware of the cultural sensitivities of the learners.

c. English for Specific Purpose

Anthony (2018) defines English for specific purpose (ESP) is an approach to teach English language to the target students according to their occupational needs in which it focuses on skills, genres, and related language expressions. In addition, he pointed out that to accommodate students' needs, the use of general or discipline-specific teaching material and methods are significantly important to address.

Another expert also views English for specific purpose is the branch of English language teaching that has focus to seek and tailor students needs both for work or study-related requirement (Basturkmen, 2021). In order to provide specific purpose of the learners' need in the context of English for specific purpose, it is very crucial that need analysis is a must to conduct (Widodo, 2016). Moreover, Simanjuntak & Ratmanida (2024) discuss English for specific purpose as a branch of English language that focuses on studying specific needs mostly by a particular group of students which aligns to the practical and professional use of English.

Therefore, it is clearly stated that English for specific purpose is an English material which function to address learners relevant to their vocations and assist them in their workplace where English is used as a medium of communication.

d. AI-Generated Audio

AI or artificial intelligence is an advanced technology that has intelligence system developed into intelligent software with capability to imitate and create various type of database. According to Sakirin & Kusuma (2023) generative AI is a type of AI tools that has ability to create or generate realistic digital content autonomously which covers in the form of audio or visual (pictures or images).

Furthermore, Lv (2023) simply defines that generative AI as the form of an advanced AI machine that is autonomously able to generate new content such as text, images, audio, and video. He adds generative AI provides innovative method to produce creative content in the metaverse which is highly customized by the users.

Other experts also argue on generative AI that it is a transformative technology that has ability to create new content, simulate human creativity, besides able to create realistic output by the creative process through advanced algorithms and deep learning technique (Takale et al., 2024). Hence, AI-generated audio can be defined as a type of generative AI which specifies to create listening audio according to designed material.

e. Vocational Education

According to Mustapa et al. (2015), vocational education is a practice-oriented approach in education field that emphasize on workplace requirement, to meet the required skills by the workplace. This education requires students to follow the lesson and skills based on the standard designed by the particular company. UNESCO Institute for Statistics (as cited in Chiang et al., 2022) define vocational education as a program of education designed for learners to gain specific knowledge, skill, and competencies toward a particular occupation.

Furthermore, Chiang et al. (2022) convey that vocational education focuses on pre-service education and training in which learners are prepared to supply market labour's needs in specific type of occupation or profession. From this definition, it can be inferred that vocational students are students who are studying and prepared to learn in the purpose of targeted goal, for example students who are studying in hospitality especially in a restaurant, they are equipped with specific skills, knowledge, and attitude about restaurant operation.

f. Perception

There are several scholars who define the definition of perception in general. (Robbins & Judge, 2013) state that perception is a process for individuals in setting and interpreting their sensory impression in order to give meaning for their environment. Furthermore, Nuraini et al. (2021) clearly argues that Perception begins when an object generates a stimulus that is received by the sensory organs. As a fundamental psychological process, perception enables individuals to interpret and respond to various stimuli and phenomena in their environment. Integrated means that all senses are cooperatively work to organize and interpret the data received.

1.9.2 Operational Definition

a. Need Analysis

Need analysis in this study is defined as the first step which aims to investigate the needs of the learners in the context of English for Specific Purposes particularly in design English for Food & Beverage Service.

b. Material Development

Material development in this study concerned on developing English module for Food & Beverage Service. The module was designed and developed through the infusion of AI-generated audio to provide the fascinating and engaging content.

c. English for Specific Purpose

In this study, English for specific purpose is defined as an English subject that specializes to provide English material developed by using generative AI that create listening audio. The English for Specific Purpose is specifically designed for Food and Beverage Service students in hospitality industry.

d. AI-Generated Audio

AI-generated audio in this study refers to generative AI tool which is implemented to produce listening audio from a provided text. The AI software is called “app.flixi.ai”, it is an appropriate tool to create authentic audio for listening content.

e. Vocational Students

In this study, vocational students refer to the learners who study about food and beverage major. They are prepared to learn deeply about particular knowledge, skills, and attitude specifically about how to operate restaurant especially become a professional waiter/waitress in the purpose of supplying the industrial demands of workers. They are expected to study off the job training in a college during 6 months and planned to conduct internship program in hospitality industry at least 6 months as a part of the requirement before applying a professional job in hospitality especially hotel and restaurant area.

f. Perception

Due to there are module development and the integration of technology, the study uses the adaptation of ESP Material Quality theory and the Technology

Acceptance Model (TAM). ESP Material Quality focuses on goal relevance, language practice opportunities, instruction clarity, engagement/interest, authenticity, and difficulty appropriateness. Besides, TAM focuses on perceived usefulness and perceived ease of use.

