

A Descriptive Study of the Development Stage in Producing a Text-to-Speech Nursing English Audio Book

Risa Arroyyani¹, Maryani², Pipin Nurhayati³

^{1,2,3}Sekolah Tinggi Ilmu Kesehatan Surya Global Yogyakarta

*Penulis Korespondensi: rissa.arroyyan@gmail.com

Received: 15/07/2026 | Accepted: 07/08/2026 | Publication: 18/08/2026

Abstract: *This study describes the development stage of producing a text-to-speech-based English for Nursing audiobook, to continue previous research of designing situational syllabus conducted within the framework of the ADDIE model. A descriptive qualitative approach is employed in this study, specifically documenting the production process during the development stage. The process involved at this stage consists of drafting and revising the dialogue script, compiling a vocabulary list and functional expressions, audio production using Natural Reader text-to-speech, editing, and presenting the final result. The content of the audiobook is structured based on the target learners' language proficiency at the CEFR Independent User B1 level. The development process for the English for Nursing audiobook resulted in audio files lasting 20–25 minutes for each chapter, with each chapter divided into five sub-sections: Introduction, Reading Text, Dialogue, Simple Expressions and Vocabulary Enrichment, and Exercise.*

Keywords : text-to-speech, audio book, English for Nursing, ADDIE, development process



Copyright © 2026, The Author(s).

This is an open-access article under the CC-BY-SA license.

How to Cite: Arroyyani, R., Maryani, M., & Nurhayati, P. (2026). A descriptive study of the development stage in producing a text-to-speech nursing English audiobook. *Journal of Innovation Research and Knowledge*, 6(3), 1849–1863. <https://doi.org/10.53625/jirk.v6i3.13220>

INTRODUCTION

Recently, English proficiency has been required across all sectors. English competency is demanded not only in fields related to education or linguistics, but also in the medical sector, particularly nursing. The global opportunities resulting from the ASEAN Economic Community (AEC) have made English proficiency a skill that nursing professionals must master. Nurses are expected to communicate effectively with patients from different backgrounds, such as social status, language, culture, and religion, across various healthcare facilities (Chetsadanuwat, 2018). Language plays a crucial role in communication, including sharing information and emotions to prevent assumptions or miscommunication. Nurses typically spend more time with patients than other healthcare providers and are often the first point of contact when patients need help. Consequently, they are expected to listen to patients' personal experiences, which affect their physical and psychological recovery (Canziani, 2021). In addition to direct communication, nurses are required to master English because it serves as a medium for transferring medical knowledge through international journals and conferences (Iwata et al., 2023). Nurses with good English proficiency have broader opportunities in their nursing profession.

However, mastering English for nursing presents several challenges, particularly in

learning medical terminology, clinical contexts, and nursing-specific registers. For nurses with heavy work schedules, learning English is not an easy task. Nevertheless, nurses need to learn English as preparation before entering the workforce. English for Nursing Purposes (ENP) is part of English for Medical Purposes (EMP), which is itself a subset of English for Specific Purposes (ESP). In ENP, English is learned based on the actual needs and situations nurses encounter in clinical settings and nursing education (Canziani, 2021). (Richards & Schmidt, 2010) stated that the content and objectives of an ESP course are tailored to the specific needs of the learner group (Mohammadi & Mousavi, 2013). The English materials taught in ESP are not general English, but are specifically designed to meet the needs of that particular group of learners (Thepseenu, 2020). ESP emphasizes topics relevant to a specific profession or discipline, using actual workplace documents such as registration forms, brochures, and instructions as learning resources. Moreover, ESP is not limited to English language learning alone, but also encompasses intercultural competence — the ability to communicate with people from different cultural backgrounds (Montasser & Althaqafi, 2023).

ESP instruction covers all language skills, such as speaking, writing, listening, and reading. Each skill plays a different role, and proficiency in one influences the others. For instance, a lack of vocabulary mastery can affect a student's confidence and motivation, eventually leading to a lack of communication skills. Obstacles in mastering vocabulary at the General English level will also affect the learning of ESP (Al Zahrani & Chaudhary, 2022; Rizkina et al., 2025). Vocabulary acquisition is closely linked to speaking proficiency. It serves as a fundamental asset for mastering a foreign language and plays a crucial role in developing both oral and written skills (Dhuli & Guduru, 2025; Le & Trinh, 2024).

Other language skills, such as listening, also play a vital role in establishing effective communication. Mastering listening skills is closely related to understanding sounds and pronunciation. Individuals who can pronounce words accurately also tend to have strong listening skills (Talib, 2018). Consequently, pronunciation-related factors such as speaking speed, accent, and clarity of articulation influence listening proficiency, particularly in the context of ESP. With effective listening and imitation skills, learners find it easier to pronounce English words correctly. However, there are several limitations in listening instruction, such as the limited availability of medical and nursing-themed listening materials. Most existing materials are designed for General English rather than for contextualized English for Nursing. Furthermore, the lack of accent or voice variety in listening materials limits students' exposure to linguistic variation, which can affect their ability to imitate pronunciation and produce spoken language.

Developing proper listening materials is crucial for enhancing other productive skills. Therefore, it is necessary to develop listening instructional media, such as audiobooks, as part of ICT development in language education, specifically ESP (Pamelasari et al., 2024). Audiobooks offer a solution for listening practice that is accessible anytime and anywhere, while also positively influencing speaking skills (Arockiasamy et al., 2025). They also provide a listening medium beyond songs or news broadcasts, which can sometimes lead to student demotivation (Susilowati et al., 2021). Due to the limited availability of nursing-based listening materials, developing audiobook media offers a way to provide a wider variety of

nursing-related expressions and conversations in specific clinical situations.

Several studies have shown the contribution of audiobooks to fluency, pronunciation, and intonation accuracy, as well as enhancing speaking ability and confidence (Guerrero, 2025; Kartal & Simsek, 2017; Kohnová & Papula, 2020). Audiobooks have also been found to help develop listening and speaking abilities (Rizal et al., 2022). In addition, audiobooks contribute to vocabulary development, familiarize students with the language used in literary works, help them predict upcoming story events, support text comprehension, and enhance language awareness (Çarkıt, 2020; Herlyanti et al., 2023).

One alternative way to develop audiobooks is by using AI voices or text-to-speech (TTS) technology, an efficient and flexible solution for audiobook production. Text-to-Speech (speech synthesis) is a type of audio medium that converts text into speech. TTS is an Artificial Intelligence product used in language learning. Various TTS services are available that feature native speaker-like AI voices with diverse accents, enabling students to develop a range of language skills such as reading comprehension, vocabulary, and listening (Bione et al., 2020; Nurhamidah, 2024; Wajdi et al., 2021; Widiana et al., 2022). Some paid TTS applications offer more comprehensive features than standard versions, such as adjustments for emotion, tone, and pitch, as well as unlimited word or sentence processing (Fitria, 2022). Students at various educational levels benefit from the use of Text-to-Speech (TTS) technology. At the university level, TTS can be used to study intonation, word stress, pitch contours, and fluency when reading English texts (Al-Jarf, 2022; Bione et al., 2020). With the help of ICT as an integral part of language education (Constantinou & Papadima-Sophocleous, 2020), audiobook production has become easier. Audiobooks created using TTS are more efficient, as they do not require significant time for recording (Kılıçkaya & Kic-Drgas, 2026). Even without native speakers, developers can create audiobooks by leveraging AI voice or Text-to-Speech (TTS) technology, adjusting features such as intonation and tone (e.g., casual, formal, melancholic, or friendly) to suit their specific needs (Salem & Levitan, 2026).

Several previous studies have discussed the use of audiobooks in English language learning. (Herlyanti et al., 2023) developed an audiobook to teach listening for senior high school students, while (Pamelasari et al., 2024) explored English audiobook development for science teachers. (Pujayanthi et al., 2026) developed English language teaching audiobooks for students with special needs, and (Imawan & Ashadi, 2019) focused on audiobook use in EFL instruction. However, few studies have focused on developing audiobook-based listening materials for English for Nursing.

This study continues a previous study (Maryani & Arroyyani, 2024) that covered the Analysis and Design stages of the ADDIE model for an English for Nursing audiobook. ADDIE is a five-step research and development models, consisting of Analysis, Design, Development, Implementation, and Evaluation. Not all studies elaborate on every step of this model; for example, (Nuranisa et al., 2024) focused only on the Analysis phase, while (Nuswantara et al., 2021) focused only on the Design phase, since ADDIE is considered a flexible model that allows each phase to be reviewed before public launch. Therefore, this study aims to provide a detailed, descriptive account of the Development stage in producing an English for Nursing audiobook using AI voice/TTS technology.

RESEARCH METHOD

This study continues previous research that reached the Design stage of the ADDIE model, proceeding to the Development stage. It uses a qualitative descriptive design that focuses on documenting the development process rather than measuring product feasibility, detailing each step of producing audiobook listening materials with AI voice/TTS technology. The object of the study is a text-to-speech-based English for Nursing audiobook, developed based on the syllabus resulting from the prior analysis and design stages.

The data source is the researcher, who acted as both developer and reflective practitioner throughout the audiobook production process. The data include various versions and revisions of the dialogue scripts, AI voice settings, and descriptions of the audio output at each production stage, using Natural Reader as the AI voice tool. Data were collected through screenshots of the production process, script versions, AI configuration logs, and a reflective journal recorded throughout production. Data analysis followed a descriptive-thematic approach organized around five stages: (1) script drafting, (2) audio production using AI voice/TTS, (3) quality control (monitoring pitch, tone, etc.), (4) editing, and (5) final output. Trustworthiness was maintained through data source triangulation, involving documents, field notes, and audio outputs, as well as an audit trail for every script and audio revision. The research procedure is illustrated in the diagram below.

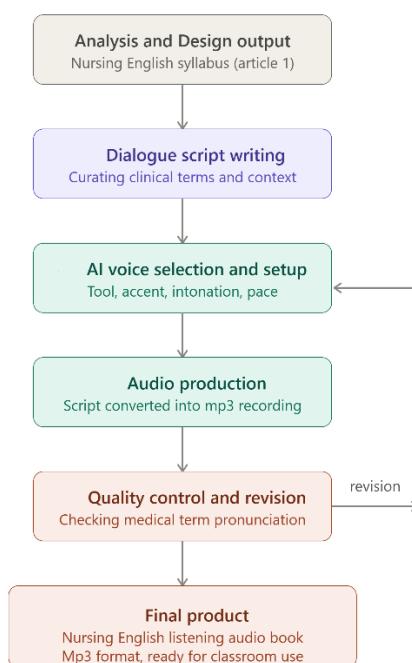


Figure 1. Diagram of Research Method

RESULTS AND DISCUSSIONS

The Development stage of the English for Nursing audiobook was carried out based on the Analysis and Design stage from the previous study (Maryani & Arroyyani, 2024), which served as the blueprint for the entire development process. That stage produced a ten-



unit syllabus based on a needs analysis: Admitting Patients, Asking Patients' Symptoms, Asking about Pain, Taking Patient's History, Hospital Staff, Hospital Departments, Instruments and Equipment, Taking Vital Signs, Assessing Patients, and Medication. Each unit comprises five subsections: (1) Introduction, (2) Reading Text, (3) Dialogue, (4) Simple Expressions and Vocabulary Enrichment, and (5) Exercises. Development was structured according to the CEFR Independent User B1 level (Modern Language Division, 2001), at which users can communicate using familiar sentence patterns of short to medium length, arrange simple phrases to describe experiences and plans, and give brief reasons for their opinions—a level appropriate for this audiobook.

The Development stage consists of four steps: (1) drafting and revising the dialogue scripts, (2) compiling a categorized vocabulary list from the dialogues, (3) audio production and editing, and (4) presenting the final output. This article focuses on the prototype of Unit 1, out of the ten syllabus units, to illustrate the audiobook development process.

Drafting and revising dialogue script

At this stage, the developer takes several steps to convert the expressions and vocabulary from the previously prepared situational syllabus into complete dialogue scripts for conversations between nurses and patients or between nurses and staff.

Tabel 1. Syllabus Unit 1 Admitting patient

Unit 1	Admitting Patients
Function:	1) Checking patient details 2) Asking patient personal information for medical record 3) Explaining a process/procedures
Expressions:	1) May I see your identification card? 2) Can you please provide me with some basic information? 3) Where do you live? 4) Who is your next of kin? 5) Who is your GP?
Vocabulary:	1) Personal information: name, date of birth, basic information, health insurance, next of kin 2) Previous and past medical history and allergies 3) Personal contact: emergency contact, family contact, family number contact.

In terms of oral communication, B1 users can communicate in everyday, routine situations and produce spoken utterances by linking phrases in a simple way to describe experiences and events (Modern Language Division, 2001). The expressions in Unit 1 represent simple, routine interrogative patterns that reflect B1-level accuracy, since routine patient admission is a predictable scenario. Furthermore, the grammatical structure and vocabulary are not highly complex, which suits the B1 proficiency level. The expressions remain consistent with terminology used in the nursing profession. Several terms were reviewed, particularly regarding redundancy and grammatical accuracy. Subsequently, the developer drafted the dialogue scripts as part of instructional content creation within the ADDIE Development stage, which includes:

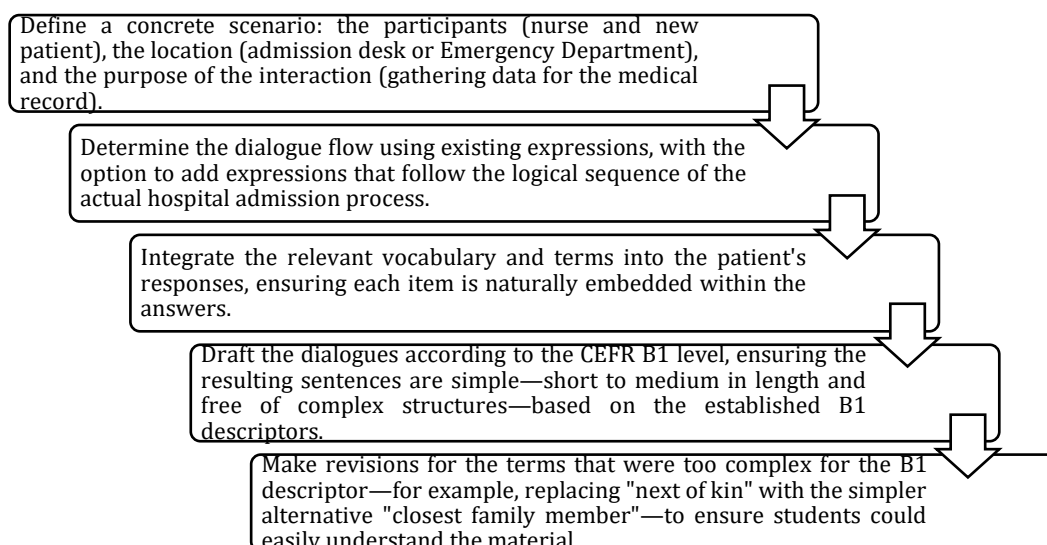


Figure 2. Stages in drafting and revising dialogue script

The following is a comparison of the initial and revised dialogues for Unit 1, "Admitting Patients," following the revision process.

Tabel 2. Initial and Revised Dialogue for Unit 1 Admitting patients

Initial Version (raw)		(Final Revised Version — CEFR B1)	
Patient	: (Walks up to the front desk where Ema, the medical record staff member, is sitting) Hello, I'm Linda. I have an appointment today with Dr. Hendra. This is my first visit.	Patient	: <i>(Walks up to the front desk)</i> Hello, I'm Linda. I have an appointment today with Dr. Hendra. This is my first visit.
Nurse	: (Smiles warmly) Hi, Linda! Welcome to our hospital. I'll be happy to assist you with the admission process. Let's get started. Can you please provide me with some basic information? I'll need your full name, date of birth, and the reason for your admission.	Nurse	: Hi, Linda! Welcome to our hospital. I'll help you with the admission process. Can you please provide me with some basic information? What's your full name and date of birth?
Patient	: Of course. My full name is Linda Wahyudi, and my date of birth is July 30, 1985. I'm here for an examination that Dr. Hendra scheduled.	Patient	: Of course. My full name is Linda Wahyudi. My date of birth is July 30, 1985. I'm here for an examination with Dr. Hendra.
Nurse	: Great, Linda! Thank you for providing the necessary information. Before we proceed, could you please show me some identification, such as your ID card, driver's license or passport, so we can verify your identity?	Nurse	: Okay. May I see your identification card, please?
Patient	: (Hands over her ID card) Sure, here's my ID card.	Patient	: <i>(Hands over her ID card)</i> Sure, here's my ID card.
Nurse	: (Checks the information and returns the license) Thank you. Now, I'll need to collect some medical history to update your records. Have you been to our hospital before, or is this your first visit?	Nurse	: Alright, let me check this. <i>(Checks the card and returns it)</i> Thank you. Now I need to ask about your medical history. Have you been to our hospital before?
Patient	: This is my first time at this hospital. I usually go to another one for my regular check-ups.	Patient	: No, this is my first time here. I usually go to another hospital for check-ups.
Nurse	: That's alright. Do you have any existing	Nurse	: I see. Do you have any medical conditions or allergies?
		Patient	: No, I don't have any medical conditions. I'm not allergic to any medicine.
		Nurse	: Alright, I'll note that down. Who is your GP?
		Patient	: My GP is Dr. Smith.
		Nurse	: And who is your next of kin?
		Patient	: My sister, Emily. Her phone number is



	medical conditions or allergies that we should be aware of?	0857-314-7865.
Patient	: No, I don't have any known medical conditions, and I'm not allergic to any medications.	Nurse : Let me confirm that number — 0857-314-7865. Is that correct?
Nurse	: Perfect. Next, I need the contact information of your primary care physician and an emergency contact in case we need to reach someone during your stay.	Patient : Yes, that's correct.
Patient	: My primary care physician is Dr. Smith, and I'll provide you with my sister's contact details as my emergency contact. Her name is Emily, and her phone number is 0857-314-7865	Nurse : Do you have health insurance?
Nurse	: Thank you. We'll make sure to keep that information on file. Now, regarding your insurance, do you have a health insurance plan that you would like us to bill for your hospital stay and procedures?	Patient : Yes, I have insurance with In Health. My policy number is 90876543890.
Patient	: Yes, I have health insurance through my company. It's with In Health, and my policy number is 90876543890.	Nurse : Okay, thank you. We'll check your insurance and take care of the payment. Do you have any questions?
Nurse	: Great! We'll contact your insurance provider to verify coverage and handle the billing. Is there anything else you'd like to ask or any concerns you have about your admission?	Patient : Not right now. Thank you, Nurse. You've been very helpful.
Patient	: Not at the moment. You've been very helpful, Nurse. I appreciate it.	Nurse : You're welcome, Linda. Let's finish your paperwork, then I'll show you to the examination room.
Nurse	: You're welcome, Linda. If you have any questions or need assistance during your stay, don't hesitate to ask. Now, let's get your admission paperwork completed, and then we'll guide you to the appropriate department for your examination.	Patient : Thank you, nurse.
Patient	: Thank you so much, Nurse. I'm glad to be here.	

The development process involved adjusting the initial dialogue according to CEFR B1 descriptors. This included simplifying the initial dialogue, which was too complex for the B1 level; replacing idiomatic phrases -such as "in case we need to reach someone during your stay," "handle the billing," and "verify coverage"- with simpler alternatives ("take care of the payment," "check your insurance") that better reflect the B1 Range descriptor of sufficient, straightforward vocabulary; simplifying sentence structures by reducing complex clauses and using only simple connectors; and removing unnecessary small talk to align with the B1 Interaction descriptor.

Listing vocabulary

The next stage is vocabulary listing. As previously discussed, vocabulary plays a crucial role in language learning (Al Zahrani & Chaudhary, 2022; Le & Trinh, 2024). Therefore, the developer included a vocabulary list in the English for Nursing audiobook to support the mastery of other language skills. At this stage, the developer categorized the vocabulary according to the thematic structure of the original syllabus and added part-of-speech markers to support sentence-level learning. This approach reflects the principle of

the situational syllabus, in which students need to memorize whole phrases rather than isolated words and understand their meaning within context. The primary categories are functional or thematic—such as Personal Information, Medical History, and Personal Contact—to help students understand vocabulary usage in context. The sub-categories consist of word types, helping students understand how each word functions within a sentence. The developer also included a column for Indonesian equivalents to help students fully understand the vocabulary list.

Table 3. Listing Vocabulary for Unit 1 Admitting Patient

Category	Word/ Phrase	Parts of Speech	IPA (UK)	IPA (US)	Indonesian Synonym
Personal Information	full name	noun phrase	/fʊl neɪm/	/fʊl neɪm/	nama lengkap
	identification card	noun phrase	/aɪ,dentɪfɪ'keɪʃn kɑ:d/	/aɪ,dentɪfɪ'keɪʃn kɑ:rd/	kartu identitas
Medical History	medical conditions	noun phrase	/'medɪkl kən'dɪʃnz/	/'medɪkl kən'dɪʃnz/	kondisi medis
	allergies	noun (plural)	/'ælədʒɪz/	/'ælərdʒɪz/	alergi
	GP (General Practitioner)	noun (abbreviated)	/,dʒi: 'pi:/	/,dʒi: 'pi:/	dokter umum
Contact & Insurance	next of kin	noun phrase	/nekst əv kɪn/	/nekst əv kɪn/	kerabat terdekat
	emergency contact	noun phrase	/'ɪmɜ:dʒənsi 'kɒntækt/	/'ɪmɜ:rdʒənsi 'kɑ:ntækt/	kontak darurat
	health insurance	noun phrase	/helθ m'fʊərəns/	/helθ m'fʊrəns/	asuransi kesehatan
Functional Expressions	May I see...?	phrase (polite request)	/meɪ aɪ si:/	/meɪ aɪ si:/	Bolehkah saya melihat...?
	Let me confirm...	phrase (confirming)	/let mi: kən'fɜ:m/	/let mi kən'fɜ:rm/	Izinkan saya memastikan...
	Is that correct?	phrase (checking)	/ɪz ðæt kə'rekt/	/ɪz ðæt kə'rekt/	Apakah itu benar?
	I'll note that down.	phrase (acknowledging)	/aɪl nəʊt ðæt daʊn/	/aɪl nəʊt ðæt daʊn/	Saya akan mencatatnya.
Verbs	to admit	verb	/tə əd'mɪt/	/tə əd'mɪt/	menerima (pasien)
	to confirm	verb	/tə kən'fɜ:m/	/tə kən'fɜ:rm/	memastikan

Audio Production & Editing Process

The audio production process involves five operational stages as follows:

In the first stage, the developer prepared the dialogue script finalized in the previous stage. The developer separated lines based on speaker roles (Nurse and Patient) and adjusted the formatting of the vocabulary list for use with the NaturalReader text-to-speech application.

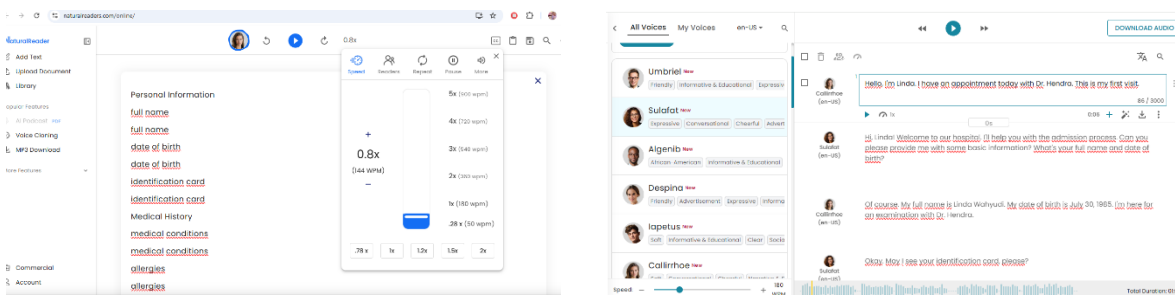


Figure 3. Comparison between Personal Version and Commercial Version of NaturalReader

In the second stage, the developer explored and tested voices using the Personal version of NaturalReader. The Personal version is suitable for standard English applications, such as presentations or news broadcasts. The developer reviewed the available voice profiles and tested parameters such as pauses, speed, and clarity of pronunciation. The Personal version is ideal for reading the vocabulary list, as it offers a more formal and authoritative tone and allows for easy adjustment of pauses between words. Meanwhile, the developer also explored the commercial version of NaturalReader for dialogue production. This version offers more comprehensive features, including a wider range of voices, tones, and emotions, as well as controls for pace and pitch at the sentence level. This feature is beneficial for practicing intonation, word stress, pitch contours, and fluency, as previously discussed (Al-Jarf, 2022; Bione et al., 2020). The commercial version of NaturalReader also allows dialogue with alternating roles to be generated within a single audio track, a feature unavailable in the Personal version. The developer experimented with the available voice options to match each role in the dialogue; for example, the nurse's role used a more formal voice, while the patient's role used a more communicative, friendly tone. Finally, the developer tested NaturalReader's voice parameter settings to align the output with the linguistic complexity required for the CEFR B1 level.

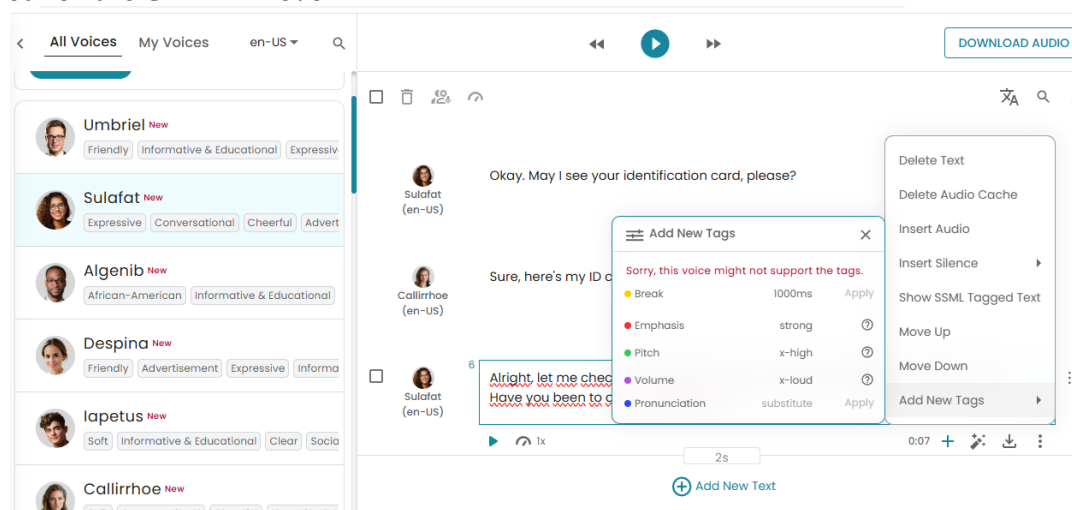


Figure 4. Exploration Process for the NaturalReader Commercial Version Dialogue Feature

In the third stage, the developer generated the audio based on the character roles and voice settings determined in the previous stage. Once the final audio for each segment was generated, the developer downloaded the files for further processing. The commercial version of NaturalReader offers more comprehensive download options than the Personal version, including a choice of file formats (.mp3 or .wav) and various quality settings. Users can also download the audio either line-by-line or as a single, combined file containing the complete dialogue.

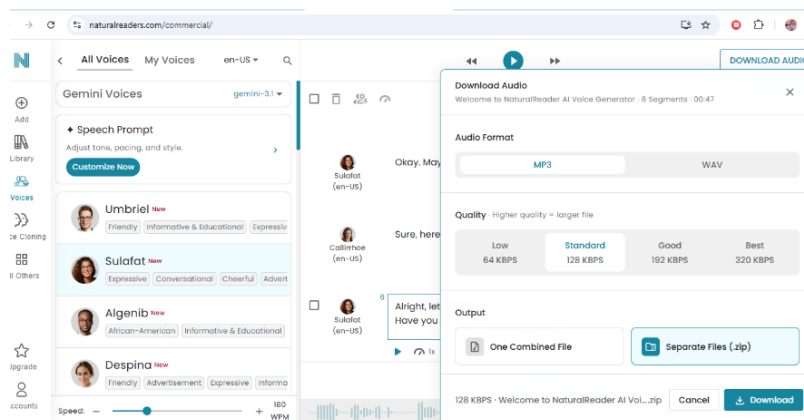


Figure 5. Downloading Process of the Commercial Version of NaturalReader

The fourth stage is audio editing, in which the developer trimmed excessive silence from the beginning, middle, and end of each dialogue line or vocabulary item generated in the previous stage. The trimmed segments were then assembled into a continuous dialogue flow. Pauses were inserted to give listeners time to process information, along with filler sounds—such as background audio (e.g., a ringing phone or the ambient sound of a hospital admissions area)—and transition sounds between units. The developer used the free audio editing service provided by Kapwing.com, and sourced free filler and transition sounds from Pixabay.com and Chosic.com.

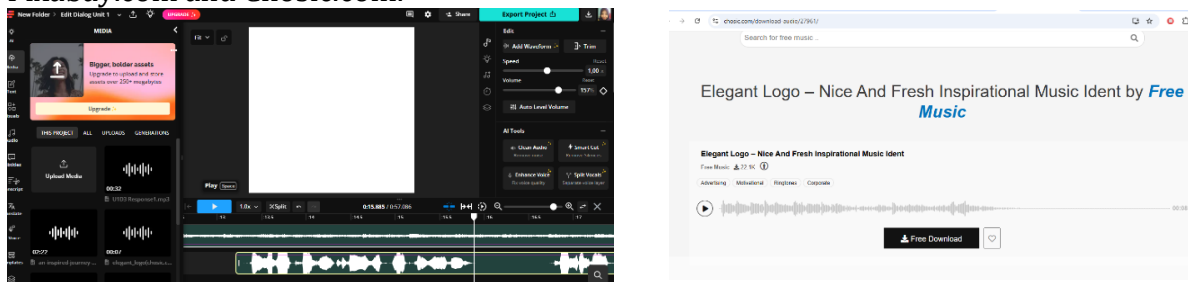


Figure 6. Sound editing process, and sound transitions source.

The fifth stage is the quality control stage. The developer reviewed the output from the previous stage to ensure that the AI voice's pronunciation, particularly of specific medical terms, was accurate. When issues such as imperfect pronunciation, excessive speed, inconsistent pauses, or mismatched voice characteristics were detected, the developer regenerated and re-edited the affected sections. This stage was repeated several times,

CONCLUSION

This study developed an English for Nursing audiobook using Text-to-Speech (TTS) technology, following the ADDIE framework and continuing from the previously completed Analysis and Design stages. The study employed a qualitative descriptive approach to provide a comprehensive account of each development stage: drafting and revising dialogue scripts, compiling vocabulary lists based on context and word type, audio production and editing, and presenting the final output. These processes reflect ESP principles, as the developer used AI voice/TTS technology to produce dialogues that closely reflect the authentic nursing contexts identified by respondents during the Needs Analysis stage.

Although this study did not reach the expert validation stage, the quality control and revisions carried out at each step served as an audit trail that helped maintain the rigor of the development process. The final product of this development stage was created in accordance with the syllabus produced during the Analysis and Design stages, adapted to the nursing context and the CEFR B1 proficiency level. Each unit of the audiobook has a duration of 20–25 minutes and is available in both .mp3 audio and .mp4 video formats.

This study has a key limitation: the audiobook has not yet undergone expert validation or pilot testing with nursing students. Consequently, its effectiveness cannot yet be determined. The next stages that are Implementation and Evaluation are needed to assess the audiobook's effectiveness in enhancing nursing students' language skills, particularly listening and pronunciation. It is also recommended that future research include expert validation to evaluate content accuracy, the clarity of the AI-generated audio, and the appropriateness of the language used.

REFERENCES

- Al Zahrani, S. M., & Chaudhary, A. (2022). Vocabulary Learning Strategies in ESP Context: Knowledge and Implication. *Arab World English Journal*, 13(1), 382–393. <https://doi.org/10.24093/awej/vol13no1.25>
- Al-Jarf, R. (2022). *Text-To-Speech Software for Promoting EFL Freshman Students' Decoding Skills and Pronunciation Accuracy*. <https://doi.org/10.32996/jcsts>
- Arini, F., & Widiarini, W. (2026). Improving Students' Listening Skills Through a Local Culture-Based Listening Module Using T2s Application. *SAWERIGADING*, 32(1). <https://doi.org/10.26499/sawer.v32i1.1650>
- Arockiasamy, J. C., Hoque, M. S., A, S. K., Kumar, J. S., & Kotte, R. (2025). The role of audiobooks in developing English listening proficiency: A study of undergraduate learner in Tamil Nadu. *International Journal of English Language and Literature Studies*, 14(4), 566–578. <https://doi.org/10.55493/5019.v14i4.5716>
- Bione, Tiago, Cardoso, & Walcir. (2020). *Synthetic voices in the foreign language context*.
- Canziani, T. (2021). English for Nurses: Needs Analysis and Syllabus. *1st Educational Sciences Conference, ESC 2020*, 2944.
- Çarkıt, C. (2020). Evaluation of Audiobook Listening Experiences of 8th Grade Students. *Educational Policy Analysis and Strategic Research*, 15(4), 146–163. <https://doi.org/10.29329/epasr.2020.323.8>
- Chetsadanuwat, K. (2018). Needs of English Skills of Thai Nurses Working in International

- Hospitals Accredited by JCI in Bangkok Area. In *Journal: Language Education and Acquisition Research Network Journal* (Vol. 11, Number 1).
- Constantinou, E. K., & Papadima-Sophocleous, S. (2020). The use of digital technology in esp: Current practices and suggestions for esp teacher education. *Journal of Teaching English for Specific and Academic Purposes*, 8(1 Special Issue), 17–29. <https://doi.org/10.22190/JTESAP2001017K>
- Dhuli, R., & Guduru, R. (2025). The Impact of High-Frequency Vocabulary Activation Strategies in Improving Speaking Skills and Vocabulary Knowledge of Rural ESL Learners. *TESL-EJ*, 29(2). <https://doi.org/10.55593/ej.29114a2>
- Pamelasari, S. D., Wusqo, I. U., & Hardianti, R. D. (2024). Developing Science Instruction in English Audiobook to Train Indonesian Prospective Science Teachers' Teaching Skills 1. *MEXTESOL Journal*, 48(2), 1–13.
- Fitria, T. N. (2022). Utilizing Text-to-Speech Technology: Natural Reader in Teaching Pronunciation. *JETLEE: Journal of English Language Teaching, Linguistics, and Literature*, 2(2), 70–78. <https://doi.org/10.47766/jetlee.v2i2.312>
- Guerrero, S. E. (2025). Audiobooks as a Tool for Improving Pronunciation and Fluency: Insights from Ecuadorian EFL Pre-Service Teachers: Audiolibros como una Herramienta para Mejorar la Pronunciación y Fluidez: Perspectivas de Docentes en Formación de Inglés como Lengua Extranjera en Ecuador. *Revista Científica Multidisciplinar G-Nerando*, 6(2), 2377–2394. <https://orcid.org/0000-0002-5619-3695>
- Herlyanti, P., Suhartono, L., Inggris, P. B., & Tanjungpura, U. (2023). *DEVELOPING AUDIOBOOK TO TEACH LISTENING COMPREHENSION*. 12, 1815–1821.
- Imawan, M. R., & Ashadi, A. (2019). Audiobooks for Assisting EFL Students in Reading Independently. *Journal of English Language Teaching and Linguistics*, 4(1), 77. <https://doi.org/10.21462/jeltl.v4i1.198>
- Iwata, Y., Iwata, Y., Iida, H., Inamori, M., & Maeda, S. (2023). Using a Smartphone Application as a Tool for English Learning Among Medical Staff and Students in Japan. *Advances in Medical Education and Practice*, 14, 167–182. <https://doi.org/10.2147/AMEP.S394625>
- Kartal, G., & Simsek, H. (2017). The Effects of Audiobooks on EFL Students' Listening Comprehension. In *The Reading Matrix: An International Online Journal* (Vol. 17, Number 1).
- Kılıçkaya, F., & Kic-Drgas, J. (2026). The audible turn: the rise of AI-led text-to-speech. In *AI and Society* (Vol. 41, Number 2, pp. 1299–1300). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s00146-025-02502-8>
- Kohnová, L., & Papula, J. (2020). Generation Z: Social media as a tool for education. *Proceedings of the 7th European Conference on Social Media ECSM 2020* [, 408.
- Le, H. S., & Trinh, M. L. (2024). An Investigation of Vocabulary Learning Strategies of ESP Students. *International Journal of TESOL & Education*, 4(1), 1–17. <https://doi.org/10.54855/ijte.24411>
- Maryani, M., & Arroyyani, R. (2024). Situational Syllabus in Vocabulary English for Nursing Book: Occupational Purposes. *Acitya: Journal of Teaching and Education*, 6(1), 34–52. <https://doi.org/10.30650/ajte.v6i1.3828>
- Modern Language Division, S. (2001). COMMON EUROPEAN FRAMEWORK OF REFERENCE

- FOR LANGUAGES: LEARNING, TEACHING, ASSESSMENT. In *Cambridge University Press* (1st editio).
- Mohammadi, V., & Mousavi, N. (2013). Analyzing Needs Analysis in ESP: A (re) modeling. *International Research Journal of Applied and Basic Sciences*, 4(5), 1014–1020. www.irjabs.com
- Montasser, R. A., & Althaqafi, A. S. (2023). Saudi Students' Perception of ESP Courses in the Preparatory Year. *Journal of Education and Learning*, 12(5), 166. <https://doi.org/10.5539/jel.v12n5p166>
- Nuranisa, N., Gailea, N., & Juniardi, Y. (2024). Negotiating Conflict, The Need Analysis of Developing Digital-Based English Teaching Materials for Midwifery Students. *East Asian Journal of Multidisciplinary Research*, 3(5), 1967–1980. <https://doi.org/10.55927/eajmr.v3i5.9150>
- Nurhamidah, N. (2024). *THE EFFECT OF TEXT-TO-SPEECH SOFTWARE ON IMPROVING STUDENTS' LISTENING SKILL*. ISLAMIC UNIVERSITY OF SULTAN SYARIF KASIM RIAU PEKANBARU.
- Nuswantara, K., Prasetyo, B., Bhawika, G. W., & Suarmini, N. W. (2021). Using ADDIE Model to Appraise MOOC English for Non-Academic Staffs (ENAS) Designed for University Non-Academic Staffs Amid the COVID 19 Outbreak. *Jurnal Sosial Humaniora*, 14(1), 74. <https://doi.org/10.12962/j24433527.v14i1.8349>
- Pujayanthi, P., Kusuma, I. P. I., Pratiwi, N. P. A., & Ana, I. K. T. A. (2026). The Quality of the Audiobook for Teaching English to Eleventh-Grade, Second-Semester Visually Impaired Students at SLB N 1 Tabanan. *Lexeme : Journal of Linguistics and Applied Linguistics*, 8(1), 103–110. <https://doi.org/10.32493/ljlal.v8i1.56648>
- Rizal, D., Masrurroh, S., Rahman Syah, R. F., Fathina, I. F., Amrullah, M. F., & Zakariyah, S. (2022). Audiobooks As Media To Increase Listening and Speaking Skills: a Qualitative Systematic Review. *ETERNAL (English, Teaching, Learning, and Research Journal)*, 8(1), 200–216. <https://doi.org/10.24252/eternal.v8i1.2022.a14>
- Rizkina, P. A., Suwartono, T., Nurhayati, S., Wijayanti, O., & Dalimarta, F. F. (2025). Creating AR Application for Advanced ESP Vocabulary Teaching in Indonesia. *International Journal of Learning, Teaching and Educational Research*, 24(5), 679–705. <https://doi.org/10.26803/ijlter.24.5.35>
- Salem, A., & Levitan, S. I. (2026). *Evaluating Question Prosody in Text-to-Speech Software*. 128–133. <https://doi.org/10.21437/speechprosody.2026-26>
- Susilowati, N. E., Luciandika, A., Ariani, D., & Martutik, M. (2021). Developing an Audiobook for Listening Courseware. *International Journal of Emerging Technologies in Learning*, 16(7), 45–60. <https://doi.org/10.3991/ijet.v16i07.21199>
- Talib, H. (2018). IMPLEMENTING LISTEN AND IMITATE TECHNIQUE TO IMPROVE PRONUNCIATION SKILL. In *Inspiring: English Education Journal* (Vol. 1).
- Thepseenu, B. (2020). Needs Analysis for ESP Course Development: Thai Civil Engineering Students' Perspectives. *Asian Journal of Education and Training*, 6(3), 433–442. <https://doi.org/10.20448/journal.522.2020.63.433.442>
- Wajdi, M., Sanjaya, I. N. S., & Sumartana, I. M. (2021). Developing a listening English learning model using text2speech application. *Journal of Applied Studies in Language*, 5(2), 274–



-
281. <http://ojs2.pnb.ac.id/index.php/JASL><http://ojs2pnb.ac.id/index.php/JASL>
 Widyana, A., Jerusalem, M. I., & Yumechas, B. (2022). *The Application of Text-to-Speech Technology in Language Learning* (pp. 85-92). https://doi.org/10.2991/978-2-494069-91-6_14

THIS PAGE IS INTENTIONALLY LEFT BLANK