Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

UTILIZING TEXT-TO-SPEECH (TTS) TECHNOLOGY IN CREATING LISTENING MATERIALS FOR ENGLISH LANGUAGE TEACHING (ELT)

Tira Nur Fitria[1)]*

[1)] Institut Teknologi Bisnis AAS Indonesia

Received on 24 December 2023 / Approved on 14 October 2024

Abstract

This research describes the use of Text-to-Speech (iSpeech) in creating English listening materials. The analysis shows that in using iSpeech Text-to-Speech, click the web https://www.ispeech.org/text.to.speech. Enter sentences or texts in the text box provided. Choose a male or female voice and choose the accent language we want to use to convert text to audio including UK (British) or US (American) English. To start the process synthesize sentences or text, press the Play button, and text that is currently sounded will be automatically blocked in yellow. Download the audio output and the audio can be played to students as listening materials in teaching English. Save the results of the synthesis into audio form in a certain format, so it can be heard again wherever and whenever by our students. With the presence of speech synthesizer technology, of Text-To-Speech (TTS) system can perform conversions from text computer-generated in the form of pronunciation (audio), where audio the resulting speed can be adjusted, intonation (prosody), as well as audio format output to be saved in the form of an audio file. English teachers/ lecturers can create listening materials according to their needs which will be taught to students.
Keywords: English Language Teaching (ELT); iSpeech; listening; Text-to-Speech

Abstrak

Penelitian ini menggambarkan penggunaan Text-to-Speech (iSpeech) dalam membuat materi mendengarkan bahasa Inggris. Analisis menunjukkan bahwa dalam menggunakan iSpeech Text-to-Speech, klik web https://www.ispeech.org/text.to.speech. Masukkan kalimat atau teks pada kotak teks yang tersedia. Pilih suara pria atau wanita dan pilih bahasa aksen yang ingin kita gunakan untuk mengubah teks menjadi audio termasuk Bahasa Inggris UK (British) atau US (Amerika). Untuk memulai proses sintesis kalimat atau teks, tekan tombol Play, dan teks yang sedang diputar akan otomatis diblokir dengan warna kuning. Unduh keluaran audionya dan audio tersebut dapat diputar kepada siswa sebagai bahan pendengaran dalam pengajaran bahasa Inggris. Simpan hasil sintesis dalam bentuk audio dengan format tertentu, sehingga siswa kita dapat mendengarkannya kembali dimana saja dan kapan saja. Dengan hadirnya teknologi Speech Synthesizer, sistem Text-To-Speech (TTS) dapat mengubah teks hasil komputer menjadi pengucapan (audio), dimana audio yang dihasilkan dapat diatur kecepatan, intonasi (prosody), serta outputnya. format audio untuk disimpan. dalam bentuk file audio. Guru/dosen bahasa Inggris dapat membuat materi listening sesuai dengan kebutuhannya yang akan diajarkan kepada siswa.
Kata Kunci: *Pengajaran Bahasa Inggris (ELT); iSpeech; listening; Text-to-Speech*

INTRODUCTION

With the rapid development of English learning materials and methods, schools are faced with the need to equip teaching materials such as CDs or English language tapes, which are relatively expensive and risk being damaged over time (Fitria, 2021b). Another important problem that occurs is teachers are less flexible in making their learning materials in the laboratory because of limited funding sources for collecting CDs/cassettes and only relying on materials that are on

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

CDs/cassettes that are already available in the market (Pudjiati & Fitria, 2022). In this case, the teacher is not optimal in providing the material that is following the needs and stages of student development.

Various problems may arise in connection with the listening if the audio is not good or not clear, the sound is too low, or even students do not understand the questions being tested. It is supported by Taylor (2009) that listening tests are almost usually necessary when evaluating speech production. Even this can be problematic, as it is exceedingly difficult to convince ordinary listeners (i.e., those who are not linguists or speech engineers) to focus exclusively on one aspect of the output. The reasons above are just problems purely technical, which can then be circumvented by providing better quality for facilities supporters while the latter, requires good skills in completing the problem.

Brown & Smith (2007) state that many students find listening to be one of the most difficult skills in English. Listening skill is one of the skills that must be mastered besides language skills. Listening is a very significant talent since it enables individuals to develop awareness, comprehension, knowledge, and performance via communication with others. Listening entails focusing on motion or tone. Listening is identifying and transforming voice sounds into words and sentences (Fitria, 2021b). It is a broad term that refers to intricate processes of attraction, perception, and behavior. While practicing listening skills, the listening component should be all ears concentrated on the listening portion. Listening is a critical component in providing the primary auditory input that is required for proper functioning. The utilization of various speech sources for the listening material will result in a variety of effects on the students' ability to comprehend what they are listening (Andriani et al., 2020).

Teaching listening is difficult (Goh & Vandergrift, 2012). Locke (2013) explains that teaching listening skills needs time and attention. Other English speaking, reading, writing, and listening skills require good hearing in the sense of not only being able to listen in general but also being able to process the information that is heard. The amount of practice is the main point to honing listening skills (Pudjiati & Fitria, 2022). For English teachers, making audio listening files as learning materials is common. Audio listening files can be used to train students in their listening skills. Listening to media is very helpful for teachers in the field of English studies in preparing listening teaching materials because teachers no longer need to have trouble recording their voices manually (Asmar & Ardi, 2013). Sometimes teachers or lecturers need to use their voices to teach listening-related material and, predictably, students will get bored or not interested in listening to it (Fitria, 2021a).

With the rapid growth of information technology with a variety of features that we can take advantage of. We as educators (teachers or lecturers) can create our audio files without involving many people and at a high cost, with the help of the Text-to-Speech (TTS) program. Text-to-speech (ITS) synthesizers, or computer programs that convert written text to spoken words, hold enormous promise for providing learners with a variety of conveniently accessible spoken language input (Kim et al., 2012). The working system of this TTS program is to convert text into sound.

In general, the task of the Text to Speech system is divided into two large parts, namely text analysis which transforms writing into linguistic representation, and speech synthesis which transforms linguistic representation into speech signals (Fitria, 2022). Of course, this system can be done with the role of advanced technology so that its use can make your work even easier. This system which can convert text into speech is generally widely used for automatic information services carried out over the telephone. So we can find it at the call center. However, what can be heard from this call center is not entirely TTS because the vocabulary is limited, whereas you can use an unlimited number of vocabulary through TTS. With the sophistication of technology that is realized in the form of TTS, we can use it for various purposes. It is hoped that all kinds of work, especially in companies, can run more easily and better so that companies can develop and experience progress.

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

There are lots of paid and free Text-to-Speech (TTS) programs. There are advantages to each program of course. Paid programs provide more complete features with premium sound while free programs provide standard features. Making audio listening files as learning material is commonplace, there are several Text-to-Speech applications to create audio listening files. Several websites and applications are available that also offer the functionality to convert text to speech (Raine, 2016).

Some of the best TTS apps include Speak It!, Verbally, Speech Assistant AAC, and iSpeech Text to Speech (Scearce, 2016). iSpeech can be accessed at www.ispcech.org. iSpeech is a Text-to-Speech application that allows users to easily convert text content in Word, and Excel. PowerPoint. PLW, blogs, RSS feeds, etc into audio. It is easy to use and has wide applications (Solomon & Schrum, 2014). Maglogiannis et al (2020) state that iSpeech is a solution that can be accessed online and is free to use for desktop and mobile applications. It utilizes a one-of-a-kind text-to-speech conversion technique that is both multithreaded and multicore, which makes it possible to use a straightforward text-to-speech method with an unlimited number of processors at the same time.

Currently, iSpeech Text-to-Speech is accessible in many languages. Users may utilize Select and Speak to have the iSpeech TTS read any chosen text from the web. iSpeech is another great tool to suggest to our students as it even supports 20 different languages (Watts, 2020). Speed and voice choices may be customized easily using the page's settings (Casterson, 2016). iSpeech is a great application that offers to translate typed text into voice or it can be said technology that can convert text into a voice with a variety of language options that we can choose from.

Several previous studies have been conducted related to making listening material or media using software or applications. First, Asmar & Ardi (2013) state that the application of “Listening Media” is very helpful for teachers in the field of English studies in preparing listening teaching materials because teachers no longer need to have trouble recording their voices manually. The use of "Listening Media" in preparing listening teaching materials is expected to help teachers achieve the goals of learning English, namely so that students can communicate using English orally and in writing. Second, Apriliana (2014) states that teachers can create content with the Longtion Autorun tool. The program uses a combination of images, audio, and video. It also can include a crossword puzzle, a matching exercise, multiple-choice questions, and fill-in-the-spaces. This program may be transferred on a DVD or USB stick and used on a PC or laptop. Third, Agustina (2017) states that designing listening media with Adobe Flash for English language learning motivates student instructors to create visually attractive listening learning material for elementary students.

Designing Adobe Flash-based learning media benefited the student teachers including increasing their knowledge of how to design listening learning media, developing their creativity and imagination, encouraging them to develop technology-supported learning materials, and encouraging them to create innovative and interesting listening learning. Fourth, Huang (2018) states that a learning system for English listening materials aimed at college students was developed using the Android software platform, which was utilized to design and develop the system. The system is capable of providing a relatively complete listening learning system that is based on the listening cognition rule, improving the intelligence test, strengthening the listening training and feedback mechanism, assisting in gradually improving the listening level, and improving the input and output processes of listening learning. Fifth, Zaim et al (2020) state that the authentic listening materials may be made with Adobe Animate CC 2019 for Android. It was validated by professionals in the fields of teaching the English language and producing media. The finished output may be put to good use by both students and instructors alike as an addition to more conventional approaches to the teaching and development of listening skills. Sixth, Fakhruddin et al (2020) that English teachers create English listening resources through YouTube video editing, in the form of audio and textbooks. They were taught how to develop English listening materials in the form of textbooks and how to edit videos in the form of audio (download, import, cut, merge, and

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

export video). The training resources include the following: (1) video search and download; (2) video editing includes the ability to import, clip, merge, and export video.

In line with what was said above, this research shares both parallels and contrasts with the preceding one. Both those earlier studies and this research focus on generating and listening through the use of software and applications; nevertheless, the purpose of research in each of these studies is different. The first study focuses on the use of apps called Listening Media, the second study focuses on the use of a tool called Longtion Autorun, the third study focuses on the use of Adobe Flash, the fourth study focuses on the use of Android software, the fifth study focuses on the use of Adobe Animate CC 2019 for Android, and the sixth study focuses on the use of YouTube for video editing. Aside from that, no study has been done about the usage of text-to-speech software in the process of planning or generating materials for English language instruction. Therefore, the purpose of this study is to investigate the application of text-to-speech software, particularly iSpeech, in the process of producing, developing, or designing listening products. iSpeech Text to Speech is one of the TTS applications, as was mentioned earlier. Therefore, the objective of this research stimulate the use of Text-to-Speech (iSpeech) in designing English listening materials for English Language Teaching.

LITERATURE REVIEW

A Text-to-Speech (TTS) system is software that converts text into speech (Stephanidis, 2011). It is similar to Filimowicz (2019) that Text-to-Speech (TTS) serves text as the input and a speech waveform is produced as output. It is supported by Reynolds et al (2018) state that Text-to-Speech is a media format created by a hardware or software system that uses a voice synthesis program to convert digital textual information (such as text in a word processing document or a web page) into audio format. Text-to-speech technology is a system or technology that converts text into sound automatically. Now many applications develop with this technology. So, we do not have to worry or feel anxious when using it because we can choose between free to paid use. This technology is also used in several well-known countries to complement daily activities, besides that, it has also been developed to meet the business world to education.

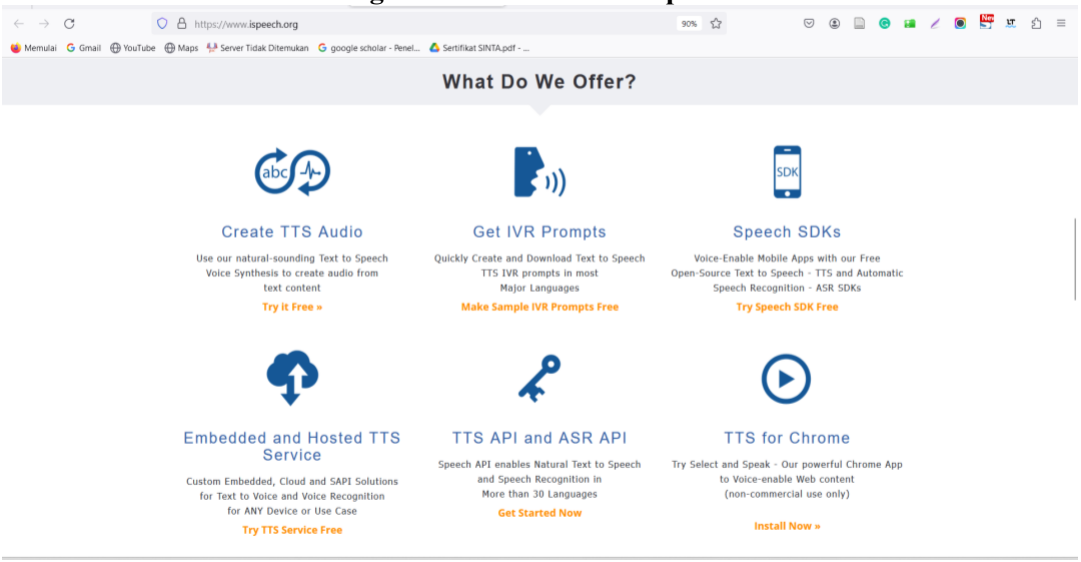
According to Yudhistiro (2016), technology that allows for the existence of The Text-To-Speech (TTS) system, also known as the speech synthesizer (synthesizer speech), is capable of converting computer-generated text to pronunciation (audio), with the pronunciation being generated at an adjustable speed, intonation (prosody), and audio output format that can be saved as an audio file. The speech synthesizer (synthesizer speech) is also known as the Text-To-Speech (TTS) system. This led to cost savings as well as increased efficiency in the usage of instruments, which would have been quickly worn out in the absence of the use of instructional media like CDs and tablets.

The text-to-Speech tool converts written text to synthetic speech, allowing the reader to listen to the text, rather than decoding the printed words (Boardman et al., 2021). It is believed that a Text-to-Speech system can utter any word since there is no limit to the number of words. The act of putting together an utterance from its component phonemes is known as phonetization. It is good knowledge that technology that converts text to speech can read any word and has an endless vocabulary. There is no way that the Text-to-speech system can function at its peak level if it is not supported by cutting-edge technology that is based on artificial intelligence. This system is capable of converting a wide variety of materials, including novels, news articles, information from websites, basic text, and so on. Text-to-speech is a technology that is based on artificial intelligence that can turn text into sound speedily and accurately. The generation of a voice that is as close to human as a text-to-speech (TTS) system can get should be the primary focus of any expressive TTS (Karpov & Potapova, 2021).

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

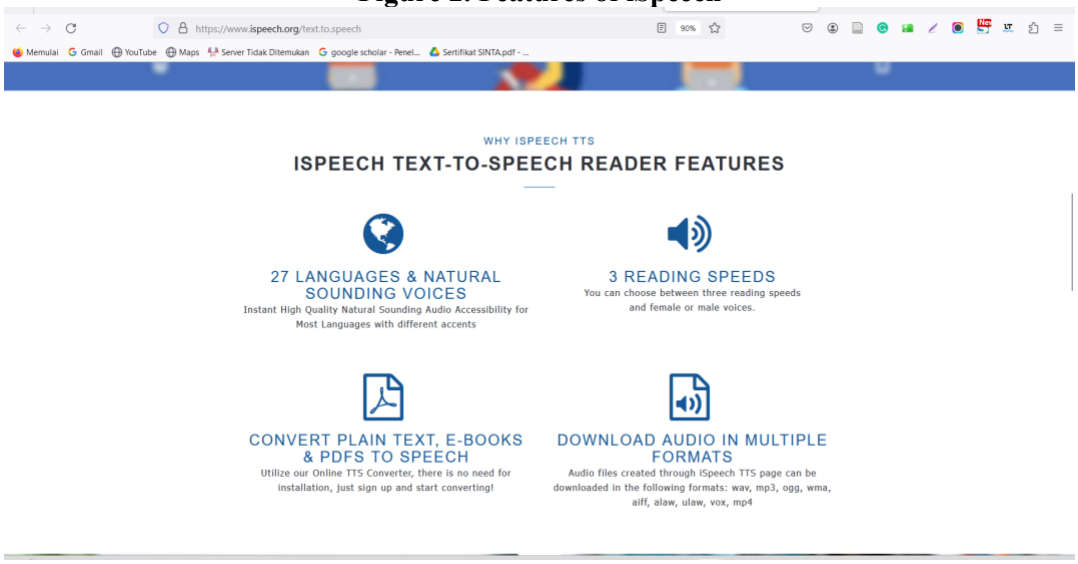
Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

Figure 1. Offer View of iSpeech



Among the many menus and offers available in iSpeech Text-to-Speech are the following: 1) Establish TTS Audio. We can generate audio from textual content using our Text-to-Speech Voice Synthesis, which produces sounds naturally. 2) Acquire IVR Prompts. Text-to-speech IVR prompts can be generated and downloaded rapidly in the majority of major languages. There are thirty speech SDKs. Utilize our frequency to voice-enable mobile applications. Text-to-speech (TTS) and Automatic Speech Recognition (ASR) software development kits (SDKs). 4) Hosted and embedded TTS service. Tailored SAPI, cloud, and embedded solutions for voice recognition and text-to-voice for ANY device or use case. 5) APIs for TTS and ASR. Speech API facilitates speech recognition and natural text-to-speech conversion in over thirty languages.

Figure 2. Features of iSpeech



Among the numerous features of iSpeech are the following: 1) Possessing a natural-sounding voice and 27 languages. Instant Access to Natural-Sounding, High-Quality Audio in the Majority of Languages with Various Accents. 2) Three reading speeds exist. There are three reading rates and male or female voices to choose from. 3) Capable of converting PDFs, plain text, and e-books to

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

speech. 4) Utilize our Online TTS Converter; registration is all that is required to begin converting; no installation is required. Download audio in an assortment of formats. The subsequent file types are available for download from the iSpeech TTS page: wav, mp3, ogg, wma, aiff, alaw, ulaw, vox, and mp4.

RESEARCH METHODOLOGY

This research is descriptive qualitative. Instead of numerical or statistical data, qualitative research emphasizes information conveyed through words, photographs, or other visuals and artifacts. (Ary et al., 2018). The objective of this descriptive study is to generate an image, description, or depiction of the investigated facts that is factually accurate. In this investigation, the result analysis is described descriptively. The method of data collection utilized in this investigation is observation. Observation is a methodical, rational, logical, and objective process of recording and observing a variety of phenomena in both natural and artificial environments to achieve particular goals (Wibowo & Cholifah, 2018). The researcher here observes the application of Text-to-Speech (iSpeech) in changing text-to-speech (audio) as listening materials.

The analysis of data in this research uses qualitative analysis such as data reduction, data display, and conclusion as proposed by (Miles et al., 2013; Miles & Huberman, 1994). During the step of data reduction, the researcher will narrow down and pick the data that will be utilized. Therefore, the researchers only looked at a small number of examples of different types of listening learning materials. During the step known as the data display, the researcher will exhibit or reveal the outcomes of the data in the form of pictures. The researcher makes use of visuals generated using Text-to-Speech here. In addition, the researchers concluded that the data gained from the analysis were explained.

FINDINGS AND DISCUSSION

Findings

Text-to-speech (TTS) is a technology that allows machines to convert text into sounds that humans can hear. TTS uses algorithms and language models to process text and produce sounds that mimic human voices. TTS applications are often used in various situations such as in navigation systems, telephones, and mobile applications. TTS can also help individuals with vision or reading problems by reading texts to them. TTS has various applications, including helping users save time and making it easier to access information.

How Text-to-Speech (TTS) works involves three main stages: voice synthesis, text processing, and speech synthesis. Here's an overview of how TTS works: 1) Voice Synthesi. TTS uses a polarized voice model to create virtual voices that represent human voices. This sound model has many customizable parameters, such as tonality, speed, and frequency. 2) Text Processing. TTS accepts text as input and processes this text into sound. This process includes breaking the text into phrases and determining the prosody (tonality, speed, and frequency) for each phrase. 2) Voice Synthesis. TTS uses voice models and prosody to generate a virtual voice that reads text. This virtual sound can be received by speakers or headphones and heard by the user. Text-to-speech has many improvements in terms of voice quality and the ability to overcome lingual problems such as accents and non-standard language. This makes TTS an increasingly useful tool and can be used in a variety of applications.

This research stimulates the use of Text-to-Speech (iSpeech) in designing English listening materials for English Language Teaching. Using iSpeech Text-to-Speech (TTS), there are several simple ways. We can click the link https://www.ispeech.org/ or https://www.ispeech.org/text.to.speech. The view of the web application can be seen below:

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

Figure 3. The View of iSpeech

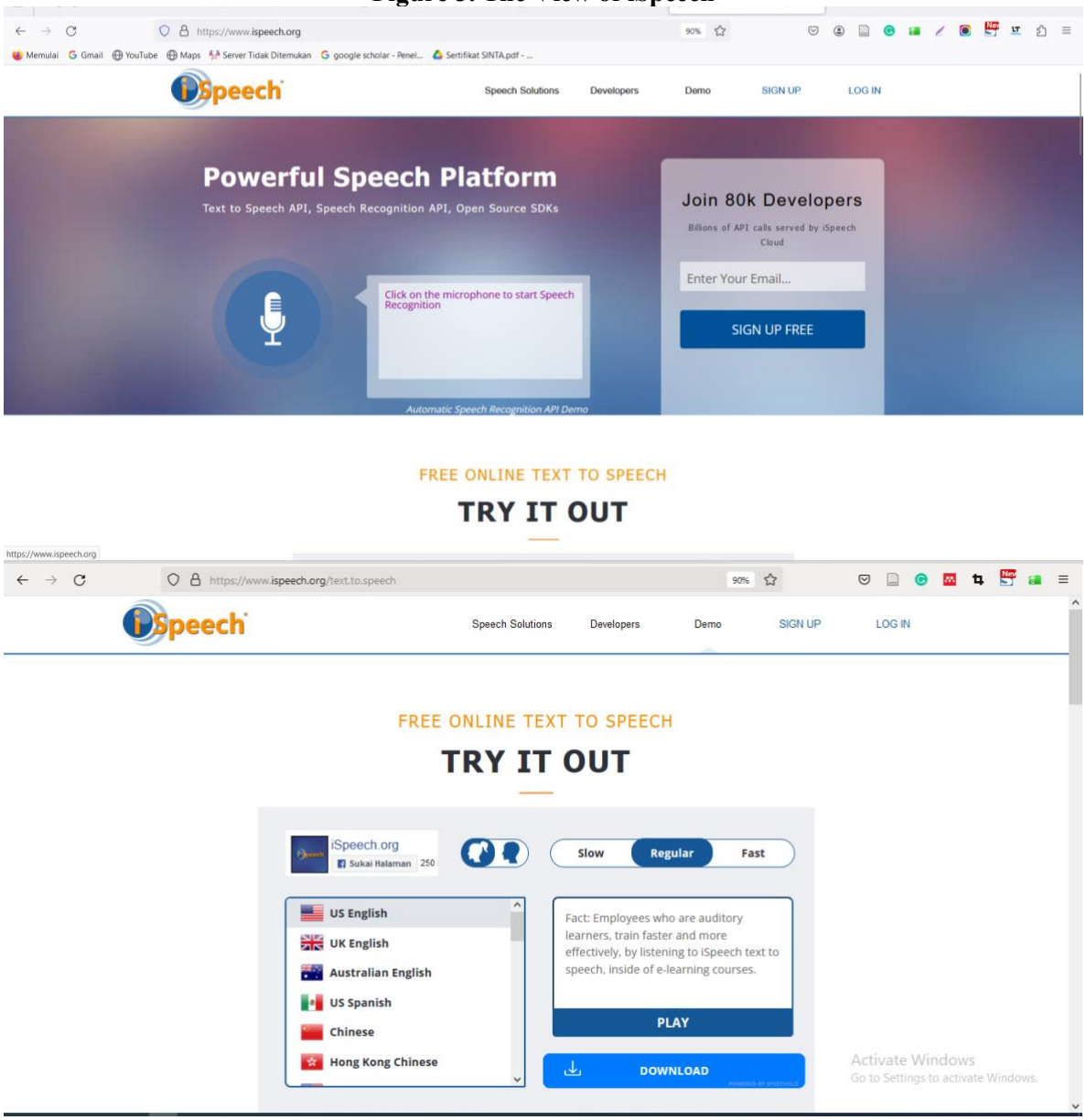


Figure 3 above shows the view of iSpeech Text-to-Speech (TTS) from the user's point of view. iSpeech Text-to-Speech possesses several useful characteristics, such as the capability to speak in a variety of languages and voices that are authentic sounding. Instant access to audio of high quality and a natural sounding for the majority of languages spoken around the world, including US English, UK English, Australian English, US Spanish, Chinese, Hong Kong Chinese, Taiwan Chinese, Japanese, Korean, Canadian English, Hungarian, Brazilian Portuguese, European Portuguese, European Spanish, European Czech, European Danish, European Finnish, European French, European Norwegian, European Dutch, European Italian, Eur. English, Hungarian, Brazilian Portuguese, European Portuguese, European Spanish, European Czech, European Finnish, and European There are also 3 reading speeds (slow, regular, and fast). Users can choose between female or male voices. There are also menus available including ”Play” to play the audio and “Download” to download the audio as output.

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

Figure 4. Creating Listening Material in iSpeech

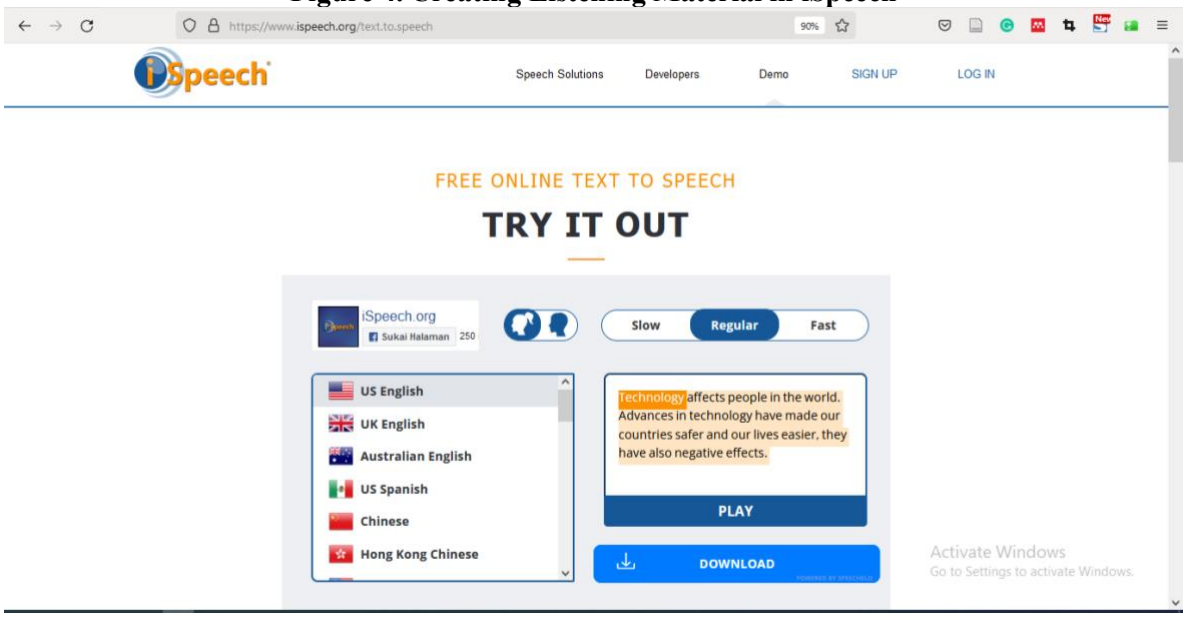


Figure 4 above shows how to create or make texts into speech. It can be used to create or design listening materials. In utilizing iSpeech, there is no need for installation, we just sign up and start converting. Users can convert plain text, e-books, and PDFs to speech. We as teachers or lecturers can type or paste the text we want to convert to audio, for example, we want to make listening materials about technology "Technology affects people in the world. Advances in technology have made our countries safer and our lives easier, they have also negative effects…". Text-to-speech technology can convert several sentences in the form of a certain language text becomes a pronunciation in a language that same. English text-to-speech will convert English text to English as well.

After we type sentences or text, we can choose one of three reading speeds such as regular (medium), and choose female or male voices. Besides, we choose one of the different accents in the world such as US English or UK English. If want to play the audio, we can click "Play", and if we want to download the audio, we click "Download". Audio files created through the iSpeech TTS page can be downloaded in mp3 or mp4. The audio that has been downloaded will automatically be saved in the download menu on our PC or device.

It is anticipated that the Text-to-Speech (TTS) technology will further expedite the process of teaching and learning, supplement the learning media, and create or produce resources for English topics, particularly in listening abilities. TTS is compatible with the vast majority of personal digital devices, such as desktop computers, mobile phones, and tablet PCs. It is possible to read aloud any form of text file, including documents created in Word and Pages. Even web pages that are viewed online can be read out loud. The text-to-speech (TTS) system uses computer-generated voices, and the reading pace may often be sped up or slowed down. With the sophistication of technology that is manifested in the form of this TTS, we as teachers or lecturers can use it for various purposes to teach and even create English materials for learners and students, especially listening skills.

Discussion

Text-to-speech technology can convert several sentences in the form of text in a certain language into a pronunciation in the same language. English text-to-speech will convert English text into English pronunciation as well. English is a language that has low regularity in the process of

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

converting text to phonemes. An English TTS is usually equipped with a database containing thousands of words and equivalent conversions of their phoneme sequences. Indonesian is a language that regulates conversion.

When teaching English listening skills, for example in a language laboratory, teachers or lecturers can use learning media application programs that can convert input in the form of English writing on a computer into sound or pronunciation automatically. Such a system of Text-of-speech (TTS) will make teaching more effective and efficient.

The application allows lecturers and teachers to enter English compositions or texts that have been modified to fit their teaching techniques or the requirements of their students so that the computer can easily read them out loud automatically. It is also possible for the teacher to change the tempo at which the text is pronounced and to halt (pause) the process of pronunciation at any point within a phrase. As a result, it will be quite useful for students who are still having trouble listening and speaking to take things slowly while they are learning. The teacher can save the pronunciation of the text into an audio format so that it can be stored on a compact disc or other storage media to be played back and is portable.

With this sound synthesis program, teachers no longer have trouble finding teaching materials, and no longer just rely on existing English learning CDs/cassettes, so teachers and lecturers can search for text-based materials or create their materials. Teachers and lecturers also can save expenses on purchasing English learning CDs/cassettes which are relatively expensive because this learning system does not use CD or cassette player hardware, it is very efficient in reducing the risk of wear or damage to these hardware devices.

Text-to-speech (TTS) software enables challenging readers to listen to the content of web pages, emails, digital textbooks, and an extensive collection of downloadable educational classes (Munday, 2016). Parris et al (2016) state that using a Text-to-Speech feature allow English learners or students to hear the entire texts or vocabulary of words or phrases on demand. The use of audio recording and Text-to-Speech software enables all learners to access content more flexibly or efficiently, as the information may be accessible from any time and location (Johnson et al., 2017).

The iSpeech App operates on the iSpeech.org Text to Speech (TTS) SaaS application programming interface. The iSpeech Text-to-Speech API is utilized daily by more than 80,000 developers and generates more than 100 million requests per month. Each call is answered in a matter of milliseconds, with no delay. iSpeech provides text-to-speech conversions using the highest-quality voices available. Simply input any text, and iSpeech will read it aloud in an instant.

Inescapable of transforming written content into spoken language, iSpeech exclusively employs high-quality voices. Simply enter any text into the text box, and iSpeech will instantly read it out loud. iSpeech makes it possible for anybody to convert text-to-speech voices, and it does so without requiring the user to download TTS voices on their device. The entire application takes up less than 3 megabytes of space, and this does not include the text-to-speech audio files that we generate by entering text. In the basic input area, we have the option of either writing our content or copying and pasting text from other documents or sources.

Text-to-speech has been completely transformed by iSpeech, which offers its service for free. Anyone may listen to any written material with very little effort, no installation of software required, and no prior knowledge of the technology required. The iSpeech App allows users to both write in text and listen to it at the same time. iSpeech is capable of converting text to speech using voices that are unparalleled in quality. Simply type any text into the box, and iSpeech will start reading the words out loud right away. Learn pronunciation, the first Text-to-Speech (TTS) software as a service (SaaS) cloud, patent; easy to use; speak anything by entering text; supports several languages; choices from almost 30 voices; rapid processing; our device will be speaking in no time at all; Pending Text-to-Speech as a Service, the best quality and most powerful TTS available anywhere, the most natural-sounding TTS audio available on the market, no hardware or software installs necessary for TTS servers, and Text-to-Speech optimized MP3 output.

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

To execute a single Text-to-Speech conversion, iSpeech possesses a one-of-a-kind multithreaded, multicore approach that is now pursuing a patent. Because of this, we can attack a single Text-to-Speech conversion with an unlimited number of processors at the same time. Therefore, a Text-to-Speech conversion that would normally take thirty minutes may be completed in substantially less than thirty seconds.

The use of iSpeech in preparing listening teaching materials is expected to help teachers achieve the goals of learning English, namely so that students can communicate using English both orally and in writing. Thus, is a free application software program that is developed to aid students in learning to listen easily. It is not difficult to use, however, it is extremely intriguing and can inspire learners or students to acquire listening skills and English pronunciation. Therefore, English instructors or educators should utilize it.

By using iSpeech Text-to-Speech, the teachers and lecturers only need to type in the text to be heard, and then this application will convert the written text into sound media. Gargiulo & Bouck (2017) state that accessing printed text via technology (listening material) can occur by using Text-to-Speech. It is similar to (Asmar & Ardi, 2013) that "Listening Media" Text-to-Speech is also very helpful for teachers in the field of English studies in preparing listening teaching materials because teachers no longer need to have trouble recording their voices manually. By using iSpeech, the teacher only needs to type in the text to be heard, then this application will convert the written text into sound media. Because it allows users to choose from seven different voice models for native English speakers, using this program is often regarded as being of great assistance to students in their efforts to acclimatize themselves to the sound of English as it is spoken by natives. As a result, downloading an application that converts text to speech to listen to assigned readings while working out might be a useful tactic (Lydia, 2015).

Using iSpeech comes with several advantages, some of which are as follows: 1) iSpeech may be downloaded for use on desktop computers as well as mobile devices (Android and Smartphone), 2) iSpeech converts written text into spoken language in 27 different languages, including: Using iSpeech allows you to avoid spending money on expensive voice talent since it creates immediate audio (TTS). Additionally, iSpeech employs voices that sound like real people and are both quick and affordable. 5) Students may acquire perfect pronunciation by speaking with the assistance of iSpeech, which utilizes voice recognition. iSpeech 6) displays highlighted words in real-time as they are being uttered, and iSpeech 7) integrates Siri-like speech interactions with mobile applications. iSpeech is a virtual teacher that can educate via spoken conversation, iSpeech makes audio for eLearning presentations, iSpeech offers audio versions of all course materials, and iSpeech enables auditory learners to listen to English while reading. In addition, iSpeech is like Siri for eLearning. iSpeech is a virtual teacher that can educate through spoken dialogue. iSpeech leveled the playing field for all students and learners by breaking down language barriers with its text-to-voice, speech recognition, and natural language understanding (NLU) features.

A good website, such as iSpeech Text, should deliver a simple, user-friendly experience. This includes those with poor eyesight, learning disabilities, and dyslexia, as well as older folks, children, and anyone who does not read in their native language. Incorporating iSpeech's Text-to-Voice Reader into our website can significantly increase accessibility. Users may read and listen simultaneously using iSpeech. iSpeech Text-to-Speech technology uses authentic, human-sounding voices to enhance the visitor experience on our website by ensuring that the content presented is comprehended effectively. iSpeech also offers a variety of other online services and applications, including iSpeech Translator which translates spoken text in 7 languages.

There are some of the advantages of the iSpeech application that we can feel. There are many choices of languages with the best voices, the languages that we can choose from about 30 different accents languages. This application can also convert text to sound for some documents such as emails, favorite news articles, and other documents. There is voice assistance that will be read aloud in the voices of iSpeech Bush and iSpeech Obama. This application also provides support with some interesting items in it, to connect we must activate it first by visiting the iSpeech

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

website itself. The iSpeech application can be used on two devices namely iOS and Android. If we are Android users we can use it for free for certain features, if we want to upgrade we have to make a purchase. As for iOS, it is used if we have made a purchase.

The increasingly widespread use of computers and the rapid development of technology open up a new stage of information exchange between users of computers with computers. Enter information into the computer through speech and convert the information in the form of text stored on the computer, in the sound form it is possible to do it. TTS (Text to Speech) which is an application in the language technology field, can convert text in the format of a language into appropriate speech by reading text in that language used. TTS can be applied in various ways fields, especially will be very useful for students or even people who have visual impairments (Pramadewi et al., 2013). It is also possible to employ TTS for activities that are geared at teaching foreigners a language. Voice, which is the output of TTS, will make it simpler for someone to learn how to pronounce a word in a certain language. This is because every language is distinct and has various rules for how words should be pronounced in the language that is being spoken.

CONCLUSION

In using iSpeech, Click https://www.ispeech.org/text.to.speech to use iSpeech. Then, copy and paste the text we wish to convert to speech into the text field and select the convert text button. Alternatively, the user can upload any supporting file, website, blog, etc., and then select the converted file or listen button. We wait and listen for several seconds. iSpeech performs all of our duties. Then, our file is converted and made available for auto-play and download. By utilizing iSpeech, English teachers and lecturers can create listening materials tailored to their students' requirements. With Text-to-Speech audio, it is anticipated that students will be trained in listening because the sound generated by iSpeech is the voice of several native speakers contained in a database system.

REFERENCES

Agustina, N. (2017). *Designing Listening Learning Media Using Adobe Flash* (Undergraduate Paper 0, Universitas Esa Unggul). https://digilib.esaunggul.ac.id/UEU-Article-NA_SIEC2016/8521

Andriani, T. N., Herawati, Y. W., & Sulistyo, T. (2020, May 25). *Text-to-Speech Application for Foreign Language Learner' Listening Comprehension in Indonesia*. ICONQUHAS 2018, Bandung, Indonesia. https://doi.org/10.4108/eai.2-10-2018.2295544

Apriliana, W. (2014). *Developing Interactive Listening Material Using Longtion Autorun Application for 10th Grade at SMA Muhammadiyah 8 Morowudi* [Undergraduate Paper, Universitas Muhammadiyah Gresik]. http://digilib.umg.ac.id/gdl.php?mod=browse&op=read&id=jipptumg--winduapril-1427&q=DEVELOPING%20INTERACTIVE%20LISTENING%20MATERIAL%20USING%20LONGTION%20AUTORUN%20APPLICATION%20FOR%2010th%20GRADE%20AT%20SMA%20MUHAMMADIYAH%208%20MOROWUDI

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2018). *Introduction to Research in Education*. Cengage Learning.

Asmar, Y., & Ardi, H. (2013). "Listening Media" Application in Preparing Listening Material. *Lingua Didaktika: Jurnal Bahasa Dan Pembelajaran Bahasa*, *6*(2), Article 2. https://doi.org/10.24036/ld.v6i2.7416

Boardman, A. G., Garcia, A., Dalton, B., & Polman, J. L. (2021). *Compose Our World: Project-Based Learning in Secondary English Language Arts*. Teachers College Press.

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

Brown, S., & Smith, D. (2007). *Active Listening 3 Teacher's Manual with Audio CD*. Cambridge University Press.

Casterson, S. (2016). *Google Chromebit: A Guide for Beginners*. Conceptual Kings.

Fakhruddin, Z., Usman, U., Rahmawati, R., & Sulvinajayanti, S. (2020). Designing English Listening Materials through YouTube Video Editing: Training for English Teachers of Islamic Junior High Schools Parepare. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, *4*(2), Article 2. https://doi.org/10.21093/ijeltal.v4i2.475

Filimowicz, M. (2019). *Foundations in Sound Design for Embedded Media: A Multidisciplinary Approach*. Routledge.

Fitria, T. N. (2021a). An Analysis of the Students' Difficulties in TOEFL Prediction Test of Listening Section. *ENGLISH FRANCA : Academic Journal of English Language and Education*, *5*(1), 95–110. https://doi.org/10.29240/ef.v5i1.2212

Fitria, T. N. (2021b). Investigating the Emergence of Digital Platforms for Listening Learning Proficiency. *Al-Lisan: Jurnal Bahasa (e-Journal)*, *6*(2), Article 2. https://doi.org/10.30603/al.v7i2.2217

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

Gargiulo, R. M., & Bouck, E. C. (2017). *Instructional Strategies for Students With Mild, Moderate, and Severe Intellectual Disability*. SAGE Publications.

Goh, C. C. M., & Vandergrift, L. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*. Routledge.

Huang, H. (2018). Design and Implementation of a College English Listening Learning System Based on Android Platform. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(07), Article 07. https://doi.org/10.3991/ijet.v13i07.8779

Johnson, D., Dobler, E., & Wolsey, T. D. (2017). *Teaching the Language Arts: Forward Thinking in Today's Classrooms*. Routledge.

Karpov, A., & Potapova, R. (2021). *Speech and Computer: 23rd International Conference, SPECOM 2021, St. Petersburg, Russia, September 27–30, 2021, Proceedings*. Springer Nature.

Kim, T., Ma, J., Fang, W., Zhang, Y., & Cuzzocrea, A. (2012). *Computer Applications for Database, Education and Ubiquitous Computing: International Conferences, EL, DTA and UNESST 2012, Held as Part of the Future Generation Information Technology Conference, FGIT 2012, Gangneug, Korea, December 16-19, 2012. Proceedings*. Springer.

Locke, A. (2013). *Teaching Speaking and Listening: One Step at a Time, Revised Edition*. A&C Black.

Lydia, K.-B. (2015). *Handbook of Research on Strategic Management of Interaction, Presence, and Participation in Online Courses*. IGI Global.

Maglogiannis, I., Iliadis, L., & Pimenidis, E. (2020, May 29). Artificial Intelligence Applications and Innovations. *International Conference, AIAI 2020*.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (Second Edition). SAGE Publication.

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). *Qualitative Data Analysis: A Methods Sourcebook*. SAGE Publications.

Munday, J. B. (2016). *Teaching a Child with Special Needs at Home and at School: Strategies and Tools That Really Work!* WestBow Press.

Parris, H., Estrada, L., & Honigsfeld, A. (2016). *ELL Frontiers: Using Technology to Enhance Instruction for English Learners*. Corwin Press.

Pramadewi, P. M. M., Kesiman, M. W. A., & Darmawiguna, I. G. M. (2013). Pengembangan Aplikasi Text to Speech untuk Bahasa Bali. *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*, *2*(3), 219–233. https://doi.org/10.23887/janapati.v2i3.9807

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com

Journal of English Language and Culture *Vol. 15 (No. 1) : 73 – 85. Th. 2024*
Versi Online: http://journal.ubm.ac.id/index.php/english-language-culture *p-ISSN: 2087-8346*
Hasil Penelitian *e-ISSN: 2597-8896*

Pudjiati, D., & Fitria, T. N. (2022). Digital literacy in learning listening skill between English and Non-English Department students through educational video. *Leksika: Jurnal Bahasa, Sastra Dan Pengajarannya*, *16*(1), Article 1. https://doi.org/10.30595/lks.v16i1.13016

Raine, P. (2016). *Fifty Ways to Teach with Technology: Tips for ESL/EFL Teachers*. Wayzgoose Press.

Reynolds, C. R., Vannest, K. J., & Fletcher-Janzen, E. (2018). *Encyclopedia of Special Education, Volume 4: A Reference for the Education of Children, Adolescents, and Adults Disabilities and Other Exceptional Individuals*. John Wiley & Sons.

Scearce, L. (2016). *Manual of Singing Voice Rehabilitation: A Practical Approach to Vocal Health and Wellness*. Plural Publishing.

Solomon, G., & Schrum, L. (2014). *Web 2.0 How-to for Educators*. International Society for Technology in Education.

Stephanidis, C. (2011). *Universal Access in Human-Computer Interaction. Applications and Services: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part IV*. Springer.

Taylor, P. (2009). *Text-to-Speech Synthesis*. Cambridge University Press.

Watts, S. (2020). *Teaching Online: Online Teaching Survival Guide: The Best Teaching Strategies and Tools for Your Online Classroom.* Wryting Ltd.

Wibowo, A., & Cholifah, T. N. (2018). *Instrumen Tes Tematik Terpadu: Untuk mahasiswa pendidikan guru Sekolah Dasar*. Media Nusa Creative (MNC Publishing).

Yudhistiro, K. (2016). Pemanfaatan Teknologi Text-To-Speech Sebagai Media Pembelajaran Pada Laboratorium Bahasa Inggris. *Jurnal Teknologi Dan Manajemen Informatika*, *2*(1). https://doi.org/10.26905/jtmi.v2i1.622

Zaim, M., Refnaldi, R., & Elfiona, E. (2020). Developing Mobile-Based Authentic Listening Materials for Senior High School Students. *Ta'dib*, *23*(1). https://doi.org/10.31958/jt.v23i1.1794

*Author(s) Correspondence:
E-mail: tiranurfitria@gmail.com