

CHAPTER IV

FINDINGS AND DISCUSSION

The researcher gather and compiled the results obtained in this chapter after finishing the research. This chapter contains essential information and implications for answering the research questions laid out in previous chapter. The discussion's objectives are to deliver explanations regarding the problem definition and the interpretation of the results in accordance with the relevant theoretical review.

A. Findings

It is imperative to point out the sources of the data that were used in this study in order to fully analyze it is findings. The teacher actively participated in the making of the dataset that was being analyzed. The data came from both structured interviews and narrative frames. In order to obtain a better understanding of the process and views faced by the participant when using TTS to create listening materials, the researcher have decided to combine interview session with the narrative frame method.

The participant were given total of six questions during the interview session, discussing the participant's general view towards TTS as a medium to create listening materials, both the positive and negative aspects of it. As for the narrative frame, its purpose is to find out the processes that went through when making listening materials using TTS.

1. The Processes of Creating Listening Materials using TTS Performed by The English Teacher

The processes of creating listening materials using TTS can vary greatly from one individual to another. Depending on what software or app is used, the method would differ in some aspect, but the core idea still remains the same, producing audio based product from bare text.

In this instance, the program that was used is Elevenlabs.io, an online based website that can output impressively human voice that does not require a subscription or any payment of some sort. The method that was used to gather the data for this section is narrative framing, and compare to data obtained with the theory presented before.

a. Scripting

In the context of creating listening materials, "scripting" refers to writing out the planned content for audio or video recordings. This script serves as a blueprint or guideline for what will be spoken or presented during the recording. To write a competent script Mulyono (2016) in his journal *“Developing Native-Like Listening Comprehension Materials: Teacher’s and Pupil’s Perceptions of a Digital Approach.”*, he has presented the necessary procedures for scripting, they are:

1) Learning indicator mapping

When a teacher prepares a lesson, they need to establish clear goals to guide the learning process. Goals are broad, overarching statements that describe the general outcomes or skills that students should achieve by the end of the lesson. They focus on the overall big picture and often align with curriculum standards or

broader educational goals, it could be about the importance of understanding proper use of an adjective, the necessity of using the correct tense in a conversation, and many more depending on the subject taught. The teacher that is being observed has stated, “Looking for the suitable materials with listening activity in the chapter of the book that I currently teach to them”.

This essentially means that they usually just use the current book that the students are using to come up with a suitable learning goals for them. Here is the learning indicator the teacher used in one of his lesson that uses TTS:

Kompetensi Dasar
3.7 Membandingkan fungsi sosial, struktur teks, dan unsur kebahasaan beberapa teks naratif lisan dan tulis dengan memberi dan meminta informasi terkait fairytales, pendek dan sederhana, sesuai dengan konteks penggunaannya.
4.7 Menangkap makna secara kontekstual terkait fungsi sosial, struktur teks, dan unsur kebahasaan teks naratif, lisan dan tulis, sangat pendek dan sederhana, terkait fairytales.

Figure 4.1. Learning Indicator the teacher used in his lesson that utilizes TTS

2) Setting objectives

Objectives, on the other hand, are more specific and detailed. They outline the precise skills, knowledge, or behaviors that students should demonstrate by the end of the lesson. This usually takes form as activities, quizzes, and games. Objectives are typically measurable and observable, allowing the teacher to assess whether students have successfully learned the material.

For example, the goal is to have the students understand how to use an expression of gratefulness, the objectives can lay down all the necessary steps for students to understand this concept. It can be the live example of a conversation, lists of key words, and a practice session for the students focused on assessment.

When asked about how the teacher decide the objective for their student when scripting, they responded by, “I make the script that related to the material being taught”.

this has a few benefits, one is that the students can relate to the lesson easier since it ties back to what they are currently learning. And the second benefit is that the teacher has better chance to make said lesson memorable for the students since it’s all ties back to the relatability of the lesson. Here is the objectives set by the teacher in their lesson that utilized TTS:

Tujuan Pembelajaran
• <u>Mengidentifikasi isi cerita teks fairytales yang didengar atau dibaca dengan tepat.</u>
• <u>Membacakan dongeng dengan ucapan dan tekanan kata yang benar </u>
• <u>Mengidentifikasi informasi rinci dari isi teks yang sedang dibacakan dengan tepat</u>
• <u>Menyebutkan bagian-bagian cerita yang memuat pesan yang disebutkan dengan baik dan benar.</u>

Figure 4.2. Objectives set by the teacher in their lesson that utilized TTS

3) Length of discourse

Determining the length of audio material used in a class depends on several factors, including the age and attention span of the students, the complexity of the content, the goals of the lesson, and the integration with other classroom activities.

When asked about how they determines the length of the discourse for their students, the teacher does not have a set length of the materials, meaning it can either be short, quick exercises or long, insightful paragraphs. It all depends on the student’s skill levels, attention span and the goals and objectives of the lesson. The teacher response go as, “I divide the script into some paragraphs to make the result end with several file of audios”.

When asked why they would do this, they elaborated that the average attention span of a middle schooler is usually around 5-10 minutes at best, so there needs to be an immediate, clear reason why they would listen to the material so that their minds do not “wonder off” to something else, and it also needs to be somewhat engaging. Several audio files can quickly indicate to the students that the material used has shifted from A to B so that they can focus on the current objective rather than mulling over the past ones, thus engaging them to be more attentive.

4) Wordings and expressions

The words and expressions used when scripting audio material for students are crucial because they determine how easily students can understand and stay interested in the content. Clear, simple language helps students grasp concepts without confusion, while engaging expressions and a friendly tone keep their attention. The script should use consistent vocabulary and be inclusive, catering to a diverse range of students, although this depends on the situation of the learning environment. When asked about how the teacher determines the correct wording and expressions used, they said, “Re-check the script to make sure that it is ready to convert”.

The importance of choosing the right kind of language to communicate with the right kind of people is big, not being able to communicate ideas just because the wordings and expressions used in the materials are too advanced is something that needs to be avoided.

Here is an example of one of the script the teacher used in his lesson that utilized TTS:

The Legend Of Bawang Merah And Bawang Putih

Once upon a time there lived a rich widow who had two children: Bawang Putih and Bawang Merah. Bawang Putih was her step daughter. Bawang Putih was a good child, while Bawang Merah was lazy.

One day Bawang Putih washed her stepmother's cloth, but the cloth was washed away in the river. Bawang Putih continued down the river until she finally met an old woman who found the cloth. The old woman would return the cloth on the condition that Bawang Putih had to help her work.

Bawang Putih helped her work. The old woman gave two gifts to choose: a small pumpkin or a large pumpkin. Then Bawang Putih took the small one.

Coming home with a small pumpkin, her stepmother grabbed the pumpkin and slammed it angrily. It turned out that the pumpkin contained gold, gems and diamonds.

The stepmother then told Bawang Merah to do as Bawang Putih did, to wash the cloth in the river. Until finally met the old woman. But Bawang Merah didn't want to help her and asked for a big pumpkin. At home she split the large pumpkin which turned out to be a poisonous snake and scorpions. Soon they realized their bad deeds.

Figure 4.3. One of the examples of the script the teacher used

B. Text-to-speech conversion

After we are done with the scripting of the materials, ensuring that it meets the goals of the lesson, the next step is to convert said script into an audio format using TTS program, the one that was being used by the teacher is called Elevenlabs.io. it is a freely available public website that allows its user to convert any text that is inputted into an audible audio format that can be listened to.

Every TTS program procedures when converting a text into an audio can vary, but in this case, we will be using the format laid by Mulyono (2016) in his research earlier in this paper to see how well the teacher met these criterias. In his paper, Mulyono (2016) divided the process into four different stages:

1) Connecting words and sentences

This is the first basic step into creating an audio from TTS. here the teacher simply copied and pasted the script that we have into the program in a duly order. It is recommended to only put several segments at a time instead of the whole script, since it allows the user to use the ‘refresh’ command to generate another sample if the resulted audio does not meet the teacher’s standards. The teacher notably shared his experience regarding this matter and said, “Choose the english language to fit the pronunciation with english script”.

Which is true, it is important to select the correct language model the TTS program will use, since it will affects the overall speech, the English text we inserted into the program will not sound correct when the language used is french for example.

2) Adjusting speed rate

The next thing to do after inserting a segment of the script is to adjust the speed of the speech, how fast the speech goes and to some extent, the pitch of the voice. This feature is available in Elevenlabs.io and was utilized by the teacher. The researcher asked the teacher about the initial resulting audio speed, whether or not it is good enough to meet the standards of the teacher and more imporatntly, the students, and whether or not they needed to re-adjust it. he quoted, “Decide the type of voice that presented on the list and set some settings in the voice settings to make the voice clearer and good”.

This is one example of creative solutions that can only be done using TTS. Since Elevenlabs.io allows you to tinker with the voice before you export it, it opens

up this kind of possibilities where one could facilitate both beginners and advanced student's levels. Here is the setting the teacher used when converting the script of "The Legend of Bawang Merah and Bawang Putih", prioritizing stability and similarity:

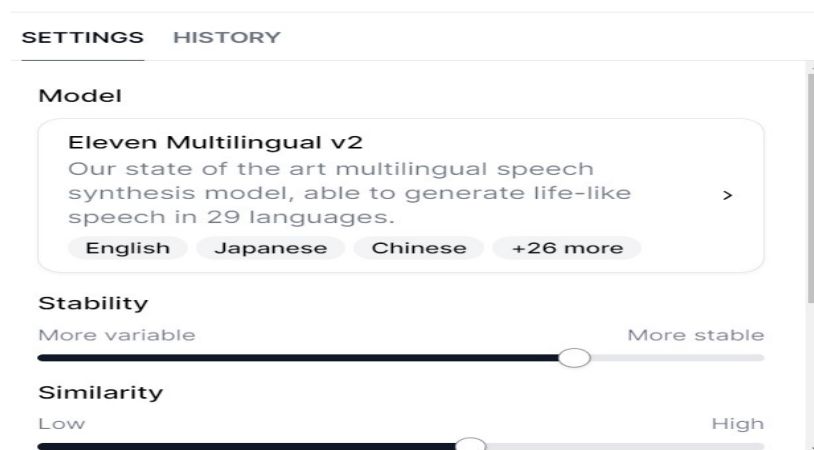


Figure 4.4. The settings used by the teacher when converting the script

3) Audio file conversion

Conversion here is referring to the end process when we have done all of the steps above, the audio is ready to be exported and downloaded by the user and then to be used freely, and fortunately Elevenlabs.io does allow its user to export the audio without paying any prices, they stated that "I convert the paragraph that I made into audio file one by one and save the files into a folder".

The Elevenlabs.io website only allows .mp3 format to be used when exporting from it, so it does limit some usage, but for the majority of cases, .mp3 is the most widely used audio format to this day. So the teacher does not think that it is a problem. After the audios are downloaded, it is then combined and stitched together in an editing software to form a single audio where it can easily played in

the classroom without the teacher having to interact with it often. The listening materials that had been created by the teacher can be examined here:



Figure 4.5. The QR Code of the video containing the audio produced by TTS that was uploaded to [YouTube.com](https://www.youtube.com)

2. The View of The English Teacher with Respect to using TTS as a Method to Create Listening Materials

The technology is important, but the opinions regarding the usability of said technology also cannot be ignored. No matter how incredible the technology is, if the usability is poor then it could not be relied upon. The teacher was also asked about their opinions and views regarding TTS and how it competes against the usual method. To accomplish this and gather the necessary information, the researcher conducted an interview with the teacher and asked six primary questions regarding the usability of TTS in regards to creating listening materials.

a. Text-to-Speech Technology as Learning Tool as a Whole

First, the researcher asked the teacher of how his impression of TTS as a tool for learning and teaching English, and he responded by, “I don’t use TTS as a whole, and only use it when I want to make my time more effective and efficient.”

They also used TTS when met with a specific situation such as, ”If I found some students still lacking in pronunciations that I already taught before”.

The teacher seems to have used TTS as a method for students to catch up with their pronunciation, as they followed the previous statement with the following, “While I focus on teaching something else, they can use the TTS if they struggle in pronouncing some words”.

This statement enlightened the researcher for the main usage of TTS for this particular teacher, it is simply a supplementary tool to help those that might struggle with the lesson.

b. TTS Can be Used to Listen to Any Text or Topic The User Desire

Following this, the researcher also asked the teacher regarding the coverage of TTS, can it contains and deliver many kinds of topics and lessons? And the teacher answered as follows, “Yes, and sometimes I make my own topic”.

It makes sense retrospectively, as TTS can output any kinds of texts, may it be historical, fiction, narration or report as few examples, since all it needs is the text for it to be able to export said texts as a audio format.

This also shows the flexibility of TTS, since the teacher also said that they would often make their own topic. This is what the researcher thinks is the main advantage of TTS. To be able to create any kinds of voiced media with just a texts

allows an individual to customize and tailor their lesson to be whatever they want, for example, to make a material consisting of the latest events in the school, the teacher could create a listening material about that using TTS with ease.

c. The User Can Adjust The Reading Speed to Suit The Students Freely When Using TTS

The next question pertains to the flexibility of TTS, can the user tinker with the speed of the text produced by TTS? And the teacher responded with, “Yes, gladly the TTS program I use is able to adjust some settings”.

Elevenlabs.io allows the user to adjust the voice’s settings such as how fast it speaks, and how much each words are emphasized. It can be made to sound so natural like a native speaker, or be made to annunciate or even exaggarate the delivery so that the lesser skilled students could still catch what is being spoken. The teacher uses this setting to adjust the audio based on their student’s level of English proficiency.

d. The User Can Easily Store or Save Any Audio Text That Has Been Processed Using TTS For Future Use

Following the topic of convenience, the researcher asked the teacher about TTS’s ability to safely store or save the produced media, is it easy for the user to save the work they have done to be used for future lessons? as to which they have this to say, “Yes, I make and prepare the audios days before i use it in the class”.

This means that the longevity of created lesson can be extended since the audio can be saved freely on your computer for future use. This obviously has the benefit of saving time when you need to reuse the materials for something else, or

give said materials to the students if they want to listen to it during their own time, or even make it into a self exercise tool for the students to practice their listening skill at home.

e. TTS Allows Its User to Facilitate Listening Materials Freely

Following this line of thought, the researcher also asked about the freedom TTS provides for the creation of original and personalized audio materials, does TTS provides the user the ability to fully express themselves when creating listening materials? (i.e. not being limited to a certain subject or topic) and they have this to say for the matter, “Yes, of course. It can be as freely as i want”.

This again goes hand in hand with the second question regarding the creative freedom TTS provides.

f. TTS as a Natural Replacement for Human Voices

And lastly, the researched wants to know the teacher’s opinion regarding the voice produced by TTS. is TTS’s generated voice are too robotic? If so, then does it impact the usability of the tool? In which the teacher answered, “No, the produced voice is suprisingly extraordinarily natural and doesn’t sound like a robot”.

The researcher further questioned the overall usability of the produced voice, which they followed with, “it will be good when used because clear sound helps a lot in practice”.

C. Discussion

In this section, all of the informations obtained during the research and displayed during the findings section will be utilized here in order to find the answer to each of the research questions. Specifically about how someone can use TTS in

order to craft or create listening materials from scratch, and how said someone feel about the technology as a whole, and compare the similarities or differences between this research and those done before it.

1. The Processes of Creating Listening Materials using TTS Performed by The English Teacher

a. The Creative Processes of Creating Listening Materials using TTS

The creative processes behind the creation of listening materials done by the teacher is one-to-one with the steps laid in Mulyono (2016)'s paper. The teacher starts by scripting the manuscript, keeping in mind what the student's needs are and sets a clear objective, the proceeds to adjust the overall length of the lesson and paid attention to the wordings used as to not confuse the students. Then during the conversion, the teacher follows the main steps in this stage faithfully, such as connecting the words and sentences from the script into the program, adjusting the rate in which the speech is performed, selecting the appropriate voices for the accent and intonation, and lastly converting it into an audble format. The creative process between the teacher and the steps included in Mulyono (2016)'s paper is similar with one another.

b. The Accessibility and Ease during The Creative Process of Producing Listening Materials through TTS

The teacher expressed an overwhelmingly positive response regarding TTS, and how one can easily and quickly learn about its usage. The conveniences it provides during the creative process easily earns their approval of the technology.

In a paper written by Moon (2020), they come to a conclusion that the vast majority of students welcome and adapted to the usage of TTS. They examined the response of the students after giving them an access to TTS program that allows them to create their own listening materials. And the response obtained by the survey of 119 students are mostly positive, A majority of learners (around 73% and 60.5%, respectively) agreed that TTS is useful and enjoyable to use, and those people also find enjoyment of the listening material production activities. While only a few (around 9.3% and 16.9%) answered that they were not useful or enjoyable.

This ties back to the accessibility and ease of TTS during the creation process. students that are given the tool managed to adapt and master it and they find it to be useful, and this sentiment echoes back to the teacher's experience. As they can easily make listening materials from only texts on the student's book and also gave their insight into the entire creative process.

2. The View of The English Teacher with Respect to using TTS as a Method to Create Listening Materials

a. The Advantage TTS Provides Compared to Using Normal Human Voices

The teacher has shared their thoughts regarding the advantages of using TTS, such as the first thing they mentioned is that TTS can save time for them, since it does not take very long to create your own original audio materials to be used in class. Andriani (2021) in her paper shared her findings about TTS, specifically how normal human voice compares to TTS generated voices in terms practical use such as listening materials. In her paper she stated that after comparing

two groups of students, one experimental and one control group, she concluded that TTS does have an edge when it comes to practicality over conventional human voices. The result of the post-test performed by Andriani (2021), the mean score of experimental group was 77.65 which higher than the mean of control group was 73.91.

This result supports the teacher's comment about the usage and viability of TTS replacing human voices altogether when using this research's context. The teacher commented about how natural the voice sounds and how good it is to be used for students to practice, since TTS offers many options such as male or female voices, british or american, how fast or slow the speech is and so on. So the teacher can use the same text, but change the speaker into a woman with a british accent and faster speaking speed for example if they wishes to do so, for variety sake.

b. Voices Produced by TTS has Reached the Level of Naturality that of a Human Being

The teacher has expressed their positive opinions regarding the naturalness of the voice produced by the TTS program, saying, "the produced voice is surprisingly extraordinarily natural and doesn't sound like a robot". This statement is in line with the findings of Nakai (2019) and Mulyono (2016), in their paper they praised TTS because of how human-like the voices are. The students in Nakai (2019)'s research has shown a boost of confidence in dictating English words. The students no longer feel scared and insecure whenever a native speaker spoke during practice ever since they implemented TTS into their study plans. This shows us how similar TTS and

native voice sounds like for even a stranger in the language could adapt more quickly using the TTS program.

And we also find similar case in Mulyono (2016)'s paper. In said paper consists of the teacher's and student's perception regarding the use of TTS as a substitute of native voice in creating listening materials. The findings shows that both the teacher and the students had a positive opinion about TTS program, both shows a somewhat enthusiastic attitude regarding their experiment of using TTS. Teachers said that the TTS materials corresponded to their classroom needs all the while the facilitating the teachers a tool to self-develop listening materials that suited the standard of competence required by the learning syllabus. Meanwhile the student's had a great experience too since they can adjust the speed to their liking and also because the materials are more relatable to them since it is custom made. A great majority of them (95.39% of the 65 students) are even excited about it.