

Learner-Generated Digital Listening Materials Using Text-to-Speech for Self-Directed Listening Practice

Dosik Moon

Associate Professor, Dept of English, Hanyang Cyber University, Korea
dmoon@hycu.ac.kr

Abstract

This study investigated learners' perceptions of using self-generated listening materials based on Text to Speech. After taking an online training session to learn how to make listening materials for extensive listening practice outside the classroom, the learners were engaged in practice with self-generated listening materials for 10 weeks in a self-directed way. The results show that a majority of the learners found the TTS-based listening materials helpful to reduce anxiety toward listening and enhance self-confidence and motivation, with a positive effect on improving their listening ability. The learners' general satisfaction can be attributed to some beneficial features of TTS-based listening material, including freedom to choose what they want to learn, convenient accessibility to the material, availability of various native speakers' voices, and novelty of digital tools. This suggests that TTS-based digital listening materials can be a useful educational tool to support learners' self-directed listening practice outside the classroom in EFL settings.

Key words: *Text to Speech, Extensive Listening, Self-Directed Listening Practice, EFL*

1. Introduction

For learners learning English as a foreign language (EFL), listening is one of the most difficult challenges to cope with primarily because listening is a complex and active process in which speech decoding and content understanding proceed in real time in the flow of sound provided continuously [1]. Learners, who are not proficient in English, should invest a plenty of time and effort over a considerable period of time to reach a proficient listening comprehension level. However, in the current English education, the time devoted to listening is considerably insufficient so that the learners need to be constantly engaged in listening practice outside of class hours as well as in the classroom.

Extensive listening (EL) can be an effective way to facilitate learners' self-directed listening practice outside the classroom. EL, the activities of listening to a large amount of materials that are easily comprehensible, is known to have a considerable benefits in improving language skills as well as listening ability [2]. However, since most of the current listening classes depend on ready-made materials selected by the teachers, it cannot effectively reflect individual learners' different English levels and interests. One possible

solution to this is to train learners to generate listening materials suitable for their own learning needs.

These days, anyone can easily make a digital listening material by using Text to speech (TTS) and digital software available on the Internet. TTS is a device capable of converting text data into a human voice featuring a wide variety of English accents and the function to freely adjust the speech rate [3]. A growing number of EFL teachers have been developing native-like listening materials using TTS to provide their students with ample comprehensible input. However, no attempts have been made to deal with learner-generated listening materials.

In this study, to support learner's self-directed EL, an online training session was developed where learners were trained to generate digital listening materials by using TTS and a video recording software. After the session, the learners were engaged in EL for 10 weeks in a self-directed way. The purpose of this study is to investigate learners' perceptions of how useful and helpful making and using the TTS-based listening materials promote and support EL outside the classroom.

2. Previous research

2.1 Extensive Listening

EL refers to any type of listening activity that is exposed to a large amount of easy-to-understand aural input that the learners enjoy listening to [2]. It is a learning method developed under the concept of 'listening is best learned through listening', influenced by the results of extensive reading research, which are known to have a positive effect on improving reading ability [3]. Unlike intensive listening, which focuses on grasping accurate sounds, words, expressions, and details, EL aims to develop "listening fluency" so that learners naturally acquire listening fluency and automatic language processing ability through a long-term exposure to a large amount of meaningful input.

A growing number of research studies show that EL has a variety of positive effects on listening improvement, particularly in the EFL context where learners lack sufficient exposure to aural input of the target language [2-4]. Engaged in extensive listening, learners can develop discernment for various intonations, acquire vocabulary, and improve pronunciation and speaking ability. It has been also found that listening while viewing the listening manuscript is more effective than just listening [5]. Furthermore, EL can promote learners' motivation and self-confidence in English, enhancing learner autonomy in learning language in and outside the classroom [6].

One important prerequisite for successful implementation of EL is to find suitable listening materials for learners. According to research, listening materials suitable for EL should be easy enough for learners to understand more than 90% of the content without the aid of a dictionary or grammar book, and to understand just by listening without stopping in the middle [6]. A possible best way to find these listening materials is to have individual learners actively involved in the selection of the materials and reflect their opinions. However, it has been a common practice in implementing extensive listening that teachers select and use ready-made materials.

This teacher-centered approach, which does not sufficiently reflect the individual differences of learners, is likely to have a negative impact on learners' motivation for learning. It is because listening materials that are interesting and understandable to one learner may be perceived completely differently by other learners due to differences in individual learners' experiences, interests, background knowledge and English proficiency. Therefore, using learners-created listening materials can be a better choice than ready-made ones in enhancing learning motivation and improving listening ability.

2.2 Text to Speech in EFL Settings

TTS is a tool that converts text input into speech output through a specially designed computer program. TTS has been recognized as having a great potential for EFL from the beginning of its development due to the advantage of being able to easily provide comprehensible language input. However, skepticism about the suitability for actual language education was strong due to the limitations that early TTS could not faithfully reproduce the naturalness, accuracy and expressiveness of human speech [7].

However, as a result of the recent rapid development of IT technology, TTS has become able to generate natural and accurate pronunciation and intonation almost similar to human voices [8]. These days, TTS is widely used for applications such as desktop voice systems, audio books and electronic dictionaries [9]. In the field of EFL, an increasing number of classes have been utilizing TTS, and research results have confirmed educational effectiveness in various areas such as word recognition, pronunciation and vocabulary learning, reading, writing, and speaking [10].

Until now, relatively few classes have applied TTS to listening education. Recently, a growing number of EFL teachers have been developing native-like listening materials using TTS. Research applying these TTS-based materials to actual classes reveals that TTS provides a lot of practical advantages for listening in EFL classrooms such as availability of various English accents, convenient speech rate control functions, cost efficiency, and simplicity of editing synthetic digital voices [11-12]. In these studies, teachers were responsible for developing listening materials using TTS, but if learners produce their own listening materials, they can be more suitable for their language levels and interests and be expected to produce much better effects than teacher-generated ones.

3. Methods

3.1 Background and Procedure

The study was conducted in an introductory English reading course at a Cyber university in Seoul. Among 263 high-beginning level students registered for the course, none of them had used TTS to learn English before. The learners learned how to make digital listening materials in an online training session, which was designed based on the following seven principles that Renandya and Jacobs (2016) proposed to consider for successful implementation of EL [3].

1. Clearly state your goals for broad listening.
2. Make sure to listen to it in large volumes.
3. Use materials within the range of learners' English proficiency.
4. Provide continuous support to learners
5. Keep the learner's motivation high in the learning process.
6. Encourage reading and listening at the same time.
7. Provide interesting and enjoyable listening activities

Considering that learners are unfamiliar with EL, the purpose and benefits of EL were presented through concrete examples at the beginning of the training session. In order to help learners select a listening manuscript that suits their English proficiency and interests, they were asked to choose paragraphs in the textbook used in the class. More than 90% of these self-selected texts are easily comprehensible to the learners because they were covered in class, so the learners can comfortably focus on listening itself without being frustrated by unfamiliar words or expressions. Before making the digital listening materials, the learners were trained to use

TTS and video recording software practice step by step at the online training session.

3.2 Text to Speech and Video Recording Software

Two types of TTS tools were used to develop the digital listening materials, Vocalware (<https://www.vocalware.com/index/demo>) and Oddcast TTS (<https://ttsdemo.com/>) released by Oddcast, Inc. The selected TTS tools meet the following criteria. First, the pronunciation is clear and natural. Second, a variety of speakers are available for learners to choose from. Third, the speech rate can be easily controlled. Fourth, at least 600 letters of the alphabet can be processed.

The selected TTS devices are two different versions of the same TTS, basically sharing the following features in common. First, 600 letters of the alphabet, which are about one paragraph in length, can be converted into audio. Both of them provide voices of American, British, Australian and Indian English speakers, as well as male and female with the functions of adjusting the speech rate.

These two versions have some different features (Figure 1). Vocalware adopts the most common form of TTS that generates voice by simply typing a manuscript in the text input window and pressing the Say It button. One downside of Vocalware is that it does not provide speech rate control function. Oddcast TTS is unique in that it generates voice through an accurate lip sync of a character as if a real person speaks, which can increase the interest of learners. The learners were given the option to choose one of these TTS tools depending on their preference.

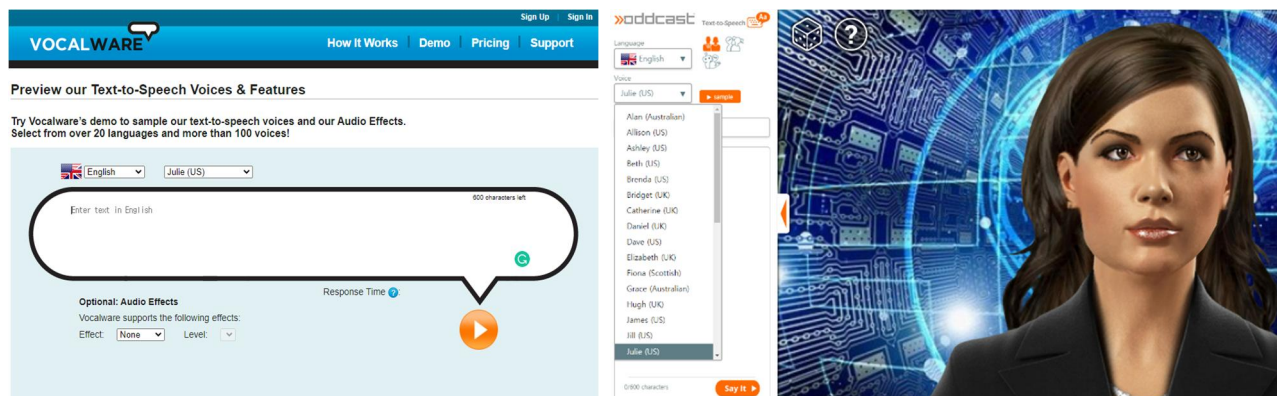


Figure 1. Vocalware & Oddcast TTS

Listening with visual support is known to be helpful to facilitate learners' listening comprehension, particularly for less proficient language learners [5]. As a way to provide visual support, video files were produced with GOMCam (<https://www.gomlab.com/gomcam-screen-recording/>), a free software with a screen recording function, and the learners saved the files on their mobile phone to access them anytime, anywhere.

3.3 Digital Listening Materials Production Process

The learners are supposed to select at least 3 paragraphs each week from the reading passages covered in class. Once they select one, they produce digital a listening material following the steps below.

Step 1: Selecting TTS and converting text input to speech

The learners choose one of the two versions of TTS, and then select the speaker and the type of English they want to listen to. They can adjust the rate of the speech if they want. After typing the manuscript on the text

input screen of TTS, the learners convert it to voice by pressing the Say IT button.

Step 2: Video recording

The learners record the TTS screen including voice by using the screen shooting function of the video recording software. Once the recording is finished, the learners save the video file on their smartphone and play and listen to it as many times as possible when they want, just enjoying the contents without paying too much attention to every word or expression. The learners repeated the same process for each new recording for 10 weeks.

3.4 Data Collection and Analysis

A questionnaire was administered to investigate the participants' reaction and attitudes towards self-generated digital listening materials. A questionnaire consists of 10 items on a 5-point Likert-scale from 1 (strongly disagree) to 5 (strongly agree). The items address three areas: perceptions of making the digital listening materials, impacts on affective domains, and impacts on language skills. Two open-ended questions were added to gain deeper understandings of learners' perceptions about the strengths and weaknesses of the listening materials. Frequency analysis was conducted on the responses to the survey, and some significant comments from the open-ended question were reported.

4. RESULTS

At the end of the semester, 119 out of 263 learners completed the survey. Items 1-3 are related to perceptions of making the digital listening materials (Table 1). A majority of learners (73% and 60.5%, respectively) agreed on the usefulness and enjoyment of the listening material production activities, and only a few (9.3% and 16.9%) answered that they were not useful or enjoyable. Concerning the easiness of making the materials, more learners answered that it was easy (44.5%) than those who answered that it was not easy (35.3%), but the difference was not so remarkable. These results show that although the learners have been trained in detail on how to use TTS and video recording software, making digital listening materials was still challenging for some learners, particularly those who are not familiar with digital software.

Table 1. Perceptions of making the digital listening materials

Item	SD	D	N	A	SA
1. Making the digital listening materials was a useful activity for learning English.	2 (1.7%)	9 (7.6%)	21 (17.6%)	68 (57.1%)	19 (15.9%)
2. Making the digital listening materials was enjoyable.	6 (5.1%)	14 (11.8%)	27 (22.7%)	59 (49.6%)	13 (10.9%)
3. Making the digital listening materials was easy.	11 (9.2%)	31 (26.1%)	24 (20.2%)	44 (36.9%)	9 (7.6%)

Items 4-7 deal with issues related to impacts on affective domains (Table 2). 60.4% of the learners responded they enjoyed listening to the material, and only 13.5% said they did not. To the questions on improving learning motivation (79.8%), reducing listening anxiety (80.6%) and increasing confidence in listening (77.2%), over 75% of the learners answered that they were helpful, while only about 10% of the learners answered that they were not helpful. These results suggest that self-generated listening materials were considerably beneficial for the learners in terms of their affective domains.

Table 2. Perceptions of impacts on affective domains

Item	SD	D	N	A	SA
4. The digital listening materials were enjoyable to listen to	4 (3.4%)	12 (10.1%)	31 (26.1%)	58 (48.7%)	14 (11.7%)
5. The digital listening materials motivated me to practice more.	4 (3.4%)	10 (8.4%)	10 (8.4%)	77 (64.7%)	18 (15.1%)
6. Listening to the digital materials was helpful to overcome anxiety in listening.	3 (2.5%)	9 (7.6%)	11 (9.2%)	75 (63.0%)	21 (17.6%)
7. Listening to the digital materials was helpful to build confidence in listening.	5 (4.2%)	8 (6.7%)	14 (11.7%)	73 (61.3%)	19 (15.9%)

Items 8-10 deal with the effects of impacts on language skill improvement (Table 3). Regarding listening comprehension ability, a majority of the learners (68%) said it was effective and only 11.7% said it was not effective. About half of the learners (53.2%) agreed on the overall improvement in English proficiency, but 20.1% did not. However, compared with the high percentage of positive reactions to the effect of learning English, relatively a small number of the learners (44.9%) wanted to learn by creating listening materials on their own in the future. Possible reasons for this result will be discussed when answers to the open-ended questions are addressed.

Table 3. Perceptions of language skill improvement

Item	SD	D	N	A	SA
8. Listening to the digital materials was effective to improve my listening comprehension ability.	3 (2.5%)	11 (9.2%)	24 (20.2%)	65 (54.6%)	16 (13.4%)
9. Listening to the digital materials was effective to improve my English ability in general.	5 (4.2%)	19 (15.9%)	31 (26.0%)	52 (43.7%)	12 (10.2%)
10. I would like to make digital listening materials for listening practice after this semester.	7 (3.4%)	23 (19.3%)	38 (31.9%)	43 (36.1%)	8 (6.8%)

Finally, Table 4 presents the four most frequent answers to the open-ended questions regarding the pros and cons of the self-generated listening materials.

Table 4. Pros and cons of the self-generated digital listening materials

Advantages	Disadvantages
Freedom to choose what they want to learn	Unstable Quality of the sounds produced by TTS
Convenient Accessibility to the material anytime, anywhere	Technical difficulties in handling software properly
Availability of various native speakers' voices	Much time consuming to make the materials
Fun to learn how to use software useful for learning English	Not fun to listen to what was already learned

Most learners regarded it as the greatest advantage that they were given freedom to choose what they want to listen to and access it anytime, anywhere by using cellphones. Some felt it beneficial to be able to be exposed to various English accents, including British, American, Australian, and even Indian English, just by clicking the button. Others liked learning how to use online software to learn English, particularly young learners who

are familiar with digital tools. These advantages of self-generated listening materials seem to have a positive effect on interest, motivation and self-confidence of the learners, contributing to creating a convenient and enjoyable learning environment.

As for the disadvantages, several learners mentioned the technical difficulties that arose from using unfamiliar online digital software. Some learners seemed to experience frustration because the voice generated by TTS was not clear enough to understand, especially when it was slowed down. In addition, complaints were raised that producing listening materials was difficult and too much time-consuming. To the survey item 10 (I would like to make digital listening materials), relatively a small number of the learners gave a positive response. Technical difficulties these learners experienced in making the listening materials seems to have influenced their negative reaction to the item 10. On the other hand, some learners mentioned that listening to the materials was not fun because they had to listen to what they had already learned in class. This response suggests that easy-to understand materials do not always meet students' interests. It seems necessary to broaden options for selecting listening materials in the next course. In sum, the survey results shows that the learners' overall perceptions of their learning experience with the listening materials was positive, finding them useful and helpful for listening practice outside .

5. CONCLUSION

This study investigated learners' perceptions of self-generated digital listening materials based on TTS. The results show that most of the learners believed that the digital listening materials were helpful to reduce their anxiety about listening and increase self-confidence and motivation, having a positive effect on improving their listening ability. The general satisfaction with learning experience can be attributed to the advantages of using self-generated listening materials including freedom to choose what they want to learn, convenient accessibility to the material, availability of various native speakers' voices, and novelty of TTS and the video recording software. This suggests that digital listening materials based on TTS can be an educational tool that effectively support learners' self-directed listening practice outside the classroom.

Despite these positive results, this study has limitations in several respects. First, since the research was conducted in a short period of time in a specific situation of a Cyber university, the results is not generalizable to other educational settings. In addition, since this study entirely relied on the learner's self-report, it did not measure to what extent does the use of TTS-based listening materials affect the listening ability. In the future, follow-up studies should be conducted applying both qualitative and quantitative research methods for learners with different characteristics in various educational settings.

References

- [1] H. Ahn, "The Relationship between phonemic awareness and listening comprehension ability: an experimental study based on the English name test," *Foreign Language Education Research*, Vol.10, pp. 39-48, 2007. <http://hdl.handle.net/10371/95926>
- [2] A. W. Renandya and T. SC. Farrell, "'Teacher, the tape is too fast!' Extensive listening in ELT," *ELT journal*, Vol. 65, No. 1, pp. 52-59, 2011. DOI: <https://doi.org/10.1093/elt/ccq015>
- [3] A. W. Renandya and G. M. Jacobs, "Extensive reading and listening in the L2 classroom," *English language teaching today*, Springer, Cham, 2016. 97-110. DOI: https://doi.org/10.1007/978-3-319-38834-2_8
- [4] Y. Vo, "Developing extensive listening for EFL learners using Internet resources," *Hawaii Pacific University TESOL Working Paper Series 11*, pp. 29-51, 2013.
- [5] F. M. Ivone and W. A. Renandya, "Extensive listening and viewing in ELT," *TEFLIN Journal*, Vol. 30, No. 2, pp. 237-256, 2019. DOI: <http://dx.doi.org/10.15639/teflinjournal.v30i2/237-256>

- [6] D. Ewert and M, Rebecca, "Extensive listening in a self-access learning environment." In Proc. the 2011 Michigan Teachers of English to Speakers of Other Languages Conference, pp. 24-40, October 7-8, 2011.
- [7] D. Moon, "Web-Based Text-to-Speech Technologies in Foreign Language Learning: Opportunities and Challenges," In Computer Applications for Database, Education, and Ubiquitous Computing, pp. 120-125, 2012.
DOI: https://doi.org/10.1007/978-3-642-35603-2_19
- [8] Z. Handley, "Is text-to-speech synthesis ready for use in computer-assisted language learning?," *Speech Communication*, Vol. 51, No. 10, pp. 906-919, 2009. DOI: <https://doi.org/10.1016/j.specom.2008.12.004>
- [9] T. Bione and C. Walcir, "Synthetic Voices in the Foreign Language Context." *Language Learning & Technology*, Vol. 24, No. 1, pp. 169-186, 2020. DOI: <https://doi.org/10125/44715>
- [10] Dosik Moon, "Development and evaluation of an English speaking task using Smartphone and Text-to-Speech," *The Journal of the Institute of Internet, Broadcasting and Communication (JIIBC)*, Vol. 16, No. 5, pp. 13-20, October 2016. DOI: <https://doi.org/10.7236/JIIBC.2016.16.5.13>
- [11] G. Sha, "Using TTS voices to develop audio materials for listening comprehension: A digital approach," *British Journal of Educational Technology*, Vol. 41, No. 4, pp. 632-641, 2010.
DOI: <https://doi.org/10.1111/j.1467-8535.2009.01025.x>
- [12] H. Mulyono and D. N. Vebriyanti, "Developing native-like listening comprehension materials perceptions of a digital approach," *Journal of ELT Research: The Academic Journal of Studies in English Language Teaching and Learning*, Vol. 1, No. 1, pp. 1-20, 2016. DOI: https://doi.org/10.22236/JER_Vol1Issue1pp1-20