The Effects of using Online Text-To-Speech Tools on EFL Students’ Perceptions in Learning Pronunciation

Tran Van Duong
School of Foreign Languages, Can Tho University, Vietnam

Abstract — The effects of Text to Speech tools on learning and teaching pronunciation have been emphasized. However, not many studies on the topic were conducted in the Vietnamese context. Therefore, the current study aims to (1) to verify whether the effects of TTS Tools into students’ pronunciation and (2) to verify how it effects on their pronunciation. The study was designed as an experimental study, following a mixed-methods approach, using questionnaire, and semi-structured interviews to collect quantitative and qualitative data. 60 high school students took part in this investigation as participants. The results demonstrated a significant improvement in EFL students’ pronunciation in the study. The participants in the also highly perceived the positive effects of on their English pronunciation accuracy. Finally, some implications were drawn from the findings so as to enhance the quality of pronunciation teaching and learning.

Keywords — CAPT, pronunciation, language teaching and learning, English as a foreign language

I. INTRODUCTION

Pronunciation plays an important role in English speaking and communication. Fraser (2000) explained that pronunciation is one of the most important sub-skills in English speaking. She argued that it was very difficult to understand the speakers with poor pronunciation, despite accuracy in other areas. In contrast, good pronunciation helps the learners contribute a successful communication, particularly intelligibility (Derwing and Munro, 2005). In fact, English as Foreign Language learners often expect to be able to have a good pronunciation that like a native speaker. However, EFL leaners usually faced with many difficulties in pronounce English in their speaking such as lack of motivation and exposure, instruction problem, teacher’s factor, etc., as mentioned in Roslaimi & Komara’s study (2020). One of the majored problems that they find difficult is the pronouncing the correct sounds of another language system. In addition, many teachers are not aware of the importance of pronunciation. According Harmer (2001) teachers focus on teaching grammar and vocabulary in learning a foreign language in order to help learners develop their listening and reading skills. Secondly, many teachers think that pronunciation study is too difficult and for learners. Moreover, Abdolmanafi-Rokni (2013) assumed that learners get difficulties with different pronunciations uttered by their teachers. To solve this problem, (Rasekhi Kolokdaragh, 2010) recommended that teachers should use ICT such as different kinds of computer software in their teaching to help students enhance their pronunciation by exposing them to authentic materials. To support this idea, Mehrpour et al. (2016) revealed that pronunciation instruction should change by using modern technologies. One of these technologies called computer-assisted pronunciation teaching (CAPT) helped teachers to provide better learning environments for their learners. In Vietnam context, Pham (2005) determined “traditional teaching methods, emphasizing the role of teaching grammar and vocabulary rather than communicative competence”. Besides, Nguyen & Newton (2020) demonstrated contextual factors appeared to affect pronunciation teaching because the teachers lacked pre-service and professional training in pronunciation pedagogy. Pronunciation pedagogical instruction is an area of TESOL that is a major factor in the problem Vietnamese learners to communicate successfully in English. Because of the benefits form TTS in language learning and the difficulties in pronunciation of Vietnamese learners, the researcher decided to conduct this study to investigate whether the effects of TTS tools into EFL students’ pronunciation accuracy in the Mekong Delta. The study is aimed to answer two research questions:

1. What are the effects of using Text to Speech on EFL students’ pronunciation accuracy?
2. What are the EFL students' attitudes toward the use of Text to Speech tools to improve their pronunciation accuracy?
II. LITERATURE REVIEW

1. Pronunciation

Pronunciation is an action in the making of English sounds, (Cook, 1996; Gilakjani, 2016). According to Yates (2002), pronunciation is the sounds system that is used for creating meaning. Furthermore, Richard and Schmidt (2002) stated pronunciation as the way of producing specific sounds.

1. Computer-Assisted Pronunciation Teaching (CAPT)

Pronunciation is an important element in second language acquisition, but it is usually unnoticed in teaching and learning (Derwing et al. 2012). In fact, the ways to give pronunciation instruction should be changed. CAPT is considered as a technique to offer learners stress-free apply on pronunciation (Neri et al. 2002). Many studies showed that EFL teachers with CAPT have more advantages than teachers with traditional methods in teaching pronunciation. Furthermore, integrating CAPT into English language courses can contribute significantly to enhance learners’ pronunciation learning. Hirata (2004) realized that CAPT was very useful in improving L2 learners’ skill to learn different features of pronunciation. Similarly, Bakla and Demirezen (2018) revealed that CAPT software provides the learners chance to self-regulate their pronunciation.

2. Text to Speech Tools

Text-to-Speech (TTS) synthesizer is a program in which the speech is convert from written text automatically. TTS technologies are now widely used with good quality and native-speakers voices such as desktop speech systems, computer voice interfaces, audio books, electronic dictionaries, etc. Moreover, a number of English teachers utilize TTS to create spoken input to the students. Recently, with development of ICT, there are many different types of web-based technologies for creating audio from texts. Among such technologies, Text-to-Speech tools are easily to convert written text into spoken words which can provide a new way for learners to experience English language input (Moon, 2012).

3. Teaching pronunciation in Vietnam

According to Cunningham (2012) Vietnamese speakers faced with various difficulties of pronunciation in English learning. There are many different aspects between the two languages that lead to the challenging Vietnamese learners. A study conducted by Nguyen & Newton (2020) showed that all the teachers tend to an intelligible pronunciation for successful oral communication rather than native-like pronunciation. However, the teachers’ self-reports in this study revealed that both were lacking or largely absent in their teacher’s training program.

4. Related study

Pronunciation is one of the key factors to contribute to intelligible communication. Nevertheless, English pronunciation teaching has not been emphasized in Vietnam’s national education system (Nguyen, 2020). Therefore, Vietnamese learners usually get difficulties in pronouncing English sounds in their learning and communication. This part would review the related studies to the effects of using CAPT in teaching pronunciation in ELT as well as teachers’ beliefs and practices in teaching pronunciation in Vietnam.

Eksi & Yesilcinar (2016) carried out a study aimed to investigate the effectiveness of text-to-speech tools used preparing for an oral achievement test. The participants included 43 junior teacher trainees at a large state university in Turkey. In conducting this study, pre-test and post-test were used as instrument. In addition, pre-questionnaire with open-ended questions will aim to explore the trainees’ opinions related to pronunciation and their practices to improve pronunciation. The results indicated that the trainees perceived a native-like accent as a measure of being a good language teacher. It was also revealed that text-to-speech websites had significant effects in improving trainees’ pronunciation.

A study to support pronunciation improvement for non-native Chinese learners’ confidence in their oral production was implemented by Soon & et al (2017). The study is conducted to investigate the impacts of Chinese pinyin text to speech system in improving students’ pronunciation accuracy for oral output. An empirical study was carried out in UiTM Terengganu including 56 students in preliminary, intermediate and advanced stages as samples. The findings of this study showed that Chinese pinyin text to speech system can assist students in developing pronunciation accuracy for their oral production. It also indicated that students of the preliminary stage had significantly more positive attitude than learners of the intermediate and advanced levels.
In 2020, Saleh and Gilakjani conducted a study to explore the impacts of Computer-assisted pronunciation teaching (CAPT) software on learners’ pronunciation. This study aimed to investigate the effects of computer-assisted pronunciation teaching (CAPT) on Iranian intermediate EFL learners’ pronunciation ability. Seventy EFL learners were selected as the participants based on their performance on Quick Placement Test (QPT). Then, they were randomly divided into two groups, one experimental and one control group. In this study, pre-test and post-test of pronunciation were used as instruments. Then, CAPT was introduced to the experimental group as the treatment, whereas the control group received no treatment. After analyzing the data, the researcher found out that CAPT had a statistically positive effect on Iranian intermediate EFL learners’ pronunciation. The results also indicated in the post-test scores of experimental group was higher than that of control group.

In Vietnam context, Tran & Nguyen (2020) carried out a study to explore the teachers’ beliefs in teaching pronunciation. Questionnaires, interviews, and observations were used as instruments in fifteen-week of the pronunciation course. Participants included seventy EFL teachers at an English language center in the Mekong Delta. The findings indicated that Vietnamese learners get trouble in pronouncing English ending sounds in their pronunciation. These problems may somehow hinder learners in successful communication.

III. RESEARCH METHODOLOGY

A. Research designs

The study was conducted with a mixed method by combining both qualitative and quantitative approaches. The quantitative including questionnaire will aim to investigate whether the effects of TTS on EFL students’ pronunciation. The qualitative data will be collected through a semi-structured interview which will collect more insightful information about EFL students' attitudes when using TTS to improve pronunciation.

B. Participants

The participants consist of 2 experienced teachers in teaching pronunciation and 90 EFL students at a high school in Mekong Delta. They are voluntary and willing to participate as population in this study. To homogenize the participants, quick survey will be administered to exclude students who are participating in the pronunciation course or using another method for improving pronunciation and 60 students will be selected as the main participants of this study.

C. Data analysis and Procedure

The researcher developed 6 sets of lesson plans based on the pronunciation section in the textbook with supporting by TTS tools to teach the students. After the intervention, a questionnaire including 34 items was aimed to investigate students’ attitude toward using TTS tools and their expectations of their pronunciation. The participants indicated their attitudes on five-point scales; (1) strongly disagree, (2) disagree, (3) neutral, (4) agree and (5) strongly agree. After that, the data was gathered and analysed by SPSS software.

In the current research, interview will be used as qualitative instrument to find out deeper information about participants’ perception and their changing. Basing on their willingness to participate in the study, five of them will be collect randomly from the two groups. The aim of using interview is (1) to verify whether the effects of TTS Tools on students’ pronunciation and (2) to verify how it affects on their pronunciation. The interview will be carried out in Vietnamese so that the interviewees can understand clearly and express their opinion freely. Finally, both of data gathered from tests and interview as the primary data will be analysed.

IV. FINDINGS

1. The effects of Text to Speech tools on EFL students’ pronunciation

This section aims to investigate how the TTS affect EFL learners’ pronunciation in their oral production. The Descriptive Statistics was conducted to compute the mean score of this section. The results from table 4.1 shows that the mean score of this section is rather high (M = 3.8). The one-sample test was run on the mean score of how TTS affect EFL students’ pronunciation and the test value 3.4.
TABLE 4.1
THE DESCRIPTIVE STATISTICS ABOUT THE EFFECTS OF TTS ON EFL LEARNERS’ PRONUNCIATION

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60</td>
<td>3.00</td>
<td>5.00</td>
<td>3.80</td>
<td>.58640</td>
</tr>
</tbody>
</table>

The result shows that there is a significant difference between the mean score of how TTS affect EFL students’ pronunciation (N=60, M2 =3.8) and the test value 3.4 (t = 3.05, df = 19, p = .00). The result suggests that TTS had a significant effect on EFL learners’ pronunciation.

As presented in Table 4.2, a great number of the respondents agreed with the statements “I would like to use the Text to Speech software in learning features of pronunciation”, “Text to Speech is effective in improving stress and intonation patterns” and “Text to Speech is effective in learning features of phonemes”. The result indicated that the TTS tools had positive effects on pronunciation features such as phonemes, word stress and intonation.

TABLE 4.2
THE EFFECTS OF TTS ON PRONUNCIATION PERCEIVED BY STUDENTS

<table>
<thead>
<tr>
<th>Statements</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>28. I would like to use the Text to Speech software in learning features of pronunciation.</td>
<td>60</td>
<td>2.00</td>
<td>5.00</td>
<td>3.65</td>
<td>.87509</td>
</tr>
<tr>
<td>29. Text to Speech is effective in improving stress and intonation patterns.</td>
<td>60</td>
<td>2.00</td>
<td>5.00</td>
<td>3.75</td>
<td>.85070</td>
</tr>
<tr>
<td>30. Text to Speech is effective in learning features of phonemes.</td>
<td>60</td>
<td>2.00</td>
<td>5.00</td>
<td>4.00</td>
<td>.85840</td>
</tr>
</tbody>
</table>

The interview data provided further explanation. Most of learners stated that TTS tools had positive effects on some features of pronunciation as accuracy, intonation, word stress and accent. In fact, participant 1 said that her English pronunciation was improved.

Most of participant agreed that TTS help the pronounce more accurately as participant 1 expressed:

“TTS software helps me form the habit of listening correctly and pronouncing more accurately through listening to native speakers' voices via TTS”.

Supporting her ideas, students 3 said

“...I recognize that my correct pronunciation has improved a lot. Even when I say it wrong, I know what word I’m mispronouncing and I need to correct it...”.

In addition, they said that some features of pronunciation such as intonation, word stress, accent were enhanced after exposing to TTS. Two interviewed EFL learners said:

“I can know the way native speakers say whole sentences...raising and falling intonation...”  
(Participant 3)

“By listening to native speakers, I can recognize that which key words they emphasize, raise their voice in some sentences, pronounce consonants and vowels clearly so you can imitate....”  
(Participant 1)

The results of the interview indicated that TTS tools have positive impacts on EFL learners’ pronunciation. It can be generally stated that TTS software improves students’ pronunciation in terms of word accuracy, intonation, accent and word stress. Accuracy and intonation are the most dominant effect on students’ pronunciation on their oral production.
2. The EFL students’ attitudes towards using Text to Speech tools in their learning

The final aim is to investigate the students’ attitudes towards using TTS tools to assist their pronunciation learning. As Table 4.3 has illustrated, the mean score of EFL students’ attitudes towards the use of TTS on English pronunciation (N=60, M3 = 3.78) is higher than the third scale (3.0) among five-point scales. The results supported the conclusion that the EFL students have positive attitudes towards using TTS to improve their pronunciation. To be specific, the mean score of all statements were high in comparison with the third scale (3.0) among five-point Likert scale (see Appendix). Then, one sample T-test was run on the mean score of EFL students’ attitudes towards using TTS to improve their pronunciation and the test value 3.4 (average). The result indicated that the mean score of how TTS affect EFL students’ pronunciation (N = 60, M3 = 3.78) and the test value 3.4 (t = 2.3, df = 19, p = .03) are difference. In conclusion, the EFL learners’ attitudes of using TTS tools in improving their English pronunciation are at the high level.

<table>
<thead>
<tr>
<th>Statements</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>22. The pronunciation program introduced by the teacher is very easy to use.</td>
<td>60</td>
<td>2.00</td>
<td>5.00</td>
<td>3.65</td>
<td>.81273</td>
</tr>
<tr>
<td>23. Computer-assisted pronunciation instruction is interesting and useful.</td>
<td>60</td>
<td>2.00</td>
<td>5.00</td>
<td>3.85</td>
<td>.87509</td>
</tr>
<tr>
<td>33. I would like to use the TTS program again in the future</td>
<td>60</td>
<td>3.00</td>
<td>5.00</td>
<td>3.85</td>
<td>.74516</td>
</tr>
</tbody>
</table>

The questionnaire data proved that students regard this software as a useful tool to develop their pronunciation. By using this software, students can practice repeatedly, which enhances positive attitudes towards learning and increases learning efficiency.

The interview data provided further explanation. Most of learners stated that they had positive attitudes to the use TTS tools. They are also willing to use Text to Speech in pronunciation learning in the future.

“I can use TTS to listen to native speakers’ voice and then repeat it many times. This makes me feel very interesting and excited because I can actively practice phonics at any time...”

(Participant 5)

Participant 3 also supported that:

“I find this app very good and I will use it to practice my pronunciation regularly in the future.”

An independent sample T-test was calculated to compare mean scores of male and female participants regarding their attitudes towards using Text to Speech in learning pronunciation. The results demonstrated that no difference between two means was observed (t = -.92; p = .37). Participants’ attitudes towards using Text to Speech in learning pronunciation are the same, taking gender into consideration.
V. CONCLUSIONS AND IMPLICATIONS

The present study offers some preliminary results on the use of text-to-speech tools by EFL learners for pronunciation learning. The literature on text-to-speech tools is mainly on the technical aspect of text and speech synthesis whereas, to the researchers’ knowledge, there are no studies on how to use text-to-speech tools for pedagogy. Based on the findings of this study, it can be concluded that theoretical knowledge about English pronunciation and phonology should be accompanied by useful web-based resources that the learners can utilise for learning pronunciation.

In sum up, the results from the collected data and analysis have indicated that TTS tools had positive effects on EFL students’ pronunciation. Students had good awareness of the importance of pronunciation in their speaking skills. They also confirmed that having good English pronunciation supports effective communication. The majority of the participants showed their positive attitude towards the effect of the TTS tools. Most of them believed that TTS is a useful tool to assist their pronunciation.

This study can have some implications not only for EFL teachers but also for EFL learners. Learners should know that the use of Text-to-Speech can give them a lot of pronunciation knowledge and use them for the learning of pronunciation. Learners should know that computer alone does not improve their pronunciation. They should exert more effort and put more time on improving their pronunciation through the use of computer. Learners should be encouraged to learn pronunciation from each other by using computer technology. Although some experienced teachers may know how to teach pronunciation by computer technology, most may need training courses in how to teach pronunciation effectively and successfully so that learners can benefit more. Additionally, EFL teachers should be aware of the advantages that Text-to-Speech brings to their curriculum and the benefits that it brings to their learners’ learning process. Generally, it can be concluded that CAPT programs can increase learners’ motivation for the better learning of pronunciation and improve their pronunciation skill.

ABOUT THE AUTHOR
Tran Van Duong is currently pursuing his Master’s at Can Tho University, Viet Nam. He completed a Bachelor’s degree in English in 2015. His research interests consist of culture teaching, teaching methodologies, and teacher and learner’s motivation.

REFERENCES


