

A Comparison Between Teacher-Led and Online Text-to-Speech Dictation for Students' Vocabulary Performance

Hui-Hua Chiang¹

¹ Central Taiwan University of Science and Technology, Taiwan, ROC

Correspondence: Hui-Hua Chiang, Department of Applied Foreign Language, Central Taiwan University of Science and Technology, 666 Buzih Rd., Beitun District, Taichung, 40601, Taiwan, ROC.

Received: December 8, 2018 Accepted: January 16, 2019 Online Published: January 30, 2019

doi: 10.5539/elt.v12n3p77 URL: <https://doi.org/10.5539/elt.v12n3p77>

Abstract

Researchers have long supported the use of dictation as a test for language learners (Fountain & Nation, 2000), and dictation has been used as a test for learners of English as a foreign language (EFL). With the advantages of productive learning and reinforcing short-term memory, dictation is a commonly used technique to develop language skills, and it can be considered to be an assessment of foreign language learning (Kazazoğlu, 2013). However, the previous research has not fully explored how technology, such as text-to-speech (TTS), can be used in EFL classrooms. To address this issue, the researcher explored the use of traditional teacher-led dictation (TLD) and TTS dictation to compare the vocabulary performance of EFL learners. Forty-two college students participated in the study. The results indicated a significant difference between TTS and TLD on the participants' vocabulary performance. Additionally, there was a correlation between the scores with TTS and TLD: the students who performed better with TLD also obtained higher grades with TTS. Based on the results, future studies and pedagogical suggestions are presented.

Keywords: text-to-speech, dictation, vocabulary learning, computer-assisted language learning

1. Introduction

The use of computers in modern education is producing gratifying and positive results with regard to improving the quality of education of the emerging and current Internet generation. Society has been subject to changes and transformations that have generated new elements in the teaching-learning process. These changes have allowed computer-assisted education to motivate emerging generations in the process of knowledge construction because these systems have a more practical dynamic for their application.

Dictation has been used as a testing device in language learning for a long time (Mohammed, 2015). Scholars have positive and negative views of dictation. Those who consider dictation to be a useful tool view it as a way to diagnose aural perception as well as English mistakes among learners (Davis & Rinvoluceri, 2002; Tang, 2012). Mohammed (2015) mentions that some people might take a stereotypical approach to dictation and consider it old-fashioned, boring, and a teacher-centered method (Mohammed, 2015). The teacher-led dictation (TLD) method has been widely used in English as a foreign language (EFL) classrooms (Alkire, 2002; Davis & Rinvoluceri, 2002; Fountain & Nation, 2000; Habibi, Nemati, & Habibi, 2012; Kavaliauskienė & Darginavičienė, 2009; Kazazoğlu, 2013; Kiany & Shiramiry, 2002; Montalvan, 1990; Morris, 1983; Natalicio, 1979; Oller, 1979; Rahimi, 2008; Sawyer & Silver, 1972), while the text-to-speech (TTS) method is a new attempt to improve language learning among EFL students (Biancarosa & Griffiths, 2012; Eksi & Yesilcinar, 2016; Tang, 2012), especially in the area of reading comprehension (Wood, Moxley, Tighe, & Wagner, 2018).

Due to the progress of science and technology, numerous studies have applied TTS to the field of language learning (Eksi & Yesilcinar, 2016; Kazazoğlu, 2013; Marzban & Abdollahi, 2013; Mohammed, 2015; Sari, Sukirlan, & Suka, 2013; Tang, 2012). However, few studies have compared the performance of TTS and TLD as a dictation method. Therefore, the purpose of this study is to compare students' vocabulary dictation performance between the traditional teacher-led dictation (TLD) method and the TTS dictation method.

1.1 Dictation in Language Learning

Computer-assisted language learning (CALL) has been applied to language skills, listening, speaking, reading, and writing. Baran (2014) claimed that students benefit from CALL mostly in relation to the language process,

especially in the field of EFL. With the development of science and technology, there will always be new software inventions that can be applied to language teaching. The TTS application is one of them.

Davis and Rinvoluceri (2002) define dictation as the decoding of sounds by learners and the recoding of them in writing. The dictation method is not new in the field of language learning. Dictation can be simply defined as “a person reading some text aloud so that the listeners can write down what is being said” (Mohammed, 2015, p. 207). According to the *Oxford Learner’s Dictionaries*, dictation is “the act of speaking or reading so that somebody can write down the words.” In fact, dictation is known as one of the oldest techniques used for testing progress in EFL. It can be traced to sixteenth-century textbooks, and it has been used in language learning as well as teaching for many years (Kazazoğlu, 2013). It is usually associated with the traditional grammar method, which emphasizes the translation and memorization of the target language’s grammar rules (Stansfield, 1985).

The dictation technique is effective for both teachers and learners (Imene, 2016). According to Alkire (2002) and Davis and Rinvoluceri (2002), the dictation technique has psychological power and is easy to manipulate, prepare, manage and fit to all proficiency levels. For teachers, even novice teachers, dictation is a useful technique for small- or large class sizes and mixed-ability groups, and it is also effective in providing individual attention, motivating self-correction, reviewing learning tasks, and preparation for oral communicative exercises. For learners, dictation provides the opportunity to practice note-taking skills, develops short-term memory, and improves unconscious thinking in the language-learning process, helping students to develop literacy (Montalvan, 1990). It can help the development of all language skills in the target language (Mohammed, 2015). Jafarpur and Yamini (1993) discussed dictation as a form of dual-access processing for learners to alter as well as harmonize their perception, conception, and expression. Despite the benefits of dictation, it is no longer as popular as it was in the past because it is considered to be boring, old-fashioned, mechanical, inauthentic, and teacher-centered (Davis & Rinvoluceri, 2002; Kazazoğlu, 2013; Mohammed, 2015), especially in ESL or EFL classrooms (Alkire, 2002).

However, due to the obvious listening and speaking features of dictation, most of the related literature has focused on applying the dictation technique as a tool for almost all language skills (Mohammed, 2015). A large number of applications are associated with listening comprehension (Davis & Rinvoluceri, 2002; Habibi et al., 2012; Kiany & Shiramiry, 2002; Marzban & Abdollahi, 2013; Sari et al., 2013; Yonezaki, 2014; Yuniarti, 2017) and language proficiency (Kavaliauskienė & Darginavičienė, 2009; Mohammed, 2015; Morris, 1983; Rahimi, 2008; Valette, 1964). Furthermore, dictation is even thought to be helpful in correcting grammatical errors (Frodesen, 1991; Kazazoğlu, 2013; Montalvan, 1990) and vocabulary (Mohammed, 2015; Natalicio, 1979; Nation & Newton, 2008), focusing on meaning (Wilson, 2003), speaking (Eksi & Yesilcinar, 2016; Kavaliauskienė & Darginavičienė, 2009), extending memory span (Oller, 1971), learning disabilities (Li, 2014; Olagboyega, 2008), and reading with disabilities (Wood et al., 2018). Although the way in which dictation facilitates language acquisition is not yet completely understood, the pedagogical value of dictation is acknowledged (Alkire, 2002).

According to Sawyer and Silver (1972), dictation can be categorized into two parts, and each part contains two types. In total, four types of dictation have been widely used in language teaching as well as learning. The first part is *phonemic dictation*, and the second is *orthographic dictation*. Phonemic dictation can be further divided into *phonemic item dictation* and *phonemic text dictation*. The former indicates that language teachers model the individual sounds of the target language for students to transcribe to increase the students’ ability to produce accurate outcomes. Similar to phonemic item dictation, phonemic text dictation extends the individual sounds into a passage. The second type of dictation is *orthographic dictation*, which can also be divided into *orthographic item dictation* and *orthographic text dictation*. *Orthographic item dictation* usually focuses on the correlation between sounds and spellings, similar to the traditional spelling test, in which teachers read individual words in isolation for transcription. Similarly, *orthographic text dictation* asks language learners to transcribe a unified text or passage for comprehension or grammatical correction. Dictation in this paper is considered under the category of *phonemic item dictation*, which is associated with word recognition. Learners write what they have heard as correctly as possible.

Huang and Liao (2015) discuss TTS programs as alternatives for reading texts aloud, on portable or desktop computing equipment as well as on smart phones and tablets. These programs help users read a text without the requirement of human intervention and sometimes with the ability to download a file to preserve audio. These applications provide good simulation of a human voice and accent; thus, language learners can use these applications to understand how to pronounce words and sentences in the learning process, thereby supporting their learning.

1.2 Text-to-Speech in Language Learning

The development of TTS technology can be traced back to the 1980s (Wood et al., 2018). Initially, TTS was adopted to help students with reading disabilities. A text was associated with either a synthesized or recorded human voice (Biancarosa & Griffiths, 2012). With the continued progress of science and technology, TTS has been utilized worldwide in diverse areas for various purposes (Wood et al., 2018). TTS usually contains voice options (male, female), custom pronunciation, text highlighting, and audio download options (Peters & Bell, 2007). The following are some TTS conversions for use in language learning.

Natural reader (<https://www.naturalreaders.com/>) (Figures 1 and 2) is an interesting application that includes a free online program for TTS conversion and allows for automated reading. It has two conversion modalities, online and a desktop application, with versions for both iOS and Android systems.

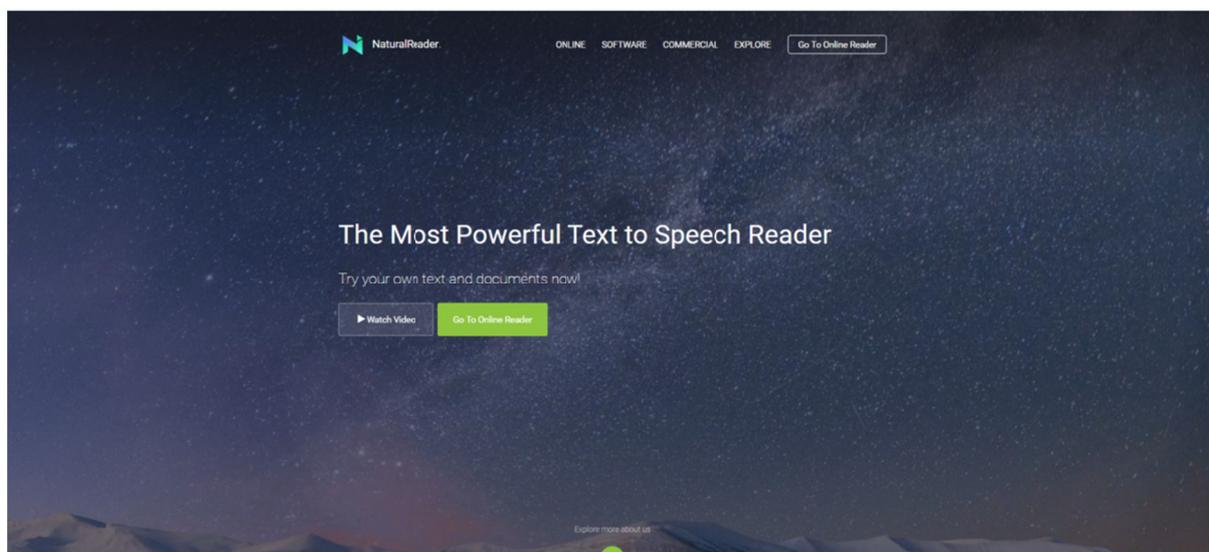


Figure 1. Natural Reader



Figure 2. Natural reader online

Both the online and the desktop versions are free, allowing for the conversion of text to reading by voice but without the ability to download the audio. For greater functionality in Natural Reader, a commercial version is also available for users who need to download the TTS audio file. For most language learners, the quality of the automated voice is acceptable in the free version.

From text-to-speech (<http://www.fromtexttospeech.com/>) (Figure 3) is a simple and user-friendly TTS online tool that charges for the download of the TTS audio file. All the user needs to do is post the text (up to 50,000 characters) in the tool box, select the language (available in US English, British English, French, Spanish, German, Italian, Portuguese, and Russian), select the voice (options of two female voices and three male voices), and select a speed (slow, medium, fast and very fast). It converts any English text into an MP3 audio file that can be played on a PC or iPod.

Convert text to speech

Simply copy and paste your text in the box below :



50000 characters remaining

Select language:

Select voice:

Select speed:

[Create Audio File](#)

Figure 3. From text-to-speech

TTSReader (<https://ttsreader.com/>) (Figure 4) is probably the most comprehensive text-to-speech reading application online. It is a web-based application that reads aloud any text, from typed text to files (including txt, pdf, e-publication, and websites), with a high-quality natural sound, including male and female voices and multiple accents. Its premium version is available for commercial use for a charge.

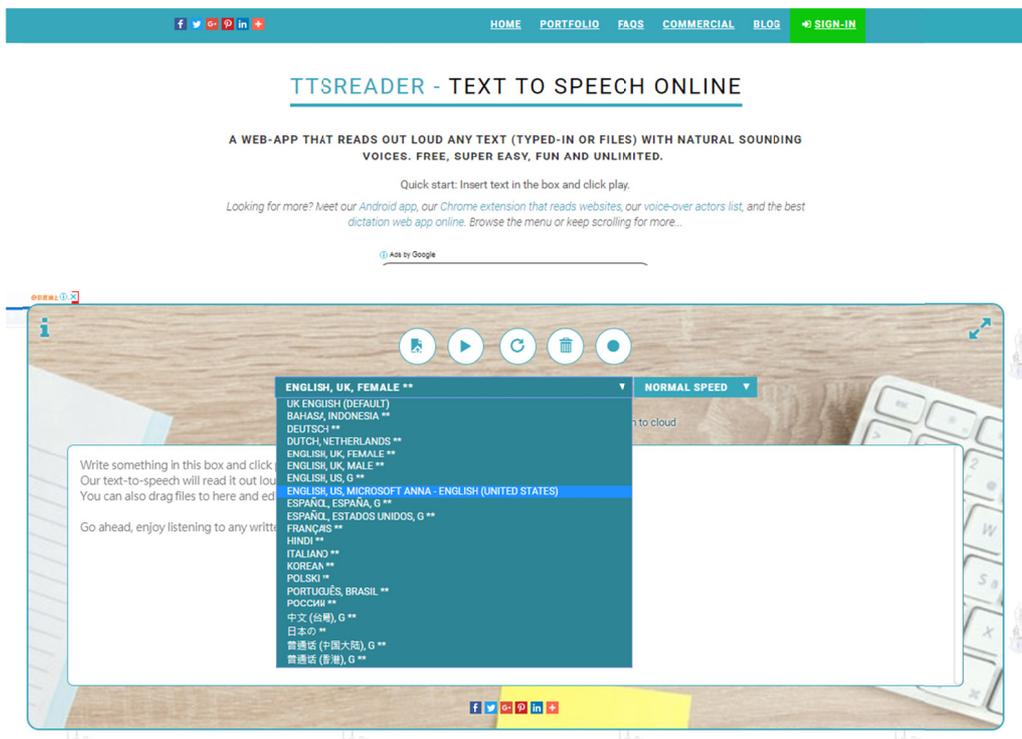


Figure 4. TTSReader

VOCALWARE (<https://www.vocalware.com/>) (Figure 5) is also an online TTS that allows users to speech-enable any online application through an online API (application programming interface). VOCALWARE supports audio effects functions for users to choose (None, Duration, Pitch, Bullhorn, Echo, Reverb, Flanger, Phase, and Whisper). Like other applications, VOCALWARE also provides more than 20 languages such as English, Chinese, French, and Korean and more than 100 voices and accents (US, UK, Australian, and Indian) from which to choose. A free demo version allows a maximum of 600 characters.

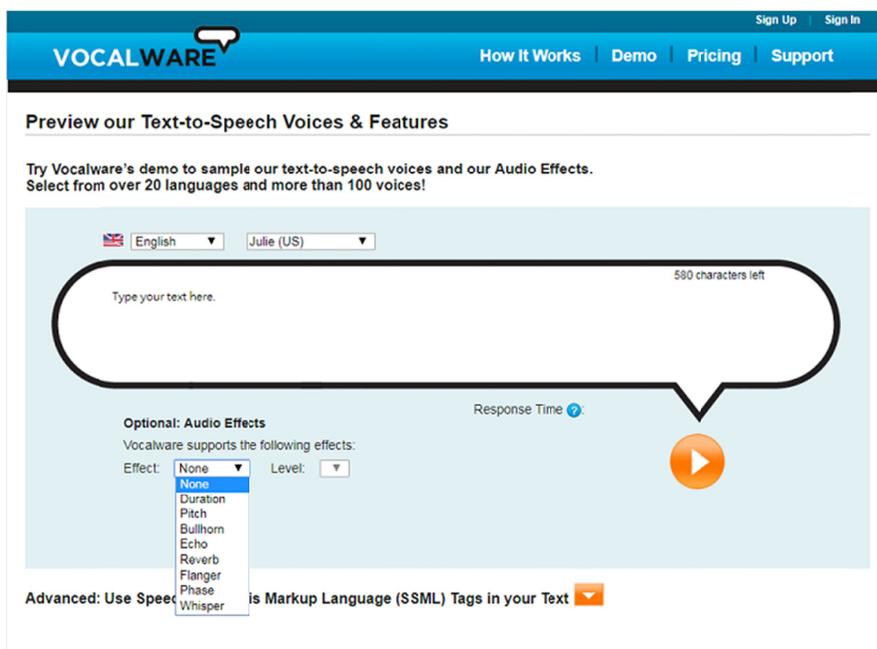


Figure 5. Vocalware

1.3 CALL-Based Vocabulary Learning

Vocabulary has always played an important role in English learning (Nation, 2015). Teaching in the 21st century has been refined by the use of supplementary technical means, and computers have played a prominent role due to the advantages they offer for the explanation of concepts as well as their applications. Effective methods for teaching have long been explored, especially as technology has progressed. Diverse applications have been associated with the teaching of various contents. Students are currently involved in changes and transformations that have generated new elements in the teaching and learning process. This has allowed computer-assisted education to motivate new generations in their process of knowledge construction and has given these systems a more practical dynamic for their application.

Dictation is a beneficial method for language learning and has been used in the classroom for centuries (Mohammed, 2015). Numerous studies have been conducted in the field of vocabulary learning as well as dictation; however, little research has examined the application of online TTS to EFL learning, especially in vocabulary. Therefore, this study aims to compare EFL college students' vocabulary performance between TLD and online TTS dictation. The research questions are listed below.

1.4 Research Questions

Few of the previous studies have focused on the use of TTS technology with dictation techniques to enhance students' vocabulary dictation performance in Taiwan.

- 1). Is there a significant difference between TLD and TTS in the performance of word recognition?
- 2). Is there a correlation between the scores of TLD and TTS?
- 3). What are students' perceptions of the use of TTS in vocabulary dictation tests compared with the TLD?

1.5 Research Null Hypotheses

To compare students' vocabulary dictation test performance when using TTS and TLD, the following null hypotheses are formulated for research question one.

H0: There is no significant difference between TTS and TLD in achievement on vocabulary dictation tests.

H1: There is a significant difference between TTS and TLD in achievement on vocabulary dictation tests.

2. Research Design and Methods

2.1 Participants

The participants for this study were 42 low- to intermediate-level students (Table 1) who majored in English in the applied foreign language (AFL) department. The students were between the ages of eighteen and nineteen years old and had studied English for at least five years before entering the AFL department. During the research period, they were taking English at least eight hours per week. A weekly dictation test was employed in the English vocabulary and reading class to develop the students' word recognition and spelling.

Table 1. Descriptive analysis of Oxford Reading Level Test (RLT)

	Frequency	Percent	Valid Percent	Cumulative Percent
-	5	11.9	11.9	11.9
A1	7	16.7	16.7	28.6
A2	18	42.9	42.9	71.4
B1	11	26.2	26.2	97.6
B2	1	2.4	2.4	100.0
Total	42	100.0	100.0	

2.2 Instruments

The *OXFORD Online English Level Test-Reading Level Test* (RLT), designed by Oxford Online English, was used in this study to better understand the participants' overall English proficiency levels. RLT contains 20 questions related to an article that is available at *OXFORD Online English Level Test* (<https://www.oxfordonlineenglish.com/english-level-test>). The students were asked to first read the article and

then finish the 20 reading comprehension questions. After they finished the 20 questions, the online system compared their grades with the Common European Framework of Reference for Languages (CEFR), which categorizes English proficiency into six levels: A1, A2, B1, B2, C1 and C2 (see Appendix A).

Text-to-Speech technology (TTS) is often used to translate written text into spoken text. The selected TTS, which is available at <https://www.naturalreaders.com/online/> (see Appendix B), provides three editions (*Web Free*, *Web Premium*, and *Commercial*) for users to choose depending on their service coverage needs. The *Web Free* edition was the one chosen by the researcher for use throughout the study. It includes 20 minutes per day of premium voices, unlimited usage of free voices, and supports PDF, Docx, RTF and TXT document upload. In the *Web Free* edition, users can upload the selected written text into the system and choose the preferred English accent (US or UK), voice (male or female), and speed (from -4 to 9, indicating extremely slow to extremely fast).

The Text-to-Speech Perception Questionnaire (TPQ), designed by the researcher, was used as a testing material in this study and graded to assist in understanding how the participants think about the implementation of TTS on a vocabulary quiz. The TPQ includes seventeen 5-point Likert-scale questions and two multiple-choice questions related to solutions for the pronunciation of unknown words. Forty-two valid questionnaires were obtained after removing invalid questionnaires.

Vocabulary Dictation Quizzes (VDQ) was used in the classroom in this study to help the researcher gather the participants' perceptions of two aspects of dictation methods: traditional TLD and the TTS method. According to Kazazoğlu (2013), the selection of a dictation text should be appropriate to the level of the learners. Dictation reinforces basic sentence structure as well as vocabulary at the intermediate level. The overall speed, pauses between material, and number of times the text is presented may influence the difficulty of the text (Kazazoğlu, 2013). Oller (1979) stated that the researcher should choose the dialect and the pronunciation with which the learners are most familiar. Hence, in this study, a US accent, a female voice, a speed of minus 2, and repeating each word twice were selected to provide the participants with a friendly atmosphere for TTS.

The vocabulary on VDQ was chosen from the *TOEIC: Vocabulary Express 3000*, from which the participants were required to acquire a certain amount of words every week during the semester.

2.3 Procedure

The participants were asked to take the online English Level Test-RLT at the beginning of the class. Forty-six students finished the RLT in a class of fifty-three students. The study began in the middle of September 2017 and was completed in January 2018. Every week, the students were assigned two pages of words in the *TOEIC: Vocabulary Express 3000*, which included 20-24 target words along with phonetic spelling and example sentences. The students took the VDQ 12 times, in which they were asked to write the words they heard. The first six VDQs were conducted through the traditional approach of TLD, and the rest of the VDQs were conducted through a TTS method in which the students listened to the words twice and an example sentence. For instance, *RESERVATION, RESERVATION, I have a RESERVATION for four under the name of Chang*. Ten words were selected from 20-24 words on every test. There was a 20-second pause between each word for the students to write the vocabulary word in English along with its Chinese meaning. Finally, ten words were played nonstop at minus one speed from beginning to end. The standard procedure for both TTS and TLD vocabulary dictation tests is shown below, taking the word "promotion" as an example (Figure 6). At the end of the semester, the text-to-speech perception questionnaire (TPQ) was issued to obtain the students' opinions on taking VDQ through traditional TLD and text-to-speech technology (TTS).

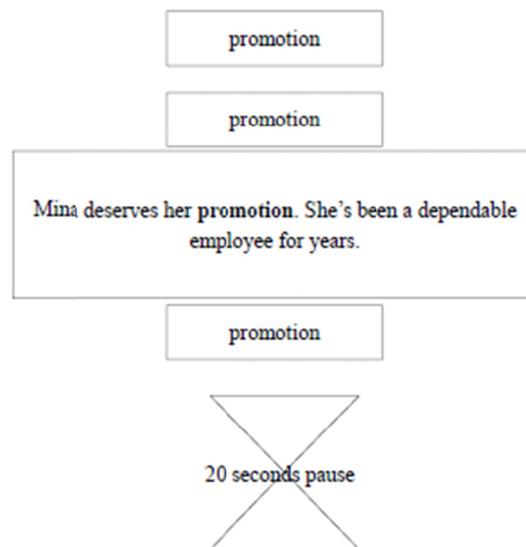


Figure 6. Standard Procedure for Dictation Tests

In summary, the study data were derived from three aspects. First, forty-six students took the OXFORD Online English Level Test-RLT to obtain the students' overall reading level. Second, the results of the VDQ conducted during one semester were used to obtain information on the students' performance on twelve VDQs. Finally, the TTS perception questionnaire (TPQ) with 20 questions was used to obtain the students' perception of the use of TTS for selected vocabulary dictation performance. The quantitative analyses were performed using statistical software and included descriptive statistics to summarize the students' perspectives on TPQ. The three open-ended questions served as qualitative data and categories for deeper interpretation. VDQ grades were used to allow the researcher to compare the students' performances acquired from the traditional method of the teacher read-aloud method and the TTS method.

3. Results

This study was designed to compare the effects of TLD and TTS technology on vocabulary dictation tests by EFL college learners. A correlation analysis was conducted on a reading class in Taiwan during 12 weeks of the academic year of 2018. All participants were sophomore students taking the same English classes in the department.

3.1 Descriptive Statistics

As the descriptive statistics show in Table 1, a total of 42 participants were illustrated. A descriptive analysis of the Oxford RLT for 42 students is shown in Table 1. For the RLT, most students were at low-intermediate to intermediate reading levels. An estimated 43% were at the A2 level and 26% were at the B1 level, which included nearly 70% of the participants. Only 17% of the participants' RLTs were below the A2 level, and their RLTs were rarely at the B2 level (2%).

To validate the comparison of the TLD and TTS scores, several statistical analyses were used. A null hypothesis was proposed for the first research question, which was that there is no significant difference between TTS and TLD in achievement on vocabulary dictation tests. The scores for the different methods of the dictation tests were examined to ensure that they were normally distributed, as shown in Table 2. Two well-known tests of normality are the Kolmogorov-Smirnov test and the Shapiro-Wilk test. As Ben-Zvi, Makar, & Garfield suggested, the Shapiro-Wilk test is suggested for small sample sizes (<50 samples) as well as large sizes. Based on the flexibility of the Shapiro-Wilk test, this study used the Shapiro-Wilk test as the numerical means of assessing normality. As shown in Table 2, the Sig. value of the Shapiro-Wilk test is greater than .05, which means the data are normally distributed.

Table 2. Tests of normality

OLT	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
		.340	5	.060	.865	5	.247
TLD Scores	A1	.250	7	.200*	.904	7	.355
	A2	.143	18	.200*	.962	18	.640
	B1	.136	11	.200*	.956	11	.725
TTS Scores		.315	5	.116	.847	5	.186
	A1	.192	7	.200*	.890	7	.277
	A2	.130	18	.200*	.952	18	.462
	B1	.159	11	.200*	.931	11	.425

a. Lilliefors Significance Correction

*.This is a lower bound of the true significance.

3.2 Paired Sample t-Test

There were 42 valid data samples for analysis. A paired sample *t*-test was used to compute and analyze the data. As presented in Table 3, the mean score for TLD was 57.04, and the mean score for TTS was 65.78. The standard error means were 3.74 and 3.19, respectively. The participants scored higher grades on the TTS vocabulary dictation tests than with traditional TLD methods.

Table 3. Descriptive statistics for the Dictation Test (Paired)

		Mean	N	SD	Std. Error Mean
Pair 1	TLD Mean	57.0476	42	24.25056	3.74194
	TTS Mean	65.7857	42	20.70644	3.19507

A paired sample *t*-test was conducted to test the research hypothesis that there is no significant difference between the TLD dictation test and the TTS dictation test. As shown in Tables 3 and 4, the *t* value of the TLD and TTS scores was 2.329 with a standard deviation of 24.31. The mean scores for the TLD and TTS were 57.04 and 65.78, respectively, which shows that the students obtained better scores with TTS than with TLD. Furthermore, the *P* value = .025 < .05 indicates that the null hypothesis is rejected, and H1 is accepted. Therefore, there is a significant difference between the mean scores with TLD and TTS.

Table 4. t-Test for the TLD and TTS

		Paired Samples Test							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference	Lower	Upper	t	Sig. (2-tailed)
Pair 1	TLD Mean - TTS Mean	-8.7381024	3.1348	3.75165	-16.31471	-1.16148	-2.32941	0.025	

3.3 Pearson Correlation Analysis

For research question 2, a Pearson correlation analysis was conducted to determine whether there is a correlation between the scores with TLD and TTS. As shown in Tables 5 and 6, the participants' TLD scores had moderate correlations with the TTS scores ($r=.424$, $p<0.01$). This finding indicates that the scores obtained by the participants with TLD were significantly correlated with their scores with TTS. In other words, the participants

who obtained better scores on TLD tests also received better scores on TTS.

Table 5. Paired sample correlation for the TLD and TTS

		N	Correlation	Sig.
Pair 1	TLD Mean & TTS Mean	42	.424	.005

Table 6. Pearson correlations among TLD and TTS Scores (N=42)

		TLD	TTS
TLD	Pearson r	1	
TTS	Pearson r	.424**	1

** Correlation is significant at the level .01 level (2-tailed).

The findings of the questionnaire are as follows. The questionnaire aimed to investigate the opinions of the participants in applying TTS technology to vocabulary dictation tests in an English class. The students were asked to rate 17 Likert-type questions about their opinions on using TTS technology in vocabulary dictation tests and three other questions about the learning of vocabulary pronunciation. The results are illustrated below in Table 7.

Table 7. Students' opinions about TTS technology and TLD on dictation and pronunciation

Survey Items	Mean	Std. dev
16. The teacher is the role model for pronunciation in the classroom.	4.24	.75
17. How often do you search for pronunciation when you can't pronounce a word?	4.21	.78
4. I think TTS technology is sometimes very mechanical.	4.14	.68
5. I think TTS technology sometimes doesn't sound emotional.	4.05	.69
2. I think it is interesting to use TTS technology.	3.95	.79
1. This is the first time I have heard about TTS technology.	3.59	1.18
15. I think using Internet resources is the best way to improve pronunciation.	3.57	.83
8. I think the speed of TTS technology is too fast.	3.50	.83
7. I think TTS technology sometimes sounds wrong.	3.33	.68
11. I think it's useful to use TTS technology in dictation tests.	3.21	.84
6. I think dictation with TTS technology will motivate me to listen to the CD in the textbook to prepare for the exam.	3.10	.93
13. In the future, I would like to use TTS technology to practice listening.	3.10	.79
10. I think TTS technology can help me learn English pronunciation.	3.05	.96
14. In the future, I would like to use TTS technology to practice pronunciation.	3.05	.79
12. I think using TTS technology in dictation tests will distract me.	3.00	.96
3. I think using TTS technology to test vocabulary dictation is more clear than the teacher-read-aloud method.	2.64	.93
9. Compared with the teacher-read-aloud method, I think TTS technology is more understandable.	2.48	.91

The students rated the survey items (questions 1 to 16) from *1-totally disagree* to *5-totally agree* and from *1-never* to *5-always* (question 17). The mean scores for each item show varying degrees of agreement. The items asked about the students' opinions of traditional TLD and TTS technology in vocabulary dictation tests and their

opinions about pronunciation. The statement about the teacher as the pronunciation role model in the classroom received the highest score. Although TTS technology is sometimes very technical ($M=4.14$) and does not sound emotional ($M=4.05$), the students agree that TTS technology is interesting ($M=3.95$). When comparing the use of TTS technology to traditional TLD in vocabulary dictation tests, the students did not think that TTS technology was more clear than TLD ($M=2.64$) and did not agree that TTS technology was more understandable ($M=2.48$). When the students do not know how to pronounce a word, they seem to always check the pronunciation ($M=4.21$).

The last two items on the questionnaire (Q18 & Q19) were additional multiple-choice questions. Question 18 attempted to determine what the students do when they encounter a word that they do not know how to pronounce. The case summary is shown in Table 8, and the frequency count for question 18 is reported in Table 9.

Table 8. Case summary for Q18

	Cases		Missing		Total	
	Valid					
	N	Percent	N	Percent	N	Percent
Q18 ^a	42	100.0%	0	0.0%	42	100.0%

a. Dichotomy group tabulated at value 1.

Table 9. How to check pronunciation of unknown words

18. What do you do when you encounter English words that you cannot pronounce? (Multiple choice)	Responses		Percent of Cases
	N	Percent	
Check pronunciation in a paper-based dictionary	2	3.8%	4.8%
Check pronunciation in a computer-based online dictionary	13	24.5%	31.0%
Q18 ^a Check pronunciation in a mobile-based online dictionary	37	69.8%	88.1%
Do not check the pronunciation, only guess	1	1.9%	2.4%
Do not check the pronunciation	0	0%	0.0%
Total	53	100.0%	126.2%

a. Dichotomy group tabulated at value 1.

As shown in Table 9, when the students encountered unknown words, they tended to check the pronunciation in a mobile-based online dictionary (69.8%) or in a computer-based online dictionary (24.5%). Less than five percent (4.8%) of the students checked the unknown words' pronunciation in a paper-based dictionary, and only a few of them guessed the pronunciation (1.9%).

Question 19 collected information on the methods by which students sought access to improve their vocabulary pronunciation. The case summary is in Table 10, and the frequency count for question 19 is reported in Table 11.

Table 10. Case summary for Q19

	Cases		Missing		Total	
	Valid					
	N	Percent	N	Percent	N	Percent
Q19 ^a	41	97.6%	1	2.4%	42	100.0%

a. Dichotomy group tabulated at value 1.

Table 11. Access to improve vocabulary pronunciations

19. What do you do to improve your vocabulary pronunciation? (Multiple choice)		Responses		Percent of Cases
		N	Percent	
	Listen to music	35	16.9%	85.4%
	Sing songs	18	8.7%	43.9%
	Record voice	1	0.5%	2.4%
	Rely on phonetic alphabet	19	9.2%	46.3%
	See movies	26	12.6%	63.4%
	Listen to podcasts	1	0.5%	2.4%
Q19 ^a	Check on mobile-based online dictionary	26	12.6%	63.4%
	Read aloud with textbooks	8	3.9%	19.5%
	Listen to websites	10	4.8%	24.4%
	Check on computer-based online dictionary	14	6.8%	34.1%
	Read aloud with teachers in the classroom	25	12.1%	61.0%
	Watch YouTube	19	9.2%	46.3%
	Other	5	2.4%	12.2%
Total		207	100.0%	504.9%

a. Dichotomy group tabulated at value 1.

As shown in Table 11, the top method for the students to improve their vocabulary pronunciation is listening to music (16.9%), followed by seeing movies (12.6%), checking a mobile-based online dictionary (12.6%), and reading aloud with teachers in the classroom (12.1%). Nearly 10% of the students rely on the phonetic alphabet (9.2%) and watching YouTube (9.2%) to improve their English vocabulary pronunciation.

4. Discussion and Conclusion

4.1 Discussion

One initial objectives of this study was to determine whether there is a significant difference between TTS and TLD performance for vocabulary dictation. It was hypothesized that there was no difference between TTS and TLD in participants' achievement on vocabulary dictation tests. In the review of the literature, little-to-no data were found to connect TTS to foreign language learning, especially in the field of vocabulary learning. The results showed that there was a significant difference between TTS and TLD with regard to the participants' dictation performance. Surprisingly, on the one hand, the participants seemed to strongly agree with the statement "the teacher is the role model for pronunciation in the classroom". On the other hand, they performed better on the TTS dictation tests. This discrepancy could be attributed to the fact that TTS native speakers of English use pronunciations closer to those of standardized English teachers as non-native speakers of English. Another possible explanation might be that the participants listened to the .mp3 tracks attached to the selected book to become familiar with the TTS pronunciations they heard during the dictation tests.

It is interesting to note that the results of this study differ slightly from those of Kazazoğlu's (2013) study, which investigated word recognition between TLD and tape-recorded dictation among 76 intermediate high school students in Turkey. Kazazoğlu's findings indicate that students made fewer word errors in TLD than in tape-recorded dictation. In other words, students seem to perform better in TLD. Kazazoğlu (2013) concluded that speed is the key to the connection with short-term memory, and it is associated with aural competence in language learning. Although Kazazoğlu's results were completely different from the results of this study, he stated that "speed" might be the key point. As technology progresses, the speaker's speed can be tuned in the TTS, which is more flexible than tape-recorded dictation.

The results also showed a correlation between the scores with TTS and TLD. It might be inferred that learners who perform better with TTS would also perform better with TLD. There are two possible explanations for this result. One possible reason is that students who perform well with both TLD and TTS may have better English abilities than others. The other reason is that students who do well with both TTS and TLD may have a strong

motivation to learn English; thus, they spend more time preparing for tests.

With regard to the students' opinions about using TTS technology, as mentioned above, the students agreed that teachers are the role model for pronunciation in the classroom. Interestingly, although TTS technology is more native-like than English teachers for Chinese as a native language in English pronunciation, the students seemed to rely on the classroom English teacher as their model for pronunciation. This may be because the TTS technology is sometimes still very mechanical ($M=4.14$) and does not sound emotional ($M=4.05$).

4.2 Future Studies

In future studies, measurements of learners' perceptions of TTS voice selections and different speeds should be assessed. User experience might also affect the effectiveness of TTS. Additionally, the application of TTS in diverse types of dictation, such as phonemic text dictation or orthographic item dictation, as well as orthographic text dictation in various fields of language learning is worth studying. The length of the text is also worth examining. Finally, including the variable of individual differences in the application of TTS in EFL learning might be an interesting angle worthy of further study.

4.3 Limitations

There are a few limitations that are worth noting. First, based on Sawyer and Silver's (1972) dictation categories, the research applied phonemic item dictation in which the teacher or TTS technology read aloud the individual target words for learners to write the words they heard. Although example sentences with target words were used, the dictation method was limited to an "item" only. Second, TTS technology has the function of speed control, from very slow to very fast (-4 to 9). In this study, a speed of minus 2 (speed -2) was used because the slower pace was expected to help learners listen more clearly and to reduce students' rejection of TTS technology.

4.4 Conclusion

The study aimed to examine the use of TTS technology and TLD in English vocabulary dictation performance and to reveal students' opinions of the TTS technology compared to TLD. The results showed that there is a significant difference between TTS and TLD in students' achievement on vocabulary dictation tests. Additionally, there is a correlation between the scores with TTS and TLD. Overall, the participants seemed disinclined to agree with TTS technology for learning English vocabulary pronunciation. In other words, they seemed not to appreciate the TTS technology in the language classroom. It seems likely that these results are due to the tradition of teacher-centered stereotypes in foreign language-learning classrooms, in which students rely primarily on language teachers' modeling of pronunciation. It is recommended that instructors implement online TTS appropriately in the language-learning classroom to promote self-learning ability, especially if the students intend to learn to listen and speak in a new language. Finally, although TTS technology is native-like, there is always room for improvement by making it more like a real person's pronunciation. Educators and software designers could consider language-learning objectives with regard to TTS dynamics to make TTS an enjoyable learning environment.

References

- Alkire, S. (2002). Dictation as a language learning device. *The Internet TESL Journal*, 8(3), 1.
- Baran, E. (2014). A review of research on mobile learning in teacher education. *Journal of Educational Technology & Society*, 17(4), 17-32.
- Biancarosa, G., & Griffiths, G. G. (2012). Technology tools to support reading in the digital age. *Future Child*, 22(2), 139-160. <https://doi.org/10.1353/foc.2012.0014>
- Davis, P., & Rinvolucri, M. (2002). *Dictation: New methods, new possibilities*. Cambridge, UK: Cambridge University Press.
- Eksi, G., & Yesilcinar, S. (2016). An investigation of the effectiveness of online text-to-speech tools in improving EFL teacher trainees' pronunciation. *English Language Teaching*, 9(6), 205-214. <https://doi.org/10.5539/elt.v9n2p205>
- Frodesen, J. (1991). Grammar in writing. In M. Celce-Murcia (Ed.), *Teaching english as a second or Foreign language* (pp. 268). Boston, MA: Heinle & Heinle Publishers.
- Fountain, R. L., & Nation, I. S. P. (2000). A vocabulary-based graded dictation test. *RELC Journal*, 31(2), 29-44. <https://doi.org/10.1177/003368820003100202>
- Habibi, P., Nemati, A., & Habibi, S. (2012). The role of listening comprehension in dictation. *Indian Journal of Science and Technology*, 5(2), 3208-3210.

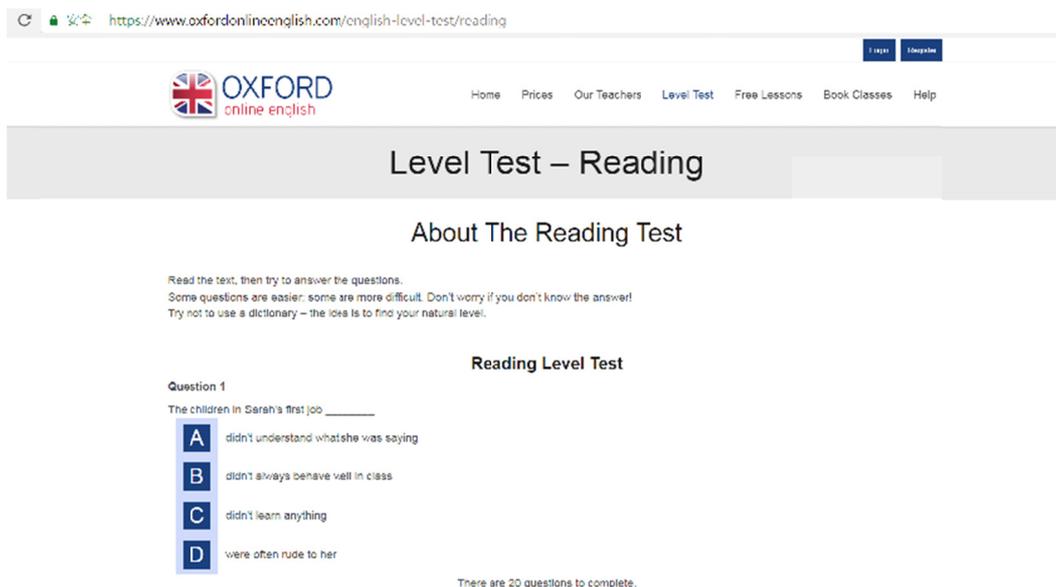
- Huang, Y. C., & Liao, L. C. (2015). A study of text-to-speech (TTS) in children's english learning. *Teaching English with Technology*, 15(1), 14-30.
- Imene, H. I. (2016). *An exploratory study on the impact of dictation as a technique in developing learners' listening* (Master's thesis). Mohamed Kheider University of Biskra, Algeria.
- Jafarpur, A., & Yamini, M. (1993). Does practice with dictation improve language skills? *System*, 21(3), 359-369. [https://doi.org/10.1016/0346-251X\(93\)90026-D](https://doi.org/10.1016/0346-251X(93)90026-D)
- Kavaliauskienė, G., & Darginavičienė, I. (2009). Dictation in the ESP classroom: A tool to improve language proficiency. *English for Specific Purposes World*, 8(23), 1-10.
- Kazazoğlu, S. (2013). Dictation as a language learning tool. *Procedia - Social and Behavioral Sciences*, 70(4), 1338-1346. <https://doi.org/10.1016/j.sbspro.2013.01.195>
- Kiany, G. R., & Shiramiry, E. (2002). The effect of frequent dictation on the listening comprehension ability of elementary EFL learners. *TESL Canada Journal*, 20(1), 57-63. <https://doi.org/10.18806/tesl.v20i1.938>
- Li, H. (2014). The effects of read-aloud accommodations for students with and without disabilities: A meta-analysis. *Educational Measurement: Issues and Practice*, 33(3), 3-16. <https://doi.org/10.1111/emip.12027>
- Marzban, A., & Abdollahi, M. (2013). The effect of partial dictation on the listening comprehension ability of Iranian intermediate EFL learners. *International Research Journal of Applied and Basic Sciences*, 5(2), 238-244.
- Mohammed, B. S. (2015). Using dictation in teaching college students. *Global Journal on Technology*, 8(3), 20-215.
- Montalvan, R. (1990). Dictation updated: Guidelines for teacher-training workshops. Retrieved from <http://dosfan.lib.uic.edu/usia/E-USIA/education/engteaching/intl/pubs/dictatn.htm>
- Morris, S. (1983). Dictation—a technique in need of reappraisal. *ELT Journal*, 37(2), 121-126. <https://doi.org/10.1093/elt/37.2.121>
- Natalicio, D. S. (1979). Repetition and dictation as language testing techniques. *The Modern Language Journal*, 63(4), 165-176. <https://doi.org/10.1111/j.1540-4781.1979.tb02439.x>
- Nation, I. S., & Newton, J. (2008). *Teaching ESL/EFL listening and speaking*. New York, NY: Routledge.
- Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136-145.
- Olagboyega, K. W. (2008). The effects of dyslexia on language acquisition and development. *Akita University Bulletin Paper*, 29(7), 23-27.
- Oller, J. W. (1971). Dictation as a device for testing Foreign-language proficiency. *ELT Journal*, 25(3), 254-259. <https://doi.org/10.1093/elt/XXV.3.254>
- Oller, J. W. (1979). *Language tests at school*. London, UK: Longman.
- Peters, T., & Bell, L. (2007). Choosing and using text-to-speech software. *Computers in Libraries*, 27(2), 26-29.
- Rahimi, M. (2008). Using dictation to improve language proficiency. *Asian EFL Journal*, 10(1), 33-47.
- Sari, F. M., Sukirlan, M., & Suka, R. G. (2013). Improving students' listening ability through dictation technique at the first year students. *U-JET*, 2(3), 1-13.
- Sawyer, J. O., & Silver, S. K. (1972). Dictation in language learning. In H. B. Allen, & R. N. Campbell (Eds.), *Teaching english as a second language* (pp. 223-229). San Francisco: McGraw-Hill.
- Stansfield, C. W. (1985). A history of dictation in Foreign language teaching and testing. *The Modern Language Journal*, 69(2), 121-128. <https://doi.org/10.1111/j.1540-4781.1985.tb01926.x>
- Tang, Q. (2012). The effectiveness of dictation method in college english vocabulary teaching. *Theory and Practice in Language Studies*, 2(7), 1472-1476. <https://doi.org/10.4304/tpls.2.7.1472-1476>
- Valette, R. M. (1964). The use of the dictée in the French language classroom. *The Modern Language Journal*, 48(7), 431-434. <https://doi.org/doi:10.1111/j.1540-4781.1964.tb04525.x>
- Wilson, M. (2003). Discovery listening-improving perceptual processing. *ELT Journal*, 57(4), 335-343. <https://doi.org/10.1093/elt/57.4.335>

- Wood, S. G., Moxley, J. H., Tighe, E. L., & Wagner, R. K. (2018). Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *Journal of Learning Disabilities, 51*(1), 73-84. <https://doi.org/10.1177/0022219416688170>
- Yonezaki, H. (2014). Effectiveness of dictation in improving english listening ability of Japanese high school students. *Journal of Research Reports of Nagaoka National College of Technology, 50*(6), 21-30.
- Yuniarti, F. (2017). Using dictation as a testing device in teaching listening. *Inovish Journal, 2*(2), 12-23.

Appendix A

OXFORD Online English Level Test-Reading Level Test

<https://www.oxfordonlineenglish.com/english-level-test>



Read the text, then try to answer the questions.
Some questions are easier, some are more difficult. Don't worry if you don't know the answer!
Try not to use a dictionary – the idea is to find your natural level.

Reading Level Test

Question 1
The children in Sarah's first job _____

A didn't understand what she was saying
B didn't always behave well in class
C didn't learn anything
D were often rude to her

There are 20 questions to complete.

Sarah is 36 years old, and she lives in Canada. She has two young daughters. She works two days a week as a teacher. Her husband's name is Nathan, and he's a sales manager. Nathan's job is very busy, so he often comes home late. On weekends, they often go driving or walking in the countryside.

Nathan was born in Canada, but Sarah wasn't. She was born in Argentina, and she moved to Canada when she was 26. When she was growing up, she was really interested in English. At first, she thought it was difficult, but when she finished school, she could already speak quite fluently and understand almost everything she heard or read. She spent a lot of time listening to songs and watching TV shows and films in English.

After she graduated from university, she decided to train as an English teacher. The certificate she needed was quite expensive and competition for places was intense, but she was determined to do it—she simply couldn't imagine doing anything else. She finished the course with distinction, which was the highest grade possible. Soon, she found work as a teaching assistant in a local primary school. She enjoyed the work, although it was often challenging—the children were not always well-disciplined, and she didn't think that the classroom teacher had enough understanding of teaching methods.

When she first went to Canada, she never would have imagined that she would end up staying there. It was supposed to be a short-term placement in a high school. She thought that she would be able to see a different part of the world and gain some useful experience, which could help her to find a better teaching position when she came back to Argentina. At first, she found living overseas much more difficult than she had expected. She felt homesick, and she had problems getting used to everything that was different in Canada—the interpersonal culture, the climate, the food... For the first three months she was there, she spent most of her free time in her room, dreaming of going back to Argentina and seeing her family again.

Over time, she adjusted to life in Canada and even started to enjoy herself a bit more. One day, she met Nathan at a party. She liked his sense of humour and how kind he was, but she was reluctant to get involved, knowing that

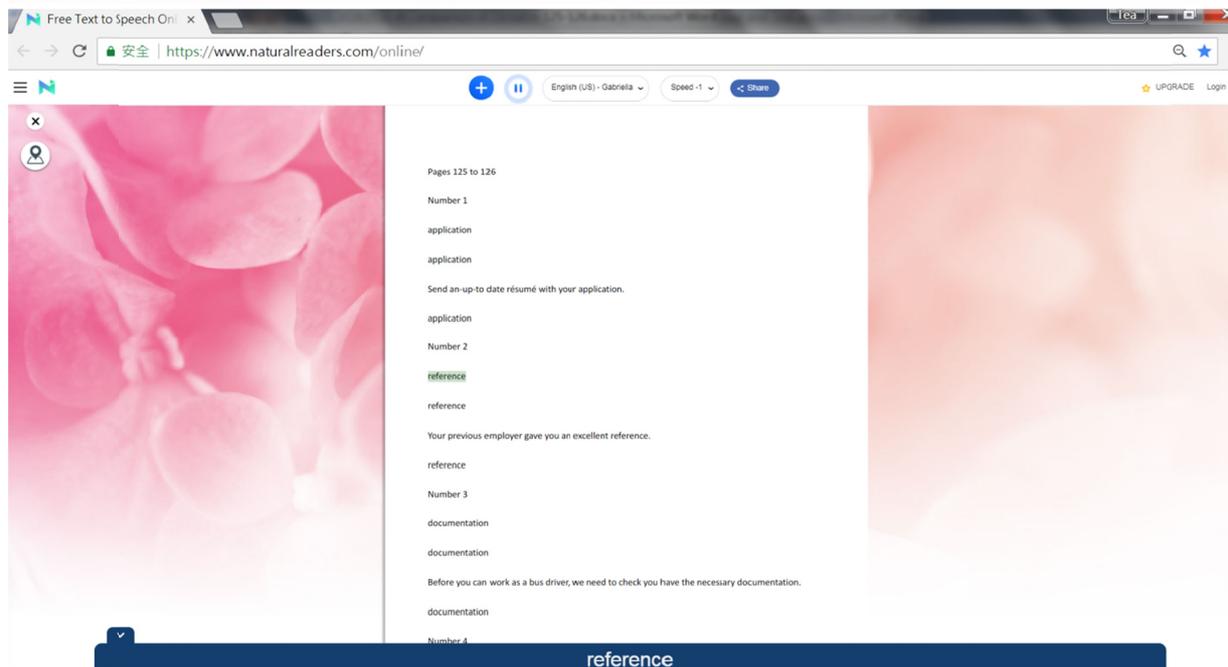
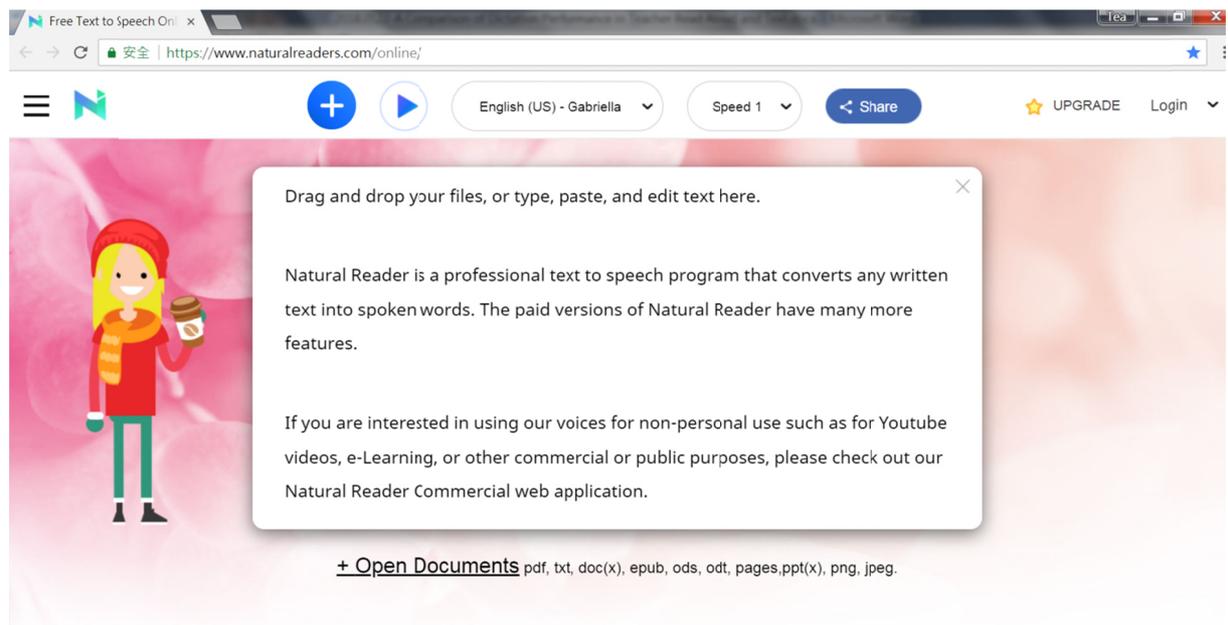
she was planning to leave in the near future. When her placement finished, he convinced her to apply for a permanent job in another school. She told herself that she would give it one more year and see how things went.

Now, Sarah is settled, although she still misses Argentina. She tries to make it back at least yearly, and she is bringing up her daughters to be bilingual, so that they can talk to their Argentinian relatives in Spanish. When she thinks back to her first few months in Canada, she can scarcely recognize herself. In some ways, she wishes she weren't so far away from her family, but at the same time, she feels that she's learned many things she never would have experienced had she stayed in Argentina. She wants to give her daughters the chance to travel and experience life in other countries as soon as she can, although of course she hopes they don't move too far away!

Appendix B

Online Text-to-Speech

<https://www.naturalreaders.com/online/>



Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).