Online Translators: Can They Help English Learners improve their pronunciation?

Yue He

A Thesis

In

The Department

of

Education

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Applied Linguistics) at

Concordia University

Montreal, Quebec, Canada

September 2021

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Yue He

Entitled: Online Translators: Can They Help English Learners improve their pronunciation?

and submitted in partial fulfillment of the requirements for the degree of

MA Applied Linguistics

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Examiner
D. Liakin

_____ Thesis Supervisor
W. Cardoso

Approved by _____
W. Cardoso Chair of Department or Graduate Program Director

_____
Pascale Sicotte, Dean

ABSTRACT

Online Translators: Can They Help English Learners improve their pronunciation?

Yue He

The literature reports a number of limitations that affect the teaching and learning of foreign languages, including a lack of teacher preparedness (Hu, 2005) and insufficient time for practice (Life, 2011). To mitigate these challenges, we proposed a self-directed learning (SDL) environment assisted by a combination of text-to-speech synthesis (TTS) and automatic speech recognition (ASR) technologies, as found in *Microsoft Translator* (MT), to examine whether this translation tool and its built-in speech features can promote the acquisition in pronunciation of English regular past tense -*ed* in a self-directed manner.

This study followed a pretest/posttest research design in which participants received autonomous but teacher-assisted TTS- and ASR-based treatment to learn about the pronunciation of English past -ed allomorphy: this suffix can be pronounced as play/d/, visit/ɪd/ and walk/t/, depending on the preceding phonological environment. We compared 29 participants' performance in past -*ed* allomorphy by assessing their phonological development in terms of phonological awareness, phonemic discrimination, and oral production, as per Celce-Murcia et al.'s (2010) framework for pronunciation instruction. The t-test results showed that there were significant improvements in participants' phonological awareness and oral production of English past -*ed* allomorphy. For the phonemic discrimination tests, the results were inconclusive: the participants only improved in recognizing the /t/ allomorph. These findings highlight the affordances of MT and its speech capabilities regarding its pedagogical use for improving second language learners' pronunciation.

Acknowledgements

First, I would like to express my sincerest appreciation and praise to my kindest supervisor — Walcir Cardoso, who has the most charming face and the greatest inner beauty at the same time. I have to admit that starting a new career in Applied Linguistics was both a surprise and shock to me. I panicked when I realized how many textbooks and papers were waiting for me every single day. Until one day, I started one of my required courses APLI 604 with Walcir, and I just fell in love with that course immediately! The interactive and caring atmosphere within the classroom aced everything. There is no way for me to forget those homemade and mouth-watering snacks we all shared at the beginning of each class. Besides these precious memories, it is you who helped me with my thesis from scratch, particularly during COVID-19, in which online communication made things even more inconvenient. Yet, you have always been super patient and generous of taking time to answer my kindergarten-level questions and discuss them with me. In addition, I would like to thank Dr. Denis Liakin for being in my thesis committee and for contributing with so many helpful suggestions. I hope I have done justice to all your insightful comments and suggestions.

Second, I owe sincere thanks to my friends Clinton Hendry and Kaylee Wu who have exhibited their talented stats skill for saving me from endless panic attacks. I could not have such a smooth data analysis without your professional and selfless assistance. You are one of a kind!

Last but not least, I truly appreciate my dear parents, participants, and two adorable bunnies for their unconditional support all the time. You all are always there for me, supporting me, and listening to my dull complaints. Your energy fulfills me every day, I am who I am because of you. I have been thinking how lucky I am to have such amazing people around me. As life goes on, I hope we can get stronger and rely on one another always and forever.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter One

Grade four was the year that entirely changed my life. I can clearly remember, when I saw the score of 59 (out of 100) showing on my English final report, I suddenly realized that I had to put more effort to learning the language. After that initial shock, I decided to take English classes in an extracurricular institution (a tutoring center). That was the very first time I learned about the International Phonetic Alphabet (IPA) and I was obsessed with those foreign sounds and their exotic spelling. As time went by, my passion for English grew and, at the same time, I started receiving compliments about my pronunciation from people around me. This inspired me to keep on improving my speaking abilities. From a psychological standpoint, the pronunciation improvements I observed in my own speech (and which I heard from others) increased my confidence and interest in learning English. From a pedagogical standpoint, I deeply realized the significance of (learner-initiated) autonomous practice, motivation, and feedback pertaining to one's language development. This encouraged me to start this research project about pronunciation learning, targeting a country where English is spoken as a *foreign* language – EFL (my native country, China) and, consequently, where learners have limited access to English.

Pronunciation Teaching and Learning

Pronunciation has always been emphasized in the realm of English learning (e.g., Celce-Murcia et al., 2010; Liang, 2015; Shi, 2019), as it is a significant component for enhancing learners' intelligibility (Seyedabadi et al., 2015), confidence, and overall language development (Derwing et al., 1998; Thompson & Gaddes, 2005; Wang & Yang, 2015; Zhang & Yin, 2019). To facilitate the teaching of pronunciation, Celce-Murcia et al. (2010) proposed a framework that consists of five stages that range from the development of phonological awareness (stage 1) to full communicative interactions (stage 5): awareness raising, phonemic discrimination,

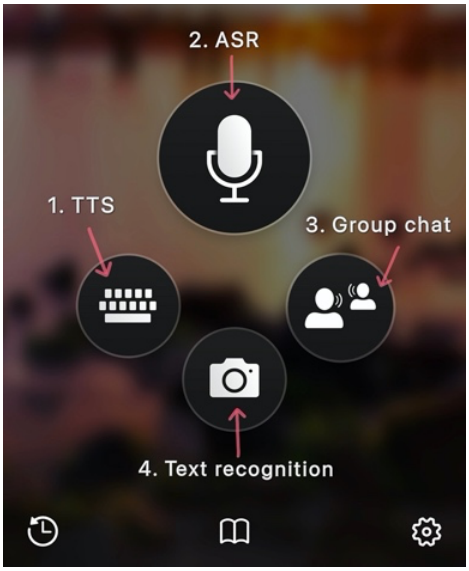
controlled practice with feedback, guided practice with feedback, and communicative practice with feedback. While designed with the goal of aiding pronunciation teaching, this framework is informed by current research in L2 phonology, emphasizing that the development of an L2 phonological system starts with the development of awareness (e.g., via noticing activities) and phonemic discrimination (e.g., via minimal pair recognition), and culminates with fluent and intelligible speech (as found in communicative interactions). Nevertheless, the literature reports a number of limitations that affect the teaching and learning of *foreign* languages, including a lack of teacher preparedness (Hu, 2005) and insufficient time for practice (Life, 2011). To mitigate these limitations, one "solution" is to encourage students to practice at their own time and pace (e.g., self-directed learning - SDL) via the use of technologies that promote autonomous learning, outside the classroom.

TTS, ASR, and Microsoft Translator (MT)

Studies have shown that technology-enriched environments are beneficial for the development of SDL behaviours (Mishra et al., 2013). Two speech technologies, text-to-speech synthesis (TTS) and automatic speech recognition (ASR), have proved to positively impact L2 learning. Via TTS, learners have access to unlimited and varied input, which can raise their phonological awareness and aural perception (Liakin et al., 2017). Similarly, ASR can increase students' self-efficacy and provide them with instant feedback, thus facilitating their pronunciation development (Celce-Murcia et al., 2010). So far, only one study has combined ASR and TTS into a single tool to explore their effectiveness in fostering L2 learning, using Google Translate (Van Lieshout & Cardoso, in press). Online translators such as Microsoft Translator (MT – adopted in our study) are tools that combine both technologies (see Figure 1).

In Figure 1, feature #1 (TTS) converts texts (input) to speech (output). Briefly, this option enables users to type words, sentences and passages and the embedded synthesizer will automatically produce the pronunciation output from the written text in whatever language is targeted (e.g., English, French, Mandarin Chinese). There are three speed levels available, which allow participants to select the one that meets their needs. Feature #2 represents the speech recognition tool, which generates texts based on users' speech input. Through this feature, users can check their pronunciation (machine-based) intelligibility and receive feedback based on the orthographic output that they receive. Feature #3 and #4 are not introduced here as they are not relevant (but see details in chapter two).

Figure 1. *Microsoft Translator*: Interface and features.



Microsoft Translator and its Speech Capabilities in Facilitating Pronunciation Acquisition

This study explores the affordances of these MT's speech capabilities for the development of English past *-ed* morphophonemics (/d/, /ɪd/, /t/ as in play/d/, visit/ɪd/ and walk/t/ respectively). It adopts Celce-Murcia et al.'s (2010) recommendations for pronunciation instruction (which also reflects phonological acquisition), starting with the development of

phonological/sound awareness (stage 1), proceeding to phonemic discrimination (stage 2), and culminating with production (controlled, guided, and communicative – stages 3, 4 and 5 respectively). Due to the duration of the study (two hours of SDL), we examined the development of past *-ed* pronunciation for the initial three stages.

The study was guided by the following research questions: (1) Can EFL learners acquire aspects of past *-ed* morphophonemics using Microsoft Translator's TTS and ASR capabilities in an SDL manner (without direct guidance from an instructor)? (2) If yes, which of the three stages of pronunciation development is affected by the proposed instruction: phonological awareness, phonemic discrimination, and/or oral production?

The implications of our findings are likely to benefit language educators who are currently struggling with teaching pronunciation due to the limitations that affect their L2 classrooms, as discussed earlier (i.e., insufficient teacher preparedness - Hu, 2005, and lack of time for practice - Life, 2011). To mitigate the abovementioned constraints, we recommend that teachers encourage their students to learn autonomously, anytime anywhere, outside of the wall boundaries of their language classroom. As such, to facilitate L2 learners' pronunciation, teachers should also consider extending the reach of the classroom by adopting speech technologies that can promote autonomous learning, such as TTS and ASR (Liakin et al., 2014; 2017; Van Lieshout & Cardoso, in press).

According to the guidelines for a manuscript-based MA thesis, the next section constitutes "a full submittable draft of a manuscript" that presents the literature review, methodology, results, and discussion of the current research.

Chapter Two

According to the British Council, English is a Lingua Franca spoken or being learned by over two billion people (Van Tol, 2016). Among these English learners, approximately 750 million speak it as a *foreign* language (EFL) while 375 million speak it as *second* language (ESL) (Beare, 2019). China is currently the largest EFL country in the world with 300 million English-speaking population (Van Tol, 2016).

Besides the large number of learners worldwide, English is also a significant asset in the international market (Van Tol, 2016), which has led to an increase in people's demand for learning the language, as attested in the proliferation of language schools worldwide. Nevertheless, the chances of learning English are not equal among places where English is a second or foreign language. Collins and Muñoz (2016), for instance, noted that learners' exposure to English in *foreign*-language contexts is much more limited than in settings where it is spoken as a *second* language (e.g., in the former, students lack an acceptable amount of in-class time for practice). Furthermore, it has been stated that the training provided to the pre-service and in-service teachers is inadequate (Hu, 2005). This results in instructors not being fully prepared for their classes (Lee & Reiko, 2001), which might trigger negative attitudes pertaining to curriculum innovation (Karavas-Doukas, 1995), consequently affecting their students' academic performance (Alrabai, 2016). These issues are exacerbated by the fact that teachers have a heavy workload (Nguyen & Newton, 2020). One way of mitigating some of these problems is by extending the reach of the classroom to an out-of-class setting, preferably in a context in which learners self-regulate their learning.

Researchers have discovered that self-directed learning (SDL), especially used alongside technology, is an effective way to promote learners' motivation (Du, 2013; Godwin-Jones, 2019;

Huang & Liao, 2015; Sade, 2011), self-efficacy (Thompson & Gaddes, 2005), and academic performance (Kennedy & Trofimovich, 2010; Rashid & Asghar, 2016). Previous studies have indicated that speech technologies such as text-to-speech synthesizers (TTS) and automatic speech recognition (ASR), which enable learners to self-regulate the learning process, are useful for facilitating their oral production and acquisition of L2 pronunciation (Chen, 2011; Handley, 2009). For instance, Google Translate (GT), which combines both TTS and ASR, has been shown to be helpful for the learning of (a small set of) phrases and their pronunciation in L2 Dutch (Van Lieshout & Cardoso, in press). We hypothesize that another online translator, Microsoft Translator (MT), can also promote students' pronunciation learning.

In the current study, we explore the effectiveness of MT in assisting Chinese EFL students in learning the pronunciation of regular past tense marker *-ed* and its three allomorphs (/d/, /t/, and /ɪd/, as in play[d], walk[t], and visit[ɪd], respectively) in an SDL environment. Based on Celce-Murcia et al.'s (2010) framework for teaching pronunciation, this study's design assumes that instruction should follow a perception-precedes-production approach that starts with awareness raising and phonemic discrimination, and then proceeds to oral production practice with corrective feedback.

Background

The English-as-a-Foreign-Language Context: Limitations

Although the terminology EFL and ESL are frequently used interchangeably when referring to English learning, they differ from each other in a very important way: the amount of input to which learners are exposed. For EFL learners, their exposure to English within and outside the classroom is much more limited than ESL learners (Collins & Muñoz, 2016), amounting to ten times more English exposure in a five-year time span than EFL learners

(Ortega, 2013). There are merits of placing learners in a high exposure environment (i.e., ESL). For instance, Longcope (2009) reported that students can have more meaningful contact with the L2 and, as a result, have more chances for access to more comprehensible input and output. In a study evaluating participants' pragmatic awareness development, Schauer also (2006) found that an ESL group located in the UK was more aware of pragmatic errors than the EFL group in Germany. There is also evidence that different amounts of exposure to the L2 can affect learners' perceptions towards learning English: Li (2014) conducted a study with 132 Chinese learners of English in China (EFL) and 122 Chinese learners of English in New Zealand (ESL), and found that the ESL learners developed more confidence and a more positive attitude towards learning English than the EFL learners in China.

The EFL environment is also affected by other constraints. It has been shown that teachers in EFL settings have limited training and pedagogical support, which in turn lead to many problems. These include inadequate class preparation (Lee & Reiko, 2001; Wahid & Sulong, 2013), lack of innovation in course implementation (Karavas-Doukas, 1995), which can consequently lead to students' low academic achievements (Alrabai, 2016). In addition, researchers have reported that EFL settings are particularly test-oriented (Barratt-Pugh, 2018; Lei & Qin, 2009; Life, 2011) and teacher-centered (Wu, 2009), which largely reduce learners' motivation to learn English. This is the case in China, a country where English is used as a *foreign* language.

An EFL Context: China

Deng Xiaoping, the Chinese chairman in the late 1970s, launched the "national modernization program", which made English a compulsory subject in Chinese public schools since it was believed that English was a significant tool for international development (Hu,

2005). Nevertheless, due to the shortage of professional teachers (Si, 2019) and immature course design, national-wide English teaching was perceived to be underdeveloped at that time. Although the proportion of qualified English teachers has been increasing and improving over time, the overall quality of teaching remains low (Xiong & Xiong, 2017; Yu et al., 2019). Despite recent educational reforms, English teaching in China currently faces three challenges: inadequate career training for pre-service and in-service English teachers (Hu, 2005), problematic curriculum design that emphasizes test performance and teacher-centeredness (Hu, 2005; Lei & Qin, 2009; Wu, 2009), and regional inequalities in the distribution of educational resources (Hu, 2002, 2003). Each of these challenges will be addressed next.

The first challenge was reported by Hu (2005), who stated that many language programs are certification-driven and, consequently, they do not emphasize teacher training on the pedagogy of L2 skills. For instance, although local teachers from Szechwan have passed the tests and receive their English teaching certifications, many cannot distinguish English /l/ and /r/ (Richards, 2012), which suggests that they might not be fully qualified for teaching English pronunciation, at least for these two specific consonants. Hu (2005) also mentioned that most of these pre-service and in-service programs provide teachers with outdated curriculum and teaching material, which considerably affect the implementation of effective language courses.

Secondly, the curriculum structure of English teaching in Chinese public schools has always been described as teacher-dominated (Wu, 2009), textbook-based, and test-oriented (Barratt-Pugh, 2018; Hu, 2005; Lei & Qin, 2009). Resembling other East Asian contexts, English teachers in China gravitate heavily towards the teaching of content knowledge (e.g., grammar, vocabulary) rather than pronunciation and oral communication (McIntyre & Foulsham, 2018).

The last challenge is the regional inequality of educational resources within China (Hu, 2002), which is unfairly distributed and concentrated in more economically developed regions. As a consequent, English learners in less economically developed regions have fewer opportunities to use the language to communicate (19.01% vs. 68.07% in economically developed regions), and they use less authentic materials for learning (17% vs. 69% respectively).

Researchers have started to explore solutions to mitigate the effects of these limitations. One of these "solutions" includes providing learners with tools for self-directed learning, assisted by technology, as will be discussed next.

Self-Directed Learning and Technology

One way of mitigating some of the challenges that EFL learners experience in China (e.g., emphasis on test performance, inequalities in the distribution of educational resources) is to encourage students to learn on their own, outside of the classroom, in a self-directed manner. Self-directed learning (SDL; e.g., Garrison, 1997) has a number of advantages: it can provide learners with opportunities to think independently (Fischer & Scharff, 1998), promote their self-control, and help them enrich their learning experience by becoming more responsible for their own learning (Benson, 2011; Carlet & Souza, 2018; Kennedy & Trofimovich, 2010). More importantly, SDL can also foster learners' motivation, goal orientation, and self-efficacy (Du, 2013; Sade, 2011). Lee et al. (2017) proposed that SDL in L2 pedagogy can be an extension of the regular classroom, especially for learners who wish to practice English but are constrained by time and limited access to speakers of the target language.

Due to the English teacher shortage in China (Hu, 2005), we need to consider another question: how can SDL be promoted to ensure that learners are motivated and acquiring

knowledge from reliable sources? One way of addressing this question is via the use of technology, which can optimize and maximize learners' motivation and self-efficacy. Technology also connects students with external resources (e.g., different platforms for practicing English) that they have never been exposed to in the regular classroom (Candy, 2004). This way, learners are able to receive L2 knowledge based on their own demands, wherever they are and whenever they need it. Several studies have shown that a technology-enriched environment is beneficial for students' development of SDL attitudes and behaviours (Guglielmino, 1977; Mishra et al., 2013), as it can train students to be not only knowledgeable about the targeted content, but also about the appropriate usage of the acquired knowledge (Du, 2013; Fahnoe & Mishra, 2013).

SDL has great potential to mitigate teachers' shortage and the regional inequalities observed in China, as described above. More specifically, technology-enhanced SDL might be an interesting approach for exploring the learning of L2 pronunciation, a language skill that remains under-investigated in EFL settings, particularly in China.

Pronunciation Teaching and Learning

Pronunciation has always been emphasized in the realm of English learning (e.g., Celce-Murcia et al., 2010; Liang, 2015; Shi, 2019), as it is a significant component for enhancing learners' communicative intelligibility (Seyedabadi et al., 2015), confidence, and overall language development (Derwing et al., 1998; Thompson & Gaddes, 2005; Wang & Yang, 2015; Zhang & Yin, 2019). There are a number of drawbacks when L2 learners have unintelligible speech, including a feeling of inferiority and lack of motivation to learn (Derwing, 2003; Munro, 2003; Thompson & Gaddes, 2005). Liang (2015) found that in China, English learners' pronunciation was positively associated with their listening comprehension competence, thus

highlighting the significance of pronunciation teaching for the learning of other English skills. However, due to the challenges previously discussed, EFL learners in China have difficulties in producing intelligible L2 speech.

To facilitate the teaching of pronunciation, Celce-Murcia et al. (2010) proposed a framework that consists of five stages that range from the development of phonological awareness (stage 1) to full communicative interactions (stage 5): awareness raising, phonemic discrimination, controlled practice with feedback, guided practice with feedback, and communicative practice with feedback. While designed with the goal of aiding pronunciation teaching, this framework is informed by current research in L2 phonology, emphasizing that the development of an L2 phonological system starts with the development of awareness (e.g., via noticing activities) and phonemic discrimination (e.g., via minimal pair recognition), and culminates with fluent and intelligible speech (as found in communicative interactions).

Because this study examines the *initial stages* of acquisition of past *-ed* allomorphy (i.e., what can be learned within the proposed two hours of SDL instruction, as will be defined later), it will focus on the initial *three stages* of Celce-Murcia et al.'s framework, namely, the development of phonological awareness, phonemic discrimination (perception), and controlled oral production. To achieve these goals in an SDL setting, the study adopts two speech technologies: one that targets the development of phonological awareness and phonemic discrimination via input exposure (Text-to-speech synthesis), and oral production via output practice (Automatic Speech Recognition).

Text-to-speech: Input Exposure

Text-to-speech (TTS) is a technology that generates speech from texts on electronic devices such as mobile phones and computers (Bione & Cardoso, 2020). Pedagogically, TTS has

been mostly used as a reading machine, a conversational partner, or a pronunciation model (Handley, 2009); however, research has also shown that its use is beneficial to language learning (e.g., Oktalia & Drajati, 2018), probably because it is free, easy to use, widely accessible (Van Lieshout & Cardoso, in press), and highly customizable (e.g., learners can customize the TTS' voice speed, accent and sex of speaker; Sha, 2010). Via TTS, learners can obtain unlimited and varied input exposure (Chapelle, 2003), which can contribute to the development of phonological awareness (Kennedy & Trofimovich, 2010; Piske, 2008). With its ability to provide learners with more exposure to spoken input, TTS has been shown to be effective not only in raising learners' awareness and aural perception of speech sounds (González, 2007; Soler-Urzua, 2011), but also in improving learners' oral production (Soler-Urzua, 2011).

In the past, one issue that hampered the pedagogical adoption of TTS in L2 settings was its unnatural voice (Nusbaum, Francis, & Henly, 1995). Current research, however, shows that current TTS synthesizers are able to generate high-quality oral output in terms of naturalness (for sentences and phrases), comprehensibility, accuracy, and intelligibility (Bione & Cardoso, 2020), sometimes sounding as natural as human voices. These findings suggest that TTS is ready for adoption in L2 English pedagogy, particularly in *foreign* language contexts (Bione & Cardoso, 2020).

Automatic Speech Recognition: Oral Production and Feedback

Conversely, Automatic Speech Recognition (ASR) converts speech into text. One of the affordances of the technology is that it encourages learners to orally produce language and, via its orthographic output, receive immediate feedback in an anxiety-free environment (Chen, 2011). In addition, ASR has the potential to increase students' self-efficacy, confidence and interest towards learning pronunciation, particularly because it can satisfy learners' individual

needs (McCrocklin, 2014) and consequently enhance their overall language learning experience (see also Chen, 2011; De Vries et al., 2014; García et al., 2020; Mroz, 2020; Van Lieshout & Cardoso, in press).

An interesting pedagogical affordance of ASR-based learning is its ability to provide immediate feedback to learners via its textual output (Chen, 2011; de Vries et al., 2014; McCrocklin, 2015; McCrocklin et al., 2019). A number of studies have indicated that these types of feedback facilitate learners' pronunciation learning (Celce-Murcia et al., 2010), as well as the overall language learning process (Lyster & Saito, 2010; Norris & Ortega, 2000). More importantly, immediate feedback can promote SDL (Liakin et al., 2017; McCrocklin, 2014; McCrocklin, 2015, 2016; Sheerin, 1997) by creating opportunities for learners to correct themselves in a stress-free environment (Neri & Strik, 2008), thus alleviating the spatial and temporal constraints observed in standard L2 learning settings of obtaining timely and sufficient feedback (Koreman et al., 2011).

Although ASR has been previously adopted as a pedagogical tool for L2 learning, very few studies have combined both ASR and TTS into a single tool to explore its effectiveness in promoting language acquisition. One tool that combines both technologies is online translators, specifically the one targeted by this research: *Microsoft Translator*.

Online Translators: Microsoft Translator

Online translators such as Google Translate (GT) are well-known and are commonly used for one particular purpose: to translate text from one language into another. However, due to its availability, accessibility, multilingualism, immediacy, and simplicity, it has great potential to be used as pedagogical tools (Niño, 2009), as attested by some current studies that indicate L2 learners' willingness to use and continue to use them to support their learning (e.g., Bahri &

Mahadi, 2016; Clifford et al., 2013; Garcia & Pena, 2011; Groves & Mundt, 2015; Jin & Deifell, 2013; Niitemaa & Pietilä, 2018; Niño, 2009; Tight, 2017; van Lieshout & Cardoso, in press). Although studies have shown that these translators may have limited ability to translate complex grammatical structures (Jin & Deifell, 2013; Josefsson, 2011; Niño, 2009), they are more capable of translating independent lexical items, which may help learners increase their vocabulary (Clifford et al., 2013; Jolley & Maimone, 2015; Niño, 2009). Van Lieshout and Cardoso (in press) conducted a study to gauge the effectiveness of using GT and its built-in TTS/ASR features to help learners acquire a set of ten simple Dutch phrases and associated pronunciation. The results confirmed that GT can be a useful pedagogical tool (e.g., participants were able to recall most of the target Dutch phrases and their pronunciation on posttests). It is predicted that other online translators such as *Microsoft Translator* will lead to similar outcomes.

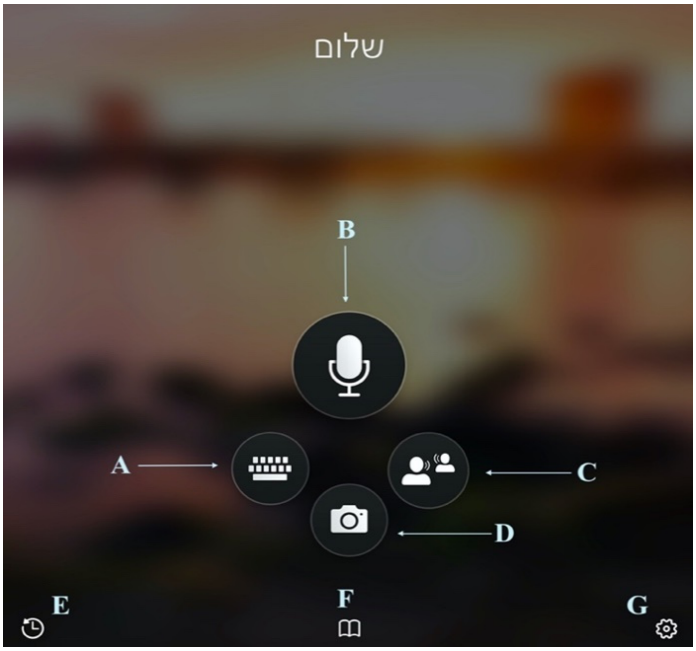
Microsoft Translator, the adopted translation tool, has at least two advantages over GT: it has a more aesthetically pleasing interface to enhance usability, and it has a higher overall rating score in both Apple App Store (4.7 vs. 4.3) and Google Play (4.6 vs. 4.5). Moreover, in the mid-2010s, MT adopted Neural Machine Translation technology, which increased the quality of translations and its synthesized voices (Microsoft, 2021).

MT is an application that can be downloaded on mobile devices such as cell phones and tablets (Android and iOS) – see figure 2 for an illustration of MT's basic features on a mobile application. Feature A enables users to translate one language to another (e.g., English, French). Basically, users can type words, sentences and passages by clicking A and the embedded TTS in the system will automatically produce the pronunciation output from the text. There are three speed levels available, which allow participants to adapt the ones that meet their needs. Feature B functions as a speech recognition tool, which generates texts based on users' speech input.

Through this feature, users can check their pronunciation intelligibility and receive feedback based on the orthographic output that they receive. Feature C is the unique "conversation" mode that allows groups of people to freely communicate using their different L1s. Feature D provides instant translation for the texts recognized in pictures taken by the user or from their photo albums. While Feature E provides the user's search history, Feature F contains practical phrases in the target language that learners are likely to encounter in their daily conversations. Lastly, Feature G is the general setting, which provides users with personalized choices such as female/male voice for the TTS. In our study, the participants are only required to use features A (TTS) and B (ASR), but they may explore the other options on their own.

Figure 2

User's Interface (UI) and features of MT



We are not aware of any studies that have explored the pedagogical effectiveness of *Microsoft Translator*'s TTS and ASR features for L2 pronunciation learning. In this study, this

translator's pedagogical effectiveness will be assessed by examining the acquisition of a rule-based feature of English morphophonology: past *-ed* allomorphy.

English Simple Past Tense Marker "*-ed*"

Our study targets the pronunciation of English regular simple past tense *-ed*: /d/, /t/, and /ɪd/ (as in *play<u>ed</u>*, *walk<u>ed</u>*, and *visit<u>ed</u>* respectively). We selected this morphophonological feature because: (1) they carry important grammatical meaning (e.g., past tense); (2) past tense and its allomorphs are hard to acquire (Collins et al., 2009), possibly because they are infrequent in the target language and are often deleted of assimilated in speech – they are opaque in the input (e.g., "ed" in *I missed the train* is often pronounced without the past /t/); (3) finally, these forms are not perceptually salient in the language (Dwight, 2012), which renders noticing difficult.

We believe that TTS is an ideal technology to deal with those problems because it can easily increase the frequency and decrease the opacity of speech input. ASR, on the other hand, is likely to encourage participants to orally practice what they are learning and receive immediate feedback in an anxiety-free SDL environment (Van Lieshout & Cardoso, in press).

Current Study

As discussed, three major issues afflict the Chinese EFL learning environment: (1) a large number of English teachers lack training (Hu, 2005), which may negatively affect language teaching (Lee & Reiko, 2001; Wahid and Sulong's, 2013); (2) the curriculum design is highly teacher-dominated (Wu, 2009) and test-oriented (Barratt-Pugh, 2018; Hu, 2005; Lei & Qin, 2009), resulting in insufficient time for student-centered communicative practice; and (3) educational resources are not equally distributed among different regions in China (Hu, 2002, 2003).

To mitigate these challenges, we proposed an SDL environment assisted by a combination of TTS and ASR technologies as found in *Microsoft Translator*. In this study, we examined whether this translation tool and its built-in speech features TTS and ASR can promote the acquisition in pronunciation of English regular past tense *-ed* in a self-directed manner. As such, the following research questions were proposed:

1. Can EFL learners acquire aspects of past *-ed* morphophonemics using *Microsoft Translator*'s TTS and ASR compatibilities on their own, without direct guidance from an instructor?
2. If yes, which of the three stages of pronunciation development is affected by the proposed instruction, considering the duration of the experiment: phonological awareness, phonemic discrimination, and/or oral production?

Methods

Participants

Twenty-nine Chinese-speaking participants were recruited via convenient sampling (15 females and 14 males; age: 16-18; English proficiency: intermediate) from Long An High School, Anyang, Henan Province, China. As English is officially taught in public systems as a foreign language since grade one in China; therefore, by the time students enroll in secondary school, they have been learning English for at least nine years.

Materials

Participants installed the application named Microsoft Translator on their mobile phone or tablet (Android or iOS operating system), which can be downloaded for free in various app stores. In this study, two English short stories were used in the treatment (sent as a docx file to

participants), each including target past tense -*ed* words and distractors. The number of each allomorph in the stories was evenly distributed (n=15, five of each allomorph; see Appendix G).

Three activities were completed by all participants. The first activity consisted of comprehension questions, which aimed to motivate listening and increase exposure to the target -*ed* forms. The second activity, filling in blanks, targeted the three past -*ed* allomorphs. Participants were asked to copy the stories from the docx file and paste them into MT. They then listened to the stories in TTS and filled in the blanks with the words they heard. The blanks in each story contained both target -*ed* allomorphs and a few distractors. The last activity involved categorizing the target -*ed* sounds: participants were asked to extract all words that end in -*ed*, listen to them, and classify them based on how the -*ed* form sounds (see Appendix G).

During the treatment, participants were also asked to use ASR to practice their pronunciation, focusing on -*ed* allomorphy (e.g., they were required to say out loud the words and sentences from the activities described above and then verify whether the ASR output reflects their intension – for example, if a participant intended to say "*walked*" but the ASR output showed "*walk*", that incorrect output would serve as feedback indicating that their attempt was inaccurate).

Instrument

Questionnaire

The questionnaire was composed of two sections: participants' language background and their experience in using online translators (Appendix A). The first section aimed to collect the participants' demographic information, including their linguistic ability, language use at school and/or work, and whether they have hearing problems. The second section provided an overview of how often and for what purposes the participants use online translators.

Pretest and posttest

Phonological Awareness. There were two tests to measure participants' phonological awareness of English past tense *-ed*. The first test consisted of a short interview-survey with three questions (Appendix B) that aimed to find out whether the participants were phonological aware of different English past tense *-ed* allomorphs. Based on the participants' answers, researchers categorized their awareness level as 0 (not aware), 1 (partially aware), and 2 (fully aware). The second awareness test was an ABX discrimination test. A total of 15 questions were given (5 questions × 3 allomorphs – see Appendix C), with each question containing three words (A, B, and X). Items under A and B differed from each other and took forms that ended with /d/, /t/ or /ɪd/, while X ended with the same allomorph as either A or B. Participants listened to the recordings (by a native English speaker) of each question twice and decided whether the allomorph of X sounded similar to A or B.

Phonemic discrimination. Two tests were designed for phonemic discrimination, which examined participants' phonemic knowledge about past *-ed* (Appendix D). The first test evaluated their ability of identifying past and non-past forms; the other test examined whether they can identify the actual allomorphs based on what they heard. The test consisted of 16 short sentences, with 4 distractors and 4 sentences per allomorph (4 × 3 = 12). Participants listened to each sentence twice and answered two related questions: (1) is the sentence in the past or not? If the sentence is in the past, how is *-ed* pronounced based on three options: /d/, /t/ or /ɪd/)?

Oral production. Two tests were used to assess the participants' *controlled* and *spontaneous* oral production (adapted from Cardoso, 2018). For test one, 30 words were given to the participants and they were invited to read each word aloud (Appendix E). The list of words consisted of 3 groups of conjugated verbs in the regular simple past (/d/, /t/, and /ɪd/) with 10

words in each allomorph group (3 × 10 =30). The second oral production test was a *spontaneous* test (in comparison with read-aloud). All participants were given a question sheet (Appendix F) and the researcher asked the questions in the left column while the participants were asked to answer them based on the hint (check list) in the right column. For instance, the researcher asked, "Did he play soccer?", if the hint was "No (baseball)", the participants were expected to say, "No, he *played baseball*." There were 12 sentences in this test, in which the three *-ed* allomorphs were evenly distributed with 4 words per allomorph (3 × 4 =12). All materials were the same for pretest and posttest, except for the order in which they were given to the participants.

Interview

The open-ended oral interview included the themes associated with the research questions and general scope of the paper to examine the participants' perceptions of the proposed self-direct, MT-assisted learning context (Appendix H). For instance, the participants were asked whether TTS and ASR had helped them notice speech errors while learning; how much progress they believed they had achieved during the treatment; their overall experience using MT and its built-in TTS and ASR features; and whether they would consider continuing to use it in the future (e.g., for language learning purposes).

Research Design and Procedure

This was a one-shot study with the duration of approximately two hours. The completion of study required participants to use their mobile devices (e.g., smart phone, iPad), which can enhance their mobility, accessibility, and consequently motivation (Cavus, 2016). The study was conducted online, and all interactions with the participants were audio recorded.

First, all participants and their parents (if underage) signed a consent form and send it to the researcher. The experiment started by explaining the study's main goals (as per the consent form), followed by a collection of the participants' background information (e.g., language use, experience using online translators). The pretest, which included the assessment of phonological awareness, phonemic discrimination, and production was implemented next.

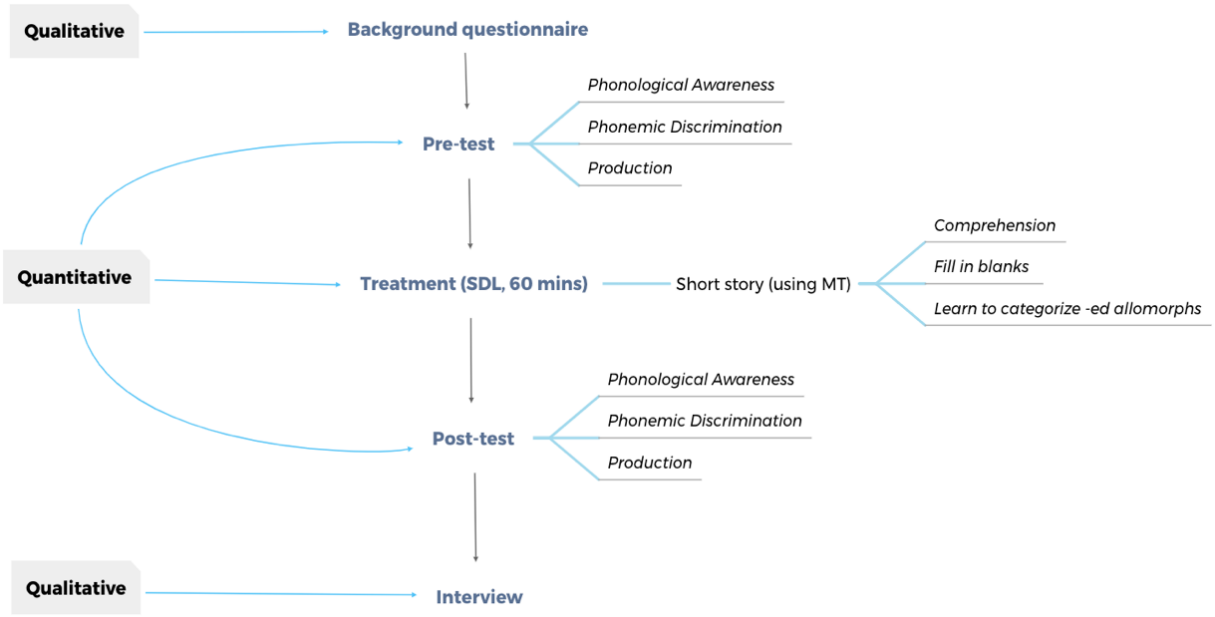
After the pretest session, the participants were trained on how to use MT for learning; for instance, they were encouraged to use TTS first to learn how to synthesize text into speech, and then utilized ASR to practice their oral production by orally repeating target words and phrases while receiving real-time (written) feedback. They were also taught how to interpret the orthographic output of the ASR. For example, if the participants said the word "touched" as "touch", the ASR would convert the speech input into text and displayed the word "touch" on the screen. In this case, the participants should consider their production as inaccurate and should return to TTS to listen to the form again and repeat the learning cycle.

After instruction, as discussed earlier, two short stories designed for the treatment period were distributed to the participants, each composed of three activities that aimed to scaffold their learning progress. After the 60-minute treatment, posttest was conducted to measure how much progress they had made and so that we could determine whether MT and its built-in speech features were effective pedagogical tools. Finally, fifteen participants were randomly selected to participate in oral interviews about their learning experience.

Figure 3 below illustrates a visual summary of the study design.

Figure 3

Design of the study.



Data Analysis

Questionnaire

The background information collected from the questionnaires was associated with other data to comprehensively analyze the findings.

Pretest and posttest

For ABX discrimination test, aural discrimination test, reading aloud, and oral "spontaneous" production test, participants' accuracy was analyzed using descriptive statistics (i.e., means and standard deviations). To better understand students' awareness and previous knowledge of English past *-ed* allomorphy, a short interview-survey was given. Based on the answers, researchers categorized students' awareness as 0 (not aware), 1 (partially aware), and 2 (fully aware). T-tests were conducted to compare pretest and posttest results. To ensure reliable

comparisons, the instruments and analytical tools applied in both pretest and posttest were the same.

Interview

The interview responses were coded using a blended approach (Graebner et al., 2012) where inductive and deductive coding were combined. The first step was to categorize participants' responses into two main themes: overall user experiences of MT and their perceptions of the tool from a pedagogical standpoint. The data were then further analyzed and allocated to new subcategories based on the findings and patterns observed, including issues related to strengths or weaknesses, learners' self-efficacy, and pedagogical usefulness. However, these findings are beyond the scope of the study and, consequently, will not be reported in detail. Instead, they will be used to support or complement our quantitative findings.

Results

This study explored two speech technologies in Microsoft Translator (i.e., TTS and ASR) to assist in L2 pronunciation learning in a teacher-guided, semi-autonomous context. Accordingly, it aimed to answer the following research questions: (1) Can EFL learners acquire past *-ed* morphophonemics (/d/, /t/, /id/) using Microsoft Translator's TTS and ASR capabilities on their own, without direct guidance from an instructor? (2) If yes, based on the framework of Celce-Murcia et al.'s (2010), which of the three stages of pronunciation development is/are affected by the proposed instruction: phonological awareness, phonemic discrimination, and/or oral production?

Phonological Awareness Test 1: Quantified answers to survey questions

For the first awareness test, the participants were categorized into one of three awareness levels based on an analysis of their answers to a set of questions (Appendix B) that probed their

phonological awareness of past -*ed* pronunciation: 0 (no knowledge), 1 (partial knowledge), and 2 (full knowledge of -*ed* allomorphy).

Table 1 and Figure 1 illustrate the number (N) and proportion (%) of participants for each awareness level before and after the treatment. Overall, the participants' performance in Test 1 after the treatment (posttest) indicate that their phonological awareness improved significantly: while cases of "not aware" and "partially aware" were significantly reduced (from 10.34% to 0% and from 72.41% to 20.69% respectively), the number of "fully aware" significantly increased (from 17.24% to 79.31%).

Table 1

Phonological awareness # 1.

	Pretest		Posttest	
Category	N	N / 29	N	N / 29
Not aware	3	10.34%	0	0.00%
Partially aware	21	72.41%	6	20.69%
Fully aware	5	17.24%	23	79.31%

Phonological Awareness Test 2: ABX

The second level of phonological awareness was assessed based on the participants' ability to aurally differentiate among the three allomorphs (/d/, /t/ and /id/), via an ABX test. In this test, participants were asked to listen to the recordings of ed-final verbs and answer fifteen questions: For each verb they heard, they had to decide on whether the -*ed* allomorph of the third word (X) sounded the same or similar to the first word (A) or the second word (B). For ease of illustration and consistency across tests, all means and standard deviations will be expressed in percentages.

The t-test results are presented in Table 2, where it is shown that the participants' overall phonological awareness significantly improved from pretest (M = 73.33%, SD = 15.22%) to posttest (M = 78.16%, SD= 15.11%), t(28) = -2.131, p = 0.042037906 < 0.05. An inspection of the results based on the individual allomorphs reveal that, while awareness of /d/ and /t/ were not significantly affected by the treatment, t(4) = -0.645, p = 0.5543 > 0.05 for /d/; t(4) = -1.616, p = 0.1813 > 0.05 for /t/, awareness of /id/ was, t(4) = -3.137, p = 0.035 < 0.05.

Table 2

Phonological awareness # 2. t-test Results

	Pretest		Posttest			
	M / 5	*SD*	*M* / 5	*SD*	*t-test*	*P value*
/d/	67.59%	16.64%	70.34%	16.28%	-0.645	0.554
/id/	77.24%	15.91%	82.76%	12.67%	-3.137	*0.035
/t/	75.17%	21.45%	81.38%	17.34%	-1.616	0.181
Total	73.33%	15.22%	78.16%	15.11%	-2.131	*0.042

*The difference is statistically significant ($p \leq 0.05$)

In summary, the findings obtained from the two awareness tests indicate that the pedagogical use of MT's speech features, as conceptualized in this study, successfully raised the participants' awareness of the past *-ed* allomorphy, particularly regarding the /id/ allomorph in the ABX text.

Phonemic discrimination Test 1: Aural discrimination between past or non-past

The first aural discrimination test (perception) asked the participants to listen to a set of 12 short phrases contained past *-ed* forms (4 sentences per allomorph and distractors) and then decide whether they were produced in the "past" or "non-past". As Table 3 illustrates, the participants did not improve significantly in their ability to discriminate past from non-past

constructions at the end of the experiment. M = 70.47%, SD = 11.81% for the pretest and M = 70.26%, 12.79% for the posttest, t(28) = 0.117, p = 0.908 > 0.05. Similarly, regarding the individual allomorphs, no significant improvements were observed, t(3) = -0.264, p = 0.809 > 0.05 for /d/; t(3) = 1.127, p = 0.342 > 0.05 for /id/; t(3) = -1.000, p = 0.391 > 0.05 for /t/.

Table 3

Perception # 1. t-test Results

	Pretest		Posttest			
	M /4	*SD*	*M* /4	*SD*	*t-test*	*P value*
/d/	62.07%	26.26%	62.93%	27.21%	-0.264	0.809
/id/	75.87%	12.27%	71.55%	12.07%	1.127	0.342
/t/	78.45%	17.67%	81.90%	13.90%	-1.000	0.391
Total	70.47%	11.81%	70.26%	12.79%	0.117	0.908

The difference is statistically significant if $p \leq 0.05$.

Phonemic discrimination Test 2: Aural discrimination between -*ed* allomorphs

The second perception test assessed the participants' ability to aurally discriminate among 12 past forms targeting the three -*ed* allomorphs (4 for each allomorph). Participants listened to the same recordings as in the previous test and made their choice among /d/, /t/ and /id/.

Table 4 shows that the participants significantly improved in their ability to aurally perceive the three -*ed* allomorphs from pretest (M = 42.67%, SD = 15.32%) to posttest (M = 47.41%, SD = 16.15%), t(28) = -2.238, p = 0.033 < 0.05. Significant improvement for /t/ was also observed when the allomorphs were analyzed individually: t(3) = -4, p = 0.028 < 0.05. However, /d/ and /id/ were not positively affected by the treatment, t(3) = -0.756, p = 0.505 > 0.05 for /d/ and t(3) = -0.214, p = 0.844 > 0.05 for /id/.

Table 4

Perception # 2. t-test Results

	Pretest		Posttest			
	M /4	*SD*	*M* /4	*SD*	*t-test*	*P value*
/d/	53.45%	27.08%	57.76%	28.49%	-0.756	0.505
/id/	22.41%	4.45%	24.14%	12.27%	-0.214	0.844
/t/	29.31%	5.97%	43.10%	5.98%	-4.000	*0.028
Total	42.67%	15.32%	47.41%	16.15%	-2.238	*0.033

*The difference is statistically significant ($p \leq 0.05$)

To summarize, although no significant improvements were observed in one of the phonemic discrimination tests (Past vs. Non-Past), findings in Test 2 show that, from a phonemic discrimination perspective, the participants significantly benefited from the MT-based pedagogical treatment, since their ability to identify -*ed* allomorphs significantly improved during the treatment, especially for allomorph /t/.

Oral production Test 1: Word-list reading aloud

The first oral production test consisted of a reading-aloud test in which the participants were asked to produce 30 *ed*-inflected forms in isolation (10 for each allomorph).

As illustrated in Table 5, significant improvements were observed in the participants' overall performance in -*ed* pronunciation, $t(28) = -5.143$, $p = 1.87493E\text{-}05 < 0.05$. However, when analyzed in isolation, the results indicate that only /t/ was significantly affected by the treatment, $t(9) = -8.703$, $p = 0.000011 < 0.05$. The production of the other two allomorphs were not affected by the treatment, $t(9) = 0.246$, $p = 0.812 > 0.05$ for /d/ and $t(9) = -0.830$, $p = 0.428 > 0.05$ for /id/.

Table 5

Production # 1. t-test Results

	Pretest		Posttest			
	M /10	*SD*	*M* /10	*SD*	*t-test*	*P value*
/d/	82.76%	13.4%	82.07%	12.14%	0.246	0.812
/id/	78.62%	12.68%	81.03%	13.33%	-0.830	0.428
/t/	42.41%	8.14%	67.24%	13.23%	-8.703	*0.000
Total	67.93%	15.54%	77.47%	13.47%	-5.143	*1.87493E-05

*The difference is statistically significant ($p \leq 0.05$)

Oral Production Test 2: Role play

Finally, the last production test evaluated the participants' ability to orally produce the target *-ed* forms in a less controlled role play activity, in which they were prompted to produce 12 *-ed* forms (4 for each allomorph).

As shown in Table 6, and similar to what was observed for the reading-aloud test, there were significant differences between the pretest and posttest results when all allomorphs were considered, $t(28) = -5.925$, $p = 0.000002 < 0.05$. However, there is no significant improvement for a specific allomorph when they are analyzed in isolation: $t(3) = 0.292$, $p = 0.789 > 0.05$ for /d/; $t(3) = -2.178$, $p = 0.118 > 0.05$ for /id/; $t(3) = -1.608$, $p = 0.206 > 0.05$ for /t/.

To conclude, the results from the two production tests indicate that there were significant improvements in *-ed* production during the treatment. Additionally, participants' overall enhancement on their oral production, under both controlled and less controlled conditions, reflects the effectiveness of the MT-based pedagogical treatment pertaining to pronunciation learning, at least in some general aspects of phonological development.

Table 6

Production # 2. t-test Results

	Pretest		Posttest			
	M /4	*SD*	*M* /4	*SD*	*t-test*	*P value*
/d/	80.17%	20.38%	79.31%	24.05%	0.292	0.789
/id/	77.59%	13.35%	88.79%	9.07%	-2.178	0.118
/t/	47.42%	18.54%	56.04%	23.62%	-1.608	0.206
Total	68.39%	16.87%	74.71%	14.86%	-5.925	*0.000

*The difference is statistically significant ($p \leq 0.05$)

Discussion

This study examined the pedagogical effects of MT's speech capabilities, ASR and TTS, in assisting learners acquire English past morphophonemics in a self-regulated but teacher-guided learning environment. Previous studies have shown that speech technologies such as TTS and ASR have great potential to facilitate L2 learners' pronunciations by complementing or extending the reach of the classroom (e.g., Cardoso, Collins & White 2012; Liakin et al., 2014; 2017; Soler-Urzua, 2011; Van Lieshout & Cardoso, in press). However, with the exception of Van Lieshout and Cardoso (in press) for Dutch, there are no studies that have explored the pedagogical use of these two technologies combined in a single application, as was the case in this study. As such, this research addressed the following two research questions: (1) Can EFL learners acquire past *-ed* morphophonemics (/d/, /t/, /id/) using Microsoft Translator's TTS and ASR capabilities on their own, without direct guidance from an instructor? (2) If yes, based on the framework of Celce-Murcia et al.'s (2010), which of the three stages of pronunciation development is affected by the proposed instruction: phonological awareness, phonemic discrimination, and/or oral production?

The t-test results showed that there were significant improvements in participants' phonological awareness and oral production of English past *-ed* allomorphy. For the phonemic discrimination tests, although participants in the first test ("past" or "non-past"?) seemed to "unlearn" to recognize the target forms (their knowledge decreased on the posttest), they did make statistically significant progress in recognizing the allomorph /t/ within past tense sentences in the second test (identification among the three allomorphs). These findings confirm the hypothesis that the pedagogical use of MT and its speech capabilities can help learners acquire "aspects" of the target pronunciation feature in all three stages of L2 phonological development. The following paragraphs provide a discussion of these findings.

Phonological Awareness

In test one, three questions were used to evaluate participants' phonological knowledge. The first question aimed to assess their knowledge of the differences in pronunciation among *-ed* inflected forms such as "walked", "played", and "visited". In the pretest, participants who were rated as *partially aware* (72.41%) formed the largest group. Participants in this group were able to notice the pronunciation differences between these words; yet, they were unable to accurately match their pronunciations with the corresponding words. For instance, one participant stated that "the endings of these three words seemed different", while another mentioned that "*-ed* is usually pronounced as /d/, but I have no idea how *-ed* should be pronounced in the given words". On the posttest, the majority of the participants improved in their phonological awareness (79.31%) and, as a consequent, received a *fully aware* rating. Specifically, they could not only identify the three allomorphs, but also correctly match the pronunciations of "walked", "played", and "visited" with their corresponding *-ed* sounds. Overall, it is possible that the increase in exposure to the target input, promoted by the technologies and activities implemented in this

study, contributed to the improvements observed in phonological awareness (Kennedy & Trofimovich, 2010; Piske, 2008; Schauer, 2006).

Regarding the ABX test, the participants' phonological awareness improved significantly after the treatment, particularly for the allomorph /id/. This founding can be attributed to the perceptual salience of /id/ in comparison with the other allomorphs (e.g., it contains a vowel, which is higher in sonority than consonants – Zec, 1995). In the literatures, these results corroborate those by Solt et al. (2004), who also found that the syllabic allomorph /id/ is better perceived by learners than the non-syllabic /t/ and /d/. In addition, orthographically, the pronunciation of /id/ is the most transparent of all past-tense allomorphs, as it closely corresponds to the suffix *-ed* (i.e., it contains a mid-front vowel followed by a consonant, just like /id/). This possibly explains why the phonological awareness of the participants increased for this particular allomorph (see also Delatorre, 2010 for similar claims).

Overall, these results indicate that the participants were able to acquire the first stage of phonological development, as per Celce-Murcia's (2010) framework of pronunciation teaching, via the use of MT and its embedded TTS and ASR capabilities.

Phonemic Discrimination

The first phonemic discrimination test examined the participants' ability to identify past from non-past constructions among a set of 16 sentences recorded by an English native speaker. Overall, the participants' accuracy was not affected by the treatment, as no significant differences were observed on the posttest results. One possible explanation for these findings might be due to the phonetic processes that naturally occur in speech (Celce-Murcia et al., 2010), which could have obscured the presence of the target *-ed* form (e.g., in oral production, the phrase *loved the* may be produced or perceived as *lov/*ð*/e*, thus rendering the original /d/

allomorph in *love/d/* opaque). For similar claims explaining why these types of phonetic processes can affect the processing and consequently learning of English past *-ed*, see Collins et al. (2009).

The second phonemic discrimination test measured the participants' ability to discriminate among the target *-ed* allomorphs (/d/, /t/, and /id/), using the same sentences from task one. Overall, their ability to discriminate among the three allomorphs improved as a result of the proposed MT-based treatment. Interestingly, when the forms were analyzed in isolation, significant improvements were only observed for /t/.

Because these two tests were designed to examine a relatively similar ability (the participant's ability to discriminate sounds), the discrepancies in results were surprising. One possible explanation for these differences might be due to a task effect (Swain & Lapkin, 2001), which are assumed to exert an influence on task performance and consequently on how L2 development is evaluated. Similar claims have been made in the variationist literature (e.g., Cardoso, 2007; Major, 2004). In fact, to address similar task affects, variationist linguists often utilize a variety of tasks in their studies, usually obeying a stylistic hierarchy ranging from most formal (e.g., the reading aloud of words) to more spontaneous interactions. Accordingly, it is also possible that the two tasks were not cognitively equal in complexity: while the first task included both an aural and a semantic component, the second discrimination test relied exclusively on the participants' ability to hear and identify allomorphs. The effects of task complexity have been acknowledged and examined in the SLA literature, resulting with the recommendation that, in L2 teaching, tasks should be sequenced in a way so that the simple ones precede the more cognitively complex ones (Robinson, 2011).

The contribution of /t/ perception for the overall findings remains inconclusive. As

previous discussed, based on the phonological salience of both /id/ (this form contains a highly sonorous vowel) and /d/ (which is more sonorous than /t/), we expected that the former would be more easily perceptible than the other allomorphs (see Solt et al., 2004 and Zec, 1995 for the rationale; and Cardoso, 2018 for an empirical study involving production).

To conclude, while the two discrimination tests were affected by a task effect, overall, we can deduce that aspects of this stage of phonological development benefited from the MT-based treatment, as the participants were able to identify the target *-ed* allomorphs in one of the tests, when they occurred in meaning-carrying sentence constructions.

Oral Production

For the oral production tests, the participants were asked to produce the target *-ed* allomorphs in two tasks: a read-aloud (of words) and a spontaneous role-play (in which they were asked to engage in a conversation about past events with the researcher). At the end of the treatment, the overall accuracy improved significantly in both tests, suggesting that the MT-based treatment was pedagogically effective and led to improvements in the participants' oral production. These findings have parallels in previous research, which indicate that speech technologies such as TTS and ASR facilitate speech production and the acquisition of L2 pronunciation (see Van Lieshout & Cardoso, in press), particularly because they provide users with ample opportunities to self-regulate their learning (Chen, 2011; Handley, 2009; Van Lieshout & Cardoso, in press).

Despite the improvements observed in oral production, when the allomorphs were analyzed in isolation, only /t/ in the read-aloud task improved on the posttest. This pattern can be explained by the ceiling effect suggested by Rifkin (2005), a phenomenon that occurs when a high proportion of participants have maximum scores on the observed variables. To illustrate,

consider the pretest results for /d/ (82.77%), /id/ (78.62%) and /t/ (42.41%) in the read-aloud test and note that the participants had already reached approximately 80% of accuracy in the production of /id/ and /d/. In this scenario, only /t/ had room for potential development in production and, as such, it was the only segment that significantly improved as a result of the proposed treatment.

In summary, our results indicate that, via the use of MT's speech capabilities in a self-directed learning environment, our participants were able to improve in most stages of Celce-Murcia et al.'s (2010) framework for the teaching L2 pronunciation: phonological awareness, phonemic discrimination, and oral production.

Conclusions

The purpose of this study was to explore MT and its built-in speech technologies (TTS and ASR) as pedagogical tools in the acquisition of L2 English pronunciation (past *-ed* allomorphy) in terms of *phonological awareness*, *phonemic discrimination*, and *oral production* in a teacher-guided semi-autonomous context. The results indicate that the technology helped improve learners' knowledge of past *-ed* allomorphy in these three components of phonological development.

Despite these optimistic findings, there are a number of limitations that will need to be addressed in future research. The first main limitation relates to the two-hour duration of this one-shot study, in which participants were asked to learn the intricacies of past *-ed* allomorphy without direct teacher intervention, in a semi-autonomous fashion. Although this format and time constraint resemble the way students complete homework assignments (e.g., they are guided by a teacher, work semi-autonomously at their own pace, within time constraints), the short duration of the study might have reduced the participants' chances of fully acquiring all aspects of what it

means to know past *-ed* allomorphy in English. In future studies, researchers should consider extending the duration of the intervention to observe the full potential of the adopted technologies for the learning of L2 pronunciation. A second limitation has to do with COVID-19, a health crisis that forced us to move all data collection to an online environment, resulting in a lack of control over equipment and other technological issues (e.g., malfunctioning microphones, faulty internet connections, small screens that may lead to fatigue). Future research should promote an environment that most likely resemble a real-life learning environment, one in which students can learn on their own, at their own pace and time. Finally, a third limitation relates to the study design, which lacked a delayed posttest. As the participants were from the same boarding school in China where course schedules were extremely strict, it was difficult to arrange another time for an originally planned delayed posttest. To address this limitation, researchers should consider a longitudinal design that incorporates immediate and delayed posttests so that we can have a more comprehensive picture of the pedagogical potential of the adopted technologies.

The present study offers some important pedagogical implications. The most important one is that it has shown that learners *can* acquire certain aspects of L2 pronunciation (e.g., phonological awareness, phonemic discrimination, and oral production) when engaged in teacher-guided semi-autonomous activities such as those that characterize homework assignments. Via technologies such as MT and its speech capabilities, teachers can mitigate the time limitation that prevents them from focusing on pronunciation instruction, and consequently extend the reach of their classroom to an environment that has potential to provide input that is abundant and varied (via TTS), with ample opportunities for production practice and feedback (via ASR).

Chapter Three

This chapter will summarize the results and conclusions that were discussed in chapter two, followed by the implications for language education and potential future directions for future research.

Summary of Goals and Findings

The purpose of this study was to explore MT and its built-in speech technologies (TTS and ASR) as pedagogical tools in the acquisition of L2 English pronunciation (past *-ed* allomorphy) in terms of *phonological awareness*, *phonemic discrimination*, and *oral production* (as per Celce-Murcia et al.'s framework of pronunciation teaching), in a teacher-guided semi-autonomous context. Specifically, we addressed the following research questions: (1) Can EFL learners acquire aspects of past *-ed* morphophonemics using *Microsoft Translator*'s TTS and ASR compatibilities on their own, without direct guidance from an instructor? (2) If yes, which of the above three stages of pronunciation development is affected by the proposed instruction, considering the duration of the experiment?

The t-test results showed that there were significant improvements in participants' phonological awareness and oral production of English past *-ed* allomorphy. For the phonemic discrimination tests, although participants' knowledge seemed to decrease on the posttest in test one ("past" or "non-past"?), they made statistically significant progress in recognizing the allomorph /t/ within sentences in the past tense.

Implications for L2 Education

The present study offers some important pedagogical implications. The most important one is that it has shown that learners *can* acquire certain aspects of L2 pronunciation (e.g., phonological awareness, aspects of phonemic discrimination, and oral production) when engaged

in teacher-guided semi-autonomous activities such as those that characterize the types of assignments that language teachers ask their students to complete on their own, at home. Via technologies such as MT and its speech capabilities, teachers can mitigate the time limitation that prevents them from focusing on pronunciation instruction, and consequently extend the reach of their classroom to an environment that has potential to provide input that is abundant and varied (via TTS), with ample opportunities for production practice and feedback (via ASR).

Limitations and Further Research

Despite these optimistic findings, there are a number of limitations that will need to be addressed in future research. The first main limitation relates to the two-hour duration of this one-shot study, in which participants were asked to learn the intricacies of past *-ed* allomorphy without direct teacher intervention, in a semi-autonomous fashion. Although this format and time constraint resemble the way in which students complete homework assignments (e.g., they are guided by a teacher and work semi-autonomously at their own pace, within time constraints), the short duration of the study might have reduced the participants' chances of fully acquiring all aspects of what it means to know past *-ed* allomorphy in English. In future studies, researchers should consider extending the duration of the intervention to observe the full potential of the adopted technologies for the learning of L2 pronunciation.

A second limitation has to do with COVID-19, a health crisis that forced us to move all data collection to an online environment, resulting in a lack of control over equipment and other technological issues (e.g., participants' malfunctioning microphones, faulty internet connections, small screens that may lead to fatigue) and, more importantly, creating a virtual learning setting that did not simulate the authenticity of anytime anywhere autonomous learning. Future research should promote an environment that most likely resemble a real-life learning environment, one in

which students can learn on their own, at their own pace and time.

Finally, a third limitation relates to the study design, which lacked a delayed posttest. As the participants were from the same boarding school in China where course schedules are extremely strict, it was difficult to arrange another time for an originally planned delayed posttest. To address this limitation, researchers should consider a longitudinal design that incorporates immediate and delayed posttests so that we can have a more comprehensive picture of the pedagogical potential of the adopted technologies.

Conclusion

Participants in this study were able to acquire aspects of past *-ed* morphophonemics (/t/, /d/, /id/) using MT's TTS and ASR on their own, as if completing a homework assignment. Additionally, three stages of pronunciation development (i.e., phonological awareness, phonemic discrimination, oral production) were positively affected by the treatment, reflected in the participants' significant improvements in most aspects of their morphophonological past *-ed* development. Despite the obvious limitations of this study, as discussed earlier, the main pedagogical implication is that instructors should encourage the self-directed use of online translators such as Microsoft Translate (as well as other poplar tools such as DeepL Translator and Google Translate) to extend the reach of the classroom (e.g., via technology-enhanced homework assignments), and thus alleviate some of the limitations that affect the EFL context, as discussed earlier.

References

Alrabai, F. (2016). Factors Underlying Low Achievement of Saudi EFL Learners. *International Journal of English Linguistics*, *6*(3), 21-37. https://doi.org/10.5539/ijel.v6n3p21

Bahri, H., & Mahadi, T. S. T. (2016). Google Translate as a Supplementary Tool for Learning Malay: A Case Study at Universiti Sains Malaysia. *Advances in Language and Literary Studies, 7*(3), 161-167.

Bahri, H., Mahadi, T. (2016). Google Translate as a supplementary tool for learning Malay: A case study at Universiti Sains Malaysia. *Advances in Language and Literary Studies*, *7*(3), 161–167.

Barratt-Pugh, L., Zhao, F., Zhang, Z., & Wang, S. (2018). Exploring current Chinese higher education pedagogic tensions through an activity theory lens. *Higher Education*, *77*(5), 831-852. https://doi.org/10.1007/s10734-018-0304-8

Beare, K. (2019). *How Many People Learn English Around the World?*. ThoughtCo. Retrieved 7 October 2020, from https://www.thoughtco.com/how-many-people-learn-english-globally-1210367.

Benson, P. (2011*). Teaching and researching autonomy in language learning* (2nd ed.). Harlow: Longman.

Bione, T., & Cardoso, W. (2020). Synthetic voices in the foreign language context. *Language Learning & Technology, 24*(1), 169–186. https://doi.org/10125/44715

Candy, P. (2004). *Linking Thinking: Self-directed learning in the digital age*. Department of Education, Science and Training.

Cardoso, W. (2007). The variable development of English word-final stops by Brazilian Portuguese speakers: A stochastic optimality theoretic account. *Language Variation and Change, 19*, 1-30.

Cardoso, W. (2011). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research, 27*(4), 433-465. doi: 10.1177/0267658311413540

Cardoso, W. (2018). Learning L2 pronunciation with a text-to-speech synthesizer. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds), *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 16-21). Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.806

Carlet, A., & Souza, H. (2018). Improving L2 pronunciation inside and outside the classroom. *Ilha Do Desterro A Journal of English Language, Literatures In English And Cultural Studies*, *71*(3), 99-124. https://doi.org/10.5007/2175-8026.2018v71n3p99

Cavus, N. (2016). Development of an Intelligent Mobile Application for Teaching English Pronunciation. *Procedia Computer Science*, *102*, 365-369. https://doi.org/10.1016/j.procs.2016.09.413

Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching pronunciation: Reference for teachers of English to speakers of other languages.* Cambridge: Cambridge University Press.

Chapelle, C. (2003). *English Language Learning and Technology*. John Benjamins Publishing Company.

Chen, H. (2011). Developing and evaluating an oral skills training website supported by automatic speech recognition technology. *Recall*, *23*(1), 59-78. https://doi.org/10.1017/s0958344010000285

Clifford, J., Merschel, L., & Reisinger, D. (2013). Meeting the challenges of machine translation. *The Language Educator, 8*, 44-47.

Collins, L., & Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal, 100*(1), 133–147.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some Input on the Easy/Difficult Grammar Question: An Empirical Study. *The Modern Language Journal*, *93*(3), 336-353. https://doi.org/10.1111/j.1540-4781.2009.00894.x

de Vries, B., Cucchiarini, C., Bodnar, S., Strik, H., & van Hout, R. (2014). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning*, *28*(6), 550-576. https://doi.org/10.1080/09588221.2014.889713

Delatorre, F. (2010). The role of orthography on the production of regular verbs ending in ed by Brazilian EFL learners. In Proceedings of the 9th Seminário do Círculo de Estudos Linguísticos do Sul (pp. 1-13). Florianópolis: Federal University of Santa Catarina.

Derwing, T.M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59*, 547–567.

Derwing, T.M., Munro, M.J., & Wiebe, G.E. (1998). Evidence in favour of a broad framework for pronunciation instruction. *Language Learning, 48*, 393-410.

Du, F. (2013). Student Perspectives of Self-Directed Language Learning: Implications for Teaching and Research. *International Journal for the Scholarship of Teaching and Learning. 7*(2), 1-16. https://doi.org/10.20429/ijsotl.2013.070224

Dwight, V. (2012). *Regular Past Tense Acquisition in L2 English: The Roles of Perceptual Salience and Readiness* (Master). Concordia University.

Fahnoe, C., & Mishra, P. (2013). Do 21st century learning environments support self-directed learning? Middle school students' response to an intentionally designed learning environment. In *Society for information technology & teacher education international conference* (pp. 3131-3139). New Orleans, Louisiana, United States.

Fischer, G., & Scharff, E. (1998). Learning Technologies in Support of Self-Directed Learning. *Journal of Interactive Media in Education, 98*(4), 1-32.

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, *24*(5), 471-487. Retrieved from http://dx.doi.org/10.1080/09588221.2011.582687.

Garrison, D.R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly, 48*(1), 18-33.

Godwin-Jones, R. (2019). Riding the digital wilds: Learner autonomy and informal language learning. *Language Learning & Technology, 23*(1), 8–25. https://doi.org/10125/44667

González, D. (2007). Text-to-Speech Applications Used in EFL Contexts to Enhance Pronunciation. *TESL-EJ*, *11*(2), 1-11.

Graebner, M., Martin, J., & Roundy, P. (2012). Qualitative data: Cooking without a recipe. *Strategic Organization*, *10*(3), 276-284. https://doi.org/10.1177/1476127012452821

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes, 37*, 112-121.

Guglielmino, L. (1977). *Development of the self-directed learning readiness scale.* (Ph.D). University of Georgia.

Handley, Z. (2009). Is text-to-speech synthesis ready for use in computer-assisted language learning?. *Speech Communication*, *51*(10), 906-919. https://doi.org/10.1016/j.specom.2008.12.004

Hu, G. (2002). Recent important developments in secondary English language teaching in the People's Republic of China. *Language, Culture and Curriculum, 15*, 30-49.

Hu, G. (2003). English Language Teaching in China: Regional Differences and Contributing Factors. *Journal of Multilingual And Multicultural Development*, *24*(4), 290-318. https://doi.org/10.1080/01434630308666503

Hu, G. (2005). English Language Education in China: Policies, Progress, and Problems. *Language Policy*, *4*(1), 5-24. https://doi.org/10.1007/s10993-004-6561-7

Huang, Y., & Liao, L. (2015). A study of text-to-speech (TTS) in children's English learning. *Teaching English With Technology*, *15*(1), 14-30.

Jin, L., & Deifell, E. (2013). Foreign Language Learners' Use and Perception of Online Dictionaries: A Survey Study. *MERLOT Journal of Online Learning and Teaching*, *9*(4), 515-533.

Jolley, J., & Maimone, L. (2015). Free Online Machine Translation: Use and Perceptions by Spanish Students and Instructors. In A. Moeller, *Learn Languages, Explore Cultures, Transform Lives* (pp. 181-200). Robert M. Terry.

Josefsson, E. (2011). *Contemporary Approaches to Translation in the Classroom: A Study of Students' Attitudes and Strategies*. Retrieved from http://du.diva-portal.org/smash/get/diva2:519125/FULLTEXT01.pdf.

Karavas-Doukas, E. (1995). Teacher identified factors affecting the implementation of an EFL innovation in Greek public secondary schools. *Language, Culture And Curriculum*, *8*(1), 53-68. https://doi.org/10.1080/07908319509525188

Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness*, *19,* 171-185.

Koreman, J., Bech, Ø., Husby, O. & Wik, P. (2011). L1-L2*map*: a tool for multi-lingual contrastive analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences.* Hong Kong.

Lee, C., & Reiko, K. (2001). A Tandem of Native and Non-Native Teachers: Voices from Japanese and American Teachers in the EFL Classroom in Japan. In *the Second International Language Teacher Educator' Conference*. Minneapolis, MN, United States.

Lei, L., & Qin, X. (2009). An Empirical Study of Success and Failure Attributions of EFL Learners at the Tertiary Level in China. *The Asian EFL Journal*, *11*(3), 29-51.

Li, Q. (2014). Differences in the motivation of Chinese learners of English in a foreign and second language context. *System*, *42*, 451-461. https://doi.org/10.1016/j.system.2014.01.011

Liakin, D., Cardoso, W., & Liakina, N. (2014). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, *32*(1), 1-25. https://doi.org/10.1558/cj.v32i1.25962

Liakin, D., Cardoso, W., & Liakina, N. (2017) The pedagogical use of mobile speech synthesis (TTS): Focus on French liaison. *Computer Assisted Language Learning, 30*(3-4), 348–365.

Liakin, D., Cardoso, W., & Liakina, N. (2017). Mobilizing Instruction in a Second-Language Context: Learners' Perceptions of Two Speech Technologies. *Languages*, *2*(3), 11. MDPI AG. Retrieved from http://dx.doi.org/10.3390/languages2030011

Liang, D. (2015). Chinese Learners' Pronunciation Problems and Listening Difficulties in English Connected Speech. *Asian Social Science*, *11*(16), 98-106. https://doi.org/10.5539/ass.v11n16p98

Life, J. (2011). Motivation and EFL University Students in North-East Asia. *The Asian EFL Journal*, *13*(3), 11-41.

Longcope, P. (2009). Differences between the EFL and the ESL language learning contexts. *Studies in Language and Culture, 30*(2), 303-320.

Lyster, R., & Saito, K. (2010). Oral Feedback in Classroom SLA. *Studies In Second Language Acquisition*, *32*(2), 265-302. https://doi.org/10.1017/s0272263109990520

Major, R. (2004). Gender and stylistic variation in second language phonology. *Language Variation and change, 16*(3), 169-188.

McCrocklin, M. S. (2014). *The potential of Automatic Speech Recognition for fostering pronunciation learners' autonomy* (Doctoral dissertation, Iowa State University, Ames, United States). Retrieved from https://lib.dr.iastate.edu/etd/13902/

McCrocklin, S. (2015). Automatic speech recognition: Making it work for your pronunciation class. In J. Levis, R. Mohammed, M. Qian & Z. Zhou (Eds). *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference* (ISSN 2380-9566), Santa Barbara, CA (pp. 126-133). Ames, IA: Iowa State University.

McCrocklin, S. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, *57*, 25-42. https://doi.org/10.1016/j.system.2015.12.013

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal Of Second Language Pronunciation*, *5*(1), 98-118. https://doi.org/10.1075/jslp.16034.mcc

McIntyre, N., & Foulsham, T. (2018). Scanpath analysis of expertise and culture in teacher gaze in real-world classrooms. *Instructional Science*, *46*(3), 435-455. https://doi.org/10.1007/s11251-017-9445-x

Microsoft Translator service - Azure Cognitive Services. (2021). Microsoft. Retrieved November 7, 2020, from https://docs.microsoft.com/en-us/azure/cognitive-services/translator/translator-info-overview.

Mishra, P., Fahnoe, C., & Henriksen, D. (2013). Creativity, Self-Directed Learning and the Architecture of Technology Rich Environments. *Techtrends*, *57*(1), 10-13. https://doi.org/10.1007/s11528-012-0623-z

Mroz, A. (2020). Aiming for advanced intelligibility and proficiency using mobile ASR. *Journal of Second Language Pronunciation 6*(1), 12-38.

Munro, M.J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*, 38-51.

Neri, A., Cucchiarini, C., & Strik, H. (2008). The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch. *Recall*, *20*(2), 225-243.

Nguyen, L., & Newton, J. (2020). Enhancing EFL Teachers' Pronunciation Pedagogy Through Professional Learning: A Vietnamese Case Study. *RELC Journal, 52*(1), 77-93. https://doi.org/10.1177/0033688220952476

Niitemaa, M. L., & Pietilä, P. (2018). Vocabulary Skills and Online Dictionaries: A Study on EFL Learners' Receptive Vocabulary Knowledge and Success in Searching Electronic Sources for Information. *Journal of Language Teaching and Research*, *9*(3), 453-462.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, *50*(3), 417-528. https://doi.org/10.1111/0023-8333.00136

Nusbaum, H., Francis, A., & Henly, A. (1995). Measuring the naturalness of synthetic speech. *International Journal of Speech Technology, 2*(1), 7-19.

Oktalia, D., & Drajati, A. (2018). English teachers' perceptions of text to speech software and Google site in an EFL Classroom: What English teachers really think and know. *International Journal Of Education And Development Using Information And Communication Technology, 14*(3), 183-192.

Ortega, L. (2013). *Understanding second language acquisition*. Abingdon, UK: Routledge.

Piske, T. (2008). Phonetic awareness, phonetic sensitivity and the second language learner. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education Vol. 6: Knowledge about language* (pp. 155-166). New York: Springer Science.

Rashid, T., & Asghar, H. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers In Human Behavior*, *63*, 604-612. https://doi.org/10.1016/j.chb.2016.05.084

Richards, M. (2012). Helping Chinese learners distinguish English /l/ and /n/. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, (pp. 161-167). Ames, IA: Iowa State University.

Rifkin, B. (2005). A Ceiling Effect in Traditional Classroom Foreign Language Instruction: Data from Russian. *The Modern Language Journal*, *89*(1), 3-18. https://doi.org/10.1111/j.0026-7902.2005.00262.x

Robinson (2011). *Second Language Task Complexity: Researching the Cognition Hypothesis of language learning and performance*. John Benjamins.

Sade, L. A. (2011). Emerging selves, language learning, and motivation through the lens of chaos. In G. Murray, X. Gao, & T. Lamb (Eds.), *Identity, motivation, and autonomy in language learning* (pp. 42-56). Bristol, UK: Multilingual Matters.

Schauer, G. (2006). Pragmatic Awareness in ESL and EFL Contexts: Contrast and Development. *Language Learning*, *56*(2), 269-318. https://doi.org/10.1111/j.0023-8333.2006.00348.x

Seyedabadi, S., Fatemi, A., & Pishghadam, R. (2015). Towards Better Teaching of Pronunciation: Review of Literature in the Area. *Mediterranean Journal of Social Sciences*, *6*(4), 76-81. https://doi.org/10.5901/mjss.2015.v6n4s1p76

Sha, G. (2010). Using TTS Voices to Develop Audio Materials for Listening Comprehension: A Digital Approach. *British Journal of Educational Technology, 41*(4), 632-641.

Sheerin, S. (1997). An exploration of the relationship between self-access and independent learning. In P. Benson and P. Voller (Eds.), *Autonomy and Independence in Language Learning* (pp. 54-65). London: Longman.

Shi, X. (2019). A Study on English Pronunciation Problems at Segmental Level of College Students from Mainland China. *Journal of Literature And Art Studies*, *9*(2), 238-244. https://doi.org/10.17265/2159-5836/2019.02.013

Si, J. (2019). English as a Lingua Franca: A New Approach for English Language Teaching in China?. *Chinese Journal of Applied Linguistics, 42*(1), 113-135. https://doi.org/10.1515/cjal-2019-0007

Soler-Urzua, F. (2011). *The acquisition of English /ɪ/ by Spanish speakers via text-to-speech synthesizers: A quasi-experimental study* (Master's Thesis). Concordia University, Montreal, CA.

Solt, S., Pugach, Y., Klein, E.C., Adams, K., Stoyneshka, I., & Rose, T. (2004). L2 perception and production of the English regular past: Evidence of phonological effects. In A. Brugos, L. Micciulla, & C. Smith (Eds.), Proceedings of the 28th Annual Boston University Conference on Language Acquisition, 553-564. Somerville, MA: Cascadilla Press.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In Bygate, M., Skehan, P., & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 99–118). Harlow: Longman.

Thompson, T., & Gaddes, M. (2005). The importance of teaching pronunciation to adult learners. *Asian EFL Journal*, *2*, 1-5.

Tight, D. G. (2017). Tool usage and effectiveness among L2 Spanish computer writers. *ELIA: Estudios de Lingüística Inglesa Aplicada*, *17*, 157-182.

Van Lieshout, C., & Cardoso, W. (in press). Google Translate as a tool for self-directed language learning. *Language Learning & Technology*.

Van Tol, C. (2016). *The global English language job market is changing: what this means for ESL teachers*. Medium. Retrieved 7 October 2020, from

https://medium.com/accelerated/the-global-english-language-job-market-is-changing-what-this-means-for-esl-teachers-92c8d87cf8eb.

Wahid, R., & Sulong, S. (2013). The Gap Between Research and Practice in the Teaching of English Pronunciation: Insights from Teachers' Beliefs and Practices. *World Applied Sciences Journal*, *21*, 133-142. https://doi.org/10.5829/idosi.wasj.2013.21.sltl.2147

Wang, Y., & Young, S. (2015). Effectiveness of feedback for enhancing English pronunciation in an ASR-based CALL system. *Journal of Computer Assisted Learning*, *31*(6), 493-504. https://doi.org/10.1111/jcal.12079

Wu, W. (2009). Criteria for Establishing an Authentic EFL Learning Environment in Taiwan. *The Asian EFL Journal*, *11*(3), 156-189.

Xiong, T., & Xiong, X. (2017). The EFL Teachers' Perceptions of Teacher Identity: A Survey of Zhuangang and Non-zhuangang Primary School Teachers in China. *English Language Teaching*, *10*(4), 100. https://doi.org/10.5539/elt.v10n4p100

Yu, L., Song, Q., & Miao, J. (2019). A Study on the Problems and Countermeasures of Oral English Teaching in Rural Junior Middle Schools under the Background of Man-machine Dialogue Examination in China. *Theory and Practice in Language Studies*, *9*(7), 810. https://doi.org/10.17507/tpls.0907.09

Zec, D. (1995). Sonority constraints on syllable structure. *Phonology, 12*, 85-129.

Zhang, F., & Yin, P. (2009). A Study of Pronunciation Problems of English Learners in China. *Asian Social Science*, *5*(6), 141-146. https://doi.org/10.5539/ass.v5n6p141

Appendices

Appendix A

Background questionnaire (背景调查)

Language Background (语言背景)

1. First name (名): _____ Last name (姓): _____
2. Gender (性别): Male (男) Female (女) Others (其他)
3. Age (年龄): _____
4. Do you have any hearing problems or speech disorder? Yes No
   你有听力障碍或语言障碍吗？ 是 否
5. Have you learned English grammar: Yes No A little
   你曾经学习过英语语法吗？ 有 没有 学过一些
6. Have you had any extracurricular English training outside school? Yes No
   你参加过英语课外辅导班吗？ 有 没有
7. Do you have knowledge of other languages other than Mandarin and English?
   除中文和英语外，你对其他语言有了解吗（如：会说，会写，或有点了解）
   Yes (是的): _____ No (没有)
8. If yes, by what means did you acquire the language(s)?
   如果是，你是怎么接触或学习这（些）语言的呢？
   Self-directed leaning (自学) Professional training (专业学习) Other (其他): _____

Online Translators Use (在线翻译的使用)

1. Have you ever used online translators such as Microsoft Translator and Google Translate in your daily life? Yes No
   你曾经在日常使用过像微软翻译或谷歌翻译此类的在线翻译软件吗？ 是 否
2. If yes, which ones do you usually use? _____
   如果用过，你经常使用的是哪个翻译器？
3. For what purpose do you use them? (you can circle more than one)
   你使用它们的目的是什么？（你可以选择多个目的）
   Text translation Pronunciation Speech recognition Other: _____
   翻译单词/文本 学习发音 语音识别功能 其他: _____
4. How often do you use the online translators?
   你使用在线翻译的频率是什么？
   Always（总是）Usually（通常）Frequently（较为频繁）Sometimes（偶尔）Rarely（很少）
5. Do you think the online translators are reliable sources of your language learning?
   在你看来，这些在线翻译对你平时的语言学习可靠/精准吗？
   Yes（可靠） No（不可靠） Neutral (not sure)（一般/不清楚）

Appendix B

Short interview about phonological awareness

1. What do you know about the pronunciation of English past tense -ed in words such as “walked”, “played” and “visited”?

   关于英语过去式 “-ed”，比如它在 “walked”, “played” 和 “visited”词中的发音你有什么了解？

   _____

2. Does the pronunciation of English past tense -ed always sound the same in different verbs?

   英语过去式 “-ed” 在不同英文动词里的发音是相同的吗？

   Yes (相同) No (不同)

3. If they sound different, can you explain how they sound? List all pronunciations you think -ed has:

   如果你认为在不同词中“-ed”发音不同，那么请列举出你认为它所有可能的发音:

   _____

(Participants’ overall awareness level: 0 = not aware, 1 = partially aware, 2 = fully aware)

Appendix C

ABX discrimination test:

**The highlights are the correct answers.*

	A	B	X
1	kicked	happened	bored
2	spotted	asked	connected
3	treated	ordered	peeled
4	suited	refused	reflected
5	pointed	walked	skipped
6	pleased	balanced	booked
7	failed	flapped	burned
8	faced	flooded	handed
9	cured	guided	closed
10	answered	counted	attracted
11	boxed	floated	hopped
12	reached	killed	talked
13	alerted	cycled	excused
14	improved	rushed	wiped
15	rotted	faxed	parted

Students' Answer:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Appendix D

Aural Discrimination Answer Sheet

In this last task, you will listen to 16 sentences. The sound target you will be focusing on is the past tense *-ed.* This sound can take one of three forms: 此任务中你将听到 16 个句子，请重点听每句话里表示过去式的 "-ed", 它的发音会是以下三者之一:

1. /t/ as in *walked*
2. /d/ as in *played*
3. /ɪd/as in *waited*

When listening to these sentences, please listen carefully and mark either PAST or NOT PAST. If you mark PAST, please also circle which of the three sounds (t, d or ɪd) that you heard. Let's practice! 认真听这些句子并且判断每个句子是过去时态 (PAST) 还是非过去时态 (NOT PAST), 若你认为是过去时态, 请圈出你听到的动词后 "-ed" 的发音 (t, d 或 ɪd). 让我们先来一起练习一下:

Practice. Please circle whether you heard the past tense *-ed* sound or not. (圈出答案)

1	PAST vs. NOT PAST	/t/ /d/ /ɪd/
2	PAST vs. NOT PAST	/t/ /d/ /ɪd/

Let's start:

(1) Circle whether you heard the past tense *-ed* sound or not.

(2) If you heard the past, indicate which sound you heard: t, d, or ɪd.

1	PAST vs. NOT PAST	/t/ /d/ /ɪd/
2	PAST vs. NOT PAST	/t/ /d/ /ɪd/
3	PAST vs. NOT PAST	/t/ /d/ /ɪd/
4	PAST vs. NOT PAST	/t/ /d/ /ɪd/
5	PAST vs. NOT PAST	/t/ /d/ /ɪd/

6	PAST vs. NOT PAST	/t/	/d/	/ɪd/
7	PAST vs. NOT PAST	/t/	/d/	/ɪd/
8	PAST vs. NOT PAST	/t/	/d/	/ɪd/
9	PAST vs. NOT PAST	/t/	/d/	/ɪd/
10	PAST vs. NOT PAST	/t/	/d/	/ɪd/
11	PAST vs. NOT PAST	/t/	/d/	/ɪd/
12	PAST vs. NOT PAST	/t/	/d/	/ɪd/
13	PAST vs. NOT PAST	/t/	/d/	/ɪd/
14	PAST vs. NOT PAST	/t/	/d/	/ɪd/
15	PAST vs. NOT PAST	/t/	/d/	/ɪd/
16	PAST vs. NOT PAST	/t/	/d/	/ɪd/

Aural Discrimination Transcripts of Target Sentences

Practice 1: I ordered a large pizza.

Practice 2: I water my garden.

1. I called my mother.
2. I visit my cousin Sam.
3. I painted some pictures.
4. I grilled many hamburgers.
5. I corrected my math homework.
6. I jumped in the fre ezing lake in winter.
7. I study English for 4 hours.
8. I invited him to dinner.
9. I finish my homework at 9pm.
10. I receive many presents on my birthday.
11. I opened the door for her.
12. I fixed the problems around the house.
13. I loved the movie.
14. I danced to the music.
15. I waited two hours for my friend.
16. I talked with Jeff in the hallway.

Appendix E

Production test: Read aloud (请读出这些词**)**

1. painted
2. picked
3. lived
4. passed
5. answered
6. attempted
7. punched
8. added
9. slapped
10. expected
11. liked
12. followed
13. guided
14. washed
15. moved
16. helped
17. divided
18. laughed
19. arrested
20. dragged
21. stayed
22. ended
23. smelled
24. noted
25. asked
26. needed
27. plugged
28. filmed
29. sniffed
30. analyzed

Appendix F

Production test: Interview

What did Kevin do during his last summer vacation?

	YES	NO
Did he play soccer?	✓	

A: **Yes**, he played soccer

	YES	NO
Did he play soccer?		✓ (baseball)

A: **No**, he played baseball

	YES	NO
Did he walk to the airport?	✓	
Did he travel by himself?		✓ (with friends)
Did he hate the weather ?	✓	
Did he visit his girlfriend?		✓ (parents)
Did he kiss his girlfriend?	✓	
Did he hug his girlfriend?		✓ (friends)
Did he check his e-mail regularly?	✓	
Did he learn French?		✓ (Russian)
Did he taste good wines?	✓	
Did he extend his vacation?	✓	
Did he work during his vacation?	✓	
Did he enjoy his vacation?	✓	

Appendix G

Short Story Exercises

Short story #1 The nasty parrot (full version)

Last summer, Jimmy received a present, a parrot! However, Jimmy noticed the parrot say some bad words.

Jimmy tried to change the bird. He thought that if he was nice to the parrot, the parrot would be nice too. So, he talked to the bird very politely, but nothing worked! He grabbed the bird and shook him, but the bird didn't stop using bad language. Finally, Jimmy decided to punish the bird. He opened the cage and the parrot jumped onto Jimmy's hand. Jimmy carried his parrot and put it in the fridge. The bird still screamed bad words and lasted for a few minutes. Then suddenly, it was completely quiet.

Jimmy was scared and opened the fridge door. The parrot stepped back and said, "Sorry that I offended you with my bad language."

Jimmy was surprised, he did not know why the parrot regretted and stopped saying bad words. Then the parrot pointed inside the fridge to the frozen chicken and said, "May I ask what the chicken did wrong?"

**15 targets (5 for each allomorph) and 5 distractors.*

Activity 1: Comprehension questions (阅读理解)

Listen to the story using Natural Reader and answer the following questions. You may answer in Chinese or English (do not worry about orthography or grammar). Provide short answers.

1) What present did Jimmy receive? (Jimmy 收到了什么礼物？)

2) Was Jimmy's parrot nice? (Jimmy 的鹦鹉友善/好吗？)

3) Why was the parrot nasty? (为什么说这个鹦鹉很粗鲁？)

4) What did Jimmy do to the parrot? (Jimmy 对这只鹦鹉做了什么？)

5) Why did the parrot stop being nasty? (为什么鹦鹉后面变的不粗鲁了？)

Activity 2: Fill in blanks

Last summer, Jimmy _____ a present, a parrot! However, Jimmy _____ the parrot says some bad _____.

Jimmy tried to change the bird. He thought that if he was _____ to the parrot, the parrot would be nice too. So, he talked to the _____ very politely, but nothing _____! He _____ the bird and shook him, but the bird didn't stop using bad language. Finally, Jimmy _____ to punish the bird. He opened the cage and the parrot _____ onto Jimmy's hand. Jimmy _____ his parrot and put it in the fridge. The bird still _____ bad words and _____ for a few minutes. Then suddenly, it was completely quiet.

Jimmy was scared and _____ the fridge door. The parrot _____ back and said, "Sorry that I _____ you with my bad _____."

Jimmy was surprised, he did not know why the parrot _____ and _____ saying bad words. Then the parrot _____ inside the fridge to the frozen _____ and said, "May I ask what the chicken did wrong?"

Activity 3: How do they sound?

The –ED sounds like:

/t/	/d/	/ɪd/

Short story #2 Baby food (full version)

A four-year-old boy named Joe was at the doctor's. He waited for his mother. He watched the clock on the wall, he was bored. Then he saw a pregnant woman, Joe stopped counting, waited some seconds, and walked to the chair where the woman was sitting. He asked, "Why is your stomach so big?"

The woman replied with a laugh, "Because I'm having a baby."

Joe was surprised and asked, "Is the baby in your stomach?"

"Of course!" said the woman. She grabbed the boy's hand and put it on her stomach, "Can you feel the baby kick?" Joe felt something moving and pulled back his hand.

"But is it a good baby?" Joe questioned.

"I'm sure it's a really good baby," added the woman. "I am sure this baby will become a good boy like you!" She repeated: "Just like you".

Joe moved back and asked, "If he is a good baby, why did you eat him?"

**12 targets (4 for each allomorph) and 4 distractors.*

Activity 1: Comprehension questions (阅读理解)

Listen to the story using Natural Reader and answer the following questions. You may answer in Chinese or English (do not worry about orthography or grammar). Provide short answers.

1) Where is Joe? (Joe 在哪里？)

2) Who is Joe talking to? (Joe 在和谁说话？)

3) Why is the woman at the doctor's? (为什么这个女人在看医生？)

4) Does the woman like Joe? (这个女人喜欢 Joe 吗？)

Activity 2: Fill in blanks

A four-year-old boy named Joe was at the doctor's. He _____ for his mother. He _____ the _____ on the wall, he was bored. Then he saw a pregnant woman, Joe _____ counting, _____ some seconds, and _____ to the chair where the woman was sitting. He asked, "Why is your stomach so big?"

The woman _____ with a laugh, "Because I'm having a baby."

Joe was surprised and _____, "Is the baby in your stomach?"

"Of course!" said the woman. She _____ the boy's hand and put it on her stomach, "Can you feel the baby kick?" Joe felt something moving and _____ back his hand.

"But is it a good baby?" Joe _____.

"I'm sure it's a really good baby," _____ the woman. "I am sure this _____ will _____ a good boy like you!" She _____: "Just like you".

Joe moved back and asked, "If he is a good baby, why did you _____ him?"

Activity 3: How do they sound?

The –ED sounds like:

/t/	/d/	/ɪd/

Appendix H

Interview questions

1. How was your **overall experience** of using Microsoft Translator? Did you enjoy using it?
2. What was the characteristic or feature that you **like** the most about Microsoft Translator?
3. What was the characteristic or feature that you **do not like** about Microsoft Translator?
4. How did you find the **naturalness, accuracy and intelligibility** of Microsoft Translator?
5. Were you able to **notice the subtle difference** between your pronunciations (inaccurate ones) and the correct pronunciations of Microsoft Translator?
6. Were you able to **recognize the errors** you have made and **correct** yourself via the feedback (via spelling) you received from Microsoft Translator?
7. Did you feel **motivated** when using Microsoft Translator to learn by yourself? Why?
8. Did you make **progress** of pronouncing English past tense -ed after this learning experience?
9. Do you think Microsoft Translator could be a **useful source** of your language learning outside the classroom?
10. Would you consider using **Microsoft Translator** as a **learning tool** in your **future** language learning?