A CASE STUDY ON IMPLICIT AND EXPLICIT ONLINE PRONUNCIATION TUTORING TO CHINESE INTERNATIONAL GRADUATE STUDENTS

BY

XINLEI CHEN

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Arts in Teaching of English as a Second Language in the Graduate College of the University of Illinois Urbana-Champaign, 2021

Urbana, Illinois

Advisers:

Teaching Assistant Professor Suzanne Carla Franks
Visiting Assistant Professor Lisa Renee Pierce
ABSTRACT

Pronunciation matters for the success of second language learners, and yet it does not receive enough attention in the language classroom (O’Brien, 2004; Sturm, 2013; Wei, 2006). Especially for the students who received their secondary and undergraduate education in China, where the negative washback from exam-driven English teaching styles is still prominent. This study compared the progress of two Chinese graduate students in a Midwestern university, one received one-on-one implicit instruction and another explicit instruction online. The measurements of progress in this case study include segmental perception and production, fluency, and intelligibility rate in both controlled and spontaneous speech. Both participants received eight weekly tutoring sessions in a semester, with one hour per session. All participants took a pre-test and a post-test the weeks before and after the instructions. The tests aimed to get and examine vowels /ʊ, /u/, /i:/, /ɪ/ and consonants /v/, /w/, /s/, /z/ and labiodental consonant, as well as temporal variables including words per minute (WPM) and filled-pause ratio (FPR). The results found that the explicit participant showed a slight improvement of their segmental perception and production of some back vowels and the labiodental consonant, as well as fluency and intelligibility rate, especially in spontaneous speech. But both instructed participants reported an elevated confidence in spoken English after the treatment.
ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Suzanne Franks, for her patient guidance and thorough support throughout my research process, who helped me start to understand and truly fall in love with English phonology. I would not have been able to pull this research off if it were not for her consistent help and teaching. Furthermore, I would like to express my deepest gratitude to Dr. Lisa Pierce. Her professional knowledge and care for students really motivated me to re-examine myself as an educator.

Also, to my parents and Sean, people who have always been there for me during my dark and bright moments, unconditionally.
TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ...........................................................................................................1

1.1 Relevance of this Study .............................................................................................................1

1.2 Personal Inspiration to this Research ....................................................................................3

CHAPTER 2: LITERATURE REVIEW ..............................................................................................8

2.1 English Pronunciation Instructions in ESL/EFL Contexts .....................................................8

2.2 Existing Research on Implicit and Explicit Instructions ..........................................................12

2.3 Measurements of Speech Quality in Previous Research .......................................................17

CHAPTER 3: METHODOLOGY .....................................................................................................20

3.1 Research Questions ................................................................................................................20

3.2 Participants .............................................................................................................................21

3.3 Materials ................................................................................................................................23

3.4 Data Collection and Procedures ............................................................................................29

3.5 Data Analysis .........................................................................................................................30

CHAPTER 4: RESULTS ................................................................................................................32

4.1 Pre-test ..................................................................................................................................32

4.2 Post-test ..................................................................................................................................36

4.3 Vowel Analysis .......................................................................................................................39

4.4 Consonant Analysis ................................................................................................................44

4.5 Fluency Analysis .....................................................................................................................49

4.6 Filled Pause Ration Analysis ................................................................................................51

4.7 Intelligibility Rate Analysis .....................................................................................................51
<table>
<thead>
<tr>
<th>Chapter/Appendix</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>CHAPTER 5: DISCUSSION</td>
<td>56</td>
</tr>
<tr>
<td>5.1</td>
<td>5.1 Results and Research Questions</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>5.2 Limitations</td>
<td>58</td>
</tr>
<tr>
<td>5.3</td>
<td>5.3 Conclusions</td>
<td>60</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>REFERENCES</td>
<td>62</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>APPENDIX A: INSTITUTIONAL REVIEW BOARD APPROVAL LETTERS</td>
<td>67</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>APPENDIX B: RECRUITMENT MATERIALS</td>
<td>68</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>APPENDIX C: CONSENT FORMS</td>
<td>74</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>APPENDIX D: SAMPLE LESSON PLANS</td>
<td>78</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>APPENDIX E: PRE-TEST AND POST-TEST</td>
<td>82</td>
</tr>
<tr>
<td>APPENDIX F</td>
<td>APPENDIX F: INSTRUCTION SCHEDULE</td>
<td>86</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 RELEVANCE OF THIS STUDY

English has come to be the most spoken language worldwide, with an estimated number of 300-400 million second language users and 500-700 million foreign language users (Crystal, 2000). These numbers combined have exceeded the number of native-English speakers (NESs), which was around 350-450 million in 2000 (Crystal).

For the past few decades, the impact of explicit and implicit teaching on learners’ grammar knowledge has been examined (Nazari, 2012) and (Krashen, 1981). But studies focusing on pronunciation (intelligibility and fluency), especially among the fastest growing EFL population – Mandarin speaking English learners, are still lacking. Current studies on Second Language Acquisition also mostly focus on face-to-face instruction, especially regarding the comparison of explicit and implicit teaching. (Toth, 2006); (Hossenini and Pourghasemian, 2019).

Pronunciation clearly matters for second language learners’ success, and yet it does not receive enough attention in the language classroom (O’Brien, 2004; Sturm, 2013; Wei, 2006). This is especially true for students who received their secondary and undergraduate education in Asian countries, where most of the classrooms are adopting a behaviorist approach focusing on exam scores toward English language education (Shih, 1999). Among them, China has been adopting such an approach on a large scale.

There are three major characteristics the behaviorist approach has that can negatively impact English learners in China, impairing their pronunciation learning, in particular. The characteristics of this approach are as follows:
(1) demanding little cognitive effort

(2) requiring grammar-translation method

(3) negative washback.

These may explain the difficulties encountered by ESL students especially when considered in light of Craig’s (1999) communicative model where communication is an active information exchange process between the sender (speaker) and the receiver (listener). For international students in the U.S., they are more likely to feel that their communication in English is being jeopardized by a lack of background cultural and linguistic knowledge, heightened anxiety, as well as unintelligibility caused by accent and delivery.

**Figure 1.** Numbers of International Students at the U.S. Universities

With one-seventh of the population in the world, China’s approach to English education has the potential to have a large impact. As **Figure 1** above demonstrates, there has been a steady and huge growth in the Chinese international student population at U.S. universities, ranking
them first among all international students, according to the Institute of International Education (2019). Therefore, pronunciation teaching targeting Chinese international students has become more relevant than ever. This issue is especially salient for Chinese graduate students because of their lack of prior pronunciation and speaking preparation in China and the student-centered, communicative classrooms. This could likely cause them to fail to meet the elevated demand for oral English competency due to various assistantships, conference presentations, and other scholarly demands.

Another factor that contributes to the importance of this study is the unpredictability for language teaching brought by the 2020 global pandemic as well as other world affairs. Because of the lockdown and limited access to international travel, synchronous computer-mediated communication (SCMC) has come to the center of attention as a trend in education formats. However, the current studies on SLA also mostly deal with face-to-face instructions, especially when comparing explicit and implicit teaching: Toth (2006), Hossenini & Pourghasemian (2019). Therefore, this case study will provide potential new insights for this new online educational trend.

1.2 PERSONAL INSPIRATION TO THIS RESEARCH

Being a primarily self-acquired bilingual English speaker, my lifelong fascination with language learning began with my own English acquisition in informal English as a Foreign Language (EFL) context in Beijing and Guangzhou, where I respectively grew up and finished my undergraduate study. While reflecting upon my own continuing oral English learning journey after adulthood, I have found my interest lies in the controversies surrounding whether neurons are still being born in the hippocampus every day among adult brains (Spalding et al., 2013),
thus the learning of novel subjects, including a new language, can still occur even with aging and lateralization. During my two years of teaching assistantship which gave me access to adult ESL students and professionals from over 10 countries, I’ve found that the confidence and comfort level with English speaking can sometimes play a bigger role in these students’ language acquisition than the learning of pronunciation rules per se. Those who felt more relaxed and thought they held more agency during the conversations tended to show higher levels of engagement and thus get more opportunities to produce oral English. Consequently, these individuals usually demonstrated higher oral English proficiency. Unfortunately, the opposite is true. For students who have not grown up with English as a part of their language identity, but rather a mere subject to study in school, their confidence in speaking the language tends to be quite low.

Due to historical, political, and socio-cultural reasons, English did not become a major subject in the Chinese public-school system until recent decades in mainland China. And even after taking English courses from an early age, Chinese students are mainly taught using a grammar-translation method which can offer negative washback likely due to the exam-driven English curriculum design.

Growing up in mainland China myself, I witnessed first-handed the resentment and fear of English learning, especially spoken English, among most Chinese students. Because Chinese public school curriculum design has been dictated by the GaoKao exam, which is the annual college entrance exam that has been deemed as the only means to change most students’ life trajectories, spoken English has been overlooked in English classes due to an absence of an oral English exam in the GaoKao. Therefore, even though the majority of Chinese students have received English education from an early age, their pronunciation and fluency are never formally
addressed. Being a member of this public education system in China, I have stopped following the instructions given by my English teachers since I was 16. I had to choose my own alternative English learning methods in order to acquire a language that could serve a more practical communicative purpose and help with cross-cultural understanding.

In retrospect, this “rebellious” decision I made about my English learning was exactly the reason why I was able to excel at it compared to my high school peers and begin the journey of becoming an English teacher.

Many US universities require that international graduate students demonstrate a high level of oral proficiency before they can become teaching assistants. Scores on standardized tests are frequently used for this purpose. Other universities have developed internal tests that are offered as an alternative way to demonstrate oral English proficiency. The English Proficiency Interview (EPI) is an example of a university-developed oral English test developed at a large Midwestern university. Since this test is offered at no cost to students, the testing unit has established some rules for test re-takers.

When I became an EPI (English Proficiency Interview) tutor for university graduate students, I learned that students who had not received a passing score on the EPI were required to engage in intensive 8-week-long tutoring to be allowed to retake the test. Having observed a tutee’s progress, I started to wonder how effective this type of short-term customized explicit instruction was, compared to an implicit instruction counterpart. Hence the idea for this case study was incubated. I wanted to discover how pronunciation tutors could help the biggest group of international students. To narrow the gap between their current English-speaking level and their desired oral proficiency.
Although the Critical Period Hypothesis (CPH) predicts little possibility for adult ESL learners to achieve native or near-native pronunciation through any kind of pronunciation training, there are still unlimited potentials in teaching English pronunciation to college-level adults because of traits that are specific to language learners of this age group. These traits could include higher motivation and self-monitoring abilities.

According to Vitanova and Miller (2002), students were excited about their improvement in segmentals, supra-segmentals, self-monitoring, and self-correcting abilities. Improvement is important and attainable even though native-speaker-like pronunciation may be impossible after a certain age. So, while younger learners may have the advantage of being able to achieve such pronunciation proficiency, adult learners are able to use their cognitive abilities to improve through self-monitoring and self-correction (Thompson & Gaddes, 2004).

Therefore, I hope to create lessons for adult ESL learners to potentially restructure their understandings of English learning, especially pronunciation at the least. I want to help them gain confidence using English to convey their ideas to the world. My aim is to help them realize that their intelligibility can be improved by effective pronunciation teaching, and that pronunciation can improve via the gradual intuitive changes brought about by real and natural interaction with English speakers.

I am aware that the biggest challenge would be helping students eliminate insecurities and self-consciousness about possible judgments of their accents and providing actionable feedback to their oral speech productions. As challenging as it might be, I still believe there is an attainable equilibrium.

Just as Fraser (1999) described, it is commonly believed that it is offensive to prescribe an 'accent norm' to which learners must assimilate, and it is true that people should be free to
express themselves in whatever accent they choose - but it is not true that this freedom is given by withholding pronunciation teaching. On the contrary, it is effective pronunciation teaching that offers learners a genuine choice in how they express themselves.

My motivation for this case study is to lay a foundation for potential future extended research that can help me further examine adult pronunciation training and develop pedagogical models that can best suit these ESL/EFL speakers’ needs for comprehensible oral production.
CHAPTER 2: LITERATURE REVIEW

This chapter will examine and summarize the existing literature from the three following areas that are relevant to this case study: a review of the history of English pronunciation instruction, implicit and explicit instructions, and measurement of speech quality.

2.1 ENGLISH PRONUNCIATION INSTRUCTIONS IN ESL/EFL CONTEXTS

In English as a second/foreign language (ESL/EFL) learning contexts, whether to improve and how to improve pronunciation have always been two critical and long-debated issues. Based on a chronological overview of the English pronunciation instruction, teaching methods began with the behaviorist theory in the 1960s. Proponents of this theory believed learning was the establishment of habits as a result of reinforcement and rewards (Rivers, 1968, p. 73), which served as the foundation to the audiolingual method that quickly gained momentum in the U.S. and Canada for foreign language teaching, and later other parts of the world including Asian countries such as China where it persists even today. Then the audiolingual approach lost favor in the late 1960s and early 1970s, mainly because it was challenged by Noam Chomsky (1959) for its unsound theoretical basis. Additionally, according to Hadley (2001), a number of language teachers and students experienced frustration with such an approach in foreign language teaching and learning. The Audiolingual Method for pronunciation instruction was not ideal due to the native principle on which it was based, and the belief that language learning is simple habit formation. It was presumed that non-native speakers could and should achieve native-like pronunciation through endless repetition with no need for explicit instruction. This idea remains permeant despite what more current research has proclaimed.
Meanwhile, in the late 1970s and early 1980s, a new *Natural Approach* developed by Krashen and Terrel (1983), along with the *Communicative Approach* (Canale & Swain 1980; Savignon, 1997), was brought to attention and started to replace the ubiquitous *Audiolingualism* because they emphasized realistic and meaningful input, and interactions (Jones, 1997). However, these new models also ignored the significance of explicit pronunciation teaching and led to diminishing interest in the method (Breitkreutz et al., 2001; Derwing & Munro, 2005; Jones, 1997). This was due to the premises that accompanied these approaches: 1. Non-children second language learners can not achieve native-like pronunciation because of the *Critical Period Hypothesis*. 2. Language pronunciation is acquired, not taught (Krashen, 1977). This general attitude towards pronunciation instruction resulted in the omission of this subject in research, teacher training programs, curriculum designs. It also ignored the communicative needs of ESL/EFL students (Munro & Derwing, 1999; Morley, 1991; Pennington & Richards, 1986).

This Native Speaker (NS)-based view on pronunciation lasted until the late 1990s and early 2000s, when a growing number of English instruction experts started to re-examine the models and goals for pronunciation teaching. Given the fact that few L2 speakers would have native-like accents (Ortega, 2009; Derwing & Munro, 2005), Munro and Derwing (1995) promoted Intelligibility in relation to Comprehensibility and Accentedness, as a foundation for an adjusted pronunciation pedagogy. Jenkins (2000, 2002) examined the matter of intelligibility through the lens of L2 English speakers who used English as an international *lingua franca* (Seidlhofer, 2001, p. 146), thus affirming that the correlation between comprehensibility and accentedness was inconsequential. This was further reinforced by Levis’s (2005) observation “communication can be remarkably successful when foreign accents are noticeable or even
strong” (p. 370). Such studies have helped place an emphasis on *Intelligibility* instead of previous NS-based pronunciation models.

Since the early 2000s, more attention has been given to NNEs due to the massive demand for English learning. As mentioned by Derwing & Munro (2005), students were often misguided by the instructions they received in ESL/EFL classes on pronunciation, focusing on targets and techniques of little practical use, such as memorizing speech templates by rote. This situation usually did more harm than good to ESL students that were studying in English-speaking countries. Meanwhile, just in U.S. universities alone, there has been an increasing in annual international student enrollment which exceeded one million during the 2019-20 academic year, as shown in Figure 2.

**Figure 2.** Student Enrollment at U.S. Universities

As O’Brien (2004) pointed out, pronunciation might be the most important element of oral communication and has a key role in learning English (Yates and Zielinski, 2009). However, as much as pronunciation is an integral part of these international students’ study and a key to their academic performance and career trajectories, it is often not taught due to inadequate teacher training, lack of resources, and NNES teachers’ lack of confidence in their ability to teach it (Murphy, 2017). Also, there were still lingering beliefs about NS-based teaching models, as noted by Morley (1994, cited by P. Gilakjani, 2012), who suggested that learners should try to get close to the standard accents because if their accent is too different there might be confusion.

Facing the aforementioned challenges and debates, researchers in the recent decade have shifted their focus from attaining a native or near-native pronunciation with an over-emphasize on segmentals to the improvement of intelligibility and comprehensibility (Kanellow, 2009; Munro & Derwing, 2011), which is more supra-segmental-orientated. This was largely due to the realization that too many L2 learners started their target language learning in adolescence or adulthood. This would be too late a start for learners to be “native”. Regardless, there is ample evidence that a large percentage of ESL/EFL learners possess that goal (Derwing, 2003; Kang, 2010; Scales, Wennerstrom, Richard, & Wu, 2006; Timmis, 2002). Language educators have come to realize that most L2 students will unlikely sound like native speakers despite the fact that good pronunciation plays a crucial role in their success in language learning (O’Brien, 2004). Therefore, in modern pronunciation classes, helping students improve their pronunciation should not simply equal getting rid out of their foreign accent, but rather assisting them to increase the understandability of their speech.

As proposed by Shahzada (2012), learners’ needs and problems in pronunciation are paramount for English instructors in order to have successful pronunciation classes. Instructors
are expected to prepare materials that are appropriate to their students’ unique needs and issues. To ensure the use of authentic materials in pronunciation class that promotes communication skills, Kolokdaragh (2010) suggested that instructors should incorporate computer technologies in class, which works well as a part of the overarching guidance for this paper. This suggestion was proven to be helpful in the past decade as the new media era requires most English courses to be conducted in computer labs. And it later became essential to online language teaching when the 2020 global pandemic hit the field of education.

2.2 EXISTING RESEARCH ON IMPLICIT AND EXPLICIT INSTRUCTIONS

There has been a considerable amount of research on implicit and explicit teaching and learning in second language education in the past few decades. Reber (1993) first provided a definition for implicit learning as “the learning of knowledge that does not involve conscious attempts to obtain knowledge.” (p. 5). Then N. Ellis (1994) further stated that implicit learning is the “acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operations” (p. 1). Dörnyei (2009) and R. Ellis (2009) later corroborated such a view on the automaticity of implicit learning.

In 2009, R. Ellis inferred that just as the manner in which L1 speakers were able to automatically retrieve implicit knowledge for spontaneous production, L2 learners also resort to their implicit knowledge when making default L2 productions. However, they usually turn to their explicit knowledge whenever they have problems executing an online task solely based on their implicit knowledge because of a shortage of the latter.
The difference between implicit and explicit learning regarding processing and accessibility is that the former is accessible via natural procedural processing, whereas the latter requires strategic learning with planned structures prepared in advance (Dörnyei, 2009). Furthermore, there were some debates on the relationship between these two types of knowledge. The first view suggested that there was no interface between these two kinds of knowledge thus one can not convert into another directly, therefore implicit instructions should only serve to fulfill the attainment of more implicit knowledge, which means that instructions that exercise explicit metalinguistic knowledge in teacher-centered learning should be avoided (Krashen, 1981). And therefore, this implicit instruction would be beneficial for students’ L2 acquisition in the long run as implicit knowledge is the primary indicator of language competence in L1 acquisitions (Chomsky, 1965).

However, a strong interface stance by scholars such as Dekeyser (1998) argued that even though these two systems have distinct characteristics, learned explicit knowledge, such as grammatical regularities, can be converted to implicit representation. And such conversion is viable through communicative practice and employment. Conversely, via conscious reflection and analysis of implicit production, explicit knowledge can be gained in the process. Although there was not a clear and precise framework to support such claims (Dörnyei, 2009), Crowell (2004) suggested that such transmissions between these two systems could be made possible with repeated exposure to the declarative network (DN) and such declarative knowledge would be proceduralized over time.

Finally, a neural view believed that there is a weak interface between implicit and explicit knowledge, which is the theoretical framework that this paper tends to be based on. This viewpoint proposed that the transmissions are possible but heavily rely on the learner’s
psychological development. Ellis (2009) mentioned that explicit knowledge needs controlled and well-planned instructions to enable the learning of implicit knowledge, but one does not automatically trigger the other. It was through the explicit instructions that sub-tasks which promote implicit learning were created, which inadvertently leads to implicit knowledge and likewise for the conversion from implicit learning to explicit learning (Dörnyei, 2009). As Dekeyser (2003) noted, explicit knowledge can be considered as “procedural knowledge that is functionally equivalent to implicitly acquired knowledge” (p. 329). This weak interface position provided excellent pedagogical implications for practical lesson planning and grammar class instruction.

One caveat worth noting is that implicit and explicit learning is closely related to but not identical to their counterpart for instruction. Ellis (2009) elucidated that implicit or explicit learning refers to learners’ standpoint, while implicit or explicit instruction can only be discussed from an external perspective (such as instructors or textbooks). In a similar vein, Housen and Pierrard (2006) defined language instruction, “any systematic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/or the conditions under which these occur” (p. 2). The nature of the instruction can influence learning differently to a certain degree, but it does not always play an absolute determining role. Consequently, one’s presence does not guarantee the occurrence of the others, and in this case, individual learner differences contributed a lot to the learning style and learning outcome resulted from either instruction.

All things considered, more research is still needed to explore the effectiveness of each instruction and thus make comparisons that are not influenced by learner variables. Based on the existing literature, implicit instructions can create opportunities for learners to unconsciously
internalize the target language and thus develop metalinguistic awareness of the rules (Ellis, 2009). Such instruction emphasizes communication while attention to linguistic forms is incidental (Housen & Pierrard, 2005). Explicit instruction stresses Focus-on-Forms (FonFs) approach (Long, 2001), where attention to linguistic forms is deliberate and not the result of organic communication. However, these discussions about implicit and explicit instruction pertaining to SLA were mostly based on the acquisition of grammar and not pronunciation (Dlaska & Krekeler, 2013), the latter has been marginalized in the modern Communicative Language Teaching (CLT) classroom (Peltekov, 2017). A popular view on pronunciation acquisition since the late 1970s and early 1980s, even still prevalent today in many language learning settings, is that a year spent in the target language country can invariably enhance one’s pronunciation. Be that as it may, Derwing and Munro (2014) argued that there is no attestation to prove such a stance is true to a significant extent solely by language exposure (p. 47). They believed that explicit pronunciation instruction is indispensable to one’s success in improving their oral production. While this stands true, it is also imperative to differentiate between mere exposure to the target language and the enriched input and well-structured practice offered by implicit instruction (Peltekov, 2017).

Recently, there has been new research on the comparison between implicit and explicit instructions for pronunciation. An article published by Medina Riveros, (2009) on a communicative approach of pronunciation tutoring for a procedural framework. This study described the interaction taking place between a tutor and her students during the online tutoring sessions (OTS) in the ALEX Virtual English Program at the Universidad Nacional de Colombia. In Bogotá Students’ surveys, chat transcriptions and tutor’s journals were used to collect data. The results showed that the methodology of this course provided a unique opportunity for
language learning through conversational topics of social and personal issues, which were prevalent in the class sessions. It was also found that the main interactions between the tutor and the learners consisted of two-way dynamics. Such results demonstrated successful implicit instruction through SCMC online learning.

Peltekov (2017) reported on the effectiveness of implicit and explicit instruction on German L2 learners’ pronunciation, further discussing the effect of explicit and implicit teaching on L2 learners’ pronunciation, not just grammatical accuracy. Peltekov (2017) examined the effects of implicit and explicit instruction on the pronunciation of beginning learners of German. Over the course of one semester, one group of learners (n=5) was taught pronunciation explicitly (i.e., using phonetic rules), another group (n=5) implicitly (i.e., without phonetic rules), and a third group (n=5) received no pronunciation instruction. A pretest-posttest design was used to measure learners’ improvement in accent and comprehensibility. A slight improvement in both variables was observed under all conditions, but no significant difference in progress was found across the three groups.

The findings suggested that some learner variables (e.g., age) might be better predictors of improvement than the type of instruction. Moreover, not all pronunciation features were equally relevant for L2 learners’ comprehensibility and accent. Although the research results indicate limited findings, the methodology of this study provided the foundation for this thesis paper to conduct a new case study which compares implicit and explicit pronunciation instructions to ESL learners to help fill in the gap for relevant research.
2.3 MEASUREMENTS OF SPEECH QUALITY IN PREVIOUS RESEARCH

2.3.1 Fluency

The measurement of fluency has been crucial in second language acquisition studies (De Jong & Van Ginkel, 1992). Tavakoli and Skehan (2005) suggested that three different aspects of fluency can be distinguished. They are breakdown fluency (number and length of pauses), speed and density per time unit (speech rate), and repair fluency (false starts and repetitions). Out of all measures, *speech rate* is judged the best predictor of subjective fluency (Cucchiarini, Strik, & Boves, 2002). And such a measure of fluency could be well predicted by the number of syllables per time unit (Kormos & Dénes, 2004).

Concerning how to measure fluency of speech, Park (2016) examined temporal variables and pausing patterns in L2 English speech to investigate fluency as a measurable component of oral proficiency. It was discussed that fluency can be defined as ‘speed and smoothness of oral delivery’. The speed of oral delivery can be measured through calculating temporal variables such as speech rate and mean syllables per run where ‘run’ is the vocal chunk between silent pauses. The smoothness of oral delivery can be measured through examination of pausing patterns by classifying the placement of pauses.

Pauses may be placed in expected positions such as clause/phrase boundaries or in unexpected positions. Pause placement in unexpected positions may reduce the smoothness of oral delivery. Park (2016) collected data via speech samples from the Oral English Proficiency Test (OEPT) and included the responses from two items, a read-aloud task (RAL) and a news passage task (NP). A total of 325 speakers across four different language groups (native speakers of Korean, Chinese, Hindi, and English) are represented across 6 proficiency levels (rated by holistic scoring based on the OEPT scale from 35 to 60).
The speech samples were transcribed manually using a computer-assisted annotation tool that allowed the capture of information about syllables, pausing boundaries, and types of pausing positions. For temporal variables, they looked at the speech rate, which is the mean syllables per run, and the number of pauses per second. They also investigated the number of unexpected pauses per second and expected pausing ratio to compare pausing patterns across proficiency levels and language backgrounds. They eventually found that there are linear relationships in temporal and pausing variables. Participants who had higher proficiency level spoke at higher rates with expected pausing patterns, and vice versa. In this thesis study, I adopted similar tools from Park’s (2016) study to measure fluency.

2.3.2 Intelligibility

There have been myriads of definitions with regards to the term *intelligibility*. This paper will only discuss the ones that are the most pertinent to the case study. Nelson (1982) gave the definition of intelligibility as “the apprehension of the message in the sense intended by the speaker” (p.63). Munro and Derwing (1995a, 1995b) further defined *objective intelligibility* as “the extent to which speaker’s intended utterance is actually understood” (p. 76, p. 291), which best befits the logistics of this study hence will be the working definition for this paper relating to the concept of intelligibility.

An important note is that accentedness does not necessarily correlate with low intelligibility, as Derwing & Munro (2006) pointed out “although some features of accent may be highly salient, they do not necessarily interfere with intelligibility” (p.184). This is relevant because this thesis paper investigates NNESs whose speech is accented to different degrees, but accentedness does not unequivocally equal intelligibility. This work considers the change
between pre- and post-test at which intelligibility is measured. And therefore, the measure of intelligibility in this study should not be entirely drawn on the accentedness.

The most used methods of assessing intelligibility was the dictation method utilized commonly by scholars such as Brodkey (1972), Derwing & Munro (1997), Bent & Bradlow (2003), and Burda et al. (2003). In their studies, the listeners heard spoken or read aloud utterances and were asked to transcribe them. Among them, Brodkey (1972) discovered that such assessment method was effective for testing “mutual intelligibility of neighboring speech communities or dialectal groups” (p. 205).

Moreover, Brodkey (1972) found that the dictation method was useful for testing “mutual intelligibility of neighboring speech communities or dialectal groups” and could identify loss of information between persons who “ostensibly speak the same language” (p. 216). This suggests that the dictation method is also likely to be useful in testing intelligibility in NNS-NNS conversations. An obvious disadvantage of the dictation method is that the context of the situation is not taken into account and that understanding individual word does not imply a general understanding (cf. Zielinski, 2004). Nevertheless, this method was found to be “a useful window on the listeners’ comprehension” (Munro, Derwing & Morton, 2006). Furthermore, Munro, Derwing & Morton (2006) observed a correlation between listener intelligibility judgments via the dictation method and the “phonological properties of the speaker’s output” (p. 113). As it can therefore be expected that the results reflect the degree of influence of segmentals on intelligibility, the dictation method was utilized in this thesis paper to analyze intelligibility.
CHAPTER 3: METHODOLOGY

3.1 RESEARCH QUESTIONS

Evidence from the existing literature (Ellis, 2009; Housen & Pierrard, 2005) suggested that implicit instruction addresses the communicative approach, which promotes communicative competences that include fluency (Hymes, 1972; Hedge, 2000). Therefore, I predict that implicit instruction could possibly help participants improve their fluency, gain incidental knowledge in pronunciation, including both segmental and suprasegmental features.

Meanwhile, because explicit instruction stresses Focus-on-Forms (FonFs) (Long, 2001), I believe it will likely provide tools for the student to improve their segmental and suprasegmental pronunciation features. This can increase the intelligibility of their speech through clearer enunciation.

The purpose of this thesis is to compare and contrast two types of instructional models currently popular in English speaking classes, which include accent reduction classes and conversational classes. In other words, this thesis will compare explicit and implicit pronunciation instructions. This project explores the effects each type of instruction has on adult ESL learners and the efficiency of each of these instructional approaches. More specifically, it seeks to answer the following questions:

(1) Will the participant who receives the explicit instruction show improvement in their fluency and intelligibility rate, including segmental productions?

(2) Will the participant who receives implicit instructions show improvement in their fluency and intelligibility rate, including segmental productions?
(3) Will both participants who receive instruction show more improvement in their speech quality, including fluency and intelligibility, compared to the participant who does not receive any instruction?

The research questions above then served as guidance for designing two tests, hosted in *Moodle*, that measured participants’ segmental and suprasegmental speech productions. The objective of the two tests is to compare these samples from participants immediately before and after the instructional sessions and to assess any progress made.

The sample screenshot of pre- and post-tests can be found in 3.3 MATERIALS and more details are attached in APPENDIX E.

3.2 PARTICIPANTS

The participants in this study were three international graduate students who received all of their education through undergraduate school in China and have spent a few years in the US. They had been studying and living in the same college town area in a Midwestern US state since they started their graduate programs, and then remained in the same physical location for the entire duration of this study. All participants were female Mandarin speakers and studied English as their foreign language in China in the traditional mainstream language classroom. The participants obtained similar TOEFL scores (see Table 1) before entering their current programs and have held research assistantships.
Table 1: Participants’ Background Information

<table>
<thead>
<tr>
<th>Name</th>
<th>TOEFL</th>
<th>Major</th>
<th>Year in UIUC</th>
<th>City of Undergrad</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>108</td>
<td>Communication</td>
<td>5</td>
<td>Hefei</td>
</tr>
<tr>
<td>P2E</td>
<td>106</td>
<td>Engineering</td>
<td>3</td>
<td>Shanghai</td>
</tr>
<tr>
<td>P3C</td>
<td>108</td>
<td>TESL</td>
<td>2</td>
<td>Beijing</td>
</tr>
</tbody>
</table>

All three participants demonstrated the desire to improve their oral English in their pre-research interviews. Their demographic questionnaires indicated that they usually spent more time speaking and listening to Chinese than English outside of their classes for both daily communication and recreational purposes.

All three participants were motivated to study English pronunciation because of their graduate research assistantships’ demand for a good oral communication. These participants were not enrolled in any ESL pronunciation class for the span of this study, for this would guarantee a certain a level of motivation while they would not receive any additional pronunciation instructions outside of the tutoring sessions. This control is a way to decrease the interference of external variables, that is, one participant might demonstrate more improvement due to instructional time outside the experimental treatment.

The first participant received implicit conversational/communicative instruction, in each session she was instructed to discuss a different topic freely and was given occasional informal recasts. The second participant received explicit instruction that consisted of session of general

---

1 These are all pseudonyms to which participants gave their consents.
2 As of the onset of this study.
introduction on the American English pronunciation system, three sessions of segmental features including vowel /ʊ:/, /ʌ/, /i:/, /æ/, /ɛ/ and consonant sounds /z/, /s/, /v/, /w/, which Mandarin speakers usually have difficulties with, as shown in previous research (Swan, 2001, pp 310), and four sessions of suprasegmental features including word stress, rhythm, intonation, and linking. The third participant did not receive any instruction throughout the whole study. For the purpose of simplicity, the remaining text will refer to the three participants as P1I (participant 1 implicit), P2E (participant 2 explicit), and P3C (participant 3 control).

Before the study began, I conducted brief interviews with all three participants to assess their speech qualities. I found that in their natural speech, P1I and P3C demonstrated a higher level of lexical complexity than P2E, while P1I and P2E exhibited a higher level of intelligibility than P3C, and P3C and P2E had a faster speech rate than P1I. Overall, despite the differences in speech quality of these participants, they were all able to provide complete and cohesive sentences to express their opinions and carry on conversations. These observations were generally consistent with their TOEFL scores.

3.3 MATERIALS

The materials used in this study consisted of a demographic questionnaire, a pre-test and post-test hosted on Moodle which both include three sections, listening, oral reading, and speaking.

In both pre- and post-tests, section 1 (S1) and section 2 (S2) each consisted of two parts, and the third section (S3) contained three parts. The first part (P1) and second part (P2) of the S1 were respectively pair-word multiple choice questions and cloze passage questions. Both pre- and post-tests had 10 questions in S1 P1 and S1 P2 (see Figure 3).
In S1 P1, participants were asked to listen to the audios embedded where they heard two minimally contrasted words. Then they had to choose what they heard from a list of four pairs. In S1 P2, participants listened to a full passage with only one blank. Then they needed to fill in the blank by choosing the word they heard from two minimally contrasted word options. In S1, participants could listen to the audios as often as they needed.

In S2, participants would read the given materials verbatim and record their speech. In both pre- and post-tests, S2 P1 included five sentences, while S2 P2 included one long passage.

The five sentences in the pre-test were:

(1) “It’s a foot in the door”

(2) “I need to sit in the seat because my feet do not fit in this shoe”

(3) “It’s okay to lose a race, it’s just a phase”

(4) “It’s said that when you’re sad all the time you might age faster”
(5) “You are going to the northwest part of the country. Please bring a vest in case it gets chilly”.

In these sentences, the target words were “foot” for /ʊ/, “shoe” for /u:/, “sit” for /i/, “seat” for /i:/, “race” for /s/, “phase” for /z/, “said” for /ɛ/, “sad” for /æ/, “west” for /w/, “vest” for /v/.

The five sentences in the post-test were as the following:

(1) “Sometimes I don’t feel like I fit in here, I need to live somewhere else.”
(2) “My mom always spies on me since she doesn’t trust me with the desserts.”
(3) “If you ask me, I bet that land is not cheap.”
(4) “Good gold should soon look good.”
(5) “We went to Wally’s volleyball event last night, it was fun!”

In the sentences above, the target words were “fit” for /ɪ/, “need” for /i:/, “spies” for /z/, “desserts” for /s/, “ask” for /æ/, “bet” for /ɛ/, “good” for /ʊ/, “soon” for /u:/, “Wally’s” for /w/, “volleyball” for /v/.

Moreover, the passage in the pre-test was as below:

“Luke had a baby wolf in his house, and he kept it as a pet. He put it in his room and fed it with some food he cooked. They both howled at the moon out of fun. One day, after a walk with the wolf, Luke tried to feed his pet, but his wolf appeared to be full. Then, Luke found some wool under his bed. The wolf ate a sheep! Luke finally understood why everyone told him that having a wolf as a pet is dangerous. He donated the wolf to the zoo and got a fish instead.”
The target words in the passage were “Luke” for /ʌ/, “cooked” for /ʊ/, “his” for /z/, “house” for /s/, “wolf” for /w/, “everyone” for /ʌ/, “pet” for /ɛ/, “having” for /æ/, “feed” for /i:/, and “fish” for /h/.

The passage in the post-test was as below:

“Everyone is a fan of Vicky because she is such a good cook. Whenever her friends need some food, she would always discover ways to keep her friends well-fed. She has seven secret recipes that can turn a pan into a magical land. Even her fussiest friends agree that Vicky is a wonderful chef. Especially Joyce, who enjoys Vicky’s food so much that she will take a ten-hour-long bus drive to join Vicky’s dinner.”

In this passage above, the target words were “Vicky” for /v/, “whenever” for /w/, “cook” for /ʊ/, “food” for /u:/, “dinner” for /ɪ/, “keep” for /i:/, “fed” for /ɛ/, “pan” for /æ/, “friends” for /z/, “bus” for /s/.

This section was where the controlled speech was obtained, and the speech from each participant will be analyzed in Section 4.3.

Finally, in S3, there were three parts that each included a prompt. Participants responded to two written prompts and one picture prompt in this section. They were required to upload recordings of their spontaneous speech in response to each prompt. Both S2 and S3 only allowed one attempt from the participants to record themselves.

In the pre-test S3, each prompt was as the following:

(1) What’s Your Favorite Hobby? Please take a moment to recall your favorite hobby. For example, do you play instruments? or sports? What do you usually do in your spare time? You have 15 seconds to get prepared and 1 minute to answer.
(2) Chinese Restaurants. Do you think we should have more Chinese restaurants around the school? If so, what restaurants would you like to see in the future? If not, why? You have 20 seconds to prepare and 1 minute to answer.

(3) Play with your pets. Please take a look at the photo (see Appendix E) and answer the prompt. You have 20 seconds to prepare and 1 minute to answer.

In the post-test S3, each prompt was as the following:

(1) What’s Your Favorite Food? Please take a moment to recall your favorite dishes. For example, the one dish from your hometown that you really enjoy and how is it different from the food you are having every day here at UIUC. You have 15 seconds to get prepared and 1 minute to answer.

(2) Please take a moment to look at the photo above (see Appendix E), and describe what you see and how you feel about it. You have 20 seconds to prepare and 1 minute to answer.

(3) That one skill that you’re good at. Everybody is naturally good at something. Some are good at artsy skills such as singing and painting, the others are good at communicative skills like facilitating discussions and making friends. How about you? Which skill are you particularly good at? Please take 20 seconds to prepare and 1 minute to answer.

This section was where the spontaneous speech was obtained, and the speech from each participant will be analyzed in Section 4.3.

All participants were enrolled to the pre-test a week before the first sessions started and were reminded to finish the test on time, which would be within 7 days. They were enrolled to the post-test immediately after the week of the last sessions and completed their test within the week that followed.
During the tutoring sessions, only the explicit participant was provided with Google Slides™ and Google Docs™ with visual and audio aids, and exercises. The implicit participant was provided with verbal conversational prompts based on a list of ESL discussion topics prepared before the lessons. See Table 2 for lesson topics for each participant.

Table 2: Lesson topics

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Implicit Instruction</th>
<th>Explicit Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>Dating</td>
<td>Introduction to American English pronunciation system</td>
</tr>
<tr>
<td>Session 2</td>
<td>Pets</td>
<td>Consonants</td>
</tr>
<tr>
<td>Session 3</td>
<td>Travels</td>
<td>Vowels</td>
</tr>
<tr>
<td>Session 4</td>
<td>Stress</td>
<td>Syllables and Word Stress</td>
</tr>
<tr>
<td>Session 5</td>
<td>Education</td>
<td>Focus Stress and Thought Groups</td>
</tr>
<tr>
<td>Session 6</td>
<td>Food</td>
<td>Focus Stress and Intonation</td>
</tr>
<tr>
<td>Session 7</td>
<td>the Internet</td>
<td>Intonation</td>
</tr>
<tr>
<td>Session 8</td>
<td>Cyberbullying</td>
<td>Linking and Fast Speech</td>
</tr>
</tbody>
</table>

All tutoring activities were recorded via Zoom and the test results were downloaded from Moodle, then transcribed on Online Microsoft Word and verified manually. The recordings were analyzed with Praat® (Boersma & Weenink, 2021) for vowel formants and consonant voicing patterns.

---

3.4 DATA COLLECTION AND PROCEDURES

Because of the 2020-2021 global pandemic lockdown, all materials utilized in this study were finished and collected electronically. The consent forms and demographic questionnaires were submitted to me by participants through email.

At the beginning of fall 2020 semester, the participants were recruited via a recruitment poster sent to the ESL writing instructors who teach at a Midwestern U.S. university. The selected participants were all given a brief interview, a demographic questionnaire, and a consent form. In addition, they all took a pre-test on Moodle the week before the first sessions began. The reason for choosing Moodle as the test platform was that at the time of this study, all participants and this researcher were very familiar with the usage of this medium and it was proven to be an intuitive method to turn in and collect audio data. For each item in the pre-test and post-test, the participants were only allowed one attempt to ensure a relatively more organic response.

After the pre-test, the implicit participant and explicit participant received a treatment of eight sessions of one-on-one tutoring, with one hour per session, while the control group received no instruction. The instructional session took place on a weekly basis with variations in their intervals. But the time gap between the sessions never exceeded seven days. The class schedule can be found in Appendix F. After eight weeks, all participants took a post-test on Moodle and the pre- and post-tests results were analyzed and compared.

When the sessions were completed, the two tutored participants were briefly interviewed on some general questions about the study and how they felt about their progress.
3.5 DATA ANALYSIS

All test items were answered by the participants and collected except for P1I’s response to the first written prompt in the pre-test S3. Because of unpredicted technical difficulties, in S2 and S3, participants did not universally record themselves directly on Moodle for all test items. Instead, they combined recording online and uploading audio files together to finish both pre- and post-tests. Even though there were some variations in the sound quality, they didn’t affect the qualitative analysis.

3.5.1 Data Collection and Variables

For the controlled and spontaneous speech, a second researcher listened to S2 P2 (the passage-reading task) and S3 P3 (the picture prompt) in the pre-test, S2 P2 (the passage-reading task) and S3 P1 (the first written prompt) in the post-test for interrater reliability of intelligibility rate. What’s more, these items shared overlapping prompt topics in pre- and post-tests, which would be ideal for comparison and observing progress.

These recordings were analyzed in Praat® for pronunciation accuracy, which involved vowel formant frequencies and intonation. These speech samples included S2 P1 (the sentence-reading task) and S2 P2 in both pre- and post-tests, S3 P2 (the picture prompt) in the pre-test, and S3 P1 in the post-test. The measurement of vowel formant frequencies involved selecting the spectrogram at the midpoint of the vowel in the target words. And then measuring the first and second formant. In Praat®, the first two formant frequency values (F1 and F2) were extracted from the exact midpoint between each vowel onset and offset. The target consonants were qualitatively assessed by the researcher who is trained in phonology. This was because of the difficulties in capturing the nuances in their manner and place of articulation in Praat® for quantifiable results.
When analyzing the controlled and spontaneous speech for fluency, this study looked at the temporal variables because temporal fluency is quantifiably measurable (Luoma, 2004). The variables in this study included three aspects: speech rate, silent pause total pause ratio, and filled pause total pause ratio.

Speech rate, measured by the number of syllables spoken per minute, is most frequently used as a proxy for fluency (Kormos, 2006). To acquire these measures, sample speeches were read in Praat® that automatically detects syllable nuclei to compute speech rate.

For the next measurement, Goldman-Eisler’s (1968) definition of a silent pause was used: a silent part longer than 0.25 seconds between runs. Therefore, combining the explanations given by Kormos (2006), the silent pause total pause ratio would be the number of pauses over 0.25 seconds divided by the total amount of time spent speaking expressed in seconds, then multiplied by 60, and filled pause total pause ratio was the total number of filled pauses such as “uhm”, “er”, “mm” divided by the total amount of time expressed in seconds and multiplied by 60.

Finally, to measure intelligibility of the sample speeches, both the researcher and the research mentor listened to the selected recordings and decided the number of words that were unintelligible, then used the total number of intelligible words divided by total words spoken.

This chapter has described the methodology used in this study including two fluency measures and one intelligibility measure. Chapter 4 will present the results of the pre- and post-tests for the three participants in this study and some relevant findings.
CHAPTER 4: RESULTS

4.1 PRE-TEST

4.1.1 Moodle perception test results

All three participants completed a 10-question listening task in the pre-test. P1I (participant 1 implicit) got seven out of 10 correct, P2E (participant 2 explicit) got four out of 10 correct, and P3C (participant 3 controlled) also got seven out of 10 correct. See Table 5 for a summary of the accuracy report on each participant.

Based on the results, it was shown that the three participants each had issues with the following minimal pairs in Table 3. While they got all the remaining pairs correct (see Table 4).

Table 3: Misheard sounds and words in S1 P1 by each participant

<table>
<thead>
<tr>
<th></th>
<th>/u:/, /o/</th>
<th>/i:/, /i/</th>
<th>/æ:/, /e/</th>
<th>/ɛ:/, /ɛ/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>/u:/kooky(/o/cookie)</td>
<td>/ɛ/tech(/æ/tack)</td>
<td>/ɛ/tech(/æ/tack)</td>
<td>/v/vick(/w/wick)</td>
<td>/s/slouce(/z/slews)</td>
</tr>
<tr>
<td>P2E</td>
<td>/u:/kooky(/o/cookie)</td>
<td>/i:/chip(/i:/cheap)</td>
<td>/ɛ/tech(/æ/tack)</td>
<td>/v/vick(/w/wick)</td>
<td>/s/slouce(/z/slews)</td>
</tr>
<tr>
<td>P3C</td>
<td>/i:/ kip(/i:/keep)</td>
<td>/æ/vat (/ɛ/vet)</td>
<td>/s/slouce(/z/slews)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Correctly heard sounds and words in S1 P1 by each participant\(^5\)

<table>
<thead>
<tr>
<th>/u:/, /ʊ/</th>
<th>/i:/, /ɪ/</th>
<th>/æ/, /ɛ/</th>
<th>/วล, /w/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P1I</strong></td>
<td>/u:/pool (/ʊ/pull)</td>
<td>/i:/sit (/ɪ/seat)</td>
<td>/æ/ /ɛ/vat (/ɛ/vet)</td>
<td>/วล/wet (/w/vet)</td>
</tr>
<tr>
<td></td>
<td>/u:/suit (/ʊ/soot)</td>
<td>/i:/chip (/ɪ/cheap)</td>
<td>/ɛ/dead (/æ/dad)</td>
<td>/w/wet (/w/wet)</td>
</tr>
<tr>
<td></td>
<td>/u:/shooed (/ʊ/should)</td>
<td>/i:/heat (/ɪ/hit)</td>
<td>/ɛ/head (/æ/had)</td>
<td>/v/vow (/w/wow)</td>
</tr>
<tr>
<td><strong>P2E</strong></td>
<td>/u:/pool (/ʊ/pull)</td>
<td>/i:/sit (/ɪ/seat)</td>
<td>/æ/vat (/ɛ/vet)</td>
<td>/วล/wet (/w/vet)</td>
</tr>
<tr>
<td></td>
<td>/u:/suit (/ʊ/soot)</td>
<td>/i:/heat (/ɪ/hit)</td>
<td>/ɛ/head (/æ/had)</td>
<td>/v/vow (/w/wow)</td>
</tr>
<tr>
<td></td>
<td>/u:/shooed (/ʊ/should)</td>
<td>/i:/kip (/ɪ:/keep)</td>
<td>/ɛ/head (/æ/had)</td>
<td>/v/vent (/w/went)</td>
</tr>
<tr>
<td><strong>P3C</strong></td>
<td>/u:/kooky (/ʊ/cookie)</td>
<td>/i:/chip (/ɪ/cheap)</td>
<td>/ɛ/dead (/æ/dad)</td>
<td>/วล/wet (/w/vet)</td>
</tr>
<tr>
<td></td>
<td>/u:/pool (/ʊ/pull)</td>
<td>/i:/sit (/ɪ/seat)</td>
<td>/ɛ/head (/æ/had)</td>
<td>/v/vow (/w/wow)</td>
</tr>
<tr>
<td></td>
<td>/u:/suit (/ʊ/soot)</td>
<td>/i:/heat (/ɪ/hit)</td>
<td>/ɛ/tech (/æ/tack)</td>
<td>/v/vick (/w/wick)</td>
</tr>
<tr>
<td></td>
<td>/u:/shooed (/ʊ/should)</td>
<td>/i:/kip (/ɪ:/keep)</td>
<td>/ɛ/dead (/æ/dad)</td>
<td>/v/vent (/w/went)</td>
</tr>
</tbody>
</table>

It was clear from the results in Table 5, three participants all demonstrated a better consonant perception than vowel perception when no contexts were involved, P2E was weaker on vowel perception than P1I and P3C, and the vowel minimal pair /æ/, /ɛ/ was the most problematic for everyone compared to other pairs.

---

\(^5\) The words in the parentheses were the incorrect answers that each participant did not choose.
**Table 5:** Accuracy rate of minimal pair perceptions in pre-test S1 P1

<table>
<thead>
<tr>
<th>Name/minimal pair</th>
<th>/u/, /o/</th>
<th>/i/, /u/</th>
<th>/æ/, /ɛ/</th>
<th>/v/, /w/</th>
<th>/z/, /s/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=4</td>
<td>N=4</td>
<td>N=4</td>
<td>N=4</td>
<td>N=4</td>
<td>N=20</td>
</tr>
<tr>
<td>P1I</td>
<td>3^6</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>85% (17)</td>
</tr>
<tr>
<td>P2E</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>65% (13)</td>
</tr>
<tr>
<td>P3C</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>85% (17)</td>
</tr>
</tbody>
</table>

S1 P2 was a cloze-passage practice that was presented to participants as a fill-in-the-blank activity. This practice contained 10 sentences with one blank in each sentence. Each sentence was presented with an embedded audio recording that the participants listened to before choosing the word that they heard to fill in the blank. In this task, P1I got eight out of 10 blanks correct, P2E got nine out of 10 correct, P3C got six out of 10 accurate.

Based on the results, it was shown that the three participants each had issues with the following minimal pairs in Table 6. While they got all the remaining pairs correct (see Table 7).

**Table 6:** Misheard sounds and words in S1 P2 by each participant

<table>
<thead>
<tr>
<th>/u/, /o/</th>
<th>/i/, /u/</th>
<th>/æ/, /ɛ/</th>
<th>/v/, /w/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I  /u/:food(/o/foot)</td>
<td>/ɛ/bed(/æ/bad)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2E  /u/:food(/o/foot)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3C  /u/:Luke(/o/look)</td>
<td>/ɛ/bed(/æ/bad)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The numerator is the times of sounds being correctly heard by the participants, while the denominator is the total times the sound appeared in all choices of 10 questions.
Table 7: Correctly heard sounds and words in S1 P2 by each participant

<table>
<thead>
<tr>
<th></th>
<th>/u:, /ʊ/</th>
<th>/i:, /ɪ/</th>
<th>/æ:, /ɛ/</th>
<th>/v:, /w/</th>
<th>/z:, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>/Luke(ʊ/look)</td>
<td>/lick(ɪ/leak)</td>
<td>/æ/spanned(ɛ/spent)</td>
<td>/v/vine(w/wine)</td>
<td>/z/sores(s/source)</td>
</tr>
<tr>
<td></td>
<td>/peek(ɪ/pick)</td>
<td></td>
<td></td>
<td>/v/vent(w/went)</td>
<td>/z/raise(s/race)</td>
</tr>
<tr>
<td>P2E</td>
<td>/Luke(ʊ/look)</td>
<td>/lick(ɪ/leak)</td>
<td>/æ/spanned(ɛ/spent)</td>
<td>/v/vine(w/wine)</td>
<td>/z/sores(s/source)</td>
</tr>
<tr>
<td></td>
<td>/peek(ɪ/pick)</td>
<td></td>
<td>/ɛ/bed(æ/bad)</td>
<td>/v/vent(w/went)</td>
<td>/z/raise(s/race)</td>
</tr>
<tr>
<td>P3C</td>
<td>/food(ʊ/foot)</td>
<td>/lick(ɪ/leak)</td>
<td>/æ/spanned(ɛ/spent)</td>
<td></td>
<td>/z/sores(s/source)</td>
</tr>
<tr>
<td></td>
<td>/peek(ɪ/pick)</td>
<td></td>
<td></td>
<td></td>
<td>/z/raise(s/race)</td>
</tr>
</tbody>
</table>

The accuracy rate of each participant in S1 P2 reported in Table 8 showed that, compared to their performance in S1 P1, now with contexts, the accuracy for P2E’s listening improved while P3C and P1I showed a decrease in their performance.

Table 8: Accuracy rate of minimal pair perceptions in S1 P2

<table>
<thead>
<tr>
<th>Name/minimal pairs</th>
<th>/u:, /ʊ/</th>
<th>/i:, /ɪ/</th>
<th>/æ:, /ɛ/</th>
<th>/v:, /w/</th>
<th>/z:, /s/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=10</td>
</tr>
<tr>
<td>P1I</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>80% (8)</td>
</tr>
<tr>
<td>P2E</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>90% (9)</td>
</tr>
<tr>
<td>P3C</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>60% (6)</td>
</tr>
</tbody>
</table>
4.2 POST-TEST

4.2.1 Moodle perception test results

The post-test follows an identical format and gradebook setup as the pre-test. All three participants also finished S1 within one attempt.

In S1 P1, both P1I and P2E got seven out of 10 items correct, and P3C got eight out of 10 correct.

Based on the results, it was demonstrated that the three participants each had problems with the following minimal pairs in Table 9. While they got all the remaining pairs correct (see Table 10).

Table 9: Misheard sounds and words in S1 P1 by each participant

<table>
<thead>
<tr>
<th>/u:/, /o/</th>
<th>/i:, /l/</th>
<th>/æ/, /e/</th>
<th>/l/, /w/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P1I</strong></td>
<td>/i:/He’s(/i/his)</td>
<td>/æ/gem(/æ/Jam)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/i:/bit(/i:/beat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P2E</strong></td>
<td></td>
<td></td>
<td>/w/worst(/v/ˈvərst/)</td>
<td>/s/niece(/z/knees)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/z/prize(/s/price)</td>
</tr>
<tr>
<td><strong>P3C</strong></td>
<td>/ɔ/hook(/u:/huːk/)</td>
<td>/æ/gem(/æ/Jam)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

36
Table 10: Correctly heard sounds and words in S2 P1 by each participant

<table>
<thead>
<tr>
<th></th>
<th>/u:/, /o/</th>
<th>/i:/, /i:/</th>
<th>/æ/, /ɛ/</th>
<th>/l/, /w/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>/o/pull(/u:/pool)</td>
<td>/i:/lip(/i:/leap)</td>
<td>/æ/beg(/æ/bag)</td>
<td>/v/advantage(/ædˈwæntɪdʒ/)</td>
<td>/s/niece(/z/knees)</td>
</tr>
<tr>
<td></td>
<td>/o/hook(/u://hu:k/)</td>
<td>/i:/eat(/i:/it)</td>
<td>/ɛ/beg(/æ/bag)</td>
<td>/w/wine(/v/vine)</td>
<td>/z/prize(/s/price)</td>
</tr>
<tr>
<td></td>
<td>/o/put(/pu:t/)</td>
<td>/i:/food(/fod/)</td>
<td>/i:/food(/fod/)</td>
<td>/i:/food(/fod/)</td>
<td>/i:/food(/fod/)</td>
</tr>
<tr>
<td>P2E</td>
<td>/o/pull(/u:/pool)</td>
<td>/i:/lip(/i:/leap)</td>
<td>/æ/beg(/æ/bag)</td>
<td>/v/advantage(/ædˈwæntɪdʒ/)</td>
<td>/s/niece(/z/knees)</td>
</tr>
<tr>
<td></td>
<td>Hook//hu:k/</td>
<td>/i:/He’s(/i:/his)</td>
<td>/ɛ/gem(/æ/Jam)</td>
<td>/w/wine(/v/vine)</td>
<td>/z/prize(/s/price)</td>
</tr>
<tr>
<td></td>
<td>/o/put(/pu:t/)</td>
<td>/i:/eat(/i:/it)</td>
<td>/i:/eat(/i:/it)</td>
<td>/i:/eat(/i:/it)</td>
<td>/i:/eat(/i:/it)</td>
</tr>
<tr>
<td></td>
<td>/u:/food(/fod/)</td>
<td>/i:/bit(/i:/beat)</td>
<td>/i:/bit(/i:/beat)</td>
<td>/i:/bit(/i:/beat)</td>
<td>/i:/bit(/i:/beat)</td>
</tr>
<tr>
<td>P3C</td>
<td>/o/pull(/u:/pool)</td>
<td>/i:/lip(/i:/leap)</td>
<td>/æ/beg(/æ/bag)</td>
<td>/v/advantage(/ædˈwæntɪdʒ/)</td>
<td>/s/niece(/z/knees)</td>
</tr>
<tr>
<td></td>
<td>/o/put(/pu:t/)</td>
<td>/i:/He’s(/i:/his)</td>
<td>/ɛ/beg(/æ/bag)</td>
<td>/w/wine(/v/vine)</td>
<td>/z/prize(/s/price)</td>
</tr>
<tr>
<td></td>
<td>/u:/food(/fod/)</td>
<td>/i:/eat(/i:/it)</td>
<td>/i:/bit(/i:/beat)</td>
<td>/w/worst(/vɔːrst/)</td>
<td>/v/vivid(/wɪvəd/)</td>
</tr>
</tbody>
</table>

Below in Table 11, it could be seen that overall P3C had the best performance on this task, which P1I had the worst accuracy rate. P1I struggled with the /i:/, /u/ differentiation the most because she did not get any test item that included this minimal pair correct. In contrast, both P2E and P3C had 100% accuracy rate on this vowel pair. P2E had the most difficulty with the /z/, /s/ differentiation because she did not get the item that contained this consonant pair correct while the other participants did.
Table 11: Accuracy rate of minimal pair perceptions in S1 P1

<table>
<thead>
<tr>
<th>Name/phonemes</th>
<th>/u:/ /o/</th>
<th>/i:/ /ɪ/</th>
<th>/æ/ /ɛ/</th>
<th>/v/ /w/</th>
<th>/z/ /s/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=4</td>
<td>N=4</td>
<td>N=2</td>
<td>N=4</td>
<td>N=2</td>
<td>N=16</td>
<td></td>
</tr>
<tr>
<td>P1I</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>69% (11)</td>
</tr>
<tr>
<td>P2E</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>75% (12)</td>
</tr>
<tr>
<td>P3C</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>81% (13)</td>
</tr>
</tbody>
</table>

S1 P2 was another fill-in-the-blank activity that follows the same format as that in the pre-test. For S1 P2, P1I checked nine out of 10 items correct, P2E checked eight out of 10 correct, and P3C only checked four out of 10 correct.

Based on the results, it could be found that the three participants each had problems with the following minimal pairs in Table 12. While they got all the remaining pairs correct (see Table 13).

Table 12: Misheard sounds and words in S1 P2 by each participant

<table>
<thead>
<tr>
<th>/u:/, /o/</th>
<th>/i:/, /ɪ/</th>
<th>/æ/, /ɛ/</th>
<th>/v/, /w/</th>
<th>/z/, /s/</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I /ʊ/ should (/u:/shooed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2E /ʊ/ should (/u:/shooed)</td>
<td></td>
<td></td>
<td></td>
<td>/z/ lies (/s/lice)</td>
</tr>
<tr>
<td>P3C /ʊ/ should (/u:/shooed)</td>
<td>/ɪ/ scene (/ɪ/sin)</td>
<td>/ɛ/ sad (/ɛ/said)</td>
<td>/ɛ/ bed (/ɛ/bad)</td>
<td>/z/ lies (/s/lice)</td>
</tr>
</tbody>
</table>
Table 13: Correctly heard sounds and words in S1 P2 by each participant

<table>
<thead>
<tr>
<th>/u:/, /o/</th>
<th>/i:/, /ɪ/</th>
<th>/æ/, /ɛ/</th>
<th>/v/, /w/</th>
<th>/z/, /s/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>/u:/Luke(/ʊ/look)</td>
<td>/i:/scene(/i/sin)</td>
<td>/æ/sad(/ɛ/said)</td>
<td>/w/Wicky(/v/Vicky)</td>
<td>/z/buzz(/s/bus)</td>
</tr>
<tr>
<td></td>
<td>/i:/leaving(/ɪ/living)</td>
<td>/ɛ/bad(/æ/bad)</td>
<td>/w/west(/v/vest)</td>
<td>/z/lies(/s/lice)</td>
<td></td>
</tr>
<tr>
<td>P2E</td>
<td>/u:/Luke(/ʊ/look)</td>
<td>/i:/scene(/i/sin)</td>
<td>/æ/sad(/ɛ/said)</td>
<td>/w/Wicky(/v/Vicky)</td>
<td>/z/buzz(/s/bus)</td>
</tr>
<tr>
<td></td>
<td>/i:/leaving(/ɪ/living)</td>
<td>/ɛ/bad(/æ/bad)</td>
<td>/w/west(/v/vest)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3C</td>
<td>/u:/Luke(/ʊ/look)</td>
<td>/i:/scene(/i/sin)</td>
<td>/æ/sad(/ɛ/said)</td>
<td>/w/Wicky(/v/Vicky)</td>
<td>/z/buzz(/s/bus)</td>
</tr>
<tr>
<td></td>
<td>/i:/leaving(/ɪ/living)</td>
<td>/ɛ/bad(/æ/bad)</td>
<td>/w/west(/v/vest)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to the accuracy rate report in following Table 14, it was clear that P3C had the lowest accuracy rate in this task overall, while P1I had the highest. Compared to the pre-test, P3C did not see any improvement in this section in general, especially with the vowel pairs /i:/, /u/ and /æ/, /ɛ/.

Table 14: Accuracy rate of minimal pair perceptions in S1 P2

<table>
<thead>
<tr>
<th>Name/minimal pairs</th>
<th>/u:/, /o/</th>
<th>/i:/, /ɪ/</th>
<th>/æ/, /ɛ/</th>
<th>/v/, /w/</th>
<th>/z/, /s/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=2</td>
<td>N=10</td>
</tr>
<tr>
<td>P1I</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>90% (9)</td>
</tr>
<tr>
<td>P2E</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>80% (8)</td>
</tr>
<tr>
<td>P3C</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>60% (6)</td>
</tr>
</tbody>
</table>

4.3 VOWEL ANALYSIS

The production tests were analyzed to observe the changes in each target vowel between pre- and post-test. Recordings from controlled speech (CS) S2 P1 (sentence reading) and S2 P2 (passage reading) in both tests were assessed. Six vowel phonemes, /ʊ/, /u:/, /ɪ/, /i:/, /æ/, /ɛ/,
were targeted in these speech samples. Table 15 below listed all the words in S2 P1 and S2 P2 where the target vowels were picked out.

**Table 15: Target vowels in recordings**

<table>
<thead>
<tr>
<th></th>
<th>/o/</th>
<th>/u:/</th>
<th>/i/</th>
<th>/i:/</th>
<th>/æ/</th>
<th>/ɛ/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-task3</strong></td>
<td>Foot</td>
<td>Shoe</td>
<td>Sit</td>
<td>Seat</td>
<td>Sad</td>
<td>Said</td>
</tr>
<tr>
<td><strong>Post-task3</strong></td>
<td>Good</td>
<td>Soon</td>
<td>Fit</td>
<td>Need</td>
<td>Ask</td>
<td>Bet</td>
</tr>
<tr>
<td><strong>Pre-task4</strong></td>
<td>Cooked</td>
<td>Luke</td>
<td>Fish</td>
<td>Feed</td>
<td>Having</td>
<td>Pet</td>
</tr>
<tr>
<td><strong>Post-task4</strong></td>
<td>Cook</td>
<td>Food</td>
<td>Dinner</td>
<td>Keep</td>
<td>Pan</td>
<td>Fed</td>
</tr>
</tbody>
</table>

To compare each participant’s vowel production performance in the pre- and post-test, the aforementioned target vowels in the target words were processed in Praat®. Formant 1 and 2 were used in plotting the vowel charts in Figure 4-6 below. Figure 4 compared P1I’s target vowels in S2 P1 (task 3) and S2 P2 (task 4), Figure 5 compared P2E’s target vowels in S2 P1 and S2 P2, and Figure 6 compared P3C’s target vowels in S2 P1 and S2 P2.
A traditional "vowel diagram" can be obtained by plotting the vowel formants in a graph where the horizontal axis is (F2-F1) and the vertical axis is inverse F1. (Goldstein, 2020)
Figure 5. Vowel Formant Plot for P2E’s speech
Figure 6. Vowel Formant Plot for P3C’s speech
Based on the vowel charts, it was apparent that in both the sentence and passage reading task, only P2E demonstrated noticeable positive changes in their high back vowels /ʊ/ and /u:/, and high front vowels /i:/ and /u/, to become closer to the target vowel positions.

In the post-test recordings, P2E showed hesitation before words such as “live” in S2 P1 (controlled speech) “..., I need to live somewhere else.”, and self-corrections with the word “could” in S3 P2 (spontaneous speech) with the sentence “...if I could choose”. These were all possible signs of self-awareness of target vowels pronunciation, which P2E was trained on during the tutoring sessions. This provides positive evidence in support of the effectiveness of the explicit instruction in this case.

As to the high front vowels /i:/ and /u/, all three participants were showing slightly more targetlikeness in their two tasks in the post-test than the pre-test, including P3C. This could also be a result of influence from different sentence and word environment, instead of any specific instruction.

Insofar as the mid-front vowel /ɛ/ and low front vowel /æ/, none of the participants showed too much positive progress between the pre- and post-test based on the vowel charts.

In sum, it could be concluded that the implicit instruction was not positively reflected in the participant’s vowel production while there was a more obvious enhancement with high back and high front vowels in the explicit participant’s target vowel production.

4.4 CONSONANT ANALYSIS

In measuring the selected consonant sounds, the voicing and places of articulations were looked at separately. The voiced /z/ sound has a lower frequency center of gravity than the voiceless /s/ sound according to Russell (2005). An example can be found in Figure 7. While the
places of articulation were more difficult to observe, the screenshot of each consonant’s spectrogram was collected and documented below in Figures 8-12 in support of the discussion of consonant changes in participants’ selected speech samples.

Figure 7. Frequencies for /s/ and /z/ (Russell, 2005)

To evaluate the target consonants, recordings from S2 P1 and S2 P2 in both pre- and post-test were selected and analyzed. For each target consonant sound, four words were chosen, with one word per task in each test. See Table 16 for the target consonants and the complete word list.

<table>
<thead>
<tr>
<th></th>
<th>/s/</th>
<th>/z/</th>
<th>/w/</th>
<th>/v/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-task3</strong></td>
<td>Race</td>
<td>Phase</td>
<td>West</td>
<td>Vest</td>
</tr>
<tr>
<td><strong>Post-task3</strong></td>
<td>House</td>
<td>His</td>
<td>Wolf</td>
<td>Everyone</td>
</tr>
<tr>
<td><strong>Pre-task4</strong></td>
<td>Desserts</td>
<td>Spies</td>
<td>Wally’s</td>
<td>Volleyball</td>
</tr>
<tr>
<td><strong>Post-task4</strong></td>
<td>Bus</td>
<td>Friends</td>
<td>Well</td>
<td>Vicky</td>
</tr>
</tbody>
</table>
Based on the productions from all participants, it could be found that even though (P2E) the explicit participant overall had more consonant issues, neither instructed participant seemed to show an improvement from their instruction.

Although both instructed participants had voicing issues, P1I only had problems with one sound. She devoiced /z/ sound in the word “his” from the pre-test when it was supposed to be voiced (see Figure 8). All of her other /s/ and /z/ consonant sounds were accurately pronounced. In comparison, the explicit participant had one more issue with voicing. P2E’s /z/s in “phase” and “his” were not voiced when they were supposed to be voiced (see Figure 9).

**Figure 8.** P1I’s “his” from task 4

**Figure 9.** P2E’s “phase” from task 3 and “his” from task 4
With regards to places of articulations, P1I overall performed better in her productions. Even though the implicit participant mispronounces /w/ in “Wally” as /v/ (see Figure 10) once, she was overall accurate with the placing of the bilabial sound /w/ and the labiodental /v/. This was because of the audible characteristics associated with these articulations found in other selected speech samples provided by all three participants.

Figure 10. P1I’s “Wally” from task 3

On the contrary, the explicit participant had overlapping voicing and articulatory patterns as P3C with the most consonant tokens. The only difference was the voicing of /z/s in two participants’ word “phase” (see Figure 11), as they were voiced in P3C’s speech and voiceless in P2E’s speech. In both of their speech samples, /v/ and /w/ in “vest” and “west” (see Figure 12) were challenging to differentiate, showing the same voicing patterns. Both /w/s in participants’ speech samples leaned more toward the unrounded labiodental sound /v/.
Based on the consonant productions, it was apparent that the implicit instruction showed little effect on P1I for consonants. The explicit instruction helped P2E slightly, with betterment.
in place articulation for /v/ and /w/. But it was not as effective when it came to the voicing for /s/ and /z/.

Overall, the results of both tests illustrated that the explicit participant exhibited appreciably more progress than the implicit participant with regards to the phonemes. There was no consistent evidence that proved the effectiveness of implicit instructions despite incidental improvements with collected data.

4.5 FLUENCY ANALYSIS

Furthermore, fluency will be assessed through appraisals of the selected speech samples from the pre- and post-tests for the temporal variables. Table 17 and 18 below included three participants’ WPM information for both controlled speech and spontaneous speech.

**Table 17: Temporal Variables – Words per Minute for CS**

<table>
<thead>
<tr>
<th>Name/tests</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>158</td>
<td>148</td>
<td>-10</td>
</tr>
<tr>
<td>P2E</td>
<td>136</td>
<td>127</td>
<td>-9</td>
</tr>
<tr>
<td>P3C</td>
<td>158</td>
<td>135</td>
<td>-23</td>
</tr>
</tbody>
</table>

**Table 18: Temporal Variables – Average Words per Minute for SS**

<table>
<thead>
<tr>
<th>Name/tests</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>100</td>
<td>94</td>
<td>-6</td>
</tr>
<tr>
<td>P2E</td>
<td>117</td>
<td>127</td>
<td>+10</td>
</tr>
<tr>
<td>P3C</td>
<td>134</td>
<td>136</td>
<td>+2</td>
</tr>
</tbody>
</table>
For simplicity, this paper only measured WPM as a fluency assessment factor. Recordings from S2 P2 were examined for the CS speaking rate; meanwhile, recordings from S3 P2 and S3 P3 in the pre-test, S3 P1 and S3 P3 in the post-test were transcribed and counted to obtain each group’s average SS rate.

The calculation for words per minute is Speaking rate (wpm) = total words/number of seconds*60. Based on The National Center for Voice and Speech (Barnard, 2018), the average speech rate for English speakers in the United States is about 150 words per minute. The progress = wpm in the post-test – wpm in the pre-test.

In the CS samples, all groups spoke at a slower rate in their post-test than in their pre-test. The control participant had the most observable decrease in the speaking rate. Because even though the other two participants were also showing a trend of decreased performance, the instructed individuals were both doing better than the uninstructed individuals, and the explicit participant had the least diminished speech rate, thus maintained better than the implicit participant.

In the meantime, during the eight-week-long sessions, the explicit participant showed a more significant improvement of speaking rate in spontaneous speech compared to the control group and the implicit participant. That is, there was only a negligible growth in the control group and a decrease in the implicit participant. In this case, the implicit instruction was less effective than the explicit instruction for participants in terms of their speaking rate.
4.6 FILLED PAUSE RATION ANALYSIS

Next, another temporal variable is examined – filled pause ratio, to measure each participant’s speech quality and observe the efficacy of each type of instruction. As shown in Table 19, the higher the number, the less fluent the speech was.

**Table 19: Temporal Variables – Filled Pause Ratio**

<table>
<thead>
<tr>
<th>Name</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1I</td>
<td>10.4</td>
<td>8.8</td>
<td>-1.6</td>
</tr>
<tr>
<td>P2E</td>
<td>5.7</td>
<td>5.1</td>
<td>-0.6</td>
</tr>
<tr>
<td>P3C</td>
<td>0</td>
<td>0.6</td>
<td>-0.6</td>
</tr>
</tbody>
</table>

This paper only measured average filled pauses in spontaneous speech from S3 P3 in pre-test and S3 P1 and S3 P3 in post-test to inspect these participants’ genuine fluency patterns. The calculation for pause ratios here is the total number of filled pauses such as “uhm”/the total amount of time expressed in seconds*60.

From the results, it was discernible that there was little advancement in reducing fillers, including unnatural fillers throughout this study across all participants. Even though the data illustrated a noticeable decrease in the filled pause ratio for the implicit group, a minor decrease for the explicit and control participants, both of P2E and P3C’s performance in the post-test was essentially maintained at the same level as in their pre-test.

4.7 INTELLIGIBILITY RATE ANALYSIS

The Intelligibility rate (IR) is calculated for all three participants by using the total number of words to deduct the number of words that are indicatable and then divided by the total
number of words in that speech. Two raters (the thesis student and a research mentor) listened to the selected speech for judgment. Rater 2 was not aware of the identity of the participants nor whether the speech samples were from the pre- or post-test. In this paper, the Cohen’s Kappa formula:

\[ k = \frac{p_o - p_e}{1 - p_e} \]

is used to measure inter-rater reliability.

Both researchers listened to task 4, which had 100 words in total, as the controlled speech in the pre-test. Rater 1 could not understand four, four, and seven words from the speech of P1I, P2E, and P3C respectively. While rater 2 could not distinguish four, four, eight words from the same three participants. Therefore, the IR offered by rater 1 for controlled speech in pre-test was 96% for P1I, 96 % for P2E, and 93% for P3C. The IR provided by rater 2 for this task was 96 % for P1E, 96% for P2E, and 92% for P3C (See Table 20). In this task, k=1 for P1I’s and P2E’s speech, k=0.93 for P3C’s speech. In this task, all k values indicate almost perfect and perfect agreement.

In the controlled speech from the post-test, there were 74 words in total. Rater 1 could not understand two, three, and four words respectively from P1I, P2E, and P3U’s speech. Meanwhile, rater 2 had trouble transcribing three, five, and six words from the speech made by the previous three participants, following the same sequence order. As a result, the IR offered by rater 1 for this post-test task was 97.3% for P1I, 95.9% for P2E, and 94.6% for P3C. Rater 2 offered an IR of 95.9% for P1I, 93.2% for P2E, and 91.2% for P3C. In this task, k=0.79 for P1I’s speech, k=0.74 for P2E’s speech, k=0.79 for P3C’s speech. And these k values all indicate substantial agreement.

---

8 Guidelines for Cohen’s kappa results (Landis, JR & Koch, GG 1977): 0.01 – 0.20 slight agreement; 0.21 – 0.40 fair agreement; 0.41 – 0.60 moderate agreement; 0.61 – 0.80 substantial agreement; 0.81 – 1.00 almost perfect or perfect agreement.
Table 20: Intelligibility Rate for controlled speech

<table>
<thead>
<tr>
<th>Name</th>
<th>Rater 1</th>
<th></th>
<th>Rater 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>P1I</td>
<td>96%</td>
<td>97.3%</td>
<td>96%</td>
<td>95.9%</td>
</tr>
<tr>
<td>P2E</td>
<td>96%</td>
<td>95.9%</td>
<td>96%</td>
<td>93.2%</td>
</tr>
<tr>
<td>P3C</td>
<td>93%</td>
<td>94.6%</td>
<td>92%</td>
<td>91.2%</td>
</tr>
</tbody>
</table>

For the spontaneous speech from the pre-test, rater 1 listened to S3 P2 and S3 P3 while rater 2 only listened to S3 P3. The following discussion will cover these two tasks from the pre-test.

In S3 P3, researcher one detected one unintelligible word out of 104 words finished by P1I, four unintelligible words out of 133 words produced by P2E, and five unintelligible words out of 113 words provided by P3C. The IR for P1I was 99%, for P2E was 97%, for P3C was 95.6%; rater 2 also detected one unintelligible word in P1I’s recording, four unintelligible words from P2E’s speech, and three indictable words from P3C’s speech. The IR for P1I was 99%, for P2E was 97%, and for P3C was 97.3% (See Table 21). In this task, k=1 for both P1I and P2E’s speech, suggesting perfect agreement, and k=0.74 for P3C’s speech, suggesting substantial agreement.

In S3 P2, rater 1 could not understand one out of 59 words produced by P1I, two words from a total of 96 words of P2E’s speech, and eight out of 113 words from P3C’s recording. The IR values for them were 98.3% for P1I, 97.9% for P2E, and 92.2% for P3C. Hence, the average IR for the three participants in the above succession was 98.7%, 97.5%, and 94.8%.
For the spontaneous speech from the post-test, the researcher listened to S3 P1 and S3 P3 while rater 2 only listened to S3 P1. This following discussion will cover both tasks from the post-test.

In S3 P1, rater 1 could not understand two out of 99 words for P1I’s speech, three out of 121 words from P2E’s recording, and five out of 134 words in P3C’s speech. The IR was 98% for P1I, was 97.5% for P2E, was 96.3% for P3C; rater 2 could not make out three words from P1I’s speech, two words from P2E’s speech, and zero from P3C’s recording. The IR was 100% for P1I, 98.3% for P2E, and 97.8% for P3C (See Table 21). In this task, k=N/A (null in the k value calculator) for P1I’s answer, k=0.80 for P2E’s answer, and k=0.74 for P3C’s answer, suggesting substantial agreement.

In S3 P3, rater 1 could not understand three out of 139 words produced by P1I, one word from a total of 160 words of P2E’s speech, and four out of 144 words from P3C’s recording. The IR for them are 97.8% for P1I, 99.4% for P2E, and 97.2% for P3C. Subsequently, the average IR of all SS tasks for P1I, P2E, and P3C was sequentially 97.9%, 98.6%, and 97.5%.

**Table 21: Intelligibility for spontaneous speech (SS)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Rater 1</th>
<th>Rater 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>P1I</td>
<td>98.7%</td>
<td>97.9%</td>
</tr>
<tr>
<td>P2E</td>
<td>97.5%</td>
<td>98.6%</td>
</tr>
<tr>
<td>P3C</td>
<td>94.8%</td>
<td>97.5%</td>
</tr>
</tbody>
</table>

Having observed the IR results between pre-test and post-test for all three participants (see Table 21), although there were occasional variations between rater 1 and 2 on the judgment of IR progress for each participant, it could be found that all three participants consistently
demonstrated more positive progress in their spontaneous speech than their controlled speech. The IR growth trend showed a subtle decrease in participants’ performance in the controlled speech, from which the implicit participant performed better than the explicit participant with a lesser degree of drop. On the flip side, most participants evinced an improvement of their IR from their spontaneous speech, even though rater 1 reported a slight regression from the implicit participant. Based on the IR percentage change, the explicit participant showed more spontaneous speech progress than the implicit participant.

It was evident that the nature of changes was similar across all three participants in both speech types. Both analyses of pre- and post-tests uncovered that overall production was of better quality in spontaneous speech than the controlled speech. Because the CS of participants were overall worse than their SS, they were able to show more improvement in their CS. All participants showed a greater intelligibility improvement in their controlled speech than their natural speech.

The unexpected results from the two fluency measurements (words per minute and filled pauses ratio) could be caused by myriads of variables, such as the difficulty level of the tasks and participants’ attitude or mood change (especially true for the explicit participant even though it happened to all participants). Nevertheless, by and large, the explicit participant demonstrated a more noticeable enhancement of their fluency measurements than the implicit participant.
CHAPTER 5: DISCUSSION

5.1 RESULTS AND RESEARCH QUESTIONS

The three research questions proposed in this study are as follows:

1. Will the participant who receives the explicit instruction show improvement in their fluency and intelligibility rate, including segmental productions?

2. Will the participant who receives implicit instructions show improvement in their fluency and intelligibility rate, including segmental productions?

3. Will both participants who receive instruction show more improvement in their speech quality, including fluency and intelligibility, compared to the participant who does not receive any instruction?

The research results turned out to only match partially with the research questions set up prior to this study. Overall, the explicit participant exhibited a noticeable and more consistent improvement in vowel perception and production tests on some vowel sounds including /ʊ/, /i:/ and /ɪ/, but this progress was not observed in other consonant sounds. In perception task S1 P1, both P1I and P3C showed regression in their overall accuracy rate, going from 85% to respectively 69% and 81%. Only P2E showed an increase of accuracy in this task, going from 65% to 75%. Based on the misheard sounds included in this task in pre- and post-test (see Table 3 and 9), it could be found that P2E showed a significant improvement on vowel perception test. She went from missing nine words containing target vowels in the pre-test to missing zero in the post test. While the implicit participant showed slight improvement in perception task S1 P2, going from 80% of accurate rate to 90%, it wasn’t evident enough to prove that she improved her segmental perception. The reason for this may be that in this task, participants could still rely
mainly on the context to pick out the correct answer, and P1I did not demonstrate any significant advancement in neither S1 P1 nor the production tests.

What’s more, P2E eventually improved her fluency rate, but P1I failed to display such improvement, which was somewhat unexpected. As predicted with the relevant and constant implicit language input and practice, the implicit participant was expected to advance their fluency level in the post-test. In the 5.2 limitation section, several possible explanations will be mentioned.

From the comparison between pre-test and post-test of IR for three participants, the explicit instruction turned out to be more effective than the implicit instruction, which was within expectations and largely followed the first research question “Will the participant who receives the explicit instruction show improvement in their fluency and intelligibility rate, including segmental productions?

Another finding was that participants performed better in their spontaneous speech with their general fluency and intelligibility rate. This could be attributable to the familiarity of the prompt topics and the difficulty level of the controlled speech. Yet even in the spontaneous speeches, P1I did not showcase an improvement in the fluency level, which results in a partial negative answer to the second research question about whether the implicit participant will show more improvement in their fluency. However, P1I did demonstrate a slightly higher intelligibility rate in her post-test despite of little improvement in her specific segmental production.

Out of all three participants, only P2E, who received the explicit instruction, demonstrated obvious attempts of self-correction in her recordings. For example, in the spontaneous speech from the post-test, P2E did a few self-corrections of the word “could”, which was not found in the pre-test, nor in speeches contributed by the other two participants. It
indicated that the explicit participant’s awareness was raising even though her production still was not target-like.

Drawing on the research results, short-term explicit instructions can help improve both perception and production accuracy to a degree, but the former is much faster than the latter. Although the implicit participant did not demonstrate an impressive change in the production tasks, both instructed participants reported an elevated confidence in speaking in the end, which was not noted by the control participant.

5.2 LIMITATIONS

Due to the constraints of time, location, and resources, this study had a very small size and short time duration. These factors could result in insufficient training for each instructed participant, and thus negatively impact the research results. Another factor was that the participants were from three different departments and did not share the same years of duration in the U.S., even though they all had arrived in the U.S. after completion of their undergraduate studies. In this study, one of the three participants had a different profile than the other two because she was more than five years younger. With the size of this study, such difference could have an impact on the study result.

As discussed by Ellis (2009), Housen and Pierrard (2006), the type of instruction does not necessarily lead to that type of learning, there might be strong influences from individual differences on each participant’s learning process caused by different learning styles. In addition, being online in a global lockdown made pronunciation teaching more challenging because of a lack of kinesthetic instructions to participants and a less ideal language environment. Due to a decrease in organic language input from face-to-face exposure to other speakers of English,
participants might also struggle with a common regression in producing oral output that is coherent and rich in lexicon diversity. The duration and intensity of one synchronous class per week might not be sufficient to see a noticeable change in speech patterns, especially for the implicit participant.

The designs of the pre- and post-tests will need further development in the future, expanded studies as the instruction and analysis targeted only 43 tasks, including 20 perception tasks and 23 production tasks. This design only provided a limited amount of valid data and thus could lead to biases. What’s more, due to technical errors, there were two perception exercise for minimal pairs /æ/, /e/ and /z/, /s/ lacking for two participants in the post-test. Because of this, there were only 16 segmental sounds included in the post-test when there were 20 in the pre-test. In the future study, it will be more ideal to include a matching number of target sounds in pre- and post.

In the analysis of phonemes, there was a lack of target phonemes in the spontaneous speech, therefore only data from the controlled speech was collected and analyzed. Even though the controlled speech could be a great indication of one’s general awareness of segmental features of pronunciation, some organic speech traits from each participant might be missing.
Lastly, the measures of progress in this paper did not include the lexico-grammatical complexities so that could potentially put the implicit participant at a disadvantage in this study because P1I was a slow-talker and her syntactical knowledge was richer than that of P2E. The choice of vocabulary and intact grammatical structure P1I was able to provide might also indicate some improvements that this study did not consider. Therefore, more extensive, and bigger-scale research will be needed in the future to explore the matter further.

5.3 CONCLUSIONS

The results of this case study were similar to what Derwing and Munro (2014) argued, sheer communicative language exposure would not lead to growth in the pronunciation of a target language. Even though the implicit instructions included enriched input and well-structured practice (Peltekov, 2017), there was not a clear positive correlation between the implicit instruction and P1I’s improvement in pronunciation. From this case study, for adult pronunciation teaching, it appears that explicit instruction is integral to ensure a more steady and controllable improvement in student performance, especially under time limitations. Not only would such instruction method lead to explicit pronunciation knowledge such as segmental perception and production, but also a leap in fluency that includes speaking rate and reducing fillers. This corresponds to the weak interface position (Ellis, 2009) as predicted. This is true because except for in rare, isolated cases, adults do not acquire a second language as organically and effortlessly as young children learn through causal daily conversations with other speakers of a target language.

However, these results do not negate the importance of systematic implicit input because it can provide opportunities to exercise knowledge gained from explicit instructions, and also
unconsciously internalize the target language feature patterns and thus develop metalinguistic acknowledgment of the rules (Ellis, 2009). Moreover, this paper did not measure lexicogrammatical knowledge in each participant’s speech. Therefore, it was unclear how much incidental learning (Housen & Pierrard, 2005) might have taken place during instructions.

What’s more, according to Dörnyei (2001) and Oxford and Ehrman (1992), learners’ attitudes toward and motivation for learning the target language are important factors for SLA. As reported by P11 in the end of the study, she believed that through the conversations that covered broad topics, she was able to understand American culture and society more and felt much more at ease discussing controversial topics in English. This was a great sign of attitude change in her English learning journey and could potentially have lasting positive effect on her future English learning.

Overall, this study might provide an insight for SLA researchers and ESL teachers who are interested in English pronunciation teaching. Based on the results, it can be found that to improve students’ oral English, the class has to incorporate explicit instructions and structured learning. Meanwhile, it is also important to keep the class materials and curriculum design relevant and thought-provoking for students to elicit genuine conversations. This way, students would likely benefit the most from attending these pronunciation courses.
REFERENCES


Heo, Y. (2016). Heritage and L2 Learners’ Acquisition of Korean in Terms of Implicit and Explicit Knowledge [ProQuest LLC]. In *ProQuest LLC*.


APPENDIX A INSTITUTIONAL REVIEW BOARD APPROVAL LETTERS

Notice of Exempt Determination

August 31, 2020

Principal Investigator	Suzanne Franks
CC	Xinlei Chen
Protocol Title	A Case Study on Online Pronunciation Tutoring
Protocol Number	Z1144
Funding Source	Unfunded
Review Category	Exempt 3 (ii)
Determination Date	August 31, 2020
Closure Date	August 30, 2025

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Office for the Protection of Research Subjects (OPRS) has reviewed your application and determined the criteria for exemption have been met.

The Principal Investigator of this study is responsible for:

- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Requesting approval from the IRB prior to implementing major modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

Changes to an exempt protocol are only required if substantive modifications are requested and/or the changes requested may affect the exempt status.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
ICRG00000014 • FWA #000008584
217.333.2670 • info@illinois.edu • oprs.research.illinois.edu
APPENDIX B: RECRUITMENT MATERIALS

1. Recruitment Poster

A CASE STUDY ON PRONUNCIATION TUTORING

4 Chinese grad students needed!

Why don’t Americans understand me when I speak English?

What are some effective ways to improve my English pronunciation?

Join us to find out!
Online Pronunciation Tutoring Study
A case study about pronunciation teaching.

Description:
This study aims at comparing different approaches on tutoring pronunciation for Chinese graduate students. Native Chinese speakers who received their education through undergraduate school in China are needed. In this study, 4 participants are selected based on a questionnaire and take a pre- and post-test on Moodle. Based on schedules and availability, 2 of the 4 participants will be given different methods of instruction for 8 sessions, with one hour per session. The other 2 participants will receive no instructions but will be asked to take both tests. All participants who take both tests will have an optional diagnostic consultation with the researcher at the end of the study. The participants who didn’t receive tutoring can also get personalized feedback after the study if needed. All sessions will be conducted via Zoom and all materials will be delivered electronically. Participants need to be in the US to join the study.

If Interested...
Please Contact Ms. Xinlei Chen at xinleic2@illinois.edu for more information.
2. Demographic Questionnaire

Questionnaire\(^9\): Online Pronunciation Tutoring for Chinese International Graduate Students

1. Name/Nickname: _______________

2. Your age: ______________

3. Your gender: ______________

4. Age you began learning English: ______________________

5. City of birth: _____________

6. City of schooling:

   - Primary school: ______________
   - Middle school: _____________
   - High school: ______________
   - Undergraduate school: ___________

7. TOEFL/IELTS/EPI score (If applicable): ______________________

8. Program and Year of study at UIUC: ____________________________

9. First language (chronologically): _____________________________

   - Second language: ____________________________

   - Other language(s): ____________________________

10. Languages of schooling

    - Primary school: ____________________________

    - Secondary school: ____________________________

---

\(^9\) Several items in this questionnaire were adapted from three sources.


Undergraduate: ____________________________________

Graduate: ____________________________________

Please circle all appropriate responses to the questions that apply to you:

11. Length of residence in Illinois

   1 year or less    2 years    3 years or more

12. Length of residence in other English-speaker environments

   1 year or less    2 years    3 years or more

   Please specify where: _______________________________________

13. Of the time you spend in a typical day in Illinois using language, approximately what percentage of time do you use English (as opposed to other languages)?

   0-20%   21-40%   41-60%   61-80%   81-100%   NA

14. Please indicate the estimated percentage of English media compared to non-English media you consume on a daily basis to obtain information about current events (academic or otherwise).

   0-20%   21-40%   41-60%   61-80%   81-100%   NA

15. Please indicate the closest percentage of English materials (books, tv shows, movies) in your overall language input for recreational purposes.

   0-20%   21-40%   41-60%   61-80%   81-100%   NA

16. Which languages are usually spoken in the following circumstances?
By you and your spouse/partner to each other. English Chinese Other: 

By you and your friends to each other. English Chinese Other: 

By you and your extended family members to each other. English Chinese Other: 

By you and your colleagues/classmate. English Chinese Other: 

By you and your college classmates to each other. English Chinese Other: 

17. a. Do you use English to express emotions orally? Such as venting to others.

b. If so, how often do you use English to express emotions? Please provide an example. *

18. How would you rate your knowledge and use of the English language for each of the following?

LISTENING COMPREHENSION near-native very good good fair poor

SPEAKING: pronunciation near-native very good good fair poor

SPEAKING: vocabulary near-native very good good fair poor
19. What do you find yourself struggling with the most in terms of English pronunciation? (check to see if their self-assessments match with mine)

   Segmental issues (consonant, vowel sounds)

   Suprasegmental issues (intonation, word stress, rhythm)

   Vocabulary

   Other: 

20. What is your strongest motivation for improving your oral English?

21. Have you received tutoring on pronunciation before? How helpful did you find it?

   ________________________________

18. If you are going to receive 8 sessions of private tutoring, what aspects of pronunciation will you expect to improve on the most? ________________________________

19. Please indicate whether you will be willing to partake in a brief interview (10-15 minutes) with the researcher before the sessions start as a diagnostic conversation.

   Yes  No
A Case Study on Online Pronunciation Tutoring

You are being asked to participate in a voluntary research study. The purpose of this study is to
discover efficient methods for short-term online pronunciation tutoring. Participating in this
study will involve a pre- and post- test, and up to 8 one-hour-long tutoring sessions. Your
participation will last 6 to 8 weeks. Risks related to this research are no greater than
experienced in everyday life. Benefits related to this research include a possible enhancement of
confidence in oral English and communicative skills during the period of this study.

Principal Investigator Name and Title: Suzanne Carla Franks, Teaching Assistant Professor
Department and Institution: Department of Linguistics at University of Illinois at Urbana-
Champaign
Contact Information: scfranks@illinois.edu, 217-300-4961
Sponsor: N/A

What procedures are involved?
All participants will complete a demographic questionnaire. Based on schedules and availability,
2 participants will be selected to receive 6-8 tutoring sessions via Zoom (the “treatment
participant”), while the others will not (the “control participant.”). The tutoring sessions will be
video and audio recorded. Further, the 2 participants in the treatment participant will be
tutored using different types of instruction. Both the treatment participant and the control
participant will engage in a pre- and post-test, that will include audio recordings, with two months between the administration of these tests. After the post-tests are completed, subjects may choose to participate in an interview with the one of the researchers in the study over Zoom. Only sections of the audio recordings will be kept for the purpose of data collection, and with your consent, possible publication. The video recordings and the remaining audio recordings will be destroyed after the study.

Please indicate that you consent to have your audio recordings published in the future by signing your initials: __________
(You may leave it blank if you don’t consent)

**Will my study-related information be kept confidential?**

Faculty, students, and staff who may see your information will maintain confidentiality to the extent of laws and university policies. Personal identifiers will not be published or presented without your explicit consent.

**Will I be reimbursed for any expenses or paid for my participation in this research?**

You will not be offered payment for being in this study.

**Can I withdraw or be removed from the study?**
If you decide to participate, you are free to withdraw your consent and discontinue participation at any time for any reason. Your participation in this research is voluntary. Your decision whether or not to participate, or to withdraw after beginning participation, will not affect your current or future dealings with the University of Illinois at Urbana-Champaign.

The researchers also have the right to stop your participation in this study without your consent if they believe it is in your best interests.

**Will data collected from me be used for any other research?**

Your de-identified information (information that won’t be traced back to you) could be used for future research without additional informed consent.

**Who should I contact if I have questions?**

If you have questions about this project, you may contact Suzanne Carla Franks at 217-300-4961 or scfranks@illinois.edu. If you have any questions about your rights as a participant in this study or any concerns or complaints, please contact the University of Illinois at Urbana-Champaign Office for the Protection of Research Subjects at 217-333-2670 or via email at irb@illinois.edu.

Please print this consent form if you would like to retain a copy for your records.

I have ready and understand the above consent form. I certify that I am 18 years old or older.

By signing my signature below, I indicate my willingness to voluntarily take part in this study:
Please send the signed consent form to xinleic2@illinois.edu along with any questions you may have.
APPENDIX D SAMPLE LESSON PLANS

1. Explicit Lessons

Pronunciation Feature(s): Syllable and Word Stress

Objectives: The student will be able to

1. Understand what a syllable and word stress is.
2. Count syllables within basic vocabulary.
3. Place the correct stress on basic vocabulary, understanding how “schwa” works.
4. Understand the general rules of word stress.

Preparation:

The students should prepare a laptop, a pair of headphones with a microphone. Make sure the internet connection is good and set up zoom in advanced.

[10 min] Warm-up conversations

1. To help student destress and warm up, greet the student and ask questions about how their week has been and whether they’ve gone to any place interesting.
2. Explain the objectives to the student briefly.

[10 minutes] Lecture and practice about syllable

1. Use the slide 4 on google slides to introduce syllables to the students.
2. Use the slide 5 to ask students to clap to the example words with different syllables.
3. Use the slide 6 to ask if the beats are all the same (implying there is word stress).

[15 min] Lecture and practice about word stress

1. Show the slide 7 to students, and pronounce “computer” with a stress on the first syllable.
2. Use the slide 8 to explain what word stress is, then use the next two slides to demonstrate how they can stress and unstress syllables.
3. Put students in pairs to discuss the examples on the slide 11 and then debrief as a class.

4. Ask students to work with the same partner to finish the handout.

[20 min] Lecture and practice about schwa

1. Use the slides 13-15 to explain what a schwa sound is.

2. Put students in pairs, ask them to read through words on the slide 16 to each other and use the next slide to debrief as a class.

3. Reassign students to groups of 3 and work on reading the sentences on the slide 17.

[5 min] Warp-up and assign homework

1. Briefly go over syllables, word stress, and schwa sounds.

2. Describe the homework to students.

Homework:

1. Record yourself reading through the schwa and word stress practice worksheet. Be ready to read to everyone in next class.

2. Play with the word stress website.

Reference:

2. Implicit Lessons

Dating Prompt

(1) What images spring to mind when you hear the word ‘dating’?

(2) What are the good things and bad things about dating?
(3) What can you remember about your first date?

(4) How would/did you feel before you went on a first date with someone?

(5) What is your idea of a great first date? And why?

(6) What is the most disastrous date you’ve been on or heard of?

(7) What should you never do on a first date?

(8) Would you ever use a dating service (online or otherwise)?

(9) Would you ever go on a blind date?

(10) Who do you think should pay on a first date?

(11) Do you think dating today is any different from dating 50 years ago?

(12) From what age is it OK to start dating?

(13) Isn’t it better to go for an arranged marriage and do away with the hassles of dating?

(14) Is it OK for older men / women to date very young partners?

(15) What do you think parents worry about when their kids start dating?
(16) Is it a better idea to date someone from your own race, religion, social status, etc.?

(17) What are the best ways of making sure a first date leads to a second date?

(18) If you could go on a date with anyone in the world, who would it be with and why?

(19) Is dating someone after two years still as exciting as the first date?

(20) Is dating a waste of time?

(21) From your observation, what differences and similarities are there between the American and Chinese dating culture?

Adapted from https://esldiscussions.com/index.html
Pre-Test For Online Pronunciation Tutoring

Question 1
Which set is correct?
Select one:
- a. chip/had
- b. cheap/had
- c. cheap/head
- d. chip/head

Question 2
Which set is correct?
Select one:
- a. kooky/heat
- b. cookie/heat
- c. cookie/hit
- d. kooky/hit

Pre-Test For Online Pronunciation Tutoring

Question 7
Which word fills the gap in the following sentence? Choose the word you hear in the audio recording.
I believe that there will be a huge ____ in our department, so please work hard so you can take advantage of the opportunity.
Select one:
- a. raise
- b. race

Question 8
Which word fills the gap in the following sentence? Choose the word you hear in the audio recording.
______, they are coming after you! Just run!
Select one:
- a. look
- b. Luke

Clear my choice
Pre-Test For Online Pronunciation Tutoring

Please record yourself reading the sentence once by clicking on the "insert audio recording" button at the very far right of the icons bar below:

**It's a foot in the door.**

Maximum file size: 1GB, maximum number of files: 1

Accepted file types
All file types

Pre-Test For Online Pronunciation Tutoring

Please record yourself reading the paragraph once by clicking on the "insert audio recording" button at the very far right of the icons bar below:

Luke had a baby wolf in his house, and he kept it as a pet. He put it in his room and fed it with some food he cooked. They both howled at the moon out of fun. One day, after a walk with the wolf, Luke tried to feed his pet, but his wolf appeared to be full. Then, Luke found some wool under his bed. The wolf ate a sheep! Luke finally understood why everyone told him that having a wolf as a pet is dangerous. He donated the wolf to the zoo and got a fish instead.

Maximum file size: 1GB, maximum number of files: 1

You can drag and drop files here to add them.
Pre-Test For Online Pronunciation Tutoring

5. What's Your Favorite Food?
Please take a moment to recall your favorite dishes. For example, the one dish from your hometown that you really enjoy and how is it different from the food you are having every day here at UIUC. You have 15 seconds to get prepared and 1 minute to answer.

Time limit
Your attempt will have a time limit of 3 mins. When you start, the timer will begin to count down and cannot be paused. You must finish your attempt before it expires. Are you sure you wish to start now?

Start attempt  Cancel

Pre-Test For Online Pronunciation Tutoring

Question 1
Not yet answered
Points out of 1.00
Flag question
Edit question

Please record yourself answering the question once by clicking on the "insert audio recording" icon at the far right corner below. You have 1 minute and 20 seconds to think and answer.

Everybody is naturally good at something. some are good at artsy skills such as singing and painting, the others are good at communicative skills like facilitating discussions and making friends. How about you? which skill are you particularly good at? Please take 20 seconds to prepare and 1 minute to answer.

Maximum file size: 1GB, maximum number of files: 1
# Pre-Test For Online Pronunciation Tutoring

<table>
<thead>
<tr>
<th>Question 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not yet answered</td>
</tr>
<tr>
<td>Points out of 1.00</td>
</tr>
<tr>
<td>Flag question</td>
</tr>
<tr>
<td>Edit question</td>
</tr>
</tbody>
</table>

Please record yourself answering the question once by clicking on the “insert audio recording” icon at the far right corner below. You have 1 minute and 20 seconds to think and answer.

Please look at the vacation photo and talk about how does it make you feel and why.

Please look at the photo above and describe what it reminds you of. You have 20 seconds to prepare and 1 minute to answer.
### APPENDIX F INSTRUCTION SCHEDULE

<table>
<thead>
<tr>
<th>Name</th>
<th>Session1</th>
<th>Session2</th>
<th>Session3</th>
<th>Session4</th>
<th>Session5</th>
<th>Session6</th>
<th>Session7</th>
<th>Session8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Week 1</td>
<td>Week 2</td>
<td>Week 3</td>
<td>Week 4</td>
<td>Week 5</td>
<td>Week 6</td>
<td>Week 7</td>
<td>Week 8</td>
</tr>
<tr>
<td>P1I</td>
<td>10/13/20</td>
<td>10/19/20</td>
<td>10/26/20</td>
<td>11/02/20</td>
<td>11/09/20</td>
<td>11/16/20</td>
<td>11/24/20</td>
<td>12/02/20</td>
</tr>
<tr>
<td>P3C</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

---

10 The pre-test took place within the week prior to the first session, and the post-test took place within the week following the last session.