4-26-2023

Persian ITAs and Speech Comprehensibility: Using CAPT for Pronunciation Improvement

Amir NA Rouhollahi

PERSIAN ITAS AND SPEECH COMPREHENSIBILITY: USING CAPT FOR
PRONUNCIATION IMPROVEMENT

by

Amir Rouhollahi

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Applied Linguistics

The University of Memphis

May 2023

Acknowledgment

I am sincerely grateful to the people who helped me throughout my graduate study and my dissertation project. I apologize if any name is missing but I will not ever forget those who taught me valuable life lessons and gave me knowledge, inspiration, fulfillment, and support.

Firstly, I would like to express my gratitude to my dissertation supervisor, Dr. Rebecca Adams who always tirelessly listened to my concerns, promptly provided me with the best feedback ever, provided me with insight and valuable advice to proceed, and encouraged me to achieve. She has been very inspirational for me in both my education and my future career, and I hope other students can be as lucky as I was to have a supportive supervisor like her.

 I would also like to express my gratitude to Dr. Romy Ghanem. I learned a lot from her. My initial understanding of my research concept happened in her Applied Phonology class where we learned to do lab work on speech analysis and do joint projects with brilliant results that were presented at several conferences. She involved me in great experiences like grant writing that I would otherwise have not been able to experience. She has always been helpful and insightful and has been ambitious to see her students' progress.

I wish to express my gratitude to Dr. Evelyn Rose Wright and Dr. Elizabeth Lane for providing me with insightful feedback on my dissertation project, especially during its early stages. Their constructive comments gave me the confidence to refine and enhance my work for its ultimate improvement.

I want to express my heartfelt gratitude to the linguistics program and English department for their unwavering support throughout my studies. Their provision of a safe and friendly environment

Abstract

It has been shown in the past that International Teaching Assistants (ITAs) struggle with phonological and communication issues in the classroom (Pickering, 1999; 2001). This issue leads to misunderstandings between ITAs and undergraduate students, frustrating them both as well as the parents of the students and the departments. However, studies have shown that with the right training, ITAs can focus on suprasegmental features, improving their speech comprehensibility and intelligibility (Gorusch, 2011). This study investigates the effect of Computer Assisted Pronunciation Teaching (CAPT) via tutorial videos and visual feedback on the improvement of ITAs' speech comprehensibility.

Across 5 US universities, 60 Persian ITAs, a video group (n=20), a visual feedback group(n=21), and a control group (n=19), completed an oral production pretest and recorded five diagnostic sentences plus spontaneous speech files. Over the next six weeks, all groups received in-person non-CAPT instruction, but the video group received and watched extra eight tutorial videos designed to target suprasegmental features and the feedback group was exposed to Praat visual feedback. Participants were also paired with a pronunciation tutor who provided instruction and feedback once a week. A perception posttest was administered, and the same 5 sentences with the spontaneous talk were once again recorded. The pre-and post-treatment sentences were then rated by 169 undergraduate students for comprehensibility.

The findings of this study provide a greater understanding of how explicit instruction of pronunciation through CAPT can improve the speech comprehensibility of ITAs. The number of international people in academic and professional contexts is rising, it is necessary to guide them through appropriate instruction to improve their communication quality. The results of this study suggest that even short intervention programs that include targeted in-person tutoring, tutorial

videos, and visual feedback may improve ITAs' communications. Results also imply the need for

pronunciation support for ITAs in their respective academic institutions.

Table of Contents

Chapter 1.

Introduction

Overview

A brief history of pronunciation instruction in language methodology

Pronunciation that was once dubbed "the Cinderella of language teaching" (Kelly,1969) to show its importance but marginalized existence has been subject to a fall from grace after being at the center of attention in L2 instruction during its heyday (Isaacs, 2018). In general, second language learning, there has been a history of extremes in the function of pronunciation (Levis, 2005) which can be compared to a pendulum that swings back and forth between periods when it was utterly disregarded and periods when it was of utmost importance (Kang & Kermad, 2017). Pronunciation has experienced a revival of interest both in pedagogy and research in the past half century. Although there has been limited teacher training for pronunciation (Baker & Burri, 2016), there are signs of its resurgence in L2 classrooms.

When we examine the many language-teaching methods that have been used at some point in the twentieth century, we must admit that older methods and approaches, like Grammar translation and reading approaches in language teaching, essentially disregard the importance of teaching pronunciation. However, through time, contact with people who spoke different languages became simpler and more frequent throughout the second half of the nineteenth century as a result of a variety of socioeconomic changes. A shift away from an exclusive focus on Latin and Greek to the European languages was brought about by factors including improved travel opportunities, a rise in the middle class, and an expansion of colleges and state-funded education (Lorch, 2016). Therefore, the first changes in language teaching theories, and pronunciation teaching in particular,

happened in the late 19th and early years of the 20th century through the direct method where pronunciation was taught through intuition-imitation. This method was grounded on the assumption that language should be learned like the way children learn their first language, in other words, the learners were exposed to a model pronunciation like a teacher without being forced to produce the language. That's why the successors of this approach are called naturalistic methods like the Total Physical approach and the Natural approach.

The Reform Movement in teaching languages happened in the 1890s when the International Phonetic Alphabet was established. With this invention, they could represent the sound of any spoken language in the world, and it was when, for the first time, linguistic/analytic teaching of pronunciation became possible. The Reform movement introduced a number of notions and practices that entailed thorough phonetics training for both teachers and learners. The Reform Movement also promoted the idea that spoken language is the most important and should be taught first (Celce Murcia, 2010). Such beliefs paved the ground for Audiolingualism (Howatt, 1984), which was one of the most significant developments in pronunciation instruction in the 20th century (Hodgetts, 2020).

The most important incident that made English speaking an immediate necessity was the outbreak of World War II which led to the establishment of the Audiolingual method where the exact production of sounds (native-like pronunciation) was of particular importance for language learners. Because it was believed that the principal practice in teaching was habit formation, pronunciation was taught through repetition. However, this method was criticized in the 1960s when the transformational-generative grammar of Chomsky and the cognitive approach gained popularity, with its proponents believing that the audiolingual method ignored the function of the mind in language acquisition. At this time, teaching pronunciation was deemphasized because the

proponents of universal grammar and mentalism believed that learning a native-like pronunciation was unrealistic (Scovel 1969) and teachers should invest in other components like grammar and vocabulary.

In the 1970s new approaches to learning languages appeared. One of them was Gattegno's Silent Way in which a teacher used a sound-color chart, and comparisons to the learners' first language sound system, to help them in internalizing the language's sound system (Larsen-Freeman, 2000). During this time, suprasegmental features of the language like stress and intonation were emphasized (Szyszka,2017), for the first time, and considered important in learning for better understanding and use of the target language.

The Communicative Approach (also known as Communicative Language Teaching, or CLT), which dates back to the 1970s, has dominated the teaching of foreign languages so far (Celce-Murcia et al., 2010). It places learners' abilities to communicate as the top priority therefore, the role of pronunciation in language teaching and learning is supportive rather than central in this approach (Szyszka, 2017). The aim of teaching pronunciation to learners is not to make them sound like native speakers of the target language, but rather to assist them in going above the threshold level so that their pronunciation won't hinder their communication. (Celce-Murcia et al, 2010)

This makes sense particularly when considering that languages don't have a singular, correct pronunciation. Kachru (1978) highlighted the diverse varieties of English spoken in three major contexts; inner circle (like the UK, and the USA), outer circle (countries with Colonial Englishes), and expanding circle (countries that use English as a foreign language). Making sure that learners can be understood should be the primary goal of pronunciation instruction. But having no first language influence (accent) is not a requirement for understanding (Kang & Kermad, 2017). Contrary to the nativeness concept, this view adheres to what Levis (2005) called the intelligibility

principle, referring to the extent to which listeners can understand L1 accented speech. One category of English language learners who need a high level of intelligibility in oral communication is international teaching assistants (ITAs) in the colleges and universities of English-speaking countries (Morley 1987).

An overview of ITAs in the United States

The Teaching Assistant (TA) system began in the US after WWII (Smith et al., 1992), when the GI bill expanded demand for higher education and universities needed instructors for teaching basic science courses. The high demand for instructors and shortage of qualified faculties made schools employ graduate students to teach. The TA system has changed in many aspects since its beginning. First and foremost, due to the lucrative job market and better financial opportunities in the 70s, native English students preferred to work out of school (Twale, Shannon & Moore, 1997) rather than pursue a graduate degree. Therefore, in the 70s, the number of ITAs grew and since then their population has been rising from 100,000 in 1979 to 377,944 in 2019 (Open Doors 2019). This increase in the number of ITAs, on a larger scale, can be due to the globalization of education, and social, economic, and institutional changes in the United States (Marvasti, 2005). Compared to the 70s, another change that took place in ITAs employment was about the variety of the subject areas they were involved in. ITAs are appointed in a much wider range of fields ranging from engineering, business administration, science, and technological disciplines to the arts and humanities. In response to Communicative Language Teaching (CLT) proponents who held the misconception that teaching pronunciation was hard and frustrating and would cause the students to lose their self-confidence (Binte Habid, 2013), Hinofotis and Bailey (1981) studied ITAs (used to be called Foreign Teaching Assistants, FTAs) for their effectiveness in the US universities. They found out that although ITAs were advanced in their overall language proficiency and their content

knowledge, they were not easily understood by the native speaker (NS) students. So, they concluded that although ITAs may not be able to pronounce like NSs, improving their pronunciation is possible and critical. Then they emphasized that the dichotomy of native vs nonnative speech be replaced with intelligible/ comprehensible vs unintelligible/incomprehensible speech. From that time on, the intelligible/comprehensible-based approach in pronunciation teaching became more important in both research and pedagogy. Munro & Derwing (1995) define "intelligibility" as *the extent to which a speaker's message is understood by a listener"* and "comprehensibility", as *the listeners' judgment regarding the effort required to process L2 speech rather than their understanding of content*. These two terms are sometimes used interchangeably (Levis, 2007). The focus of the current study is comprehensibility of the ITAs' speech.

1.2 Statement of the problem

In the 80s, complaints started arising from native English students and parents against ITAs' classroom oral performance and the adverse effects it could have on educational outcomes. Based on this, states (Oklahoma being the first one in 1982) began to establish laws creating baseline requirements for the communicative competence of all international teachers (Smith et al., 1992). Graduate schools all over the United States have been required to demonstrate via English proficiency test scores or mandatory oral proficiency exams that they are able to communicate effectively enough in English to teach (Johncock, 1991, Plough, Briggs, & van Bonn, 2010). Now, it has become quite common practice for American research universities to depend on ITAs for a variety of tasks particularly teaching undergraduate students. However, school departments are still informed of some negative beliefs and attitudes towards ITAs' performance in their teaching. Therefore, several scholars (Tapper & Kidder, 2006; Gorsuch, 2006; Papajohn, 2006)

advised researchers and practitioners to focus on empirical evidence obtained through studies to guide the design and implement ITA training programs.

The so-called problem of ITAs has been investigated from different perspectives. Some studies explored the perception of undergraduate students (Rubin 1992; Kang & Rubin 2009, Subtirelu, 2015), which shows their biases and experiences affect their speech perception. Some other studies have focused on the ITAs' cultural adjustment (Canagarajah, 2018a, 2018b: Lo Castro & Tapper, 2006), believing that ITAs may have a hard time adjusting to new culturally based educational norms. For instance, rather than adapting to the typical North American classroom norms, studies suggest that ITAs largely follow the standards they experienced in their home countries (Davies & Tyler 2005). Other researchers have investigated the role that ITAs' identities and attitudes have in their miscommunication. When undergraduate students, researchers, and journalists write about ITAs, their status as second-language English speakers is prominent and negatively represented because it is a crucial component of many studies' definitions of ITA identity (Borjas 2000, Clayton, 2000, Finder, 2005; Fitch & Morgan, 2003 in Looney & Bhalla, 2019).

Investigating ITAs' language abilities is especially important when they identify their English language competence as a source of concern (Ashavskaya, 2015). As a result, research has focused on specific areas of linguistic competence. For example, some studies examined the ways that ITAs used discourse markers, which are an important element of cohesion (Looney 2015; Looney et al 2017). Among their findings, they discovered that while ITAs are aware of discourse markers, they need time to prepare lectures in order to maximize their use. The grammatical and lexical constructions of ITAs speech were the subject of certain studies. For example, their findings revealed that ITAs use fewer questions (Myers, 1994) and modal constructions (Reinhardt, 2010;

Tapper, 1994) compared to their L1 English-speaking counterparts in interactions with undergraduate students.

Inspiring research into the pronunciation of ITAs, Hinofotis and Baily (1980) discovered undergraduates frequently complained that ITAs were boring, which was interpreted as being monotonous in their speech (Lindemann & Clower, 2020). Therefore, more recent research has addressed the comprehensibility of ITAs speech. Although some researchers emphasize the role of segmentals in speech comprehensibility (Sereno et al, 2016), others stress the importance of suprasegmental elements (Anderson-Hsieh et al, 1992; Derwing et al 1998; Pickering 2004; 2018) which have been discovered to be relevant in ITA pronunciation research. ITAs have been seen to misplace lexical stress and to lack control of intonation in English (Hahn, 2004; Pickering, 2001,2004).

There hasn't been much focus on teaching pronunciation in general, and among the many factors that are discussed for this inadequacy in pedagogy, a lack of teaching resources is a significant factor. This issue multiplies when the target learners are ITAs who have a busy graduate life and urgently need efficient pronunciation instructions. Literature on research about teaching suprasegmentals on the improvement of ITAs comprehensibility has been quite scarce. To the best of the author's knowledge, only one study (Lima, 2020) evaluated the comprehensibility of 12 ITAs from various first-language backgrounds after suprasegmental training. ITAs received an online pronunciation course for suprasegmental features and then their pre and post-treatment oral productions were rated by undergraduate students. The ratings showed that Just 4 out of 12 ITAs had significant improvement in their speech comprehensibility. That study demonstrated that ITAs improved individually but not collectively. The goal of the current study is to determine whether (1) a short-term suprasegmental-based pronunciation program can promote the comprehensibility

of ITA speech and (2) whether computer-assisted pronunciation teaching (CAPT) methods could enhance the effects of instruction. The current study used tutorial videos and visual feedback as CAPT for a large sample size of ITAs (n=60) from the same L1 background (Persian) in the United States.

1.3 Motivation of the study

Firstly, universities and colleges may have many international teaching assistants, including Persian ITAs, who provide valuable support to professors and help native English speaking undergraduate students (and international undergraduate students) learn and understand the course material. However, the language barriers that can arise due to differences in pronunciation may hinder effective communication between Persian ITAs and students, which can have a negative impact on the educational experience for everyone involved.

Most ITA pronunciation research has considered speakers of East Asian languages including Korean and Chinese (CITE SOURCES). Persian has different phonological characteristics that lead to accent features for Persian learners of English. For example, Persian (Farsi) words generally do not have many consonant clusters, and those that do occur are usually borrowed words from other languages. Persian language (Farsi) is also a syllable-timed language (although there are exceptions) as opposed to English, which is a stressed-timed language. This means that Farsi learners of English are prone to spend the same amount of time uttering every syllable in their speech, which makes the rhythm of their speech unexpected and difficult for English speakers to parse. Therefore, Persian ITAs might subconsciously transfer the qualities of their L1 prosody into English which might cause them not to fulfill the temporal expectation of native English undergraduate students. This fact is manifested in anecdotal accounts of Persian ITAs who receive

comments in teacher evaluations suggesting that they have monotonous or boring speech. Some ITAs in the current study even expressed that they were aware that their prosody does not meet the expectations of native English speakers, but they did not have any tools, resources, or systematic tutoring to overcome such concerns.

I, as a person who shares the same L1 with participants and as a person who has been teaching EFL in Iran for years, believe that pronunciation, in general, is not always given sufficient attention in language teaching in Iran, and that this is particularly the case for instruction on suprasegmentals. Teachers focus more on grammar and vocabulary and are not inclined to spend time on pronunciation. Moreover, teachers in Iran often rely on traditional teaching methods which might not help learners develop their pronunciation skills. Therefore, in this study, both explicit instruction and CAPT were used. The use of CAPT can complement class instruction by promoting autonomy, individualized learning, and accessibility. This may be particularly helpful for teaching pronunciation, an area where many classroom teachers feel unprepared to teach.

1.4 Purpose of the study

The main purpose of this study was to investigate the effect of instructional tools on the speech comprehensibility of ITAs in the US. Specifically, this study aimed to explore how ITAs grow their speech comprehensibility through 3 different instructional tools: CAPT tools including (1) tutorials and (2) visual feedback as well as (3) in-person, pronunciation teaching. Therefore, 3 groups of ITAs were randomly selected for this study, each group receiving different instruction. Since one argument against teaching pronunciation in general has been lack of teachers' confidence and teaching materials, this study can give insight into which educational resources work better to increase the speech comprehensibility of ESL students generally and ITAs specifically. Therefore, the results of this study would render insight and confidence to researchers,

teachers, and material developers in their investment in pronunciation. Additionally, this study aims to determine whether ITAs speech comprehensibility only enhances the production of prefabricated sentences or if it extends to the production of spontaneous speech.

1.5 Research Context

The data for this study was collected from Persian ITAs, across several graduate schools in the US and from freshman native English undergraduate students enrolled in freshman year composition classes in the English department at the University of Memphis. The data collection was through Zoom with the ITAs meeting with their instructor once a week and through online Qualtrics surveys distributed among undergraduate students.

1.6 Research questions

This study aims to find answers to the following questions:

1. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' oral sentence production as rated by native English undergraduate students?
2. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' short-sentence oral production as rated by native English undergraduate students?

1.7 Organization of the Study

There are five chapters in this work. The study's topic is introduced in Chapter 1, which also gives an overview of pronunciation teaching history, ITAs in the United States, the statement of the problem, the purpose of the study, research context, research questions, and organization of this study. Chapter 2 provides a review of the literature related to the current study exploring the main topics of communication breakdown between ITAs and undergraduate students, targeting suprasegmental features in pronunciation teaching and Computer Assisted Pronunciation Teaching

(CAPT). Chapter 3 describes all details about the participants, the design of the study, research materials and tools that were created or used during the research, and the procedure for collecting the study data. Chapter 4 represents the quantitative analysis of the results. It begins with the results of a pilot study that was carried out prior to the main study, then it describes the effort for finding appropriate inferential statistics for analyzing the data, shows the results of descriptive and inferential statistics that were run and answers the research questions. Chapter 5 discusses the findings of this study and contextualizes the findings in relation to previous research. Chapter 6 describes the conclusion of the study and provides pedagogical implications plus recommendations for future research.

Chapter 2

Literature Review

2.1 Breakdowns in communication between ITAs and their students

The fact that ITAs teaching has become an issue is clear to both TESOL researchers and native undergraduate students (Hoekje & Williams, 1992). Many researchers have attempted to find out why ITAs cannot be easily understood. The findings show that there is no single answer to this question because there are several factors such as social, cultural, linguistic, and pragmatic issues involved in ITAs' teaching undergraduate students (LeGros, & Faez, 2012; Hanayeen, 2018; Agostinelli, 2019; Looney & Bhalla, 2019; Adebayo & Allen, 2020).

This chapter begins with literature about communication breakdowns between ITAs and undergraduate students from non-linguistics and linguistics perspectives and then it draws on why there is a need for pronunciation instruction to ITAs, how segmental and suprasegmental features of pronunciation affect speech comprehensibility and then it continues with CAPT and how visual feedback and tutorial videos can be helpful for pronunciation instruction. At the end, the statement of the problem in this study is discussed in this chapter.

2.2 Breakdowns due to nonlinguistic factors

ITAs and their students are grown up in different sociocultural environments therefore, their preferences, values, and decision makings are always influenced by their cultures which can be one of the main problems that exist between ITAs and their undergraduate students. ITAs in a training workshop at the University of Georgia were found to be unwilling and uncomfortable with participative learning (Rubin, 1992) while the cultural norms in North American schools demand more freedom exercise for both instructors and students to openly and collaboratively communicate their ideas (Baily, 1982, 1984).

ITAs who are teaching in the US can gain higher scores in their teaching evaluations if they have better interactive behavior with their students, particularly through nonverbal behavior and positive affect (Baily, 1984; Inglis, 1993; Staple et al., 2014). To decrease such cultural difference effect between ITAs and NS undergraduate students, pragmatic competence in ITAs needs to be developed. They should know how to talk in different settings for different purposes, but this quality is sometimes missing in ITAs' communication. For example, they need to know that in an office-hour interaction with an undergraduate student, a conversation begins with "creating the right atmosphere" through greetings, and small talk but a classroom discourse is less interactive, and greetings may often be skipped (Axelson & Madden, 1994)

Another communication breakdown between ITAs and Undergraduate students can be due to listeners' biases. A speaker's speech will be seen as less understandable if the listener has a negative opinion of them. On the other hand, more positive opinions of a speaker will lead to speech that is easier to understand.

Rubin's (1992) widely recognized study, in which two groups of American undergraduate students listened to the same recorded lecture delivered by a native speaker of "Standard American English," provides the clearest illustration of how racial embodiment is related to comprehensible speech. The lecture for the first group had an image of a White face, but the second group's session featured an Asian face. Rubin discovered an interesting difference between the two groups' levels of lecture comprehension. Additionally, the second group thought they detected an accent from a different country. Although both groups received the same recorded lecture, Rubin's findings imply that white people are regarded as representing speech comprehensibility whereas Asian (and other racialized) ITAs are perceived as embodying speech incomprehensibility.

2.3 Breakdowns due to linguistics factors

There is not just one factor impeding the communication between ITAs and undergraduate students, but the linguistic factor seems to be the most important one, especially in the minds of native-speaking undergraduate students. In a narrative study of ITAs' identity construction, Fitch and Morgan (2003) had 900 undergraduate students tell their stories about their communication with ITAs. 71% of the stories were negative and, as a consistent theme among these stories, the students emphasized linguistic misunderstanding responsible for their interaction breakdowns with ITAs. Even some of the students made extreme remarks like blaming the school for employing ITAs with poor English-speaking skills to cause the students to fail and retake the course which would benefit the school financially (Fitch & Morgan, 2003).

While the focus of the current study is pronunciation issues relating to ITA comprehensibility, it should be remembered that other aspects of linguistics can cause miscommunications in classroom settings. For example, a linguistic source of miscommunication can be the use of discourse markers (DMs) in ITAs' speech. DMs, according to Yurng et al. (2016), are important for the structure of human thinking as well as for the clear understanding of human messages. DMs, especially those in spoken language, aid the listener in deciphering meanings. Each segment of a discourse indicates a unique link. The relationships between the segments range from switching topics to contrasting, elaborating, and drawing conclusions. The hearers can infer the signal of linkages between segments by relying on DMs. Mc William's (1992) study showed if ITAs prepare their lectures or explanations, they are able to maximize their use of DMs because they are aware of the need for DMs and they can use them in a more systematic way but in real-time, their state of L2 development, personal attitude, and contexts of interactions hinder their ability to use DMs (Looney, 2015, Looney et al, 2017). Likewise, Tyler (1992) studied the use of sequential DMs (e.g., first, finally) by ITAs. The results showed that although ITAs are good at employing certain DMs

14

at first, much like how native English speakers do, their usage of these markers becomes unclear and inconsistent as the lecture progresses. Sometimes ITAs misuse the DMs as shown by Looney's (2015) study investigating the use of *so*, *okay,* and *okay so* among ITAs in the physics lab. After analyzing a lengthy interactional sequence, multiple errors were discovered. For example, ITAs used so in the form of questions instead of drawing a conclusion which demonstrated genuine misunderstanding.

2.4 The need to target pronunciation for ITAs.

Bailey (1983) stated that ITAs had been speaking English in their home country for several years before moving to the US, therefore they did not have any modeling or correction. That is why their speech is sometimes characterized by unintelligibility and therefore incomprehensibility. ITAs are categorized among adult ESL learners. There have been teaching approaches (e.g., Silent Way, Total Physical Response, and Suggestopedia) that emphasized pronunciation for adult ESL learners but the concept of considering them as learners with fossilized pronunciation, who cannot produce and do not need native English pronunciation, came forth in the 1980s (Acton, 1984; Morley, 1986). Considering the important role that ITAs play in the universities across the US, they are among the professions next to others (e.g., foreign-born professionals working in business and industry in English-speaking countries, business professionals and diplomats who need to use English as their working language) that require urgent attention to their speech intelligibility and comprehensibility(Morley, 1989) . Therefore, discovering the target areas of pronunciation that should be addressed to meet this need is crucial.

2.4.1 Targeting segmental features.

Considering that comprehensible pronunciation is the purpose of successful L2 communication (Derwing & Munro, 2005), the new trends of teaching pronunciation are encouraged to prioritize

this aspect of pronunciation in pedagogy. For this purpose, suprasegmentals have been the focus of pronunciation research for the past 25 years (Kang, 2010) but what about segmental features.? What roles can they play in speech comprehensibility and how much have they been discovered as important for this purpose? Jenkin (2000, 2002) identified the role of segmental importance particularly in Nonnative English (NNE)-NNE communication or otherwise called English as an International Language (EIL) communication. Jenkins (2003) also emphasized "main core items" which are five categories to be prioritized in facilitating the pronunciation teaching/learning of English. Four of them are segmental features like the production of various consonants, voiced and unvoiced consonants, consonant clusters, and the production of specific vowels, and one of them were word stress.

For example, for some languages like Japanese which has an L1 phonetic system that differs greatly from that of English, learners tend to have many L1 to L2 transfers. Saito (2011) examined the relative influence of eight segmental sounds of [æ, f, v, θ, ð, w, l, ɹ] produced by Japanese English learners on two domains of comprehensibility and accentedness as rated by NE speakers. 20 native Japanese participants read two types of sentences: sentences with and sentences without these sounds. Then they were all rated by Native English-speaking raters. The results showed that these eight segmentals determine significantly impacted the comprehensibility and accentedness rating of NEs' speech.

Munro and Derwing (2006), for the first time, empirically tested the theoretical notion of functional load (FL), which determines which consonant distinction has the largest impact on learners' perception of accentedness and comprehensibility. 23 Cantonese-accented sentences with varying combinations of high and low FL errors were evaluated by 13 native English speakers. They discovered consonant substitutions with a high functional load had a bigger influence than

16

those with a low functional load. For instance, in /l/and/n/ vs /ð/ and /d/, the first pair holds a

higher functional load than the other pair because minimal pairs that are distinguished by /l/ and

/n/ contrast either in word-initial or word-final are more in number than the /d/ and / ð/ contrast.

Therefore, the role of segmental features in speech comprehensibility cannot be denied but before

we make any decision about developing instructional materials and providing curriculum for ESL

learners particularly adult learners like ITAs, we need to see the role of suprasegmental features

in speech comprehensibility as well.

 2.4.2 Targeting suprasegmental features.

 In answering the question of whether segmental (single sounds or phonemes) or suprasegmental

(e.g., pitch, word stress) features of language should be addressed first, Fraser (2001) believes that

comprehension of the listeners can tell us which feature is more important. The earliest

investigation of ITAs' communication problem used the comments of many undergraduate

students on ITAs at the University of California. The results showed that the most frequent

comment was the flat and boring tone of ITAs' voices (Hinofotis & Bailey, 1980). When a

speaker's voice's pitch doesn't vary, the listeners are less likely to maintain interest in their talk.

Given that a considerable portion of nonverbal communicative cues (40%) come from the prosodic

features (Mehrabian,1968), pitch variety is very influential in communicating feelings, and

intention of the interlocutors, for example, an instructor can use the high or low pitch and align it

with appropriate speech rate to make a distinction between more important and less important

points in his lecture. Many ITAs, as opposed to TAs, fail in their interactions with the students

because they dointentions not implement an appropriate tone to show the key points of their

presentations to the attention of their students (Rounds, 1987; Pickering, 2001). Lectures,

homework assignments from several classes, and other commitments always compete for

17

undergraduate students' attention even in a single classroom session; therefore, communicative cues like an engaging voice are a critical means to keep the students focused on the classroom. By comparing the recorded oral class performance of Chinese ITAs with that of their native counterparts, Pickering (2001) concluded that ITAs use frequent and longer silent pauses that are often misplaced, and it causes undergraduate students to have a negative perception of them. Considering that intonation is closely associated with feelings and emotion, Chinese ITAs were rated as uninvolved and unsympathetic because they couldn't, like native TAs, align their tonal structure toward informational and social convergence. For better interaction with their students, ITAs should be given the awareness or trained to use appropriate prosodic cues on which the native undergraduate students depend to understand discourse.

Although adult ESL learners can benefit from both segmental (vowels and consonants) and suprasegmental (e.g, lexical stress, prominence, rhythm and intonation) features of language in their pronunciation practice, current pronunciation scholars (Derwing & Munro, 2005; Celce Murcia et al, 2010) unanimously believe that it is suprasegmentals that have a significant effect on both intelligibility and comprehensibility of L2 speakers. Second language learners who receive suprasegmental training can perform better in real life context. Derwing et al (1998) compared the instructional gain of learners in 3 groups: segmental-based, suprasegmental-based and non-specific or control group. Although both experimental groups enhanced their intelligibility for controlled tasks, improvement in less controlled narrative tasks was observed just in the suprasegmental group. The fact that suprasegmental training causes the learners to perform better in their real-life context can be explained by the contextualized instruction they receive versus distinct sound practice that rarely goes beyond the lexical level in segmental-based teaching.

Suprasegmentals also lend themselves very well for short pronunciation courses (McNerney and Mendelsohn,1992) which can aptly accommodate the needs of adult ESL learners who either may not have enough time to spend for their learning due to other obligations or the class time may be too long for them. Teachers still have a great deal of confusion about what is desirable in teaching pronunciation and what is possible (Derwing & Munro 2005). This uncertainty can inform on several areas that are not yet investigated in pronunciation teaching; however, there have been quite definitive research findings on suprasegmentals and the improvement of L2 learners' intelligibility. To have a better understanding of the role of suprasegmentals in speech comprehensibility, we need to learn about types of suprasegmental features which are word stress, rhythm, and intonation. The relevant research findings related to each of them are discussed below.

2.4.1 Word stress

One of the prosodic features that function as a navigational guide for the listeners is word stress. If L2 speakers place stress on a wrong syllable, their speech intelligibility will be impaired (Benrabah, 1997). In English, word stress can be indicated by a variety of prosodic signals, such as syllable length (also known as duration), pitch (also known as fundamental frequency), and loudness (i.e., amplitude). Any of these indicators can distinguish differences in word stress (Zhang and Francis, 2010). Since they are the most noticeable characteristics of a stressed syllable that facilitate simpler word perception, proper word stress is essential for accurate speech processing by the listener (Celce-Murcia et al., 2010).

The importance of stress can be partially explained by the fact that its placement within a word can result in different meanings (Sa'di et al., 2022). For instance, if we move the stress from the first syllable of the word "produce," which means "things grown by farming," to the second syllable, /proDUCE/ which means "to make or manufacture something," the word's class and

meaning change. There are reasons why such a prosodic feature is highly critical in enhancing intelligibility: lexical stress represents the profile of a word, the native listeners recognize the words by a greater reliance on the stressed syllable (louder, longer and higher pitch) than by unstressed syllable, and the stressed syllable of a word produces a code that links directly to the representation of the lexicon in the mind of the listeners (Field, 2005). For instance, much as the stressed syllable / tɑːɡ / directs the search for the word photography, the stressed syllable / næʃ/ directs the search for the word international.

Stress is not always used to indicate word prosody. Some languages have no apparent word prosody (e.g French), while others feature tone (e.g., Chinese, Thai), or pitch accents (e.g., Japanese, Swedish). Some of the languages that have word stress have fixed stress, which means that the same syllable is stressed in every word (for example, Polish and Hungarian). Other languages (e.g English, Spanish, German) are characterized by free word stress, which means that for some words, the stress comes first, for others, it comes last, and for yet others, it comes on the penultimate syllable (e.g., PHOto, phoTOgraphy, photoGRAPHic) (Ghosh & Levis, 2021). L2 learners whose L1 is characterized by much different word stress patterns may need to invest more in learning the prosodic features of the target language. In other words, cross-linguistic temporal feature variations in L1 and L2 can potentially cause problems for L2 and particularly adult learners.

Native listeners are highly perceptive of the temporal features of speech produced by second language learners. Lower degrees of first language influence were perceived by the native listeners when Chinese English learners produced a larger vowel duration distinction between the word pairs such as "beat/bead" or "bat/bad" (Flege 1993). Spanish learners of English were also rated more intelligible when they produced the stressed syllables and unstressed syllables distinctively

(Hutchinson,1973). Therefore, L1 suprasegmental features can inform both the L2 needs of the learners and the curriculum that should be developed for them. For example, this matter is very true for Persian ITAs whose L1 phonological rules are different in timing influencing their English word stress and sentence stress. Persian language is characterized by a pitch accent that is applied to the last syllable of phonological words (Zadeh et al, 2011). In nouns, adjectives, prepositions, adverbs and infinitive verbs, stress invariably falls on the last syllable. Therefore, L1 Persian learners of English may fall back to their L1 phonological habit and misplace word stress on English words. Moreover, Since the Persian language is a syllable timed language, L1 Persian learners of English might apply the same duration in producing all the syllables in a sentence leading to indistinguishable prominent words which may cause them to be perceived as incomprehensible.

2.4.2 Rhythm

   Rhythm in English is the realization of an alternation in stressed and unstressed syllables (Celce-Murcia, Brinton, & Goodwin, 2010). Languages are different in their rhythm patterns. English is a stressed timed language which means there is equal time distance between the start of a stressed syllable and the beginning of the next stressed syllable in an utterance which can be a word, phrase or a sentence. Persian and French languages are syllable-timed meaning that each syllable takes the same duration to be pronounced and the longer an utterance is, the longer it takes for the speakers of these languages to say that utterance, but this is not the case in English. For example, the utterances *dog chase cat, dog chased the cat, the dogs chase the cats, the dogs will chase the cats,* and *the dogs will be chasing the cats* all take the same time to say.

In English, the stressed syllables rather than the number of syllables determine the rhythm and timing of utterances. Compared to word stress, which requires one to learn the primary stress in a multisyllabic word, rhythm is rule-based and easier to learn (Saito & Saito, 2017)

Native English listeners rely on rhythm in processing and segmenting the speech. Thus, the listeners may fail to comprehend speech if the rhythm of a phrase or sentence is distorted (Murty et al, 2007). From a psycholinguistic viewpoint, the rhythmic patterns of a language (whether it is stressed-time, syllable-time or mora-based as in Japanese) informs the native listeners about the word and syntactic boundaries when processing continuous speech (Cutler & Clifton, 1999).

In a study carried out by Tajima, Port, and Dalby (1997), the speech rhythm of Chinese learners of English were synthetically manipulated to approximate the rhythmic patterns of native English speaker. To do so, they added or deleted sounds, and modified the length of segments. Similarly, they adjusted native English speech samples to reflect the Chinese tonal system. The ratings from 36 native English listeners showed that the modified Chinese speech samples increased in intelligibility, but the native English speaker intelligibility score declined markedly. The finding of this study implies the importance of teaching rhythm to ESL learners for the improvement of their comprehensibility. In another study, Hahn (2004) had 90 undergraduate native English speakers listen to 3 types of speech stimuli made by a Korean ITA. The stimuli were a lecture with 3 types of word stress patterns: correct primary stress, incorrect primary stress and missing primary stress. The undergraduate listeners had two tasks to do: they were to listen and remember and also, by a click of a mouse, they had to show a reaction to a background sound while they were listening to the lecture. The results revealed that the students who listened to the lecture with correct primary stress had a shorter reaction to the background sound and performed better in listening and remembering. The result of this study contributes very much to the importance of learning timing

features of a second language and their role in improving the interaction between native and non-native speakers.

Additionally, Levis (2018) emphasizes the value of instruction and offers suggestions for teaching suprasegmentals. He thinks that when it comes to rhythm, perception should be the primary focus, with exposure to both naturalistic and controlled samples concentrating on the identification of keywords, weak forms, the identification of the relationship between word class and stress and features of connected speech (such as linking). Assimilation, elision, and linking are phonological phenomena in a connected speech that should be taken into account since they are acquired by native English speakers and may therefore be crucial components of suprasegmental instruction. In English, assimilation is most noticeable when a word's last consonant is impacted by the word's subsequent consonant. Common instances are the change from the consonant/d/ to the consonant/b/ in the words "bad" and "bad boy." Similar changes can be heard in the phrase *good girl*, where the final/d/ sound is changed to an initial/g/ to blend with the *girl's* initial/g/. (Roach, 1983, p. 126 cited in Hodgetts, 2020). Elision in connected speech refers to the deletion of specific sounds that would typically be present if the word was uttered alone. The deletion of the /v/ sound in of, for instance in the phrase a pack of wolves, is one of the most prevalent instances of elision.

However, Levis (2018) recommends teachers to concentrate on the vowel length of stressed and unstressed syllables while teaching rhythm for oral production. When deciding what should be covered in pronunciation instruction for foreign learners, it may be crucial to take into account these related speech elements because they are significant prosodic features that native speakers acquire and employ.

2.4.3 Intonation

Intonation, which is another prosodic feature of a language, is defined differently by scholars. Allen (1971) defines intonation as "produced by tonal height and depth along with stress, volume and varying lengths of pause". Levis (1999) has a shorter and more interesting definition for intonation. He defines it as a "significant, linguistic use of pitch". Pike (1945) emphasizes that intonation signals attitudes and it is "a rich enough system to capture different degrees of attitudes thought to be communicated".

Intonation plays several functions in English. Roach (2009, as cited in Low, 2014) classified intonation functions into four groups: attitudinal, accentual, grammatical, and discoursal. The attitudinal function of intonation, as its name suggests, aids the speaker in expressing their attitude when speaking and, more precisely, different emotions that follow their speech. For example, the utterance *I passed* conveys surprise if it takes a rise-fall tone while it indicates the speaker is asking for information if it takes a rising tone. The distinction between some emotions through the tone choice of the speaker is easy for some emotions while it is difficult for others. According to Pittam and Scherer (1993), fear, happiness, and contempt are considerably harder to distinguish than anger and sadness. As a result, recognizing an emotion's consistent effect on voice can be both simple and challenging at times.

In order to direct the listener's attention to the main syllables in an utterance, the accentual function is utilized to identify which syllables are significant or emphasized. Prominence that is also known as tonicity or focus word (Celce-Murcia, 2010) is the higher pitch that is placed on the word the speaker wants to emphasize. For example, considering the same sentence *the only thing I am interested in is completing this project on time,* there are two prominent words: "interested and time" which are the last content words in each thought group.

The grammatical function aids the listener in recognizing assertions from inquiries. Regarding the grammatical relationships of tone choices, Wells (2006) introduced an approach for teaching tone choices. He stated that although there is no straightforward relation between the tone choice and grammatical structures, it is important, by default, to apply a tone choice for each sentence type (e.g a rising tone is used for *yes/no* questions, a falling tone is used for *wh* questions, statements, exclamations, and commands). Therefore, pitch contours are among the rule-based learning (Saito & Saito, 2017) and the learners can grasp the application of basic intonation patterns.

Another aspect of the grammatical function of intonation is helping to clarify sentences that would otherwise be misunderstood. For example, in the sentence *I gave her dog biscuits*, two kinds of tone unit boundaries can be created by applying different intonations and therefore two meanings can be conveyed. If *dog biscuit* is a tone unit in this sentence, then it means someone received dog treats from the speaker but if *biscuit* by itself is a tone unit, it means the speaker gave biscuits to someone's dog.

To clarify tone units further, Celce-Murcia (2010) states that intonation includes thought group, prominence and pitch contour. Thought groups that are sometimes called intonation units, tonality, pause, or tone units play an important role in the intelligibility of speakers. Tone units allow the speaker to organize the groups of words that make up a single idea and they help the listeners to receive your speech as organized comprehensible packages that can be easily processed (Grant, 2010). For example, in the sentence *The only thing I am interested in is completing this project on time*, the speaker can produce two thought groups by inserting a very short pause right before the word *is.*

The control of conversational behavior is the discourse function of intonation which is defined as 'language above the sentence', and 'language in use' (Wichmann, 2014, p. 2). One technique to do

this is to utilize a dropping tone, which can convey to the other person that the speaker has finished speaking and that it is now their turn to talk. Another example is controlling the listener's response by using a rising tone on the final syllable of the sentence *do you like sugar or cream?* which is not limiting the available options for the listener but if the same utterance is mentioned by adopting a falling tone on the last syllable, the speaker is limiting the options for the listener.

Pedagogy is inspired by another intonation function which is proposed by Levis, & Wichmann (2015). They state intonation suggests information structure which means important information is communicated not only by the choice of tonal contours, such as rising, falling, or falling rising but also by its placement. In English, the placement of emphasis communicates an utterance's information structure, or how the speaker has organized the information within it in reference to the listener's prior knowledge. By default, the last lexical word of a phrase—or, more precisely, the stressed syllable of that word—is where nuclear prominence is used to draw attention to what is novel in the utterance. If the final syllable in an utterance is not given prominence, it is viewed as given information, and the new information is instead indicated elsewhere in the phrase or speech. For example, in the conversation that is provided below.

A: Please <u>give</u> me some cash.

B: Well, I can <u>lend</u> you some <u>cash</u>

The word *cash* is given information, but the word *lend* (probably with a fall-rise intonation) suggests a contrast with the word *give* and therefore the hearer can be pointed to a different focus.

Speakers who adopt a monotonous voice in their speech without any pitch variation are not giving enough structural cues to their interlocutors (Hincks and Edlund, 2009), and this can potentially lead to misinterpretation (Ohata 2004) and decreased comprehensibility (Pickering, 2004). By raising the pitch of a word at the beginning of one's speech, an interlocutor can help a listener to

orient himself in the mainstream of a talk (Hincks, 2005). The listeners can perceive the post-lexical level or sentence-level pragmatic meaning of utterances (Ladd, 2008). Raising the pragmatic awareness of different speech tone choices is very critical in establishing effective communication. Using the intonation model of Brazil (1997), Pickering (2001) compared international teaching assistants with native teaching assistants in terms of their tone choice in their lecture presentations. She discovered that the international teaching assistants underused rising tones which would leave a detached and unattractive impression on their listeners, native undergraduate students.

Other scholars (Jiang & Chun, 2021) investigated the effect of online instruction in English discourse intonation on the improvement of spontaneous speech among Mandarin Chinese speakers in China (as the EFL context) and in the United States (as the ESL context). The participants were 12 learners (7 in the US, 5 in China) in the experimental group and 12 learners (6 in the US and 6 in China) in the control group. The treatment was in 4 weeks and their activities were listening and speaking exercises as well as meta-instruction videos. It also included rapid feedback on visual pitch contours and personalized evaluation using the Praat program. For their pre and post-tests, students used videoconferencing to deliver a one-minute spontaneous speech on a particular subject. The students were evaluated on their speech comprehensibility besides other speech features by four native English speakers. The experimental group made statistically significant gains in their speech comprehensibility and speaking confidence, according to the results of a two-way ANCOVA test. The control group, on the other hand, exhibited no improvement. There was no interaction between the participants' training and their residence. Exploring the challenges of the Chinese learners in this study, based on the learners' self-evaluations and raters' judgment, researchers discovered that the learners struggled more with

thought groups and prominence than they do with tone choices. The learners also preferred receiving personalized feedback from the researcher to making their own visual feedback with Praat.

For teaching intonation to language learners, there have been a few suggestions from scholars, for example, McGregor et al (2016) recommend TED talks because they are authentic, engaging, and tailored to address the audience. Other scholars recommend explicit teaching of suprasegmentals (Derwing et al,1998) and others believe that teachers should target those aspects that affect comprehensibility and fluency (Kang et al., 2010) because native-like pronunciation is not seen as being as important as intelligibility and because teachers and students can only dedicate a limited amount of time on pronunciation in or outside of the classroom (Derwing & Munro, 2015). Although there is a lack of resources to teach intonation due to the complex structure of intonation, lack of emphasis on pronunciation in general, or a lack of teaching knowledge (Foote et al,2016), speech visualization tools can be effective in teaching intonation (Yu et al., 2016; Derwing & Munro, 2015)

2.5 Technology to improve comprehensibility.

 Technology can be of great help to L2 teachers and learners to achieve realistic global intelligibility. There are many software programs that are used in pronunciation labs and classrooms. Despite many calls for increased use of technology (Levis, 2007) in pronunciation teaching and research, only small progress has been made.

Technology in teaching / learning a target language pronunciation has many valuable and practical features which can benefit ELLs. Language learners can become autonomous and self-monitor themselves making them more confident in L2 communication (Tanner & Landon, 2009;

Olson,2014) They can record their voice while imitating a speech model and listen to their recorded voice for accuracy as many times as they like without needing to wait for a teacher's confirmation (Wallace & Lima, 2018). They can try this in their own privacy several times without feeling shy about their first few failures and without having to speak English with a teacher until they feel they are confident to talk. Technology can increase the amount and variety of input for the ELLs (McCrocklin, 2019). They can learn through videos, music, movies, audio recording and even through   social technology like Skype.

Two common terms that are often referenced in research and pedagogy are Computer Assisted Pronunciation Teaching (CAPT) and Automatic Sound Recognition (ASR). CAPT includes all uses of computer and software that are implemented in teaching and learning pronunciation. It can range from something very simple like recording voice to produce speech files to something complicated like analyzing the acoustic features of a speech file or converting it to a visual for teaching and learning. ASR refers to a more complex aspects of CAPT. ASR software uses the speakers' speech as input and analyzes it and then converts it to different forms and representations (e.g pitch tracings, spectrograms, and waveforms,) which can be used as feedback for the program users (Dixon, 2018).

2.5.1 Visual feedback

For pronunciation teaching, visual tools are very effective particularly for displaying the similarities and differences in the learners' speech prior to and after their instructional treatment. Visuals are also a good tool to compare the speech samples made by the learners and the ones made by native speakers (Hardison, 2004). For the first time in the 1960s, tone analyzer, a visual technology used for language acquisition of deaf children, was introduced to L2 tone learning. There was an assumption that it would be very difficult for the learners to grasp the tones of a

second language if they do not have excellent listening ability but by converting the auditory information into visual information, the learners could successfully decode the speech. By the 1980s, Kees de Bot systematically provided a set of observable tone elements like changes in pitch, speed and position. Then in the late 1980s and early 1990s, the application of visual technology was realized in many software and hardware programs for L2 pronunciation teaching. This technology was recommended in foreign language teaching particularly for ITAs by academic researchers like (Anderson- Hsieh 1989, 1992).

Among many common visual technologies for pronunciation (e.g, better accent tutor, Speech Analyzer, and speech visualization software), Praat (Boersma & Weenink, 2015) is very often used and therefore empirically more validated especially in speech assessment. In the past, Praat was mainly used by researchers to analyze speech spectrums, labeling signals, and producing text reports but in recent years, there has been a growing interest in using it in teaching English pronunciation. Since Praat is a free computer software program that continues to be updated and due to having many features serving targeted purposes for the teachers, it seems a very practical tool for learners and researchers. Praat and other visual models practicing its principles can provide the ground for understanding the subtlety of intonation in the target language and using less of the "listen and repeat" pronunciation technique which would present just a hazy illustration of speech features like intensity and pitch.

Fig.1 shows the 3 most used acoustic features of Praat for pronunciation learning. The upper band is the waveforms displaying the amplitudes or emphasis of a syllable or word pronunciation. Comparing the waveform band with the transcription or lower band, language learners can easily understand which syllable in a word or which word in a sentence takes stress. The middle band is

an intonation contour which represents the change of the pitch in a sentence. This acoustic feature informs the learners of different tones like falling, rising and level tones.
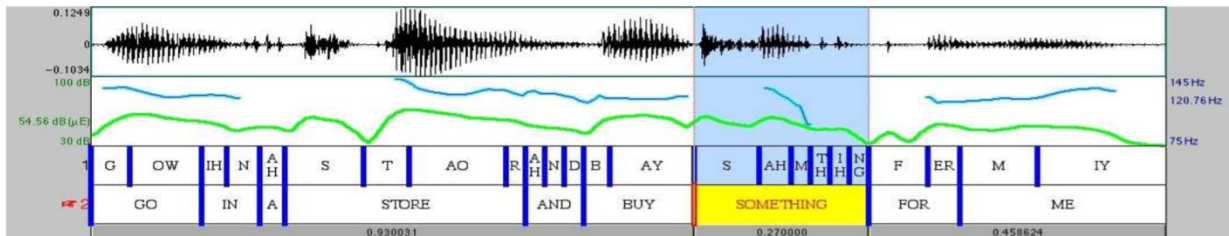

Figure 1. Speech amplitude, pitch contour and transcription (Hamlaoui, & Bengrait, 2016)

Pronunciation visual feedback as a self-learning tool especially for suprasegmental features is relatively new in research because it has been mostly used for assessment and error analysis goal (Boersman,2001; Abker, 2019). Recently studies are emerging with a focus of using Praat for individualized learning. In a control-experimental group research design, Liu et al (2018) trained Grade-one students in a Chinese senior high school to use Cool Edit (CE), as a tool for their English intonation visual feedback. The students in the experimental group used both the traditional "listen and repeat after the teacher" and their self-learning of waveform visual feedback which provided them the opportunity to compare their own pronunciation visuals with that of a model. The results proved that after a 3-month instruction period, the experimental group improved significantly in their perception and production accuracy of tone units. This implies if visual feedback can improve the target prosodic knowledge of even the learners of lower proficiency, it may enhance the speech intelligibility of more advanced learners like ITAs who have already passed an English proficiency test before their admission to American schools.

Recently, many other software programs are used in research. Many of them incorporate visualization technology similar to that of Praat. In a study investigating the impact of CAPT on the EFL learner's pronunciation of 20 single words, Saleh & Gilakjani (2020) used *Accent Master*

31

software in 5 sessions to teach 32 learners in the experimental group while the control group (38 people) received traditional instruction of listening and imitating. This software allows the learners to watch detailed video instructions, use wave graphs to compare their own productions with that of a model, play interactive listening games and listen to thousands of words and sentence pronunciation models. After having their speech productions rated by 2 raters on a 1-7 Likert scale, the researchers discovered learners in the experimental group could significantly outperform their counterparts in the control group.

In a qualitative study, Bozorgian & Shamsi (2020) investigated to what extend the EFL learners used suprasegmental features after experiencing *My English Tutor* software. This software provides the learners with prompt visual feedback of their productions and the learners can compare it with a speech model's. The feedback briefed them about loudness, pitch, timing/rhythm and the waveform of their productions. Researchers also gathered data through the learners' reflective notes, interviews and direct observations of learners' perception and production behavior while using *My English Tutor.* The results revealed that CAPT improved EFL learners not only in their use of suprasegmentals but also, as their reflective notes revealed, in their attitude towards CAPT especially due to the confidence and autonomy that it brought for them. They also mentioned that traditional pronunciation teaching is boring, frustrating and inadequate for them.

2.5.2 Tutorial Videos

Media technology has become an inseparable element of every person's daily life in different aspects (Burnett and Merchant 2015). In terms of language learning, a wide range of learning resources has appeared with much flexibility and diversity (Burnett and Merchant, 2015) which provides learners with autonomous learning opportunities, using these materials at their own pace, time and place even outside of their classrooms. Therefore, learners can address their own interests

and need much better this way (Richards, 2015). Among these learning technologies are videos that students can watch in advance before having a class with their teachers face to face, something that is called flipped or blended learning (Rotellar, C., & Cain, J., 2016). Since excessive use of open-source videos can lead to unstructured learning for the students (Jackman and Roberts, 2014), blended learning can effectively function as a pedagogical solution for this problem, offering video-based learning online plus face-to-face learning (Kinash et al., 2015, Yousef et al., 2014). This way of teaching exposes the students to the lessons at two different times, giving them the chance to digest and think about the tutorial video subject lesson before attending a real class where they will be more perceptive to what their instructor teaches them.

Videos are beneficial in promoting cognitive learning, helping learners retain knowledge, changing attitudes and learning skills (Taslibeyaz et al., 2017). Videos can demonstrate concepts that would be difficult to address verbally or in a written text (Rasi and Poikela, 2016). Regarding teaching prosodic features of a target language, pitch and loudness on stressed syllables may seem difficult for the learners to grasp therefore, using relevant notions and graphics like small, medium and large bubbles over the syllables of a single word can help the language learners clearly understand and differentiate primary, secondary and tertiary stressed syllables. Likewise, they will have a better understanding of the final tones when they watch a tutorial video with the speaker using relevant facial and hand gestures to show, for instance, the falling and rising intonation. Another advantage of learning with videos is that the learners are not limited with time or place of learning; they can watch the videos in between other tasks, whenever it is convenient for them and as many times as they wish (Taslibeyaz et al., 2017)

The engagement of the students with the tutorial videos can inform the researchers and teachers to a great extent about the effectiveness of these instructional tools and whether they intend to

complete the course (Martinez, 2001). Feredricks et al (2004) created a three-factor model discussing the engagement of the learners with teaching materials. The learners show their involvement through behavior (e.g attendance and participation), cognition (being attentive to what the instructor teaches) and emotion (e.g students' feelings and interest in their learning).There are several studies investigating the satisfaction of the students with videos as instructional tools (Kay,2012; Kay & kletskin, 2012; Yousef et al., 2014; Giannakos et al., 2016). These studies prove that learners are satisfied with their learning experience since videos are motivating, autonomous, thought-provoking, easily accessible online, and convenient for learners in terms of time and the place of learning.

Sometimes, videos can be integrated into online tutoring. Lima (2015) carried out a study with twelve ITAs from different L1 backgrounds, focusing on the improvement of their intelligibility. The videos, that were used beside online reading passages and exercises, included simultaneous lecture speech files on visuals on suprasegmental features like word stress, rhythm and intonation. This online tutoring proved engaging enough for the learners to complete the course.

As discussed above, there are many benefits in learning through videos but most of the research studies about this educational technology has been in the fields of science and, to the knowledge of the authors, literature is scarce about any study focusing on the effect of pronunciation tutorial videos on the gains of adult language learners. Therefore, the present study aimed to investigate how tutorial videos, together with personal tutoring, can help adult ESL learners (ITAs here) improve their perception of the suprasegmental.

Statement of the Problem

The arrival of ITAs to US schools began almost half a century ago and their number has been increasing ever since. Despite all their contribution to research fields, ITAs' interaction with native undergraduate students has always been a matter of concern for parents, students and the school departments. Among all the issues causing this concern, linguistic factors are recognized as the most important subject to address. Although graduate schools all over the USA mandated English proficiency acceptable test scores and sometimes interviews before giving admission to the applicants, linguistic communication breakdowns due to ITAs' first language influence are still frequently observed between ITAs and undergraduate students. This concern has led to negative beliefs and attitudes towards ITAs' performance in their teaching. Unfortunately, ITAs are generally too busy with their graduate life to spend time attending pronunciation courses. Therefore, this study is aiming to resolve this problem by making Persian ITAs improve their pronunciation with effective technologies. Due to a lack of time, skill and experience, Iranian EFL teachers didn't like to pay enough attention to pronunciation in general and using technologies in particular (Abdlmanafi-Rokni, 2013).

ITAs are categorized as one of the four categories of language learners whose oral communication needs mandate a high level of intelligibility and therefore require special assistance with pronunciation (Morley, 1987, mentioned in Celce Murcia, 2010). Despite all the studies using technology for teaching and learning a second language, no study, at least to the knowledge of the author, has ever addressed the pronunciation needs of international teaching assistants, using visual feedback and tutorial videos.

The present study aims to investigate the effect of pronunciation instruction on the oral production of Persian ITAs, using two pronunciation technologies: Praat program and tutorial videos. Therefore, the following questions will be answered in this study:

1. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' short-sentence production as rated by native English undergraduate students?

2. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' spontaneous speech production as rated by native English undergraduate students?

Chapter 3

Methods

This study aimed to investigate the effect of computer-assisted pronunciation teaching (CAPT) on the improvement of Persian ITA's speech comprehensibility as intuitively rated by native English-speaking undergraduate students' perception. The following sections will briefly explain the participants, research instruments, and the data collection procedure for this study.

3.1 Participants

There were two groups of participants in this study: ITAs and undergraduate students.

3.1.1 ITA

As Table 1 shows, ITA participants (n=60) were male and female Persian graduate students studying and teaching at two Southern and three Northeastern universities in the USA, with the majority of them being within the 25-35 age range. These participants were recruited through electronic flyers, social media, and through word of mouth. Iranian associations on the campus of these schools were also very instrumental to distribute this project's advertisement among Persian ITAs. After collecting their informed consent, participants were assigned into 3 groups: control group (n=19), feedback group (n,=21) and video group (n=20). Since this study aimed to investigate the effect of suprasegmental instructions on ITAs' speech comprehensibility, the ITAs who had such instruction ever before were excluded from participation. To control for interlanguage effects, ITAs with the same L1 background, the Persian language, which is also the researcher's L1, were recruited for this study. Theycame from variety of disciplines with a particular emphasis on engineering fields like civil engineering, computer science, mechanical engineering and biomedical engineering.

Table 1 ITA's Demographic Information

	ITAs	Gender	L1	Age	The average years of residence in the US	The average number of languages they understand	The average years of working as an ITA
Control Group	19	63.1% M, 39.9% F	Farsi	(25-35)	3.8	2.1	1
Feedback Group	21	52.4% M, 47.6%F	Farsi	(25-35)	3.5	2	1.1
Video Group	20	55% M, 45% F	Farsi	(25-32)	2.25	2.1	0.8

3.1.2. Undergraduate students (raters)

As shown in Table 2, undergraduate students (n= 169) in this study were males and females who were either freshmen or sophomores from a southern University in the USA. They were all taking first year composition courses, which are college-level writing courses offered in many universities in the US for undergraduate students. Therefore, these participants came from a variety of disciplines. ITAs as instructors are a relatively new experience for these students. With this in mind, their selection contributes to a better reliability of the results of this study. The majority of the undergraduate students were within the 18-20 age range, and they were all native speakers of English. The data from those students who were not native English speakers were excluded during data cleaning. These participants were all recruited through an advertisement which was emailed to them via their professors. Those who were willing to participate clicked

on a Qualtrics survey link which was at the end of the advertisement. That survey started with a

consent form and then to survey pages where they rated the speech files.

Table 2 Undergraduate students' demographic information

Undergraduate students	The average age	Gender	L1	Number of languages you understand	Number of classes you had with ITAs
169	20.5	57% F, 2% non-binary, 41% M	English	1-2	0-2

3.2 Design

The intervention for this study was 6 weeks as shown in Table 3.

Table 3 The Design of the Intervention

Weeks	Activities	Control group	Feedback group	Video group
1	Diagnostic test and consent forms	Pretest	Pretest	Pretest
2	Word stress	Non-CAPT	Non-CAPT plus visual feedback	Non-CAPT plus Tutorial video
3	Prominent words and rhythm	Non-CAPT	Non-CAPT plus visual feedback	Non-CAPT plus Tutorial video
4	Intonation	Non-CAPT	Non-CAPT plus visual feedback	Non-CAPT plus Tutorial video

5	Review	Non-CAPT	Non-CAPT plus visual feedback	Non-CAPT plus Tutorial video
6	Diagnostic test	Posttest	Posttest	Posttest

The present study had an experimental design including 3 groups of participants. All these groups received constant non-CAPT pronunciation instruction. The control group received no extra instruction, the feedback group was taught with the software program Pratt, providing them with visual feedback on their suprasegmental features and the video group received tutorial videos besides their shared instructions with other groups.

3.3 Materials

This section will discuss various instruments, materials, and tests that were used to conduct this study.

3.3.1 pre-and post-test speech samples items

In order to determine whether treatments influenced pronunciation, pre, and post-treatment speech samples were collected from all participants. First, the speech samples are discussed below then the instruments that were used to record and prepare the speech files are discussed.

3.3.1.1 Speech Files

Two types of speech files were produced by ITAs. The short sentence speech files and spontaneous speech files. Each ITA read and recorded 5 short speech files in the diagnostic pre-test and recorded their voice for the same sentences in the diagnostic post-test, with each speech file lasting 3-5 seconds. For the longer speech files, ITAs answered a TOEFL speaking question where they listened to a conversation between a professor and a student (campus talk) then they

produced a one-minute oral summary about that conversation while recording themselves. There was a total of 120 spontaneous speech samples (60 pretest and 60 posttest samples) for all ITAs, 5 randomly selected extracts from each speech sample were taken for Qualtrics surveys. It was not possible to produce extracts with durations that were the same because this would have produced utterances that weren't always beginning or ending at clausal or phrasal borders. Therefore, the length of spontaneous speech files varied slightly; the mean length was 10 words.

3.3.1.2 Voice Recording Apps

For both pre- and post-diagnostic tests, and for convenience and to avoid any technical glitch, ITAs were asked to use their own Android or iPhone apps to record their voice based on the researcher's instruction. These apps are free and user-friendly and ITAs could easily share the produced speech files with the researcher.

3.3.1.3 Audacity

It is a free, and open-source audio software having many advanced functions in recording, editing, and analyzing speech files. In preparation for the survey for undergraduate students to rate ITAs' speech files, this program was used by the researcher, for optimizing the ITAs' speech files by removing static, hiss, hum, or other constant background noise.

3.3.2 Testing

3.3.2.1 Pre and posttest for ITAs' production

After providing all materials discussed above, pre- and post-diagnostic tests were created. These tests were administered two times, before and after the treatment. They were intended to examine the ITAs' demonstrated abilities in the production of suprasegmental features. They were of two types: controlled-sentence production tests and spontaneous production tests.

3.3.2.1.1 Controlled-Sentence Diagnostic Test

This was a 5-item test (Appendix A), with each item beginning with a context to provide enough background knowledge for the ITAs before producing the target sentence speech file that was rated later by undergraduate students. The written passage of this test was adapted from the book *Delta's Key to the TOEFL iBT* and modified by professional linguists to fit this test. The content of the written passages in this test are conversations between a professor and a student discussing different topics.

3.3.2.1.2 Spontaneous Diagnostic Test

This was the oral performance of ITAs on narrating the story of a campus talk. ITAs listened to a listening file (Appendix B) taken from the book *Delta's Key to the TOEFL iBT*. This listening file did not have any special or difficult terms to be confusing for ITAs. They first listened to the audio file and took notes, then they narrated the summary of the conversation while having their voice recorded.

3.3.2.2   Speech Evaluation Survey (Qualtrics survey)

20 Qualtrics surveys were created with each one containing 40 questions. Each survey began with a background questionnaire that collected general information from undergraduate students before listening to speech samples and rating ITAs' speech files. In addition to age, gender, and L1 background, the questionnaire gathers information like participants' familiarity with other languages and the number of courses they have had with ITAs. Each item of these anonymous surveys (appendix C) consisted of a coded speech file followed by a 9-point Likert scale adapted from Munro and Derwing (1995); number 1 suggests that the speech file is hard to understand and number 9 is easy to understand. Undergraduate students listened to the speech file and rated it at

least. Each rater needed a time between 10 and 12 minutes to complete the survey. No survey contained all the speech samples, as this would have made the surveys much too long. Speech samples were distributed through the surveys to ensure that each sample was rated five times.

3.3.3 Instructional treatment

This section discusses the pronunciation instruction that each group in this study received.

3.3.3.1 Non-CAPT Suprasegmental Instruction

This pronunciation instruction was inspired by two books. The first one was *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages* (Celce-Murcia, 2010) which was created as the foundation for a thorough course in pronunciation pedagogy for ESL/EFL instructors. The second book was *Delta's Key to the TOEFL iBT* (Gallagher, N. 2012) which contains many campus conversations and academic talks and is intended for the TOEFL preparation courses. Lessons were created for each of the targeted suprasegmental features (Appendix D shows the lesson plan for intonation module) and are described in detail below. In general, every class began with a 5-minute review, then 10 minutes of practice and finally 10-15 minutes of focus on student questions and/or additional practice. Based on the communicative framework of pronunciation teaching (Celce-Murcia, 2010), every module of word stress, rhythm, intonation and review included description, listening discrimination, control & guided practice and finally communicative exercise. For example, for their rhythm lesson, their teacher described content vs function words, then they listened to a few examples where they recognized the speakers' voice rising up in pitch and increasing in length on content words compared to function words. Then their teacher read sentences while tapping or clapping on content words and they did the same after their teacher. Then they were asked to imagine some professions for themselves and play roles in a conversation like this:

A: WHAT do you DO?
B: I am a DOCtor and I WORK in a HOSpital.
A: WHAT do YOU do?
B: I am a proFESsor and I LECTURE at the uniVERrsity.

 For their communicative practice, a student paired up with another student or the teacher. One of them took the role of the host and the other the role of a guest. They were given a role card with identification on it when it was their turn to be guests. The role on the card was Guinness World Records-inspired identities. For example, the guest card read " you are the first person to cross the Pacific Ocean in a hot air balloon" then another person asked questions like " what is your name? What did you do? what record did you break?" The instructor provided them feedback after monitoring their rhythm and precise placement of stress.

3.3.3.2 CAPT Suprasegmental Instruction

3.3.3.2.1 Tutorial Videos

I produced eight 15-minute tutorial videos (Appendix E) under the supervision of a professional linguist (one of the committee members) for the video group. First the content and scripts for each module were created while observing the principle of beginning the lessons with listening, discrimination, guided and controlled practice, and finally communicative activities. For better illustration, several shapes, forms and symbols like bubbles, arrows, sinuous lines together with color choices were used to show a better representation of utterances' prosody like loudness and pitch. Then all the videos were recorded on PowerPoint slides and edited for their sound quality. For each suprasegmental feature, a pair of videos were created: word stress (2 videos), prominent word and rhythm (2 videos), intonation (2 videos), and reviews (2 videos). In each pair, one video provided the ITAs with the theoretical explanations of a certain suprasegmental feature and the other presented relevant exercises to that lesson. The two review videos (one theoretical review

and one review exercise) were produced to help ITAs to reinforce their learning and assess their understanding of the suprasegmental lessons. Based on the principle of flipped learning by which learners watch the instructional materials before coming to class, participants in this group received the videos for word stress one week before their class and they were asked to watch them. As they attended their first session on week 1, they were first exposed to non-CAPT instruction, as explained above, for 30 minutes. Shortly after, all ITAs and the instructor altogether watched the 15-minute theoretical explanation video pertaining to week 2 , followed by 5 minutes of Q&A Upon leaving the class, the exercise video for week 2 was immediately distributed to the participants electronically. Therefore, in preparation for the non-CAPT lesson on week 2, the participants were given a week to watch the theoretical explanation and exercise videos at least a couple of times before attending the class on week 2. This approach was maintained throughout the subsequent sessions.

3.3.3.2.2 Praat Visual Feedback

Praat is a software program with many functions of voice recording, editing and acoustic analysis. It provides the learners with visual feedback on their oral productions. The software features that were used in this are shown in Appendix F. There are 3 bands shown in this program. The upper band is the waveforms displaying the amplitudes and duration of a syllable or word pronunciation. Comparing the waveform band with the transcription or lower band, language learners can easily understand which syllable in a word or which word in a sentence takes the stress. The middle band is the intonation contour which represents the change of the pitch in a sentence. This acoustic feature informs the learners on different tones like falling, rising. and level tones

3.4. Procedure

In this study, a pre-test/post-test design was used to measure the extent to which different types of pronunciation instructions contribute to the improvement of ITAs' suprasegmental production and speech comprehensibility. For this purpose, this research was completed through a 5-phase procedure as shown in Fig.1.
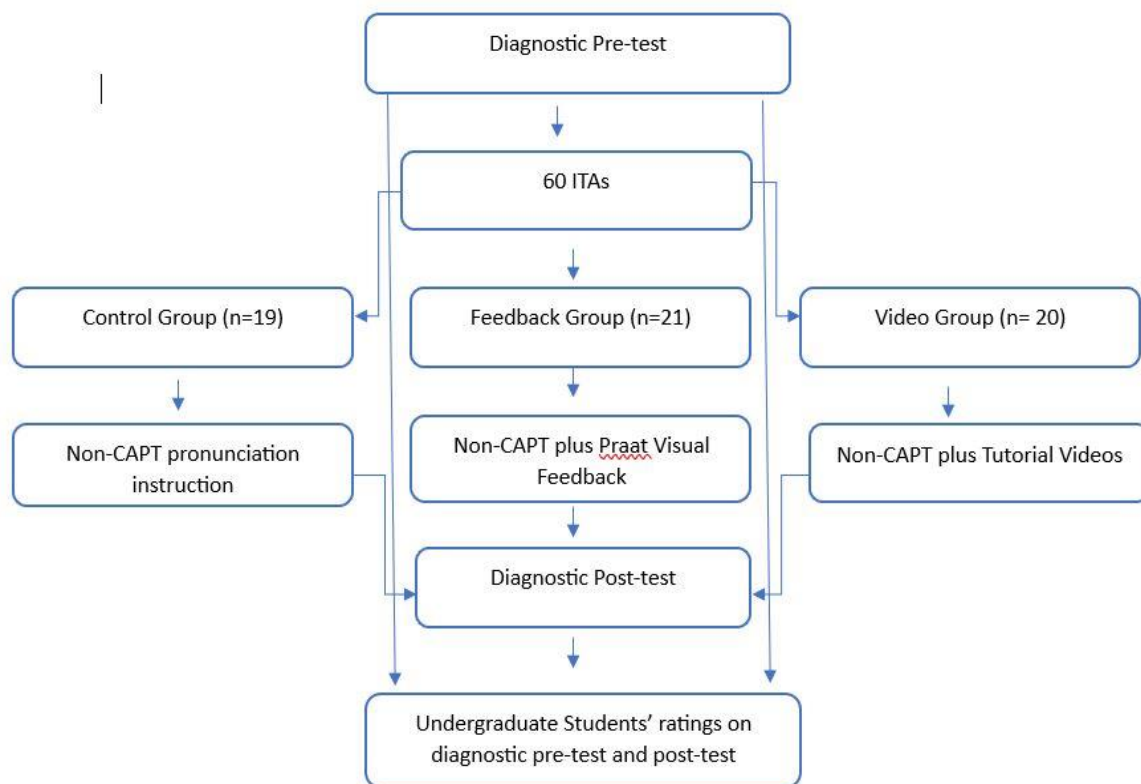
Figure 1 Design of the study

Firstly, all ITAs (n=60) recorded speech samples for the diagnostic pre-tests. Secondly, they were randomly assigned to three groups: the control group (n=19), the feedback group (n=21), and the video group (n=20). All these three groups received non-CAPT pronunciation instruction (Celce-

Murcia,2010) for one 30-minute session each week. Both experimental groups received additional pronunciation instruction immediately after their non-CAPT instructions. The video group was provided with tutorial videos for the week's topic while the feedback group received visual feedback on their suprasegmental production through the Praat program (more detail on instruction and treatments is below). Finally, all the participants took the diagnostic post-tests and had their speech recorded. Finally, All the speech files were coded, randomized, and incorporated into Qualtrics surveys and were distributed among 206 native undergraduate students to rate the speech files for their comprehensibility. The following section will describe the procedural steps taken in this study.

3.4.1 Diagnostic pre-test

After having their informed consent forms signed, ITAs received a conversation audio file. They listened to it and prepared to speak and record the summary of the conversion story. On another day in the first week of the experiment, they took the second diagnostic test which was 5 short sentence oral productions. They were given enough time to study the short context (1-3 lines) for each target sentence and once they were ready, they read that sentence aloud and had their voice recorded. Once they were done with this diagnostic test, they were randomly assigned to three groups, one control group (n=19) and two experimental groups (n= 21+20).

3.4.2 Suprasegmental instructions

The total instruction treatment took 6 weeks, a 45-minute pronunciation instruction each week. All groups in this study experienced non-CAPT pronunciation instruction for suprasegmentals. This instruction followed the principles of the book *A Reference for Teachers of English to Speakers of Other Languages* (Celce-Murcia, 2010). Relevant exercises were given during teaching for all

groups, and they were expected to complete them at home. However, this study did not focus on data about time spent on tasks or data on effective task completion but ITAs were trusted in stating that they had reviewed the material and completed the activities and were encouraged to express any questions they had.

3.4.3 Non-CAPT Suprasegmental Instruction

The control group and both experimental groups had shared pronunciation instruction supplied with relevant exercises. In each session, they were exposed to instructional materials in this order: identification, description, controlled and guided practice, and communicative exercise (Celce Murcia 2010). In other words, they were first warmed up with listening and recognizing the target suprasegmental features then the teacher explained the target suprasegmental lesson for that week. The teacher's description contained the rules and examples of word stress, rhythm, and intonation. Then the teacher provided the chance for the learners to produce those target suprasegmental features through easy-to-difficult exercises designed with controlled and guided practice and finally, they were taken to more independent production practice through communicative exercises. For example, for teaching word stress, the teacher read a list of campus and academic words from the book *Delta's Key to the TOEFL iBT*. ITAs listened to the teacher and identified which syllable in every word was longer, louder, and higher in pitch. Sometimes they showed their understanding by moving their hands up and down and sometimes they categorized the words in a table based on the number of stressed syllables. Then, the teacher explained the word stress rules to them. For instance, he explained some two-syllable words function as both verbs and nouns. If they are nouns, the first syllable takes the stress but if they are verbs stress shifts into the second syllable. Then, through controlled and guided practice, he asks them to fill out the following blanks using the word in parentheses: *if you have a parking …. , they ….. you to park your car on campus*

(permit) and then asked them to read their sentences aloud. For their communicative practice, they were invited to play a game while having to use some target words. For example, Students drew an index card with professions (e.g politicians) or academic field names on it. Then they produced as many derivative words (policy, political, politics) as they could. Each derivative word they produced brought them one point. Then they defined one or two words (e.g *a politician is a person who is professionally involved in politics*) and for each definition, they saved one point. This game was done with time limitations. At the end, the teacher gave them feedback on their word stress.

As noted above, the control group only participated in the instructional sessions. The two experimental groups additionally received CAPT instruction.

3.4.5 Tutorial videos

 The video group were given access to 8 tutorial videos, two for word stress, two for prominent words & rhythm, two for intonation, and two tutorial videos as lesson reviews. Each pair of videos was released to the ITAs one week before the relevant suprasegmental lesson was taught to them so that they could have enough time to see the video before their class. In the class, the teacher first taught the non-CAPT lesson in 25-30 minutes, which was common for all groups. Then he spent 5 minutes answering students' questions about the videos that they had watched, and the rest of the class time was spent watching the video for the following lesson. This was done on purpose to ensure that they watched the videos more than one time before they came to class. The social media app Telegram was used for this group for two purposes: first, the videos were uploaded for them on this platform, and second, they used the platform to report for each video that they had watched at least once before the relevant instructional session.

3.4.6 Feedback Group

In a 10–15-minute session, the feedback group was trained to install the Praat program after their diagnostic test and before their first session depending on their convenience, and they received instructions on how to work with key features of this software program. They learned how to record and display the acoustic features (e.g., pitch contours, intensity, and wavelength) of their oral productions. For example, they were told pitch contours are the visual representation of intonation showing the rise and fall of frequency and wavelengths can represent the duration of stressed syllables in a word (Figure 2)
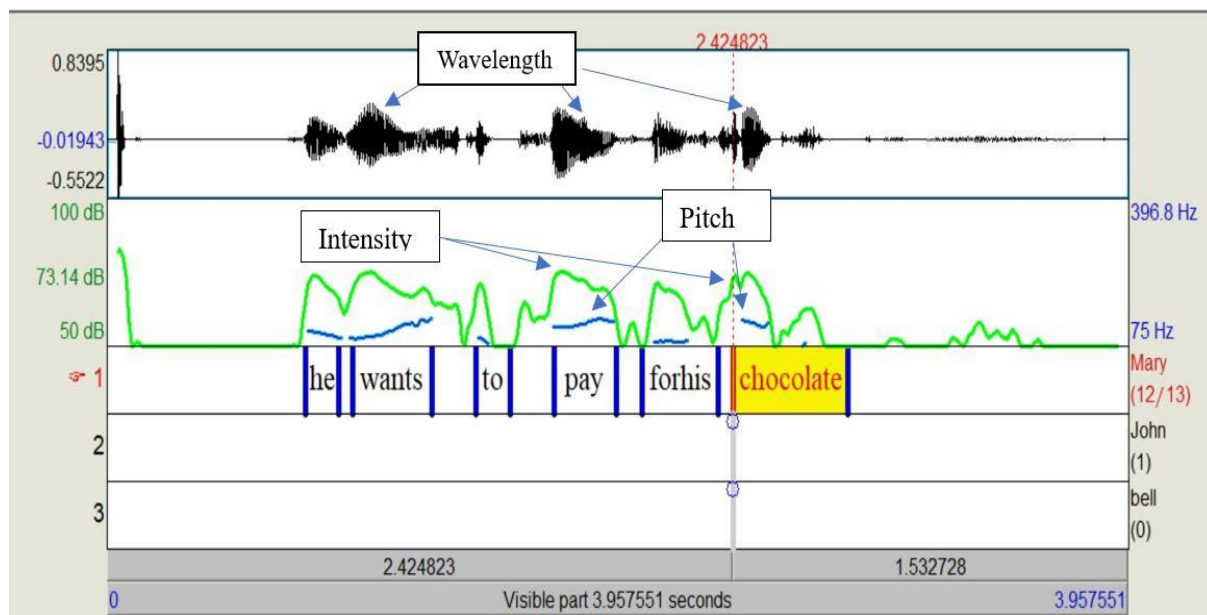
Figure 2 representation of wavelength, intensity and pitch of stressed syllables in Praat

ITAs first received non-CAPT pronunciation instructions for 25 minutes and then they discussed and analyzed several examples of that day's lesson besides presenting and sharing the visual feedback from their homework. For example, for their word stress lesson, they were given speech samples for several of the examples they learned in their nonCAPT lesson. They were supposed to review the word stress rules they had learned then orally produce the words and finally compare and juxtapose their production visuals with those of a speech model over the same utterances

(Figure 3) and then they shared their performance with their teacher and classmates in the following session. Figure 3 shows the attempt of an ITA to compare the visuals for the word "industry". Most of the participants orally reported they did not know the correct stressed syllable of this word until they saw the visual of the speech model. The left side of this figure shows the dashboard through which the speech files can be either directly recorded (ITA's speech) or can be uploaded from their computer (speech samples). Then, they made the visuals appear through the view and edit option. There are two visuals in this figure which show just the wavelength (duration) and blue line pitch (frequency) of the word syllables, but the word "industry" is added in this figure for clarity. The upper visual on the right side is what the ITA produced, and the lower one is the sample visual.
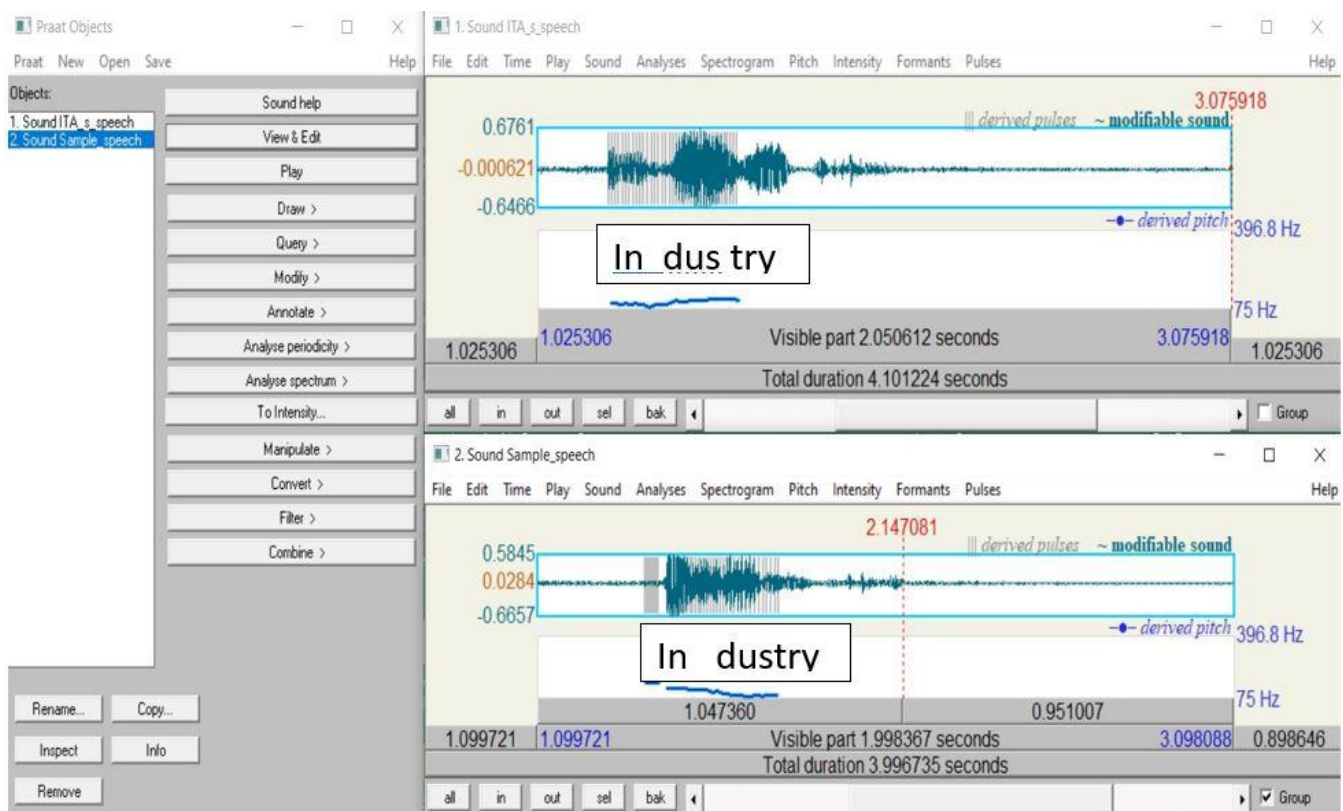
Figure 3 Comparison of an ITA's visual with that of a speech sample

3.4.7 Diagnostic post-test

On the last week of the experiment, all ITAs from the control group and the experimental groups were asked to take the same diagnostic tests that they had for their pretest. They all had their voices recorded.

3.4.8 Native speaker undergraduate ratings

Request emails were sent to the civil engineering, mechanical engineering, and English departments and their professors, requesting their assistance in distributing the surveys among their students for their rating. Freshman and sophomore native English speaker undergraduate students were recruited through their school departments. The surveys were distributed to their emails. On the first page of the survey, they were given their informed consent then they filled out a 1-page background questionnaire and finally they answered the survey questions. Each survey took around 12 minutes for the students to complete answering. They listened to each speech file and rated the comprehensibility of that production on a 9- point Likert scale.

3.4.9 Pilot study

Prior to conducting the main study, a pilot study was done with two Persian ITA participants in each of the three groups (control, feedback, tutorial videos; n=6) to ascertain the practicality of this research. These ITAs were recruited from a northeastern university and their speech files were distributed among 31 undergraduate students in a southern university to rate them for speech comprehensibility. The result of that study, which is discussed in the following chapter was presented at AMTESOL 2022 conference. The only change that was provided in the main study after this pilot study was reducing the duration of each tutorial video from 20 minutes to 15 minutes while still maintaining the essence of the videos. This decision was made because ITAs in the pilot study mentioned that they preferred the videos to be shorter.

3.4.10 Data Analysis

After ITAs' pre- and post-test speech comprehensibility ratings were collected through Qualtrics surveys, the data were prepared for analysis. To understand the effect of CAPT on the speech comprehensibility improvement of ITAs, descriptive statistics were performed first to calculate the mean and standard deviation scores for each of the control, feedback, and video group. Then inferential statistics were conducted through repeated analysis of variance (ANOVA) and where appropriate, post hoc analyses were carried out to determine the specific differences between them.

Chapter4:

Result

The following chapter presents the statistical analysis for each research question for the current

study.

3. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve
   ITAs' short-sentence production as rated by native English undergraduate students?

4. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve
   ITAs' spontaneous speech production as rated by native English undergraduate students?

4.1 Pilot study

Since the pilot study was carried out on a limited scale, inferential statistics were not used. The

descriptive data are described below.

Table 4 Descriptive analysis for short sentence production

	N	Pretest		Post-test	
		Mean	**SD**	**Mean**	**SD**
Control	2	7.08	0.66	7.21	1.2
Feedback	2	7.37	1.27	8.14	0.9
Video	2	7.27	0.95	7.43	0.09

Table 5 Descriptive analysis for spontaneous production

	N	Pretest		Post-test	
		Mean	**SD**	**Mean**	**SD**
Control	2	7.67	1.33	7.42	0.65
Feedback	2	7.79	0.14	7.71	0.83
Video	2	6.44	1.01	7.76	0.16

The pilot study provided confidence about the feasibility of the research approach. In this small study, the video group improved, while the feedback and control group did not, as shown in figures 4 and 5.
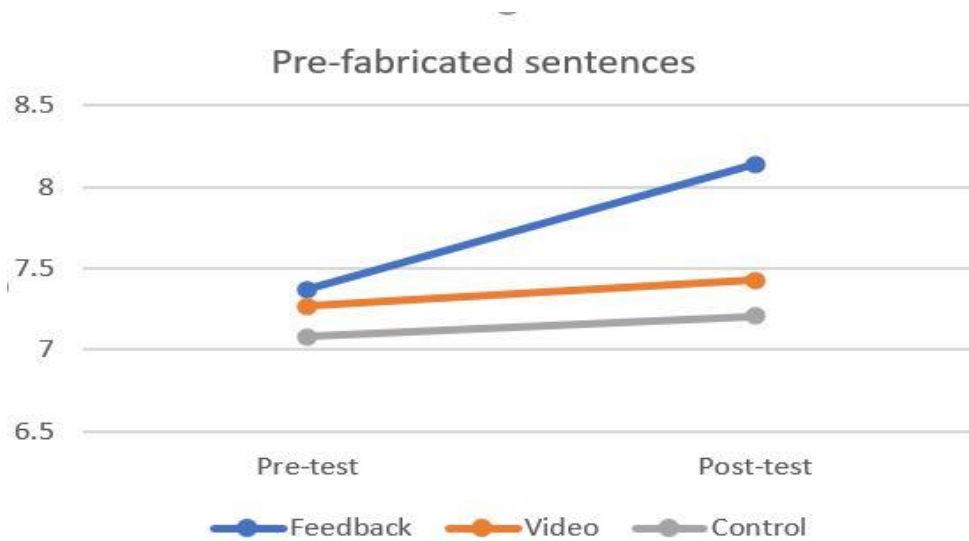
Figure 4 Performance on short sentence production
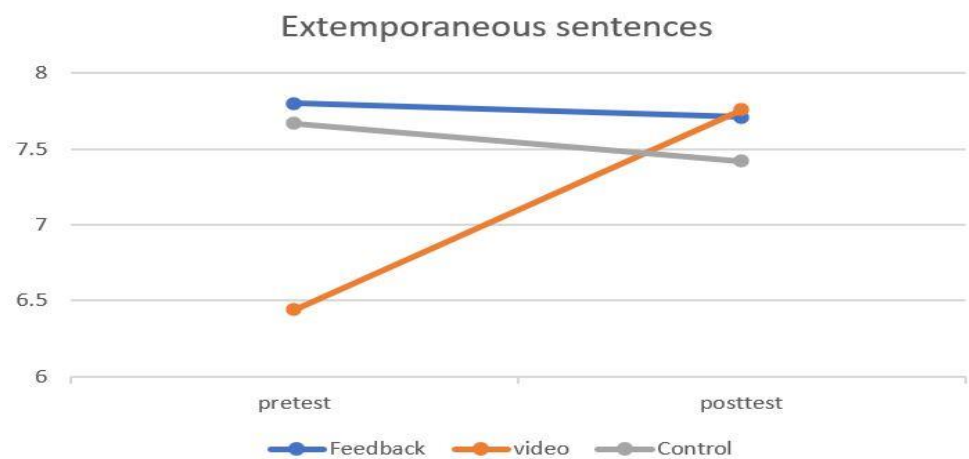
Figure 5 Performance of ITAs on spontaneous production

Descriptive Statistics

The main study was done over 6 weeks. The speech samples were collected and randomized into qualtrix surveys and were distributed among native English undergraduate students to collect their rating data. Descriptive statistics for the final data set are shown in Table 6

Table 6 Descriptive Statistics

	Sentence production pretest	Sentence production posttest	Spontaneous production pretest	Spontaneous production posttest
Mean	7.129	7.529	6.575	7.244
Std. Deviation	0.755	0.610	0.758	0.532
Shapiro-Wilk	0.871	0.953	0.982	0.969
P-value of Shapiro-Wilk	< .001	0.021	0.538	0.130
Minimum	4.500	6.100	4.800	5.700
Maximum	8.500	8.490	8.100	8.110

As shown above, ITAs' speech comprehensibility in sentence oral production increased from the pretest (M=7.12, SD=0.75) to the posttest (M=7.52, SD=0.61) as did their spontaneous speech from the pretest (M=6.57, SD=0.75) to the post (M=7.24, SD=0.53).

Inferential Statistics

In general, omnibus testing (where all the data is analyzed in a single statistical model) is preferred if possible. In this case, repeated measure MANOVA would be used. First, data were examined to see if they met the assumptions of repeated measures MANOVA The following table shows the criteria for this assumption juxtaposed with the quality of the current data.

Table 7 Confirming the suitability of the repeated measures MANOVA

Confirmation	Criteria assumptions for repeated measures MANOVA	Current data
Agree	Two or more interval-dependent variables	pre and post-tests for sentence and spontaneous oral production
Agree	The Independent variable is categorical	groups are categorical variable
Agree	Adequate sample size	There are more participants per group than the number of dependent variables. The dependent variables are 4
Agree	No univariate outliers	Only 3 in the whole data set (of 240 values)

There were two more assumptions, multicollinearity, and multivariate linearity, that needed to be tested before choosing MANOVA as a good fit for data analysis. For this purpose, Pearson correlation analysis was used to determine if the data were linearly related at each level of the independent variable. The results are shown in tables 8-10 below.

Table 8 Pearson's Correlations for the Control Group

Variable		Sentence production pretest	Sentence production posttest	Spontaneous production pretest	Spontaneous production posttest
1. sentence production pretest	Pearson's r	—			
	p-value	—			
2. Sentence production posttest	Pearson's r	0.298	—		
	p-value	0.215	—		

57

Table 8 Pearson's Correlations for the Control Group

Variable		Sentence production pretest	Sentence production posttest	Spontaneous production pretest	Spontaneous production posttest
3. Spontaneous production pretest	Pearson's r	0.108	0.042	—	
	p-value	0.661	0.863	—	
4. Spontaneous production posttest	Pearson's r	0.241	0.335	0.375	—
	p-value	0.320	0.160	0.114	—

Table 9 Correlations for the Feedback Group

Variable		Sentence production pretest	Sentence production posttest	Spontaneous production pretest	Spontaneous production posttest
1. sentence production pretest	Pearson's r	—			
	p-value	—			
2. Sentence production posttest	Pearson's r	**0.552**	—		
	p-value	0.010	—		
3. Spontaneous production pretest	Pearson's r	**0.580**	0.248	—	
	p-value	0.006	0.278	—	
4. Spontaneous production posttest	Pearson's r	0.074	0.357	0.271	—
	p-value	0.751	0.112	0.235	—

Table 10 Correlations for the Video Group

Variable		Sentence production pretest	Sentence production posttest	Spontaneous production pretest	Spontaneous production posttest
1. sentence production pretest	Pearson's r	—			
	p-value	—			
2. Sentence production posttest	Pearson's r	0.360	—		
	p-value	0.119	—		
3. Spontaneous production pretest	Pearson's r	0.172	**0.557**	—	
	p-value	0.467	0.011	—	
4. Spontaneous production posttest	Pearson's r	0.233	**0.559**	0.395	—
	p-value	0.322	0.010	0.085	—

Significant correlations are in bold. While there was no evidence of multicollinearity (defined as correlations above .90), there were also very few linearly related distributions, as required for a MANOVA. Based on this, the decision was made to conduct separate repeated measures ANOVAs models for (1) sentence production and (2) spontaneous production.

Using repeated measures ANOVA

4.3.1 Analysis of oral sentence production

Alternatively, repeated measure ANOVA was used for this analysis. For this purpose, the descriptive analysis of the data for sentence production and spontaneous speech was run and the results are as follows.

Table 11 Descriptive Statistics

| | N | Pretest | | Post-test | |
		Mean	SD	Mean	SD
Control	19	7.33	0.79	7.32	0.56
Feedback	21	6.93	0.89	7.34	0.58
Video	20	7.15	0.50	7.92	0.50

 As the descriptive table for sentence production shows, the pretest-posttest ratings increased in the scores of both feedback groups from (M=6.93, SD=0.89) to (M=7.34, SD=0.58) and the video group from (M=7.14, SD=0.50) to (M= 7.92, SD=0.50), however, there was almost no change in the performance of control group from their pretest (M=7.32, SD=0.79) to their post-test (M=7.31 SD=0.56). Another representation of the trends is shown in figure 6.
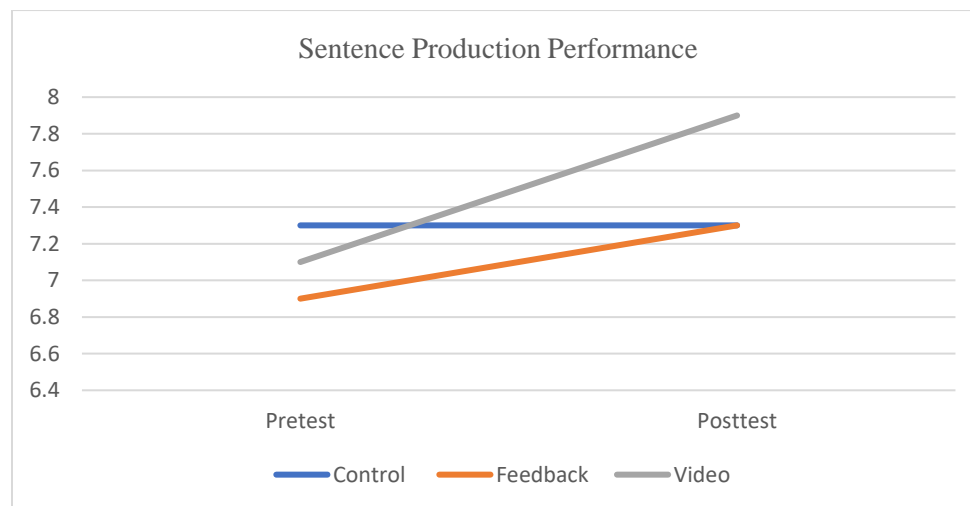
Figure 6 Sentence Production Performance

 As it is evident, both experimental groups had a rising trend with the video group higher in their performance. To see the significance of these differences within and between subject effects, ANOVA was run. The ANOVA included one within-subjects factor (Time) with two levels (pre-treatment and post-treatment) and one between-subjects factor (Group) with three levels (one for each treatment group).

Table 12 Within and between-subject effects

Within Subject effect	Cases	Sum of Squares	df	Mean Square	F	P	Effect size
	Time	4.62	1	4.62	17.79	<.001	0.238
	Time*Group	2.96	2	1.48	5.7	0.006	0.167
	Residuals	14.83	57	0.26			
Between Subject effect							
	Group	3.22	2	1.61	2.66	0.079	0.085
	Residuals	34.57	57	0.60			

The main factor effect for time (pre-test/post-test) was significant (p<.001) and had an effect size of 0.23, The interaction effect for the time x group was also significant (p<.006), the main factor for group approaches but does not reach significance. When we have a significant interaction effect, we don't interpret the main factor effects.

We used Levin's tests to see the homogeneity of variables. If the P value for the Levene test is greater than .05, then the variances are not significantly different from each other meaning that the equality assumption is met. As shown in Table 13, P values for both the sentence production pretest (>0.50) and sentence production post-test (>0.54) are larger than the standard (0.05) P value. This gives us some evidence that the variances are equal in the population.

Table 13 Test for Equality of Variances (Levene's)

	F	df 1	df 2	p
Sentence production pretest	0.701	2	57	0.500
Sentence production posttest	0.615	2	57	0.544

Since ANOVA for the interaction of the time* group produced a p-value (<0.006) less than our significance level, a post hoc test was used to find out where the difference truly came from.

*Table 14 Post Hoc comparison Group*Time*

		t	p_{holm}
Control, PreTest	Feedback, PreTest	1.885	0.561
	Video, PreTest	0.849	1.000
	Control, PostTest	0.032	1.000
	Feedback, PostTest	-0.079	1.000
	Video, PostTest	-2.827	0.069
Feedback, PreTest	Video, PreTest	-1.040	1.000
	Control, PostTest	-1.860	0.561
	Feedback, PostTest	-2.602	0.118
	Video, PostTest	-4.809	**< .001**
Video, PreTest	Control, PostTest	-0.824	1.000
	Feedback, PostTest	-0.951	1.000
	Video, PostTest	-4.806	**< .001**
Control, PostTest	Feedback, PostTest	-0.105	1.000
	Video, PostTest	-2.851	0.069
Feedback, PostTest	Video, PostTest	-2.818	0.069

As evident in the table above, the true difference can be seen across the pre-test to post-test for the video group (p <0.001), but this is not the case for the control or feedback group. In other words, the ITAs who received tutorial videos as their treatment had significant improvement in the comprehensibility of their sentence production.

4.3.1 Analysis of spontaneous speech production

To see the improvement of ITAs in their spontaneous speech comprehensibility, the descriptive analysis of the data was run for their spontaneous speech production.

Table 16 Descriptive Statistics

	N	Pretest		Post-test	
		Mean	SD	Mean	SD
Control	19	6.71	0.74	7.18	0.51
Feedback	21	6.41	0.81	7.00	0.52
Video	20	6.62	0.71	7.55	0.40

As the descriptive table for spontaneous production shows, the pretest-posttest ratings increased

in the scores of both feedback groups from (M=6.41, SD=0.81) to (M=7.00, SD=0.52) and the

video group from (M=6.62, SD=0.71) to (M= 7.55, SD=0.40). The mean score for the control

group also improved from their pretest (M=6.70, SD=0.74) to their post-test (M=7.18 SD=0.51).

The following figure shows the improvement in the performance of ITAS' spontaneous speech.
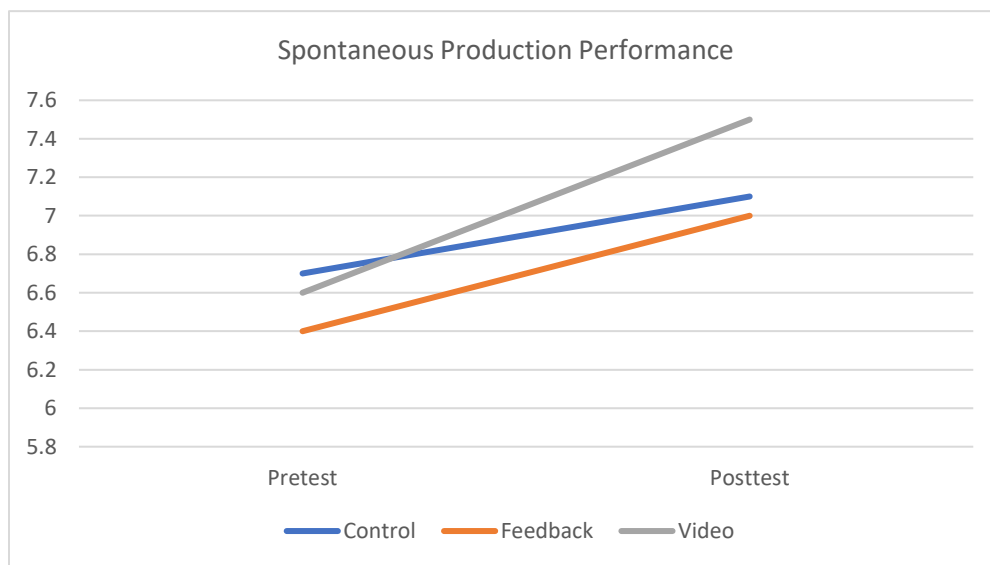
Figure 7 Spontaneous Production Performance

As is evident, all groups had a rising trend with the video group as the highest in their

performance. To see the significance of these differences within and between subject effects,

ANOVA was run.

Table 17 Within and between-subject effects

Within Subject effect	Cases	Sum of Squares	df	Mean Square	F	P	Effect size
	Time	13.32	1	13.32	47.09	<.001	0.45
	Time*Group	1.11	2	0.555	1.96	0.150	0.064
	Residuals	16.12	57	0.28			
Between Subject effect							
	Group	3.02	2	1.51	2.84	0.067	0.091
	Residuals	30.31	57	0.53			

The main factor effect for time (pre-test/post-test) was significant (p<.001) and had an effect size of 0.45, however, the interaction effect for the time x group is not significant (p<.15), the main factor for the group also approaches but does not reach significance.

Post hoc analysis (Table 18) shows that there are significant differences across pretest to posttest for both feedback (p <.001) and video groups (p <.001) However, the performance of the control group across pretest to post test approached but was not significant. Likewise, for the main factor effect of the group, the difference between the feedback and video groups approached, but did not reach significance, with the video group once again outperforming the feedback group.

Table 18 Post Hoc Comparisons- Group Time*

		Mean Difference	SE	t	P
Control, PreTest	Feedback, PreTest	028	0.20	1.40	0.71
	Video, PreTest	0.08	0.20	0.39	0.92
	Control, PostTest	-0.48	0.17	-2.80	0.06
	Feedback, PostTest	-0.29	0.20	-1.47	0.71
	Video, Post Test	-0.85	0.20	-04.18	<.001
Feedback, PreTest	Video, PreTest	-0.20	0.19	-1.02	0.92
	Control PostTest	-0.76	0.20	-3.79	0.003
	Feedback PostTest	-0.58	0.16	-3.55	0.008
	Video PostTest	-1.13	0.19	-5.71	<.001
Video PreTest	Control PostTest	-0.56	0.20	-2.75	0.06
	Feedback PostTest	-0.37	0.19	-1.89	0.42
	Video PostTest	-0.93	0.16	-5.55	<.001
Control Post Test	Feedback PostTest	0.18	0.20	0.91	0.92
	Video PostTest	-0.37	0.20	-1.81	0.43
Feedback PostTest	Video PostTest	-0.55	0.19	2.78	0.06

As shown in Table 19, Levene's test was used and there was a homogeneity of variances, for competence, p >. 05. This provides some evidence that the population's variances are equal.

Table 19 Test for Equality of Variances (Levene's)

	F	df 1	df 2	p
Sentence production pretest	1.08	2	57	0.34
Sentence production posttest	0.30	2	57	0.73

Summary

As for oral sentence production, both feedback and video group were rated higher in the post-test compared to their pretest but only the performance of the video group was significantly higher on the post-test. While all groups performed better in their posttest in terms of spontaneous speech production, the performance of the video group closely approached the level of statistical significance. This did not happen for the other groups. In conclusion, the data suggest that all three conditions led to improvements, but the video treatment outperformed the other two treatments.

The research questions posed for this project will be addressed in the following chapter. The next

chapter will likewise provide an interpretation of the most recent study findings.

Chapter 5

Discussion

The results of the research questions are explored in more detail in this chapter. A general discussion about overall results will be presented first, followed by an expansion on results considering other relevant studies' findings on this subject.

5.1 Summary of the findings

In this study, two main questions were addressed: 1. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' short-sentence production as rated by native English undergraduate students? and 2. Compared to instruction alone, does CAPT via visual feedback or tutorial videos improve ITAs' spontaneous speech production as rated by native English undergraduate students?

Although the discussion chapter covers all the findings of the current study, its content is presented in an order different from the research questions just for ease of presenting ideas in this section, I will present and discuss the results of the experiment, starting with the control group who received explicit instruction alone or non-CAPT instruction. Next, I will discuss the results from the first experimental group who were exposed to visual feedback and non-CAPT instruction, followed by a discussion of the findings from the second experimental group who received both tutorial videos and non-CAPT instruction for their instructional treatment. By presenting and discussing the data in this order, I hope to provide a clear and comprehensive understanding of the effects of the experimental manipulations on the outcome variables.

As noted in the previous chapter, all groups (except for the control group in their sentence production performance) demonstrated progress in the comprehensibility of their speech, with some groups experiencing greater improvement. This is illustrated by Figure 8 that almost all ITAs in the feedback and video group demonstrated progress in their speech comprehensibility to varying degrees.
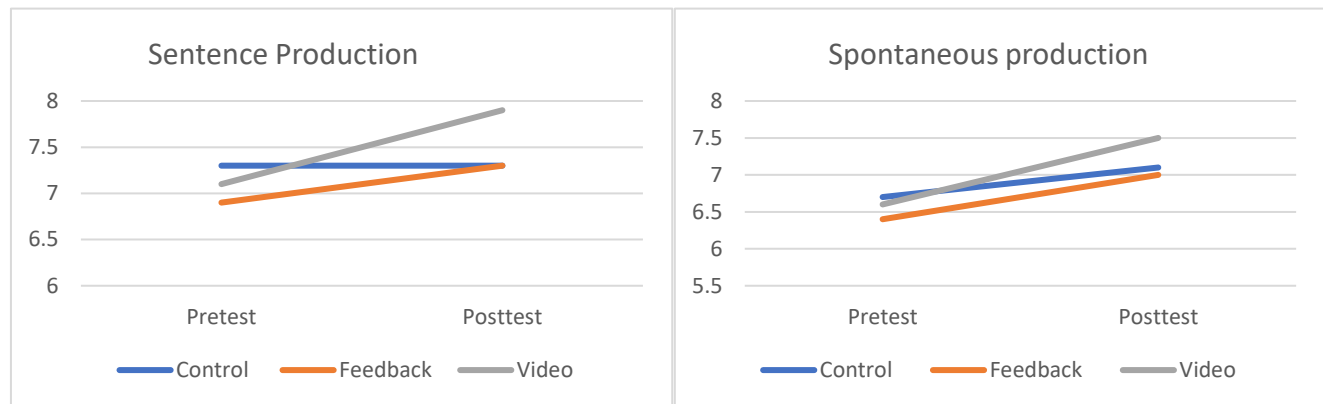
Figure 8 ITAs' comprehensibility improvement in experimental groups

Based on the figure shown above and the detailed explanations explored in the following sections in this chapter, it is evident that almost all groups showed improvement in their speech comprehensibility. These findings suggest that explicit instruction, in general, is an effective tool for pronunciation instruction. However, it is important to note that the magnitude and nature of improvement varied across the groups. Furthermore, the types of instruction approach used in the experimental groups, particularly for the video group, proved to be more effective in enhancing speech comprehensibility. Overall, these results indicate that explicit instruction can be a useful tool for improving pronunciation. This is similar to the findings of Gordon et al's (2013) study, which looked at the impact of short explicit vs. unexplicit and segmental vs. suprasegmental instructions on the speech comprehensibility of ESL learners. The findings demonstrated that

learners in the experimental group who were given explicit lessons on suprasegmental features had a significant improvement in their speech comprehensibility.

Moreover, the finding of the current study also supports the results of Saito & Saito's (2017) study where they investigated the effect of explicit teaching of suprasegmental features to Japanese EFL learners in a short course of six weeks. Being rated by the intuitive judgment of L1 speakers, the participants showed significant improvement in the comprehensibility of their oral production.

5.2 CAPT findings

All CAPT groups in this study had improvement in their speech comprehensibility either for their short sentences or for their spontaneous speech. The pretest-posttest ratings increased in the scores of feedback group from (M=6.93, SD=0.89) to (M=7.34, SD=0.58) for their sentence production and from (M=6.41, SD=0.81) to (M=7.00, SD=0.52) for their spontaneous speech. ANNOVA and then Post hoc analysis revealed that the difference between pretest and posttest was significant just for ITAs' spontaneous production.

Video group gained rising scores for their speech comprehensibility ratings across pretest-posttest which was from (M=7.14, SD=0.50) to (M= 7.92, SD=0.50) for their sentence production and from (M=6.62, SD=0.71) to (M= 7.55, SD=0.40) for their spontaneous production and their post hoc analysis revealed that all the differences, between pretest and posttest scores, for both sentence and spontaneous productions were significant.

One main reason for the significant improvement in CAPT groups can be motivational factors as several scholars (Reed & Levis, 2019) believe CAPT can make learning maximized and personalized. Therefore, ITAs in this study must have had the chance to select the exercises to have their needs met without being limited to any time, space, or teacher availability. As a result,

the control group, which in the current study was not exposed to CAPT, might not have done repeated practice and might have done the exercises just once, on the due date, rather than taking them progressively over the course of the week. Results also indicate that autonomy in learning plays a key role in pronunciation learning. This result had been demonstrated in other studies (e.g., Khoi and Aghabeig, 2009) that showed CAPT decreases the dependence on teachers. Khoii and Aghabeig (2009) specifically showed that the use of Rosetta Stone software gives Iranian EFL learners autonomy and motivation in learning pronunciation. Similarly, autonomy in receiving feedback is proved effective in learners' quality of pronunciation in some other studies (e.g Hincks, 2005) that demonstrated Swedish EFL learners who used Talk to Me software for their pronunciation feedback performed significantly better than those who received teachers' feedback. Learners in this study particularly reported a high level of motivation and engagement with computer-based activities. These findings suggest that the lack of improvement in the control group's sentence production or lack of significant improvement in their spontaneous production can be attributed to the absence of CAPT intervention provided for the experimental group. In other words, they did not have the Praat visuals to continue practicing and tried several times to produce a speech sound like the speech models, nor did they have access to the videos to watch them in their free time. Therefore, the lack of motivation (and therefore consistency in) ITAs' practice might be an explanation for why the speech comprehensibility of the control group in the current study did not improve.

5.3 Feedback group

ITAs in the feedback group improved in their speech comprehensibility both in sentence production and spontaneous production but their improvement was significant in spontaneous production with their average test scores increasing from   6.41(SD= 0.81) at the pretest to 7.00

(SD= 0.52) at the post-test, which represents a statistically significant improvement (p < 0.001). One possible explanation for why visual feedback for speech comprehensibility was not statistically significant in their sentence production can be the lack of instant feedback from Praat program. Studies like Bozorgian and Shamsi (2020) suggest that instant feedback plays an important role in the oral production of suprasegmental features. The EFL learners in Bozorgian and Shamsi's study received instant visual feedback on their emphasis (loudness), pitch (intonation), and rhythm (the rate of syllable production during speech) through MyET program. Faster feedback encourages and allows the learners to respond to the existing problem right away (Mason & Bruning, 2001). However, Praat program's slight complexity with visual feedback for the learners (Kolat & Morgan 2018; Wang & Young, 2015), might be the reason for why the participant's performance improvement on sentence production in the current study, did not reach significance. Moreover, the choice of data analysis method can play a crucial role in determining whether speech comprehensibility has been significantly improved. While the current study relied on the intuitive judgments of native English undergraduate students, it is important to note that a more objective analysis such as the use of ASR-based software to assess suprasegmental production, as done in the Bozorgian and Shamsi (2020) study, might have yielded different results.

In addition, findings suggest that reading aloud anxiety might have negatively affected ITAs' sentence reading performance in the feedback group. This result has been demonstrated by other studies (e.g. Zemni & Alrefaee, 2021; Zhou, 2017) which demonstrate uncertainty about pronunciation and obsession with the thought of making errors are among the main sources of anxiety. These findings are similar to the current study, particularly regarding the fact that ITAs, in their instructional treatment, monitored their own speech production acoustically on Praat and

had to try a couple of times to adjust their speech pitch, loudness, and intensity to produce qualities like those of the speech model. Finally, they had to prepare and apply their recently learned knowledge of prosody to reading sentences aloud, something which might have been accompanied by anxiety and therefore lower performance in speech comprehensibility.

However, the feedback group (from M=6.41, SD=0.81 to M=7.00, SD= 0.52) like the video group (from M=6.62, SD=0.71 to M=7.55, SD=0.40) both significantly improved (p<0.001) in their spontaneous production. These results are similar to Jian and Chung's (2021) study which implemented the use of videos and Praat contours. Similar to the results of the current study, Jian and Chung found Mandarin Chinese English learners' speech comprehensibility improved significantly after receiving suprasegmental instruction through CAPT as assessed by a one-minute presentation on a topic both prior to and after the instructional treatment. In both their study and the current study, the data related to the evaluation of speech comprehensibility were collected through native English speakers' ratings.

 Furthermore, the study's results suggest that the amount of time invested in improving pronunciation is a critical factor. This finding is consistent with prior research, such as Lima's (2021) study, which found that participants who did not dedicate sufficient time to practice were less likely to improve their speech comprehensibility. For example, Lima observed that only 4 out of 12 ITAs demonstrated improvement, and this was attributed to their failure to fully utilize the suprasegmental instruction they received. This aligns with the current study's results, as participants in the CAPT group had access to tools that provided them with greater autonomy and accessibility. As a result, they were more likely to watch the instructional videos and use Praat more frequently.

5.4 Video Group

  ITAs significantly improved in speech comprehensibility in video groups both at the sentence level ($p<.001$, $d=0.23$) and spontaneous oral production ($p<.001$, $d=0.45$). This result is in line with other studies (Hasan et al, 2018; Wang & Chen 2020) that emphasize tutorial videos' crucial role in L2 learners' academic progress and engagement. This observation is particularly consistent with the work of Wang and Chen (2020), which highlights the value of self-regulated learning that can be achieved by increasing learners' autonomy and flexibility. Likewise, the tutorial videos in the current study were readily available to the participants; they could easily click a link to watch them using either a computer or even their mobile devices. They could watch, and review, the videos whenever they wanted, and at whatever pace they chose. It is likely that this increased participants' engagement with the   tutorials.

The current study's findings suggest when learning anxiety decreases, learners may feel more motivated to engage with the lessons. This finding aligns with prior research such as Ebru Atak Damar's (2014) study which found that 83% of the participants believed learning through videos decreased their learning anxiety and increased their motivation in both the perception and production of word stress and intonation. This result is in agreement with the current study as ITAs could watch the tutorial videos at their own pace in a comfortable environment. They could pause and rewind the videos as many times as they needed to fully understand the concept without feeling rushed or embarrassed. Moreover, contrary to a traditional classroom where learners may feel pressure to perform in front of their peers or to keep up with the pace of the class, by watching the videos before the class, ITAs learned in a more private and low-pressure environment which can reduce stress and anxiety.

The study result also suggests when the learners find the content of the videos interesting and relevant to their needs, it activates their engagement meaning that it helps them to understand it better and retain it longer. This finding is consistent with previous research like Al-Domi's (2017) study which found that EFL learners who were exposed to authentic videos featuring the everyday life of the English language community reinforced their prosodic skills and motivation in their oral productions. This aligns with the current study as the videos used for ITAs' instructions featured campus-related content and examples, such as office hour short conversations or academic talk content.

The study result suggests that with using videos in blended learning, much of confusion and frustration that may arise, when learners feel rushed or overwhelmed in a traditional classroom, can be avoided and class time can be invested for other activities. This finding is consistent with previous research like Zhang et al's (2016) study in which flipped learning through videos is shown as a time-saving opportunity for interactive, collaborative, and applied learning activities. Similarly, The ITAs in the video group of the current study had the chance to watch the videos before their class. They prepared themselves to ask their questions from their instructor during the class which gave them the upper hand over the control group who faced all instructions for the first time and might have to keep their questions so that they could find a chance to ask them later, or they might have needed to tolerate any confusion until it was resolved by further examples or their peer questions.

Blended learning provided a rich, motivating, and ubiquitous learning environment for the learners (Caulfield, 2011). This method prevents learners' isolation and reduces the number of dropouts (Victoria Lopez-Perez et al. (2011), as there was no dropout report neither in the current study nor in the studies discussed above. This means that learners were well engaged and therefore they had

positive feelings about their learning (Zhang & Zhu, 2018; Linder, 2017), as it was shown in all relevant mixed methods discussed above. One of the reasons learners do not follow lessons or sessions is that they lose track of the lessons either due to their unresolved confusion or because of their life commitments that make them skip the classes. However, with technology-assisted language learning, particularly with blended learning, their access to and awareness of the courses rises, and they maintain motivation. Blended learning can be one possible explanation for the reason the video group in the current study produced better speech comprehensibility scores in both sentence and spontaneous oral production compared to the feedback group.

In conclusion, data presented in the current study demonstrated that explicit instruction of suprasegmentals can improve the speech comprehensibility of ITAs. According to the findings such improvement proved significant if instructors employ CAPT, particularly tutorial videos and visual feedback, in their pronunciation teaching. The following chapter will provide an overall conclusion, pedagogical implication, limitations of this study, and recommendations for further research.

Chapter 6

Conclusion

This chapter sums up the dissertation with a summary of the key findings, along with theoretical and pedagogical implications, limitations, potential future research ideas, and the study's contribution. The theoretical implications go over how this study advances our knowledge of the pronunciation teaching trend and speech comprehensibility principle. Developers of curricula and instructional materials, teachers, ITA training programs, and the general community of language learners are all addressed by the pedagogical implications. While the perspective for future research indicates various opportunities for expanding pronunciation instruction, the limitation of the study identifies the deficiency of the current study. There will also be a discussion of what this study contributes to the corpus of prior research.

6.1 Summary of the Dissertation

This experimental study followed a pretest-posttest design to explore the effect of digital instructions or CAPT in the form of visual feedback and tutorial videos on ITAs' speech comprehensibility. 60 Persian ITAs were the participants who received pronunciation instructions in 3 groups of control, visual feedback, and video groups and 169 native English undergraduate students were the participants who rated the pre-treatment and post-treatment speech samples of ITAs for their comprehensibility or ease of understanding. For ITAs' pronunciation instruction that was in 6 weeks, all 3 groups received non-CAPT instructions which covered word stress, rhythm, and intonation but the other two experimental groups received extra training in form of visual feedback for one of them and tutorial videos for the other. All speech files from ITAs' pretest and posttest were randomized, coded, and incorporated into Qualtrics surveys. Then they were distributed among Native English undergraduate students to give their ratings on the speech

comprehensibility of ITAs' speech files. After the data were collected and analyzed, significant differences were found in ITAs' speech comprehensibility for short sentences and spontaneous oral production in the video group and spontaneous oral production in the visual feedback group. These findings suggest that CAPT can be an efficient medium of instruction in improving the speech comprehensibility of second language learners in general and field-specific occupations.

6.2 Theoretical Implications

Teaching pronunciation in various language teaching methodologies and approaches has undergone constant changes. For example, while the Reform movement and audiolingual method elevated pronunciation teaching to its maximum level of significance, others like the cognitive approach almost completely devalued it. Despite all these changeable trends, pronunciation teaching has mostly remained the same since the 1990s. Two important debates that have been addressed by scholars during this time have been nativeness vs. intelligibility and segmental. vs. suprasegmentals. The advocates of the nativeness principle believe that pronunciation teaching should be for the purpose of achieving a nativelike proficiency of the target sound system. However, the proponents of the intelligibility principle revisit the goal of pronunciation teaching/learning by maintaining speakers should just be easy to understand. They have several justifications for such a viewpoint, for example, the population of non-native speakers (750 million) is larger than English native speakers (300 million) (Morley 1991), and acquiring a native-like pronunciation is neither desired nor feasible for all students because all foreign or second language learners just need a functional comprehensible speech (Ketabi & Saeb, 2015).The question of whether segmental (single sounds) or suprasegmental (intensity patterns, location of stress, and rhythm in spoken language) is more crucial in generating comprehensible speech has been debated for a long time in pronunciation instruction. Some scholars believe that segmentals

are more important (e.g., Jenkins, 2002; Saito & Lyster 2012; Zielinski, 2015), while others believe suprasegmentals should be given priority in pronunciation teaching (e.g., Kang et al, 2010; Issacs & Trofimovich, 2012; Derwing et al, 2014). Considering the two above-mentioned pronunciation debates, the current study supports the intelligibility principle by focusing on the comprehensible speech of ITAs and emphasizes the role of suprasegmentals through CAPT to achieve such comprehensible speech.

This research also highlights the significance of blended learning. Technology has created significant changes not only in face-to-face but also in virtual classes leading to changes in theory and pedagogy. The foundation for blended learning, which is defined as combining several elements (teaching and/or technology) in an effort to improve the teaching and learning process (Marques, 2011, p. 83), was created by these advancements and the ease of gaining access to broadband Internet. The Flipped Classroom is a technique that is gaining popularity within Blended Learning approaches. In a traditional classroom, the teacher typically uses class time to discuss theoretical and conceptual material while leaving practices and exercises for after-school activities. According to the tenets of the "flipped classroom" (Bergmann & Sams, 2014), students complete the theoretical portion of the course at home utilizing multimedia materials, while the instructor oversees practical activities and exercises during class time. The current research is supported by this theory because computer technology is used for both experimental groups and flipped learning is particularly used for the video group where participants watched the content of the upcoming sessions in advance. Because the video group was particularly successful in enhancing the comprehensibility of their speech, these findings provide support for the use of flipped classroom approaches to pronunciation teaching.

6.3 Pedagogical Implication

One of the pedagogical implications of this study (for curriculum developers also for ESL/EFL teachers in general and pronunciation teachers in particular) is the need for pronunciation training to be explicitly included in L2 English pronunciation lessons, giving students knowledge that is both concrete and descriptive. For this purpose, they can include visual reinforcement for pronunciation teaching, as it was in this study through featuring acoustic analysis of learners' speech and through tutorial videos. Such visual reinforcement can take other forms like charts, diagrams, use of flashcards, or wall charts to name a few.

Foote et al. (2016) discovered that teachers were not paying as much attention to pronunciation as they thought they were and that there was a glaring absence of suprasegmental training. Therefore, when creating their pronunciation curriculum, teachers should focus more on speech comprehensibility than correctness. To do this, educators must concentrate on the sounds and prosodic qualities (such as stress, intonation, and rhythm) that contribute to comprehensibility. According to the literature (e.g., Hahn, 2004; Kang et al., 2010), learners' capacity for communication is greatly hindered when they are unable to control suprasegmental aspects well. Prosodic elements must be viewed as a crucial part of oral proficiency and what English language learners should be taught.

Moreover, for the purpose of teaching pronunciation in CAPT, a communicative framework is preferable. This method of teaching is similar to a developmental process, which is more effective in preparing students to acquire and use their information in speaking. As pronunciation training progresses, it should shift from controlled (e.g., listen and repeat) to more spontaneous (e.g., role-playing) activities (Celce-Murcia et al, 2010) as it was in this study. All three groups in this study received instruction in each module of word stress, rhythm, and intonation, starting with

79

cognitively easy exercises (listening and recognizing), moving on to more difficult exercises (spontaneous exercises, such as solving a problem through a set of questions and answers with an interlocutor while the teacher was monitoring them). Since significant results were achieved for CAPT groups, such a way of treatment can be insightful for those who develop digital instructional materials.

The findings of this study have some implications for the instruction of English pronunciation and the pedagogical use of CAPT. As evidenced by the study's findings, CAPT appears to be a useful technique for facilitating L2 pronunciation. Utilizing this application, language instructors can create lessons that encourage L2 pronunciation. These tasks can be completed both within and outside of the classroom. Students can work through these exercises at their own pace and in accordance with their degree of proficiency. However, additional consideration should be given when using CAPT. Given the variety of computer software programs on the market, it is essential that teachers know the benefits and drawbacks of the programs and select the one that will best serve their students' needs and help them develop their communication and English pronunciation skills. As further clarification, among several reasons, two of them are critical to consider for why Praat visual feedback and tutorial videos were used in this study. ITAs are competent graduate students with a large number of them coming from engineering programs therefore it is easy for them to learn and work with Praat which is not recommended for young ESL learners as they might not have computer skills to handle it. Regarding the tutorial videos which produced the best significant results in this study, they are the most user-friendly teaching tools that are omnipresent for ITAs who live a hectic life on campus and cannot spare any time for in-person pronunciation classes.

Another pedagogical implication is the idea of blended (flipped) learning that was applied by using the tutorial videos in this study. The learners had the chance to watch the videos of their new lesson before they received that instruction in their in-person class. This was helpful because it gave confidence and motivation to the learners to attend the class with an open mind. It also eased the teaching for the instructor, especially in terms of saving time by emphasizing the lesson's most crucial components. Establishing and equipping pronunciation labs to enhance ITAs' speech comprehensibility is worthwhile in higher educational settings where many students are accepted each year as graduate students who teach undergraduate courses for their graduate assistantship. It is also worthwhile hiring linguists (phonologists) to supervise these labs and teach pertinent courses. This initiative is essential because it can enhance communication between undergraduate students and ITAs and benefits departments and the university overall, especially in light of the widespread knowledge of the "Oh No I'm Having an International Professor" syndrome among American undergraduate students and international teaching assistants (Rubin, 1998).

6.4 Limitations

There are several limitations to the research that must be addressed in this section. First, the ITA participants for this study shared the same L1, Persian language. Although this control decision was made for the purpose of strengthening the result, it does not reflect accurately the real-world ITAs because L1 background is diverse among ITAs who might respond to this treatment differently. ITAs were not screened for their English proficiency, as it was assumed that they had already passed the TOEFL or its equivalent test, which is a required document for their admission to graduate school. However, passing the TOEFL or its equivalent test does not necessarily indicate that all ITAs have the same level of English proficiency, as some graduate programs may mandate a certain skill score for their admission. Considering the variation in English proficiency

among ITAs could have an impact on the results of the study, and the findings might differ if this factor were taken into account. Moreover, ITAs play different roles on campus, which could impact their level of interaction with undergraduate students. Some ITAs work in the lab, while others assist professors by grading, and some even give lectures or act as the main instructor in class. This study did not control the varied experiences of ITAs, including their level of interaction with undergraduate students. This lack of control could have influenced the results of the study.

Although they have many benefits, CAPT software programs have significant drawbacks. The reason is that educational values are chosen over technological advancement (Neri et al 2002). Praat was chosen for this study because it produces visual outcomes that are more sophisticated and polished than those produced by any other application. This program has a number of benefits, including the ability to upload external speech files, record voice, and compare visuals from various speech files on it. It also comes with a range of coloring choices and zooming-in and-out options for better examination of utterances. However, Praat has its own limitation; it is not instant. It takes time and a few mouse clicks to upload a speech file, compare the visuals, and record voice on it and users sometimes have to adjust the settings on it.

Another drawback is that Praat occasionally doesn't provide the same visual for appropriately pronounced utterances produced by different speakers. Since the physical quality of the pitch varies from person to person, this is largely an issue with pitch representation. Therefore, users are readily fooled into rerecording and even going out of their way to approximate their oral production to that of a sample if pronunciation teachers do not make them aware of this issue.

The instruction session's maximum of four sessions was another restriction. The reason was the hectic schedules of ITAs, who in addition to their research projects had to manage tasks related to their undergraduate students and prepare for their courses. Even though the study's instructor

meticulously planned the sessions and went through all pronunciation lessons, some students could have needed more practice time to completely grasp the new lessons and apply them in their speech. It's possible that this was the case with the control group, who only got non-CAPT pronunciation coaching. Their short-sentence performance did not significantly improve, and their performance did not improve at all during spontaneous speech.

Another limitation is the negative attitude that native English speakers might have toward accented speech (Shuck, 2004). The raters' bias was not under control in this study. Therefore, the native English-speaking undergraduate students who rated the speech files may have had negative opinions toward the non-native accents they listened to, which may have affected their judgment. The outcomes of the speech ratings were not objective; rather, they were subjective judgments. Since the real-world audience of ITAs, undergraduates, are accent-biased, an objective assessment of speech comprehensibility was necessary for this study. Despite the limitations that were discussed above, this study is insightful in terms of ideas for future research.

6.5 Recommendation for future research

According to the current study, developing CAPT courses for ITAs with the goal of enhancing their speech comprehensibility has the potential to provide positive outcomes. Future researchers can employ certain tasks for investigations on suprasegmental explicit instruction and speech comprehensibility. For example, the instructional intervention could be expanded by creating more examples of suprasegmental features, more tutorial videos, and by lengthening the time and increasing the number of pronunciation instruction sessions in order to achieve more extensive and significant results for the speech comprehensibility of ITAs in the future research.

As was already noted, the findings of this research were based on undergraduate students' intuitive judgment. In order to more accurately measure the progress in speech comprehensibility of ITAs, researchers might in the future examine the acoustic analysis of their pre-and post-intervention oral productions.

Since the current study showed promise in short-sentence and spontaneous oral productions, future research can investigate the long-term impact of explicit suprasegmental instructions via CAPT on speech comprehensibility.

There are now at least four categories of English language learners that need extra help with pronunciation because their oral communication requirements call for a high degree of intelligibility (Morley 1981 mentioned in Celce-Murcia,2010) and therefore comprehensibility. ITAs are one of these categories that were addressed in this study so future research can address the need of other professions for speech comprehensibility. Those professions are foreign-born professionals working in business and industry in English-speaking countries, business professionals and diplomats who need to use English as their working language, and refugees (adult and adolescent) enrolled in resettlement and career development programs who want to go to English-speaking countries. Celce-Murcia (2010) adds two more professions to this list: teachers of English as a foreign language who are not native English speakers and who anticipate becoming their students' primary role models and sources of information in the language, and those employed in non-English-speaking countries as tour guides, waiters, hotel staff, customs officers, and similar positions who utilize English to communicate with foreign tourists.

6.6 Contribution of the Study

Considering the effort undertaken and the questions that were answered in this research, this study has some contributions to offer. First, it fills a gap in research because, to the best of the author's knowledge, no study has ever been carried out with a such large number of ITAs to assess their speech comprehensibility through explicit CAPT instructions. As a clarification, explicit instruction in this study included description and awareness for the students in terms of what word stress, rhythm, and intonation are and why they are important and then the participants were taught through a communicative framework of pronunciation instruction that is suggested by the most recent pronunciation scholars (Celce-Murcia, 2010).

This study is insightful for curriculum developers and syllabus designers. Recent research has shown that pronunciation teaching is still neglected in L2 classrooms (Foote et al., 2016), despite the well-documented value of pronunciation skills for effective communication in the L2 (Derwing & Munro, 2015). There are several reasons for this issue: There is a widespread dearth of structured pedagogical resources for teaching pronunciation, which serves to further marginalize pronunciation instruction in second-language classrooms (Lee, Plonsky & Saito, 2019). Also, due to factors including a lack of time or a lack of teacher preparation to teach pronunciation, instruction is frequently ignored in L2 classrooms (Darcy, 2018; Foote, Holtby, & Derwing, 2011). This study has much to offer in terms of not only producing pronunciation instructional materials but also making the learners independent by CAPT, something that can address both the issue of insufficient preparation and time pressure for teachers.

ITA training programs that have become popular in most US schools can benefit from the results of this study. These programs are designed and developed to promote the linguistic and cultural knowledge of ITAs and prepare them for better communication with undergraduate students. As

85

far as the linguistic part of these programs is concerned, the current study results are very insightful for the stakeholders in these programs. Why not spend more time recruiting specialists in this field, creating discipline-specific materials for ITAs, and equipping pronunciation labs for them given that the short teaching period of four weeks improved the speech comprehensibility of ITAs in this study? This work will be fruitful since it allies the concerns of many undergraduates, their parents, and departments over misunderstandings brought on by ITAs' incomprehensible speech.

References

Abker, I. A. A. (2019). Pronunciation Problems of Words' Stress Placement in English by Saudi Students at Albaha University, Saudi Arabia. A case Study at Almandag. International Journal of Applied Linguistics and English Literature, 8(6), 103-113.

Abdolmanafi-Rokni, S. J., Jalal, S., & Rokni, A. (2013). The effect of listening to audio stories on pronunciation of EFL learners. *MJAL*, *5*(2), 69-85.

Acton, W. (1984). Changing fossilized pronunciation. *TESOL quarterly*, *18*(1), 71-85.

Adebayo, C. T., & Allen, M. (2020). The Experiences of International Teaching Assistants in the US Classroom: A Qualitative Study. Journal of International Students, 10(1), 69-83.

Agostinelli, A. (2019). Second Language Identities of International Teaching Assistants in the US Classroom. Louisiana State University.

Al-Domi, I. (2017). The effect of using authentic videos on training center and community service students' prosodic competence and motivation. *Arab World English Journal (AWEJ) Volume*, *8*.

Allen, V. F. (1971). Teaching intonation, from theory to practice. *TESOL quarterly*, 73-81.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentais, prosody, and syllable structure. *Language learning*, *42*(4), 529-555.

Ashavskaya, E. (2015). International Teaching Assistants' Experiences in the US Classrooms: Implications for Practice. *Journal of the Scholarship of Teaching and Learning*, *15*(2), 56-69.

Axelson, E. R., & Madden, C. (1994). Discourse strategies for ITAs across instructional contexts. Discourse and performance of international teaching assistants, 153, 187.

Bailey, K. M. (1982). The classroom communication problems of Asian teaching assistants. Selected papers in TESOL, 1, 19-30.

Bailey, K. M. (1984). Foreign Teaching Assistants in US Universities.

Baker, A. A., & Burri, M. (2016). Feedback on second language pronunciation: A case study of EAP teachers' beliefs and practices.

Bergmann, J., & Sams, A. (2014). *Flipped learning: Gateway to student engagement*. International Society for Technology in Education.

Benrabah, M. (1997). Word-stress-a source of unintelligibility in English. iral, International Review of Applied Linguistics in Language Teaching, 35(3), 157.

Binte Habib, R. (2013). Intelligibility in EIL pronunciation: An analysis. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, *14*, 21-26.

Boersma, P. (2004). Stemmen meten met Praat. *Stem-, Spraak-en Taalpathologie*, *12*(4).

Boersma, P., & Weenink, D. (2015). Praat, doing phonetics by computer (v. 5.4).

Bozorgian, H., & Shamsi, E. (2020). Computer-assisted pronunciation training on Iranian EFL learners' use of suprasegmental features: A case study. *Computer-Assisted Language Learning Electronic Journal*, *21*(1), 93-113.

Brazil, D. (1997). *The communicative value of intonation in English book*. Cambridge University Press.

Burnett, C., & Merchant, G. (2015). The challenge of 21st-century literacies. *Journal of Adolescent & Adult Literacy*, 59(3), 271-274.

Canagarajah, S. (2018). Translingual practice as spatial repertoires: Expanding the paradigm beyond structuralist orientations. *Applied Linguistics*, 39(1), 31-54.

Canagarajah, S. (2018). Materializing 'competence': Perspectives from international STEM scholars. *The Modern Language Journal*, 102(2), 268-291.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). Teaching pronunciation hardback with audio CDs (2): A course book and reference guide. Cambridge University Press.

Chiang, S. Y. (2009). Dealing with communication problems in the instructional interactions between international teaching assistants and American college students. Language and Education, 23(5), 461-478.

Chung, B., & Bong, H. K. M. (2019). A Study on the Relation Between Intelligibility and Attitudes. ENGLISH TEACHING (영어교육), 74(2), 103-123.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. *The neurocognition of language*, 123-166.

Damar, E. A. (2014). Task-based video use for the improvement of English stress and intonation. *Journal of Educational and Social Research*, *4*(2), 227.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. Journal of phonetics, 11(1), 51-62.

Davies, C. E., & Tyler, A. E. (2005). Discourse strategies in the context of crosscultural institutional talk: Uncovering interlanguage pragmatics in the university classroom. In *Interlanguage pragmatics* (pp. 133-156). Routledge.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. Language learning, 48(3), 393-410.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. TESOL quarterly, 39(3), 379-397.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins Publishing Company.

Derwing, T. M., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning*, *64*(3), 526-548.

Dixon, D. H. (2018). Use of technology in teaching pronunciation skills. The TESOL encyclopedia of English language teaching, 1-7.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. TESOL quarterly, 39(3), 399-423.

Fitch, F., & Morgan, S. E. (2003). "Not a lick of English": Constructing the ITA identity through student narratives. *Communication* Education, *52*(3-4), 297-310.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada journal*, 1-22.

Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, *44*(2), 181-196.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. The Journal of the Acoustical Society of America, 93(3), 1589-1608.

Fraser, B. J. (2001). Twenty thousand hours: Editor'introduction. *Learning Environments Research*, *4*, 1-5.

Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. Trends in cognitive sciences, 10(6), 244-249.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

Fuse, A., Navichkova, Y., & Alloggio, K. (2018). Perception of intelligibility and qualities of non-native accented speakers. Journal of Communication Disorders, 71, 37-51.

Gallagher, N. (2012). *Delta's Key to the TOEFL iBT [R]: Advanced skill practice*. Delta Publishing Company. 1400 Miller Parkway, McHenry, IL 60050.

Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2016). Exploring the relationship between video lecture usage patterns and students' attitudes. *British Journal of Educational Technology*, *47*(6), 1259-1275.

Ganie, R., Maulana, W., & Rangkuti, R. (2019). ERRORS IN PRONOUNCING ENGLISH PHONEMES: A PRAAT ANALYSIS. Language Literacy: Journal of Linguistics, Literature, and Language Teaching, 3(1), 49-63.

Ghosh, M., & Levis, J. M. (2021). Vowel quality and direction of stress shift in a predictive model explaining the varying impact of misplaced word stress: Evidence from English. *Frontiers in Communication*, *6*, 628780.

Gorjian, B., Hayati, A., & Pourkhoni, P. (2013). Using Praat software in teaching prosodic features to EFL learners. Procedia-Social and Behavioral Sciences, 84, 34-40.

Gorsuch, G. J. (2006). Discipline-specific practica for international teaching assistants. *English for Specific Purposes*, *25*(1), 90-108.

Grant, L. E. (2010). A corpus comparison of the use of I don't know by British and New Zealand speakers. *Journal of Pragmatics*, *42*(8), 2282-2296.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL quarterly*, *38*(2), 201-223.

Hanayeen, N. (2018). Undergraduate Students' Perceptions of Their International Teaching Assistants (ITAs) and Perceptions of Themselves in a Course (s) Instructed by an ITA.

Hamlaoui, N., & Bengrait, N. (2016). Using BetterAccent Tutor and Praat for Learning English Intonation. Arab World English Journal (AWEJ) Special Issue on CALL, (3).

Hardison, D. M. (2004). Generalization of computer assisted prosody training: Quantitative and qualitative findings. Language Learning & Technology, 8(1), 34-52.

Hincks, R. (2005). Measures and perceptions of liveliness in student oral presentation speech: A proposal for an automatic feedback mechanism. *System*, *33*(4), 575-591.

Hincks, R., & Edlund, J. (2009). Promoting increased pitch variation in oral presentations with transient visual feedback. *Language Learning & Technology*, *13*(3), 32-50.

Hinofotis, F., & Bailey, K. (1980). American undergraduates' reactions to the communication skills of foreign teaching assistants. On TESOL, 80, 120-133.

Hodgetts, J. (2020). *Pronunciation instruction in English for academic purposes: An investigation of attitudes, beliefs and practices*. Springer Nature.

Hoekje, B., & Williams, J. (1992). Communicative competence and the dilemma of international teaching assistant education. Tesol Quarterly, 26(2), 243-269.

Howatt, A. P. (1984). Language teaching traditions: 1884 revisited. *ELT Journal*, *38*(4), 279-282.

Chen Hsieh, J. S., Wu, W. C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, *30*(1-2), 1-21.

Hutchinson, S. P. (1973). The Learning of English Suprasegmental Rules for Stress and Final Syllables by Spanish Speakers.

Inglis, J. (Ed.). (1993). Traditional ecological knowledge: Concepts and cases. IDRC.

Isaacs, T. (2018). Shifting sands in second language pronunciation teaching and assessment research and practice. *Language Assessment Quarterly*, *15*(3), 273-293.

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34*(3), 475-505.

Jackman, W. M., & Roberts, P. (2014). Students' perspectives on YouTube video usage as an E-Resource in the University classroom. *Journal of Educational Technology Systems*, 42(3), 273-296.

Jenkins, J. (2003). World Englishes: A resource book for students. Psychology Press.

Jiang, Y., & Chun, D. (2021). Web-based intonation training helps improve ESL and EFL Chinese students' oral speech. *Computer Assisted Language Learning*, 1-29.

Johncock, P. (1991). International Teaching Assistants Tests and Testing Policies at US Universities. *College and University*, 66(3), 129-37.

Kachru, B. B. (1978). Toward structuring code-mixing: An Indian perspective.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301-315.

Kang, O., & Kermad, A. (2017). Assessment in second language pronunciation. In *The Routledge handbook of contemporary English pronunciation* (pp. 511-526). Routledge.

Kang, O., & Rubin, D. (2014). 11. Listener expectations, reverse linguistic stereotyping, and individual background factors in social judgments and oral performance assessment. Social dynamics in second language accent, 10, 239.

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. Journal of Language and Social Psychology, 28(4), 441-456.

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*(4), 554-566.

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, *28*(3), 820-831.

Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers & Education*, *59*(2), 619-627.

Kelly, L. G. (1969). 25 Centuries of Language Teaching.

Ketabi, S., & Saeb, F. (2015). Pronunciation teaching: Past and present. *International Journal of Applied Linguistics and English Literature*, *4*(5), 182-189.

Kinash, S., Knight, D., & McLean, M. (2015). Does digital scholarship through online lectures affect student learning. *Journal of Educational Technology & Society*, *18*(2), 129-139.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

Larsen-Freeman, D. (2000). *Techniques and principles in language teaching*. Oxford University.

Ladefoged, P., & Bladon, A. (1982). Attempts by human speakers to reproduce Fant's nomograms. Speech Communication, 1(3-4), 185-198.

Lee, B., Plonsky, L., & Saito, K. (2020). The effects of perception-vs. production-based pronunciation instruction. *System*, *88*, 102185.

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. *Procedia cirp*, *16*, 3-8.

Le, H. T., & Brook, J. (2011). Using Praat to teach intonation to ESL students. Hawaii Pacific University TESOL Working Paper Series, 9(1), 2.

LeGros, N., & Faez, F. (2012). The intersection between intercultural competence and teaching behaviors: A case of international teaching assistants. Journal on Excellence in College Teaching, 23(3).

Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL quarterly*, *33*(1), 37-63.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL quarterly*, *39*(3), 369-377.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. Annual Review of Applied Linguistics, 27, 184.

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Levis, J. M., & Wichmann, A. (2015). English intonation–Form and meaning. *The handbook of English pronunciation*, 139-155.

LoCastro, V., & Tapper, G. (2006). International Teaching Assistants and teacher identity. *Journal of Applied Linguistics*, *3*(2).

Lima, E. F. (2020). The Supra Tutor: Improving speaker comprehensibility through<? br?> a fully online pronunciation course. *Journal of Second Language Pronunciation*, *6*(1), 39-67.

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. Language in Society, 419-441.

Lindemann, S., & Clower, A. (2020). Language attitudes and the 'ITA problem': Undergraduate reactions to instructors'(non) nativeness and pitch variation. *International Journal of Applied Linguistics*, *30*(1), 127-143.

Lippi-Green, R. (2012). Teaching children how to discriminate. Rosina Lippi-Green, English with an accent, 101-129.

Liu, X., Zhu, C., Jiao, J., & Xu, M. (2018). Promoting English pronunciation via mobile devices-based automatic speech evaluation (ASE) technology. In *Blended Learning. Enhancing Learning Success: 11th International Conference, ICBL 2018, Osaka, Japan, July 31-August 2, 2018, Proceedings 11* (pp. 333-343). Springer International Publishing.

Looney, S. D. (2015). Interaction and discourse markers in the ITA-led physics laboratory. *Talking Matters: Research on Talk and Communication of International Teaching Assistants*, 75-108.

Looney, S. D., Jia, D., & Kimura, D. (2017). Self-directed okay in mathematics lectures. *Journal of Pragmatics*, *107*, 46-59.

Looney, S. D., & Bhalla, S. (Eds.). (2019). A Transdisciplinary Approach to International Teaching Assistants: Perspectives from Applied Linguistics. Multilingual Matters.

Lorch, J. M., Knowles, S., Lankton, J. S., Michell, K., Edwards, J. L., Kapfer, J. M., ... & Blehert, D. S. (2016). Snake fungal disease: an emerging threat to wild snakes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1709), 20150457.

Low, E. L. (2014). *Pronunciation for English as an international language: From research to practice*. Routledge.

Marques, B. P., Villate, J. E., & Carvalho, C. V. (2011, June). Applying the UTAUT model in engineering higher education: Teacher's technology adoption. In *6th Iberian Conference on Information Systems and Technologies (CISTI 2011)* (pp. 1-6). IEEE.

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation*, *5*(1), 98-118.

McGregor, A., Zielinski, B., Meyers, C., & Reed, M. (2016). An exploration of teaching intonation using a TED Talk. *Pronunciation in Second Language Learning and Teaching Proceedings*, *7*(1).

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language learning*, *44*(2), 283-305.

McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. Teaching American English pronunciation, 185-196.

Martinez, M. (2001). Key design considerations for personalized learning on the web. *Journal of Educational Technology & Society*, *4*(1), 26-40.

Marvasti, A. (2005). Being Middle Eastern American: Identity negotiation in the context of the war on terror. Symbolic interaction, 28(4), 525-547.

Mehrabian, A. (1968). Some referents and measures of nonverbal behavior. *Behavior Research Methods & Instrumentation*, *1*(6), 203-207.

Morley, J. (1987). *Current Perspectives on Pronunciation. Practices Anchored in Theory*. TESOL, 1600 Cameron Street, Suite 300, Alexandria, VA 22314.

Morgan, B. M. (2003). Cooperative learning in higher education: Undergraduate student reflections on group examinations for group grades. *College Student Journal*, *37*(1), 40-50.

McWilliam, E. (1992). Towards advocacy: post-positivist directions for progressive teacher educators. *British Journal of Sociology of Education*, *13*(1), 3-17.

Mehrabian, A. (1968). Some referents and measures of nonverbal behavior. Behavior Research Methods & Instrumentation, 1(6), 203-207.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. Language learning, 45(1), 73-97.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*(4), 520-531.

Murty, L., Otake, T., & Cutler, A. (2007). Perceptual tests of rhythmic similarity: I. Mora rhythm. *Language and Speech*, *50*(1), 77-99.

Myers, C. (1994). Question-based discourse in science labs: Issues for ITAs. *Discourse and performance of international teaching assistants*, 83-102.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer assisted language learning*, *15*(5), 441-467.

Opendoors (19, February 2019). *New International Student Enrollment*. Open Doors. Retrieved February 2, 2022, from https://opendoorsdata.org/data/international-students/new-international-students-enrollment/

Olson, D. J. (2014). Phonetics and technology in the classroom: A practical approach to using speech analysis software in second-language pronunciation instruction. Hispania, 47-68.

Ohata, K. (2004). Phonological differences between Japanese and English: Several potentially problematic. *Language learning*, *22*, 29-41.

Papajohn, D. (2006). Standard Setting for Next Generation TOEFL Academic Speaking Test (TAST): Reflections on the ETS Panel of International Teaching Assistant Developers. *TESL-EJ*, *10*(1), n1.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *Tesol Quarterly*, *35*(2), 233-255.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for specific purposes*, *23*(1), 19-43.

Pickering, L. (2009). Intonation as a pragmatic resource in ELF interaction. Intercultural Pragmatics, 6(2), 235-255.

Pickering, L. (2018). *Discourse intonation: A discourse-pragmatic approach to teaching the pronunciation of English*. University of Michigan Press.

Pike, K. L. (1945). The Intonation of American English.

Pittam, J., & Scherer, K.R. (1993). Vocal expression and communication of emotion.

Plough, I. C., Briggs, S. L., & Van Bonn, S. (2010). A multi-method analysis of evaluation criteria used to assess the speaking proficiency of graduate student instructors. *Language Testing*, *27*(2), 235-260.

Reed, M., & Levis, J. M. (2019). *The handbook of English pronunciation*. John Wiley & Sons.

Renner, F., Kersbergen, I., Field, M., & Werthmann, J. (2018). Dutch courage? Effects of acute alcohol consumption on self-ratings and observer ratings of foreign language skills. *Journal of Psychopharmacology*, *32*(1), 116-122.

Rasi, P. M., & Poikela, S. (2016). A review of video triggers and video production in higher education and continuing education PBL settings. *Interdisciplinary Journal of Problem-Based Learning*, *10*(1), 7.

Reinhardt, J. (2010). Directives in office hour consultations: A corpus-informed investigation of learner and expert usage. *English for Specific Purposes*, *29*(2), 94-107.

Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *Relc Journal*, *46*(1), 5-22.

Richards, K., Williams, J. M., Smith, T. E., & Thyer, B. A. (2015). Financial video games: A financial literacy tool for social workers. *International Journal of Social Work*, 22-35.

Roach, P. (2009). *English phonetics and phonology paperback with audio CDs (2): A practical course*. Cambridge university press.

Rotellar, C., & Cain, J. (2016). Research, perspectives, and recommendations on implementing the flipped classroom. *American journal of pharmaceutical education*, *80*(2).

Rounds, P. L. (1987). Characterizing successful classroom discourse for NNS teaching assistant training. *Tesol Quarterly*, *21*(4), 643-671..

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. Research in Higher education, 33(4), 511-531.

Sa'di, R. A., Sharadgah, T. A., & Yaseen, M. S. (2022). An Investigation of Common Stress Placement Errors in English Word Roots and their Suffixed Derivatives by Arabic-Speaking EFL Learners.

Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: The case of Japanese learners of English. *RELC Journal*, *42*(3), 363-378.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of/ɹ/by Japanese learners of English. *Language learning*, *62*(2), 595-633.

Saito, Y., & Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research*, *21*(5), 589-608.

Saleh, A. J., & Gilakjani, A. P. (2021). Investigating the impact of computer-assisted pronunciation teaching (CAPT) on improving intermediate EFL learners' pronunciation ability. *Education and Information Technologies*, *26*, 489-515.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance 1. *Language learning*, *19*(3-4), 245-253.

Sereno, J., Lammers, L., & Jongman, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics*, *37*(2), 303-322.

Shuck, G. (2014). Language in Human Life: A GE Course Targeting English Language Learners.

Smith, R. M. (1992). Crossing Pedagogical Oceans: International Teaching Assistants in US Undergraduate Education. ASHE-ERIC Higher Education Report No. 8, 1992.

Staples, S., O. Kang, and E. Wittner. 2014. 'Considering interlocutors in university discourse communities: impacting U.S. undergraduates' perceptions of ITAs through a structured contact program,' English for Specific Purposes 35: 54–65.

Subtirelu, N. C. (2015). "She does have an accent but…": Race and language ideology in students' evaluations of mathematics instructors on RateMyProfessors. com. *Language in Society*, *44*(1), 35-62.

Szyszka, M. (2017). Pronunciation learning strategies and language anxiety. *Switzerland: Springer*, *10*, 978-3.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. Journal of Phonetics, 25(1), 1-24.

Tanner, M., & Landon, M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology*, *13*(3), 51-65.

Tapper, J. (1994). Directives used in college laboratory oral discourse. *English for Specific Purposes*, *13*(3), 205-222.

Tapper, G., Kidder, K., Kaufman, D., & Brownworth, B. (2006). A research-informed approach to international teaching assistant preparation. *Professional development of international teaching assistants*, 17-33.

Taslibeyaz, E., Dursun, O. B., & Karaman, S. (2017). Interactive video usage on autism spectrum disorder training in medical education. Interactive Learning Environments, 25(8), 1025-1034.

Twale, D. J., Shannon, D. M., & Moore, M. S. (1997). NGTA and IGTA training and experience: Comparisons between self-ratings and undergraduate student evaluations. *Innovative Higher* Education, 22(1), 61-77.

Tyler, A. (1992). Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *Tesol Quarterly*, *26*(4), 713-729.

Wells, J. C. (2006). *English intonation PB and Audio CD: An introduction*. Cambridge University Press.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. The Journal of the Acoustical Society of America, 113(2), 1033-1043.

Liu, X., Wu, D., Ye, Y., Xu, M., Jiao, J., & Lin, W. (2020). Improving Accuracy in Imitating and Reading Aloud via Speech Visualization Technology. International Journal of Emerging Technologies in Learning (iJET), 15(08), 144-160.

Wallace, L. R., & Lima, E. F. (2018). Technology for Teaching Pronunciation. *The TESOL Encyclopedia of English Language Teaching*, 1-7.

Wells, J. C. (2006). *English Intonation HB and Audio CD: An Introduction*. Cambridge University Press.

Wichmann, A. (2014). *Intonation in text and discourse: Beginnings, middles and ends*. Routledge.

Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). Video-based learning: A critical analysis of the research published in 2003-2013 and future visions. In *eLmL 2014, The Sixth International Conference on Mobile, Hybrid, and On-line Learning* (pp. 112-119).

Yu, J., & Wang, Z. (2016, December). A realistic and reliable 3D pronunciation visualization instruction system for computer-assisted language learning. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 786-789). IEEE.

Yung, F., Duh, K., Komura, T., & Matsumoto, Y. (2016, August). Modelling the usage of discourse connectives as rational speech acts. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning* (pp. 302-313).

Zadeh, V. A., Gussenhoven, C., & Bijankhan, M. (2011). A Pitch Accent Position Contrast in Persian. In ICPhS (pp. 188-191)

Zielinski, B. (2015). The segmental/suprasegmental debate. *The handbook of English pronunciation*, 397-412.

Zhang, H., Du, X., Yuan, X., & Zhang, L. (2016). The Effectiveness of the flipped classroom mode on the English pronunciation course. *Creative Education*, *7*(9), 1340-1346.

Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, *38*(2), 260-271.

Zhu, Y., & Bresnahan, M. J. (2020). A thematic analysis of international teaching assistants' stigma experience in a US university: English-proficiency determinism. Journal of International and Intercultural Communication, 1-18

Appendices

Appendix A: short sentence Diagnostic test

1. You are not certain if you set a deadline for your students' submission, so you say:

   "I set the deadline on Tuesday at midnight, didn't I?"

2. Your students want you to postpone the mid-term exam to the following week. You have no problem with giving them more time, but you are wondering if there would be other factors- like the department administration- that would allow this extension. So, you say:

   "I personally don't mind giving you an extension but …."

3. A student asks a question, but you want to tell him that the answer is in that day's lesson, and he needs to be patient for the answer. You say "Well! That's what we are trying to figure out."

4. A student must miss your class for an athletic event and asks to make up the scheduled exam. You say "Sure! But you need to be in my office at 3:30, alright?"

5. A student didn't do well on his last exam. He asks you for help after class. You say: "I know you had trouble on this exam. Come to my office during my office hours, and we can talk about it.

Appendix B. Spontaneous production diagnostic test (transcript)

Student: Excuse me, Professor Thompson. I know your office hours are tomorrow, but I was wondering if you had a few minutes free now to discuss something.

Professor: Sure, John! What did you want to talk about?

Student: I have some questions about how to write up the research project that I had this semester about climate variations.

Professor: Oh, yes. You were looking into climate variation in the Grand City area. Right? How far have you gone?

Student: I only got my data, so I am going to summarize it now. Preparing graphs and stuff. But I'm just. I'm looking at it and I'm afraid that it's not enough but I'm not sure what else to put in the report.

Professor: We hear the same thing from every student. You know. You have to remember that you are the expert on what you have done.So think about what you need to include to explain your research project to someone with general or casual knowledge about this subject like your parents. That's usually my rule of thumb with my parents understanding this.
Student: Good, I got it.

Professor: I hope you recognize it by my saying how much you do know about the subject

Student: Right, I understand. I was wondering if I should also include the notes from the research journal you suggested I keep.

Professor: Yes, definitely. You should use them to indicate what your evolution and thought was through time. So just set up what was the purpose of what you were doing to try to understand the climate variability of this area and what you did and what your approach was.

Student: Ok, so for example I studied meteorological records, I looked at climate charts, used different methods for analyzing the data, like certain statistical tests then I discussed the results. Is that what you mean?

Professor: Yeah! That's right. You should include all of that. The statistical tests are especially important and also be sure you include a good reference section where all your published and unpublished data came from because you have a lot of them as published climate data.

Student: Mmmm. Something just came out of my mind and went to another side.

Professor: Huu.. That happens to me a lot. I have come up with a pretty good management tool. I carry a little pad with me all the time and jot down the ideas that I don't want to forget. For example, I went to a doctor with my daughter and her baby son last week and we knew we would not remember everything we wanted to ask the doctor, so we actually made a list of five things we wanted answers to.

Student: Notepad is a good idea. Since I'm so busy at the end of the semester. I'm getting pretty forgetful these days. Ok, now I just remember what I wanted to ask before.

Professor: Good, I was hoping you come up with it.

Student: Yeah, it ends up that I have data on more than just the immediate Grand City area, so I also included some regional data on the report. With everything else, it should be a good indicator of the climate in this part of the state.

Professor: Sounds good! I'd be happy to look over a draft version before you hand in the final copy if you wish.
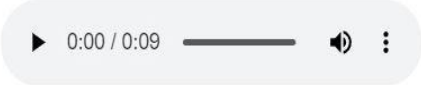
Student: I plan to give you a draft of the paper by next Friday. Thanks very much. See you.

Professor: OK

Appendix C: Speech Evaluation Survey

Instruction: Listen to the following audio file and then select from 1-9 how easily you have

heard the speech.

1. Ce31f8.

	1	2	3	4	5	6	7	8	9
1= hard to understand, 9= easy to understand	O	O	O	O	O	O	O	O	O

Appendix D: Lesson plan for intonation

Description

Intonation

What is intonation?

Intonation is often called the melody of language since it refers to the pattern of pitch changes that we use when we speak. If you listen to someone speaking, you will notice that there are many changes in pitch. These pitch changes are called intonation patterns and play an important role in conveying meaning.

Final intonation

Rising-falling intonation Listen to yourself when you say the following sentence:

Susan bought a new sweater.

Notice that the pitch of your voice rises at the major sentence stress, the first syllable of the word 'sweater', and falls over the second syllable of this word. The pitch of the entire sentence is referred to as the intonation pattern. The pattern in this sentence is rising falling. It is the most common intonation pattern in English and is characteristic of simple declarative sentences, commands and questions that begin with a wh-word, such as 'who', 'what', 'when', 'where', 'why', or 'how'. Say the sentences below, concentrating on the pitch change at the word receiving major sentence stress.

He wants to GO home.

She gave him five DOLlars for it.

What do you want to DO with it?

Rising intonation

Listen to the pitch following sentence:

Did Susan buy a new sweater?

Notice again that the pitch of your voice rises at the major sentence stress. However, rather than a sharp decline in pitch level after the stressed syllable, as with the rising-falling intonation pattern, the voice continues to rise. The intonation pattern in this case is rising and is characteristic of questions that require a simple yes or no answer.

Does he want to go HOme?

Did she give him five DOLlars for it?

Do you want to GIVe it to him?

This intonation contour is also used to express doubt.

He left town.

Non-final intonation

Rising- falling intonation

Complex sentences often have two separate intonation patterns. An example of such a sentence is provided below.

Because of his athletic ability, he was given a scholarship.

Here the pitch rises and falls on the word 'ability', and on the word 'scholarship'. on 'scholarship', the pitch drifts to the bottom of the pitch range' while on 'ability', the pitch does not fall nearly as far. The intonation contour on the first half of the sentence is a non-final rising-falling contour. The following sentences usually have two intonation patterns – the non-final contour on the first phrase and the final contour on the second. A few examples are provided below:

When he left her office/ it was raining.

After they completed the survey/ they went out for a walk.

The man you said you met yesterday/ has left town.

Continuation rise

Say the following sentence.

Susan bought a new sweater, new shoes, and a new dress.

The intonation contour of this sentence is termed a continuation rise and is often used with lists. The pitch of the voice rises slightly on each noun of the list, indicating that we are not yet finished speaking. On the final noun of the list, we find the familiar rise-fall.
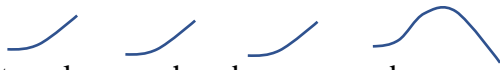
He bought apples, peaches, bananas, and oranges.

You should study the sources, outline the main points, write the summary, and submit it to the teacher.

Tag question

Tag questions can display either final rising-falling or final rising intonation contours. Their meaning will differ depending on which of these contours is used. Pronounce the following tag questions with the intonation contours indicated:

Deana's helpful, isn't she?

Deana's helpful, isn't she?

The first sentence, with a rising contour, indicates that the speaker genuinely does not know whether Deanna is helpful and wants the listener to provide this information. The second sentence, with a rising-falling contour, indicates that the speaker believes that Deanna is helpful and is merely eliciting confirmation from the listener.

Speaker attitude

Speaker attitude can be signaled through the use of pitch variation in intonation patterns. For example, if we raise our pitch, we may be indicating surprise; if we lower our pitch, we may be indicating anger. If we expand our pitch range (that is, if our high pitches become higher and our low ones become lower), we may be indicating deference. If we narrow our pitch range, we may be indicating boredom. Given the role of intonation in conveying speaker attitude, there is great potential for ESL students to be misunderstood if their intonation Patterns are too dissimilar from the English ones. For example, many languages of the world display less pitch variation than English. Thus, learners who speak these languages may unwittingly convey boredom or lack of interest through the use of too narrow a pitch range.

Thought groups

Thought groups allow one to organize his/her speech into groups of words that make up a single idea (Grant, 2010). They help the listener(s) better understand the information in interlocutors' speech as speakers organizetheir ideas into comprehensible "packages" that are easy to process (Grant, 2010).

Keys to identifying thought groups.

Sentences are made up of grammatical units. Pausing between larger grammatical units creates a natural rhythm and more fluid speech.
Noun phrases
A tall teacher [article+ adjective+ noun]
Subjects
A tall teacher/ entered the room.//
A tall teacher/ and his assistant/ entered the room.//
Verb phrases
Write clearly [verb+ adverb]
Send an email [verb+ object]
Had been studying [auxiliary verb+ main verb]
Prepositional phrases
On that page
Clauses
If there are technical words, / they are explained in a glossary. //
Students,/ who wear uniforms,/ identify themselves with their school.//
Don't break up short statements.
Examples:
I ran.
Who jumped?
Did you dance?
It's Jennifer's composition.

If you need to speak slower for your listeners, break sentences into shorter thought groups.

Listening discrimination

Listen to the following conversation as your teacher reads it out. What is the final tone of each line?

- A: Finished?

- B: Almost.

- A: Five minutes?

- B: No.

- A: When?

- B: Later.

Listen to three ways of answering the question "Could you please help me"?

Could you please help me?

Yeah   (certain to help)

Yeah   (affirmative and enthusiastic = I would be happy to)

Yeah   (hesitant = I probably agree but I need to know exactly what I can do)

Listen to the following speech files. Both are reading the following passage. Which one is harder to follow? Why? (Listeners expect to hear appropriate thought groups from speakers)

 When I was a girl, I skipped down path, danced in my bedroom, spun in circles under the sky, jumped rope with my friends and squatted to smell flowers in my mother's garden.

Listen to the sentences in the box. Check the box to that corresponds to the pattern used.

Question or statement		
Utterance	Question	statement
1. He left already.		
2. Sally's moving.		
3. John missed his flight		
4. Its snowing in NY.		

Sure or unsure		
Utterance	Unsure	Sure
1.You passed the test, didn't you?		
2.The library is closed, isn't it?		
3.You should type it, shouldn't you?		
4.The test is easy, isn't it?		

Sure or unsure		
Utterance	Yes-No	Choice
1.Are you coming Friday or Saturday?		
2.Can you meet us at 8 or 9?		
3.Would you like beer or wine?		
4. will you graduate in spring or summer?		

Listen to the first 2-3 minutes of President Obama's speech through the following link. What aspect of his speech impresses you? Is it easy to follow? Why?

https://www.youtube.com/watch?v=cVW6jBbD5Q8

Controlled and guided practice

The first speaker mentions one of the following ten statements. Then he adds a tag question either with a falling or a rising intonation. If it is falling intonation, the second speaker should just confirm it. If it is a rising intonation, the speaker should provide further information.

1. He goes to see his professors.
2. She got a low grade on the last test.
3. Our professor will give us an extra credit assignment.
4. The professor emphasized this point.
5. The experiment takes a long time.
6. He accepts just one over-due assignment.
7. Our class starts at 8:00 AM
8. There is a time conflict for 2 of your courses.
9. You can solve this problem in the registrar office.
10. This class meets once every week.

Example:

A: They close the library at 6:00 PM tonight, don't they?

B: Yes, they do.

A: We can continue our discussion in the University Center, Can't we?

B: they just allow you in today if you have reservation.

The following poem, Uphill, by Christina rosette (1862) contains alternative questions and answers. Pair with a classmate and read it aloud together.

Does the road wind up-hill all the way?

   Yes, to the very end.

Will the day's journey take the whole long day?

   From morn to night, my friend.

But is there for the night a resting-place?

   A roof for when the slow dark hours begin.

May not the darkness hide it from my face?

   You cannot miss that inn.

Shall I meet other wayfarers at night?

   Those who have gone before.

Then must I knock, or call when just in sight?

   They will not keep you standing at that door.

Shall I find comfort, travel-sore and weak?

   Of labor you shall find the sum.

Will there be beds for me and all who seek?

Yea, beds for all who come.

First study your line to identify the thought groups then with your peer play the role of the student and the professor.

Student: Hi Professor Taylor.

Professor: Hi Jack.

Student: I was hoping that I could talk with you for a few minutes. It's about the test.

Professor: Oh, okay.

Student: well, I've never taken an open-book test, and I just don't know what to expect. Does that mean I can use my book during the test . . . as a reference?

Professor: Exactly. And you can use your notes and the handouts, too.

Student: Really?

Professor: Yes, but Jack, since you've never taken an open-book test, I should warn you. It isn't as easy as it seems.

Student: Because?

Professor: Because you don't have enough time to look up every answer and still finish the test.

Student: Oh.

Professor: That's the mistake that most students make. You see, the purpose of an open-book test is to allow you to look up a detail or make a citation. But the students who are looking up every answer spend too much time on the first few questions, and then they have to leave some of the questions at the end blank.

<center>Communicative exercise</center>

There are two pictures that are basically similar but there are a few differences. Pair with your friend and try to ask each other questions to discover the differences. (Conversations will be monitored by the teacher for the intonational patterns that are implemented).
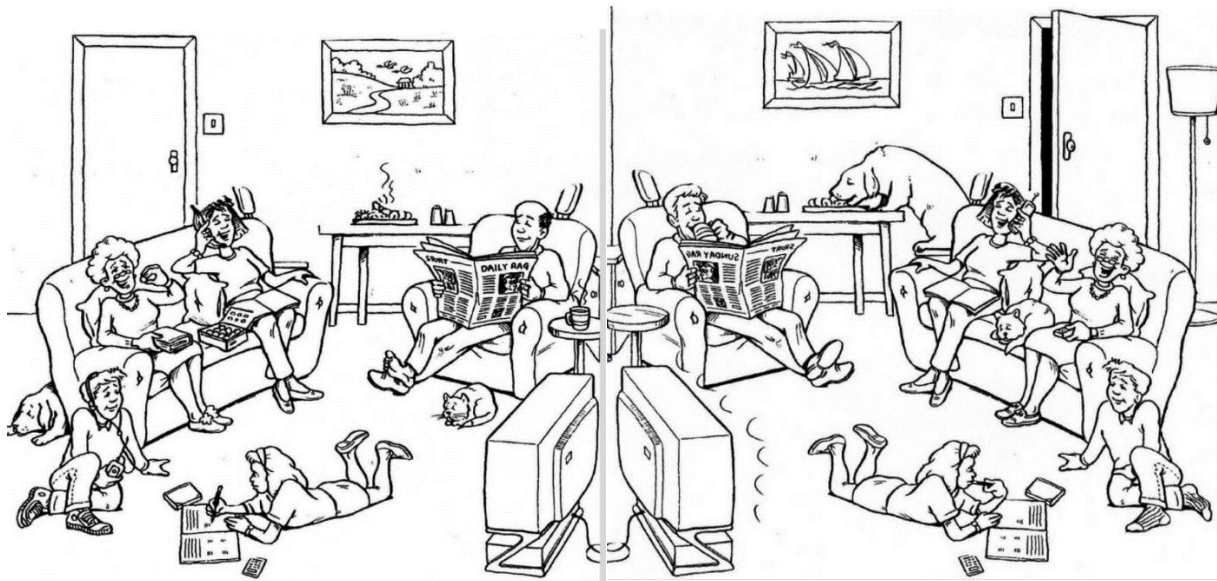
Example:

A: I have a picture of a room with five people, a cat and a dog. Each person is doing something.
B: I have a similar picture. A man, a girl, a boy, and two ladies.
A: what is the man doing?
B: He is reading newspaper?
A: Is he also drinking something?
B: No, he is not.

Yes/ no question game (1)

Pair with a classmate. One of you chooses an animal but does not say what it is. The other classmate should be able to guess that animal's name by asking at most 20 yes-no questions.

Example:

A: Does it live in water?
B: No, it doesn't.
A: Does it eat meat?
B: No, it does not.
A: Do people raise them for their milk?
B: No, they don't.

Yes/ no question game (2)

Pair with a classmate. One of you thinks of good/bad news about your education or courses but do not say what it is. The other classmate should be able to guess that news by asking at most 20 yes-no questions.

Example:

A: Is it a social problem?
B: No, it isn't.
A: Is it a financial one?
B: Yes, it is.
A: Do you need a loan to pay your tuition?
B: No, I don't.
A: Have you recently lost your money on campus?
B: No, I have not.

Appendix E: Tutorial videos

Word stress1: https://youtu.be/q7zUuTYjKV8
Word stress 2: https://youtu.be/rvJ7P21bIrM
Rhythm 1: https://youtu.be/AUOmxnYCE0c
Rhythm 2: https://youtu.be/Gm0OzcCwMTM
Intonation 1: https://youtu.be/M3xq6QQx8Sw
Intonation 2: https://youtu.be/GHP3085s5e0
Review1: https://youtu.be/xLfT5-o5AxI
Review 2: https://youtu.be/wRjpWyTlkVU

Appendix F: Praat Visual