

## **Computer-Assisted Pronunciation Training on Iranian EFL Learners' Use of Suprasegmental Features: A Case Study**

Hossein Bozorgian (h.bozorgian@umz.ac.ir)  
University of Mazandaran, Islamic Republic of Iran

Esmat Shamsi (shamsiesmat@gmail.com)  
Rodaki Institution of Higher Education, Islamic Republic of Iran

### **Abstract**

In this paper, we investigated the extent to which Computer Assisted Pronunciation Training (CAPT) improved five Iranian English as Foreign Language (EFL) learners' use of suprasegmental features. The My English Tutor (MyET), computer software, was utilized to train English as Foreign Language (EFL) learners for using suprasegmental features of English pronunciation (e.g., stress, rhythm/timing, and intonation). Data from a two-month direct observation, the second researcher's and participants' reflective notes, the feedback and scores provided by MyET, and oral interviews were collected and analyzed. The results revealed that CAPT improved EFL learners' use of suprasegmental features. Moreover, the oral interview analyses indicated that EFL learners had a positive attitude toward CAPT, which helped them to be autonomous and confident in learning how to pronounce English words and phrases.

*Keywords:* Computer-Assisted Pronunciation Training (CAPT); English as a Foreign Language (EFL) learners; pronunciation; suprasegmental features

### **Introduction**

Computer-Assisted Pronunciation Training (CAPT) was originally developed and used as a clinical instrument by language pathologists to treat speakers' communication disorders. Only later did CAPT begin to be applied in pronunciation instruction (Pokrivcakova, 2015). The use of computers is supportive to learn pronunciation skills because it can provide individualized instruction and instant feedback on learners' speech (Levis, 2007; Tanner & Landon, 2009; Yoshida, 2018). Moreover, they usually comprise fun activities with games, animation, and role-play to motivate young users (Neri, Mich, Gerosa, & Giuliani, 2008). In most CAPT systems, Automatic Speech Recognition (ASR) technology automatically transcribes students' voice recordings into written words. The speech visualization technology integrated into the system visually shows the deviation of students' pronunciation from that of native model speakers. This speech model can provide learners with particular metacognitive strategies to facilitate learning the target language sound patterns (Elimat & Abuseileek, 2014; Neri, Cucchiaroni, & Strike, 2002; Tsai, 2019).

All CAPT systems are supposed to provide abundant oral input, the opportunity to receive and store learner's output, and giving instant feedback on the quality of learner's speech compared to a native-like paradigm (Neri, Cucchiarini, Strik, & Boves, 2002). Providing instant feedback on learners' pronunciation in EFL classrooms with too many students is an intensive and time-consuming activity requiring one-to-one work. Furthermore, many English language teachers have many difficulties in giving proper feedback, due to lack of confidence, deficiencies in teacher training courses, and uncertainty about the correct form of articulations (Breitkreutz, Derwing, & Rossiter, 2002; Levis, 2007; Warren, Elgort, & Crabbe, 2009). In this context, CAPT can aid learners with working on their pronunciation skills independently and privately whenever they want and wherever they are.

Pronunciation experts have considered improved intelligibility as the most important goal of pronunciation teaching (Celce-Murcia, Brinton, & Goodwin, 2008; Derwing & Munro, 2015; Huensch, 2018). The intelligible pronunciation makes the speaker powerful in communication, be understood properly, and succeed in making an effective connection as well. A fairly good agreement among experts in foreign language learning education is that in the communicative curriculum, suprasegmentals are more critical for intelligibility and comprehensibility than segmentals.

For instance, Anderson-Hsieh, Johnson, and Koehler (1992) compared the degree of intelligibility by suprasegmentals, segmentals, and syllable structure. Among 11 different language groups, they found that the score for suprasegmentals was most significantly associated with the overall degree of intelligibility. This is in line with the findings of Ginther, Dimova, and Yang (2010) that suprasegmental measures alone explain over half of the variance in oral proficiency ratings. Their findings showed that Oral English Proficiency Test (OEPT) scores extremely corresponded with the suprasegmental measures of mean length of run, number of silent pauses, speech time ratio, and speech rate. Accordant with these notions, Saito, Trofimovich, and Issacs (2016) stated that suprasegmental errors impede listeners' evaluation of non-native speakers' speech more directly than segmentals do. Furthermore, suprasegmentals such as word stress and intonation are equally important at every stage of L2 oral ability learning from beginner to advanced, while segmental accuracy is related to higher-level oral development.

These studies indicated that improving dominant features of prosody supplies greater benefits for intelligibility and comprehensibility than focusing on segmental features. This feature directs the listener's attention to important information in the act of communication. It may sufficiently prove the necessity of teaching and learning suprasegmentals in the EFL context where the oral input is limited to a teacher's speech and audio files and learners suffer severely from the lack of intelligibility and comprehensibility.

## **Review of literature**

A number of studies have reported the effectiveness of CAPT programs in improving pronunciation. It seems that qualitative inquiries can indicate participants' problems and preferences for using a CAPT program and explore requirements during the practice. Very few studies were conducted with this approach. For instance, Kim

(2012) investigated the effectiveness of incorporating CAPT in pronunciation to two Korean male ESL students. Results showed that learners who prefer visual learning style may benefit more from CAPT. Moreover, the visual feedback used in this study enhanced the learners' self-awareness and self-monitoring and led to improved pronunciation, native-like quality, and fluency. He also claimed that it is easier to enhance suprasegmentals than segmentals at a short training period.

Calvo-Benzies (2017) in her study on the contribution of new technologies to the English pronunciation teaching provided a comprehensive review of the newest CAPT software, as well as an empirical study on the EFL learners' experiences of using CAPT in different kinds such as Apps, Blogs, Websites, and Social Networks. The learners stated that all types of technological resources they experienced were easy to use and they enjoyed applying them. The learners also commented that they would use similar tools again because they are motivating, engaging, and funny. The learners even preferred working with the software to their textbook's pronunciation tasks.

Some studies have particularly focused on the effects of computer-assisted pronunciation training on suprasegmental features of pronunciation. Hardison (2004) conducted a study focusing on both quantitative and qualitative aspects of computer-assisted prosody training. Results obtained from the quantitative study revealed noticeable improvement in the learners' use of suprasegmentals. Interpreting qualitative data showed that for some learners, technology holds greater interest than others and it depends on learners' preference and learning style.

Focusing on the function of intonation in discourse vs. sentence level, Levis and Pickering (2004) conducted a study on the effects of using visualization technology on the interpretation of pitch patterns. According to the findings of this study, they declared that effective intonation practice without computers is limited to teachers and students with good listening skills. The use of visualization technology helps those students without the ability to identify pitch changes. Moreover, providing visual feedback encourages sentence level practice and a focus on grammatical form.

Similarly, Abuseileek (2007) in his experimental study assessed the efficacy of computer-based pronunciation instruction for enabling EFL learners to perceive and produce correct "stress patterns". Results showed that the students who received computerized training achieved better results in their tests. Through an interview that was conducted after training sessions, the students expressed that they had a positive attitude towards CAPT program methods and activities because they could use the instructions efficiently by themselves. He also recommended that the learners' attention should be directed to overall utterances in connected speech, not words or segments in isolation.

In an experimental study, Tanner and Landon (2009) examined the effects of CAPT on ESL learners' use of prosody and their overall comprehensibility and found that computer-assisted treatment had a noticeable effect on the perception of word-stress and pausing. The treatment had also a significant positive impact on the participants' total pause marking errors and the number of missing pause marks. Regarding stress, a noticeable decrease in the number of stress marking errors and stress placement errors could be observed.

Munro and Derwing (2015) noted that individual differences in L2 perception and production are under-researched. Most pronunciation studies have reported only group means, but individual learning patterns may shed light on pronunciation learning

strategies. The present study aimed to investigate the extent to which CAPT improves five Iranian EFL learners' use of suprasegmental features and addressed the following research question:

To what extent does CAPT improve Iranian EFL learners' use of suprasegmental features?

## Method

### Participants

This study has used a case study to explore the bits and pieces of EFL learners' improvement in using suprasegmental features in English. Participants for this case study were sampled from among the ten Iranian students of a private English language institute. After a conversation with the administrator, the researchers held a meeting with all ten students. Five out of the ten agreed on taking part in this study voluntarily. The consent forms were signed before the beginning of the study. The participants were informed about the purpose of the study and were told that they would quit the project whenever they want.

The subjects of this study were five Iranian students of Islamic Azad University of Tonekabon in various fields of study. The group was made of two male and three female students aged 19 to 25 years. All of the participants began to receive formal English education as a subject focusing mainly on reading through the Grammar Translation Method in Middle high school at the age of 12. They were all from Iran and studied English as a foreign language. Their purpose in English language learning was to continue their further education or to find better job opportunities in one of the English language countries. Their proficiency level was B1 according to an Oxford Placement Test (OPT), conducted by the researchers before the training session about the Common European Framework, a level of independent users. The researchers used pseudonyms for the participants of this study and Table 1 illustrates the participants' demographics for more details.

**Table 1**

*Participants' demographics*

Participant	Age	Gender	Major	University level	L1	L2	Level of proficiency
Arad	22	Male	Civil engineering	Junior	Persian	English	Lower intermediate
Daniel	21	Male	Accounting	Freshman	Persian	English	Lower intermediate
Negar	19	Female	Law	Freshman	Persian	English	Lower intermediate
Sanaz	25	Female	M.B.A	Senior	Persian	English	Lower intermediate
Tara	22	Female	Midwifery	Junior	Persian	English	Lower intermediate

### Instruments

This study has used the following instruments for data collection: Direct observation, My English Tutor (MyET) application, researcher/teacher's and learners' reflective notes, and interview.

### 1. Direct observation

There were eight observation sessions fulfilled in this research and the participants were observed simultaneously. The setting was in a language lab at an English language institute. All the participants had their laptops with microphones and headphones. The observed sessions lasted from 60 to 70 minutes in length and were observed openly. Field notes were completed during the observations based on the participants' activities, statements, and behaviors besides the second author's reflections. Observation notes and the second author's reflective notes were taken to keep a detailed record of the computer-assisted pronunciation training activities covered the following areas: participants' attitude and reaction towards the sentence models; difficulties that they experienced during the perception and production of the models; participants' success in producing speech compared to the native-speakers' speech model provided in the application (Kim, 2012).

### 2. MyET

The training software used in this program was MyET, that is, the software appropriate to all ages of English learners and the best-known pronunciation application in Iran, which is classified to different levels of proficiency from elementary to advanced level. Each lesson has been divided into five parts: a conversation sample, role-play, reinforcement, self-test, and repeat. The feedback and scoring instances can be observed in Figure 1. On the screen, learners' performance and scores of three main components of pronunciation including emphasis (i.e. loudness), pitch (i.e. intonation), timing/rhythm (i.e. the rate of syllable production during speech), plus the waveform of standard speech and learners' speech are shown.

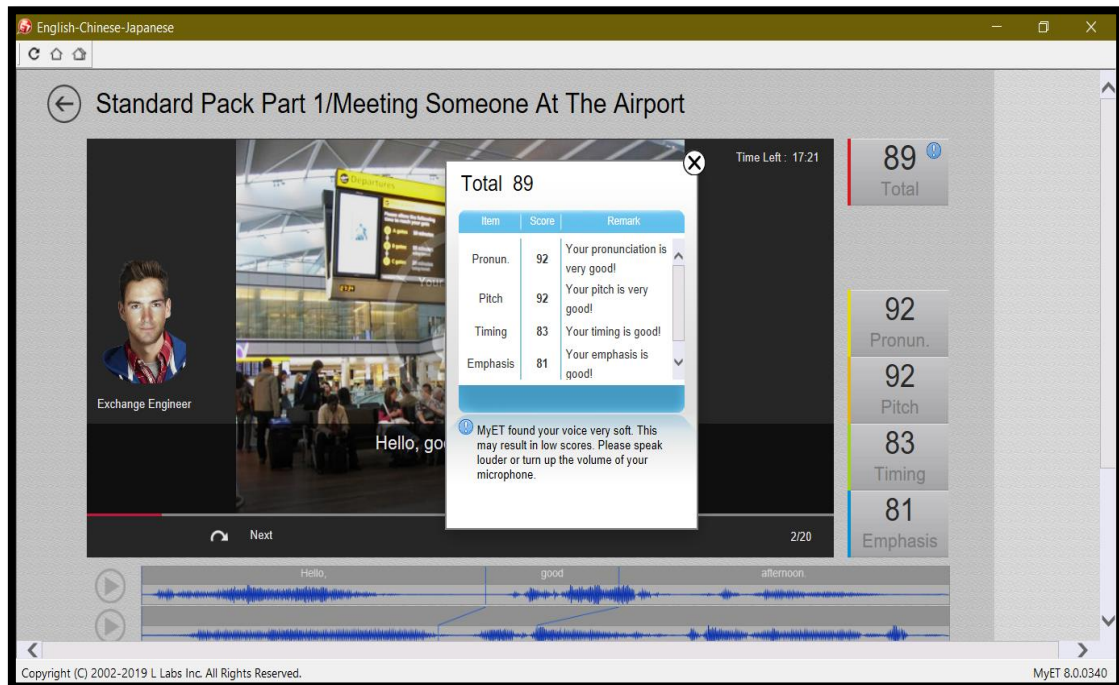
Audiolingual theorists have presented theoretical underpinning of Computer-assisted pronunciation training: drills, repetitions, and mimicry; and the communicative language learning theories are not put into practice in this area (Reeds & Levis, 2015). Like other pronunciation training Apps, MyET's instruction is based on listening and repeat drills. The correct form is presented by the App and the learner tries to repeat it properly. The focus is on successful imitation and producing desirable output. Learning the correct forms of sounds and chunks can be reinforced by practicing different exercises embedded in the App.

### 3. The learners' reflective notes

To encourage learners to participate in the learning process and enrich the evidence, they were asked to write down their reflection notes during the training period. They penned their reflections, attitudes, and experiences on working with MyET and its particular feedback system (in terms of suprasegmental features of pronunciation i.e. emphasis, pitch, and timing).

**Figure 1**

*MyET instant feedback on pronunciation, pitch, timing, and emphasis*



#### 4. Interview

Interviews are appropriate approaches to produce a method in qualitative research in education (Merriam, 2007). In this study, a semi-structured interview was conducted towards the end of the program. All five participants took part in this phase of the project. The interview consisted of six questions (see appendix A) focused on their opinions, feelings, and experiences of using MyET as well as the participants' difficulties, challenges, and problems with activities. It took approximately one hour for each participant. All interviews were recorded to ensure accuracy of reporting on a digital voice recorder, which made it easier to upload to a computer for further transcription.

#### Procedure

The training sessions period lasted two months. The class met once a week for nearly 60 to 70 minutes. Before training began, an OPT was conducted to identify participants' level of language proficiency and aided the teacher to select the appropriate material. The teacher followed Tsai's (2019) study during the intervention.

#### Intervention

The teacher/researcher divided the training session into three stages: introduction, perception, and production. In the introduction stage, an oral illustration of how features like stress, intonation, and rhythm were used. This stage helped learners develop basic

concepts about English suprasegmental features and realize that the feedback the system provides during practice with MyET.

Through training sessions, the teacher provided sufficient information about “stress”, its meaning, functions, kinds, and how to produce it in using the English language. In this sense, she noted that the words, which carry the newest information in a sentence, are usually stressed. Moreover, stress is usually placed on interrogatives, adverbs, demonstrative, and possessive pronouns. For example, in the sentence ‘*Jean sells cars*’, Jean and cars are regarded as new information which is stressed, but in the sentence, She *sells* them sentence stress shifts to the verb *sells*.

To introduce intonation, the teacher referred to the falling, rising, and level intonation and difference between intonation in wh-questions and yes/no questions. She stated that in yes/no questions, there is a long pitch from the beginning to the end of sentences (e.g., would you like more /milk?) but in wh questions, there is an abrupt pitch falls on the last stressed word of the sentence (e.g., /when we came /in, we had /dinner.)

In rhythm and timing area, the teacher introduced the concepts like syllable-timed vs. stress-timed languages, length, and weak vs. strong syllables. For example, in these sentences: ‘They eat them, They’re eating them, they might have eaten them’ despite the difference in the number of the words in sentences, all of them are equal in timing and take the same amount of time to say since they have the same number of stressed words. Also, the teacher taught the learners how to use MyET and interpret feedback and waveforms provided by the system.

In perception and production stage, for practice reading on the MyET system, eight conversations for eight sessions, based on-campus courses include “living with a host family, apartment renting experience, campus meal plan, student apartment activity, maintenance service, register for a new semester, meeting the academy advisor, and grocery shopping” were selected. They were selected for prosody practice within longer discourse rather than only at the sentential level (Hardison & Sonchaeng, 2005). In addition, the dialogues were short and full of varieties of suprasegmental features. Since the participants were university students, they could benefit from the practice with dramatic intonation and become more aware of the uniqueness of English suprasegmentals in everyday situations in the campus.

The teacher followed suggestions made by Chun (2002) for teaching suprasegmentals. For instance, regarding the “thought group,” the teacher printed a few dialogues from the learners’ textbooks on paper sheets and distributed them among the learners. Then, she wrote the dialogues on the board and read them aloud. Simultaneously, the learners inserted slash marks (/) in dialogues printed on the papers to indicate thought groups so that they could enhance their ability to define this concept. The similar practice used to the instruction on intonation. Choral and individual reading of the dialogue followed, with the teacher noting to the intonation contours. Other instructions such as tapping and clapping were applied for teaching rhythm and timing, for the learners to notice that English is a rhythmic and melodic language and differs from their native language completely.

Through the production stage, MyET provided numeric scores on the learners’ performance. The teacher recorded their highest scores in her weekly journal and uploaded them in a data processor (Microsoft Excel) to create tables and graphical representation for further analysis. The learners wrote down their reflection notes on

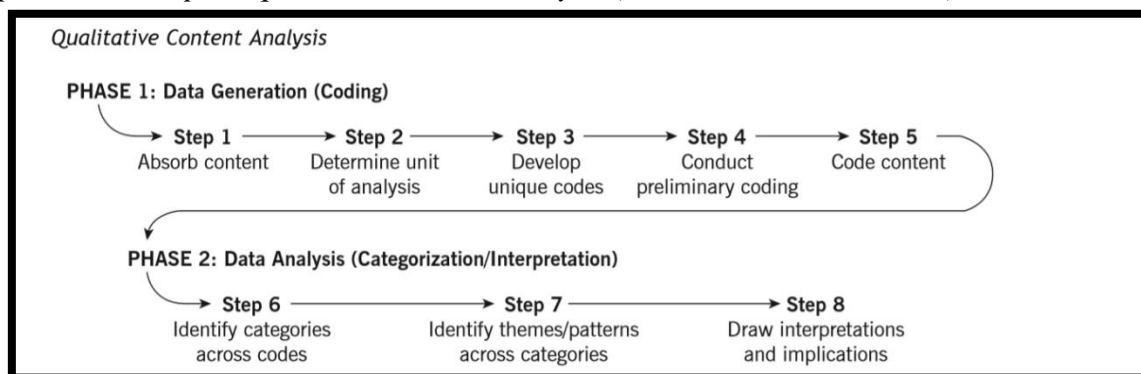
their attitudes and ideas toward practicing with MyET. They spent 30 minutes practicing MyET and 20 minutes writing down their reflection notes used for further analysis. In the last session, the researchers conducted a semi-structured interview to extract the learners' comments, suggestions, and opinions.

## Data analysis

Qualitative data analysis is multi-phased and involves several stages of coding from initial or open coding, through category creation, to further conceptual development (Hahn, 2008). Kuckartz (2014) believed that in content analysis, the coding scheme is based on categories designed to find major themes in a text. In this method, researchers began by analyzing the text itself and allowed the theme to appear. According to Roller and Lavrakas's (2015) qualitative data analysis model (see Figure 2), the early stages of analysis involves the careful reading of the data sources (e.g., observation notes, reflection notes, interview transcripts) and developing initial codes. The next stage of qualitative data analysis involves the development of categories as a means of bringing order and organization to the initial codes. The final stage of data analysis involves the further refinement of codes and categories, which are then arranged around the broad themes emerging from the analysis. The process ends with interpreting the emerged themes and drawing implications. The researchers utilized this approach with the aid of qualitative coding software named NVivo (10.0) for storing all the transcriptions, creating codes, and grasping the themes of the notes.

**Figure 2**

*phases and steps in qualitative content analysis (Roller & Lavrakas, 2015)*



## Findings

To respond to the only research question of this study, 'To what extent does CAPT improve Iranian EFL learners' use of suprasegmental features?' the participants' learning behaviors were observed and field notes were taken during training sessions. The quality of their suprasegmental features was measured and recorded at the first, third, fifth, and seventh sessions. Moreover, in considering the themes from each of the participants in this case study, the researchers sought similarities expressing the



significant factors of CAPT that influenced learning and using suprasegmental features. These themes were grouped according to similar ideas and then five different themes emerged from collected data. The following presents the themes and narrative of the subjects' performance.

- **Instant and visual feedback as the prerogative of CAPT**

Students receive maximum benefit from feedback only if it is given promptly and the visual representation of the learner's articulation and that of a native speaker are best shown simultaneously. Delayed positive feedback reduces the encouragement and reinforcement, and delayed negative feedback affects the crucial knowledge a student must master (Hattie & Timperley, 2007). The visual feedback provided by Apps such as MyET requires no special access to in-class technology and demotes much of the pronunciation practice as homework. Therefore, not only does this particular feedback system produce perceptible benefits for learners, but also it allows for practical combination into the L2 general skills classroom without any special modifications or need for spending large amounts of in-class time (Olson, 2014).

Almost all the participants believed that instant and visual feedback helped them detect their mispronunciation at the very first stage. Sanaz noticed the necessity of providing adequate feedback from all the language learning software:

There should be a correction and feedback system in every CAPT system and then it will be complete and perfect. Since we don't know what aspect of our pronunciation is wrong, we should see our errors in detail and realize why they are wrong and how they should be corrected.

In some sort of CAPT programs, which are equipped with Automated Speech Recognition (ASR) technology, the participants' utterances are juxtaposed with the standard speech samples provided by the system. They are displayed as the waveforms so that the learners can compare them and notice the most relevant features they should match. Moreover, all the segmental and suprasegmental features of learners' speech are analyzed in detail and the system provides an accurate report on their performance and determines the aspects that should be corrected. Arad made the following comments in referring to this beneficial option:

'...the program puts your speech and standard form together... it can be analyzed by yourself easily .... it helps you to find out which parts of your pronunciation should be modified.'

- **Fostering motivation, autonomy, and privacy**

Traditional pronunciation teaching in the classroom can be boring, frustrating, and inadequate; students lose interest and motivation in learning. CAPT programs can provide students new ways to learn pronunciation through games, animation, and role play, which can make drills more interesting. Moreover, appropriate CAPT systems can allow training to be personalized and maximized. Special exercises can be selected to meet learner's specific needs. Besides, the opportunity to practice is not limited to the

time a teacher is available. Since computers are tireless, learners can increase the time on tasks infinitely. It may help learners to work on pronunciation learning individually and ubiquitously (Reed & Levis, 2015).

Teachers need to pay much more attention to the less motivated and low proficiency learners. They would use the computer for fun such as playing games, chatting, or Web surfing. Asking them to achieve a small part of their learning goal (the correct pronunciation, timing, and rhythm of one or two sentences) and focus on them until they learn completely can help learners to enhance their motivation. They can gradually increase the time and amount of their practice. Furthermore, these students need the teacher's constant reminding and supervision. (The teacher/researcher's self-reflection 1).

Tara had this to say about her opinion toward the role of CAPT in increasing motivation and interest:

For me, it was a wonderful opportunity to receive English instruction in a casual stress-free environment where I was not graded for my performance. MyET uses various interesting and encouraging methods for teaching pronunciation. It makes you speak more, compare your pronunciation with native-speakers, and be proud of your progress.

Being aware of their level of competence, and weaknesses in particular, as explained by Daniel below, was an important factor:

I knew like when I started that my English pronunciation is not as good as I want to be and then when I worked with CAPT, I felt really good. It didn't disappoint or frustrate me at any stage. It was a nice eye-opener and an encouraging tool ... yeah, it really works for those who want to be an independent learner and have their private learning environment.

Almost all the participants mentioned that CAPT can increase the learner's self-confidence as well. In this sense, Arad expressed his view in the following statements:

'What should be noted is the learners use the program by themselves and improve their self-monitoring skills. In MyET, tasks and exercises have been provided by different levels of difficulty and it would increase the learners' self-confidence.'

In Tara's opinion, privacy is a very important factor and she referred to it in her reflective notes in this way: 'Computer feedback is precise, immediate, personalized, and private that is only visible to me. Nobody but me can see my errors and it reduces my stress and fear of making mistakes in front of other students.'

- **Saving time, money, and energy**

Using CAPT enables learners to choose appropriate timing for learning. It provides the opportunity of learning anywhere, anytime, and anything they want which is very true for computer-based instruction in CAPT. Learners are given a chance to study, exercise, and review the materials as many times as they want without any

limitation. Moreover, there is no need to register inexpensive language institutes and pay high tuition.

CAPT also cost-effective in the sense that there is no need for the learners to travel to language institutes. They don't need to spend their time in heavy traffic to attend English classes in specific institutes, which are located far from their place. Furthermore, it is cost-effective in the sense that it provides opportunities for learning for the maximum number of learners with no need for many buildings. Sanaz expressed her view that CAPT could help her to manage her time and save money and energy. In a part of her interview, she pointed to this issue:

I can use it whenever and wherever I want. I can listen to native-like pronunciation over and over, and repeat the sentences until I learn them perfectly. As a married woman, I should spend more time at home and work on my language skills in my spare time. I can simply turn on my laptop and have access to my tutor...Your computer would be your private teacher...

- **No feelings, sympathy, or flexibility**

The learning situation that a foreign language learner encounters is varied and dynamic. CAPT programs utilize artificial intelligence, and it cannot completely deal with unexpected learning problems, meet the learners' emotional and affective needs, and respond to learners' questions immediately as a human teacher does. Moreover, each learner has her/his unique and specific learning style that is in opposition to predetermined learning patterns in computer-assisted learning. Negar uses very interesting expressions about this issue:

We can't interact with the virtual teacher directly and ask our questions. In addition, there is no face-to-face communication or extra information about that specific area of pronunciation. Every message is like a command or recommendation and to a computer, it doesn't matter if you learned the lesson or not. You should arrange your learning plan according to the software, but we know that every student has her particular learning style.

Arad believed that computer-assisted instruction for learning suprasegmentals did not give him very clear and adequate advice and it should be more similar to instruction that is provided by human teachers. Daniel also agreed with Negar and Arad and mentioned the human teacher knows their students' weaknesses and strengths well; therefore, she can propose them the best method of learning. He referred to other differences between CAPT and a human teacher and stated:

My teacher reads the text slowly so I can realize the content easily, but the speaker in MyET speaks a little faster. Furthermore, my teacher gives extra information about the text, makes jokes, and explains the meaning of the words in Farsi whenever it needs. She also reads long sentences separately so that we can repeat them readily.

It seems that CAPT is more suitable for adult learners with a higher level of proficiency. They need teacher's applause, emotional and affection care, or sympathy less than younger learners do. A primary computer operation skill also is needed for utilizing the software and interpreting the waveforms (the teacher's reflective notes 2). Students' use of CAPT with peers, groups, and teachers in the classrooms may decrease the limitations of language learning method, compensate deficiencies in providing feedback, and help the user with interpreting feedback and waveforms. This approach can also increase their interest and motivation but decrease the frustration of using technology tools.

- **Interruption in the learning process due to the breakdown in software, hardware or poor internet connection**

Sometimes computers fail to function properly. To prevent many of these failures, all the computer programs and anti-virus software should be kept up-to-date. Hardware can also fail and it is inevitable. Furthermore, in a few countries, the slow speed of the Internet may cause an interruption in using web-based programs. Unlike other participants who did not refer to any breakdown in software/hardware, Arad mentioned two problems with his computer and network connection. He made the following comments in summarizing his troubles of using computer technologies for pronunciation learning:

Actually, I had some problems with my computer, software, and the Internet connections. The applications such as MyET need the Internet connection and you can use them just in on-line mode. Sometimes the Internet speed was low so the program wouldn't run. Rarely did the program not make any response to the clicks and I had to rerun it and subsequently, all the history and previous exercises and scores got lost. Besides, I had some problems with my headphones and the battery of my laptop.

Network connection problems or hardware/software breakdown accrues rarely; therefore, they may be ignored in CAPT programs. Through this research, the participants just encountered network problems once or twice. The problems were solved easily and they did not threaten the training process seriously (the teacher's reflective notes 4). However, this issue should have been mentioned here as a disadvantage of using CAPT in different regions or situations. Moreover, using web-based programs needs Wi-Fi connection and it may cost the users and some of them cannot afford it. Inaccessibility to Wi-Fi particularly out of homes and labs is another limitation of web-based programs.

### **The narrative on participants' performance**

Below is a brief report on the participants' performance progress evidence according to the researcher's observation, field notes, and the statistical analysis recorded by MyET is in three areas (stress, intonation, and timing) on a scale from 0 to 100. The researchers compared participants' improvement with their performance more clearly and precisely as the study progressed.

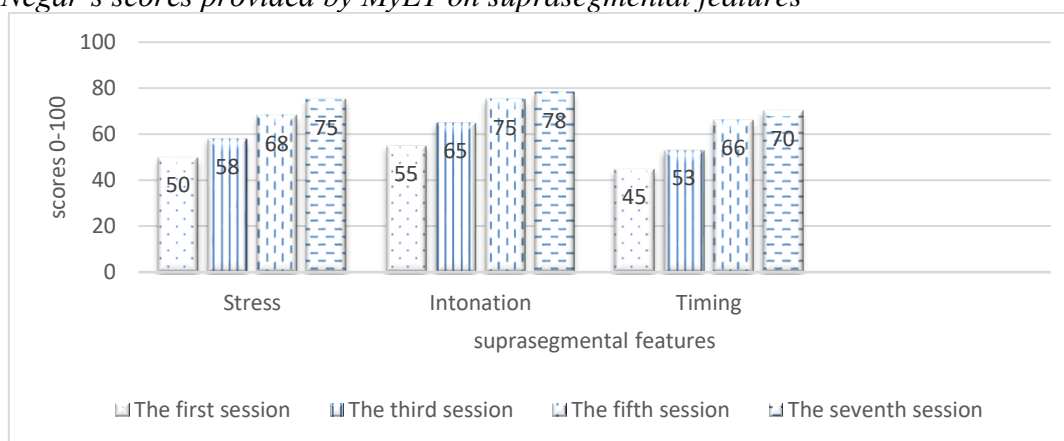
## Negar

Negar was the youngest participant as the most motivated and the keenest member of the group. She was a freshman in law and had taken a course on computer science before the project began. In her interview, she stated that she spent three to four hours working on suprasegmentals with MYET at home per week and found it constructive and absorbing. She also mentioned that she enjoyed working with the App and using technology tools for learning suprasegmental features. During training sessions, it appeared that with appropriate teacher guidance, Negar quickly began developing her self-monitoring, self-correction, and problem-solving skills.

Analyzing her speech showed that her segmental pronunciation became accurate, however, she had some problems with consonant clusters and short vowels. At the early stages of the study, the speed of her speech was a little slower than normal pace, but when the study progressed a significant tune in the speed of speech, was obvious. Unwanted pauses and fillers like ‘uh’ or ‘um’ had been decreased and the learner sounded more fluent and natural. Based on direct observation, the scores presented by the App, and her noticeable progress, it can be guessed that motivated and eager learners with a higher level of proficiency in using computers are more successful in utilizing CAPT programs. Figure 3 shows her performance and noticeable improvement throughout the training period. As can be seen, she made the most progress during the study in her timing and stress skills (25%) which is from 50% to 75% considerably consistent with qualitative evidence obtained from direct observation.

**Figure 3**

*Negar's scores provided by MyET on suprasegmental features*



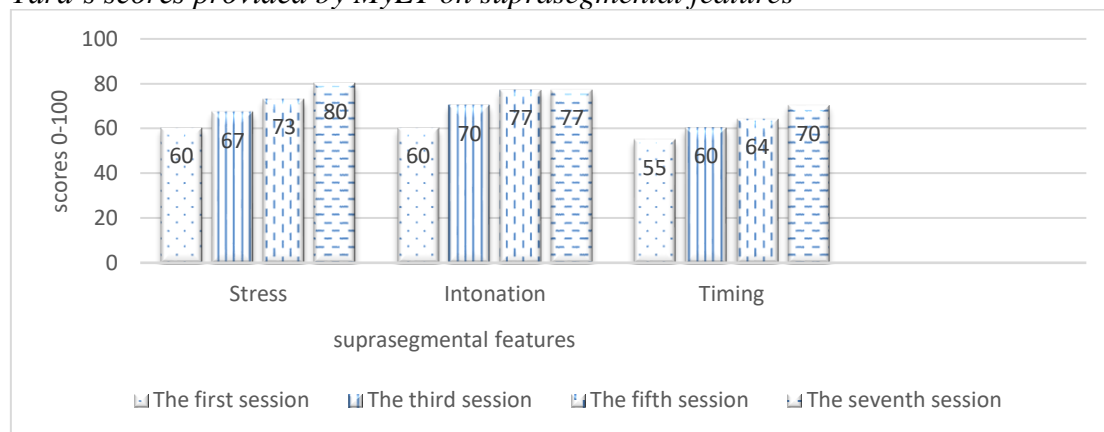
## Tara

Another participant, Tara, also made evident progress. She was a junior student of midwifery and spent a lot of her time in hospital. Nevertheless, she attempted to make time to work on her pronunciation skills. Initially, her speech sounded monotonous with limited use of falling/rising intonation. Her yes/no questions were not raised high enough and statements did not fall completely. Her use of word stress and sentence stress was relatively satisfactory. Therefore, she was supposed to focus on learning and using different forms of intonation.

At the final stages of the study the researchers noticed that although her segmentals were not accurate, her speech was perceived to be more comprehensible due to the use of appropriate intonation and stress patterns after receiving treatments. It can prove the claims of the priority of suprasegmentals over segmentals. Her performance in the three areas of suprasegmentals examined in this research can be seen in Figure 4. As can be seen, she showed the most improvement in stress and intonation areas by 20% (from 60% to 80%) and 17% (from 60% to 77%), respectively. These scores properly confirm qualitative data gathered from direct observation.

**Figure 4**

*Tara's scores provided by MyET on suprasegmental features*



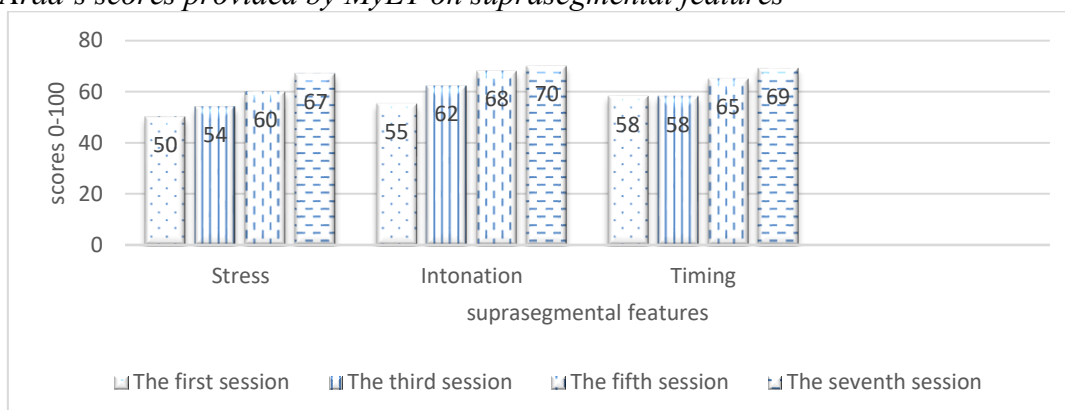
## Arad

Arad was a junior student in civil engineering. He declared that he spent only two to three hours per week working on his pronunciation skills at home. He seemed a little unmotivated and it might be the reason for his insignificant improvement throughout the project. He needed to be reminded constantly and the teacher tried to use more interesting and varied activities to raise his motivation and whole the group's as well.

Initial analysis of his speech revealed that due to his inaccurate use of rhythm, he pronounced long and short vowels incorrectly. His rhythm and timing were generally fine, but sentences lacked speed and sound too slow. He showed slight progress in the CAPT activities that he worked on. His communicative proficiency showed improvement and his listening comprehension also improved. He became more aware of his speech, overly self-conscious of making pronunciation errors, and paid more attention to the word stress and sentence stress. Yet, there were times when he frequently made the same mistakes. His gradual improvement is obvious in Figure 5. He showed the most progress in his stress skills by approximately 17% accretion, that is, from 50% to 67%. It may confirm that the program was effective for him to enhance this skill more than other skills.

**Figure 5**

*Arad's scores provided by MyET on suprasegmental features*



## Daniel

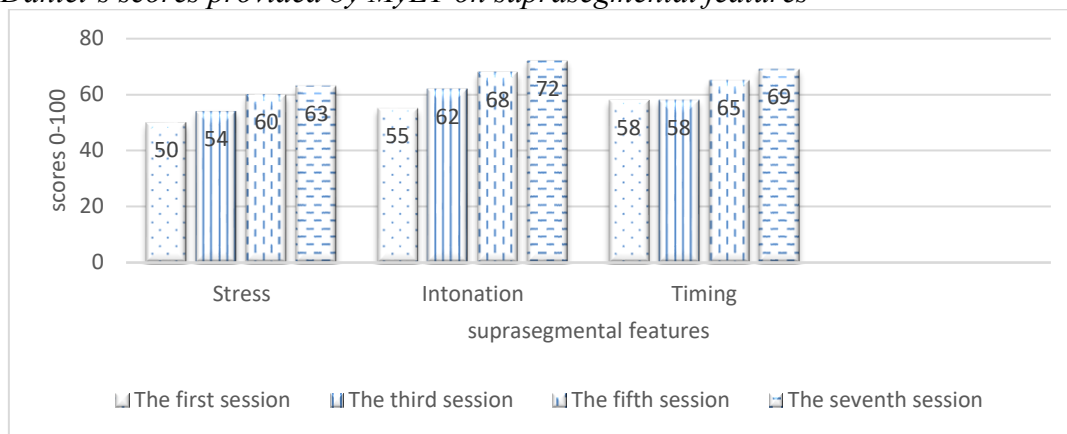
Daniel, a freshman in accounting, initially showed some problems in falling and rising intonation and rhythm of the speech. He stated that he regularly watched English movies and series, listened to English songs, and enjoyed them. In his reflective notes, he stated that it was his first time to learn pronunciation using a computer and showed great interest in MyET program. Moreover, he tended to practice conversational English and learn common expressions.

Analyzing his speech initially revealed that he failed using intonation correctly in the different forms of sentences and there was a general tendency to falling intonation in statements as well. Furthermore, he had some difficulties in detecting the words that carry the prominence stress in sentences. At the beginning of the CAPT program, he worked on those skills that he experienced the most difficulties with. Since he had non-standard stress patterns, he worked on the Syllable Stress activities. Often he could not correctly identify certain phrases from the model sentence, which led to mispronunciation. Either he did not know the meaning of the phrases or he did not know how the phrases were pronounced in English.

Noticing his speech during the program indicated that the participant gradually realized how to produce rising-falling intonation patterns while reading aloud wh-question and rising intonation patterns while reading aloud yes/no questions relatively. Additionally, analyzing his performance showed that he became highly aware of his misuse of stress patterns and making pronunciation errors. He commented that when using a dictionary, he pays more attention to the stress of words. The improvement in producing appropriate intonation and stress patterns might be attributed to the effect of receiving focused instruction of intonation and practicing with computer-assisted pronunciation training software. His gradual improvement is obvious in Figure 6. According to the results, the participant made improvements in his intonation skills (from 55% to 72%) more than other skills, stress (13%), and timing (11%) and it is in line with the results gained from direct observation.

**Figure 6**

*Daniel's scores provided by MyET on suprasegmental features*



### Sanaz

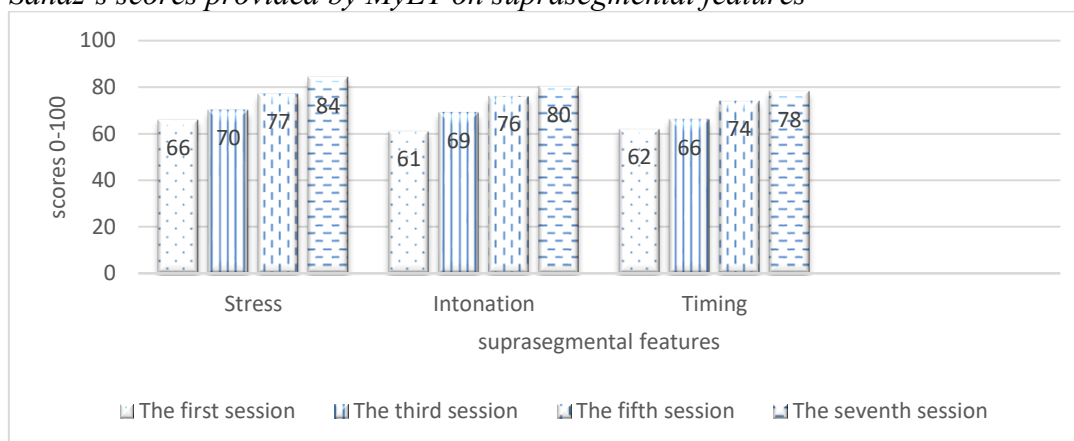
Sanaz, a senior MBA student, was very eager to improve the quality of her pronunciation. She was married and had a two-year-old child but it could not affect her involvement in the project. During training time, she always listened carefully, took notes of important points, worked hard on tasks, and asked several important questions. She seemed to be more active in the learning process than other participants.

In the initial stages of the training process, she suffered poor stress, wrong use of intonation, and wrong pause patterns. For example, “keep it up to date” should be pronounced like “keepitup” to date but she pronounced it disconnected and word by word like a robot. She also incorrectly stressed keywords such as “Internet, develop, interesting.” Moreover, the rhythm was ripply with inappropriate pauses and her stressed syllables were not strong enough. The further analysis of her speech indicated that the participant significantly improved in using suprasegmentals. Her rhythm and stress got fairly better and whenever she made an error she immediately corrected it. Figure 7 presents her performance and scores on suprasegmental features throughout the study. According to the scores, she made considerable enhancement in all three areas. The scores provided by the App show a 19% growth in intonation, which is noticeable progress. Her stress and timing skills also exhibit 18% and 16% enhancement, respectively. This numerical data supports the mentioned observation results about this participant.



**Figure 7**

*Sanaz's scores provided by MyET on suprasegmental features*



## Discussion

According to the research question, this study focused on the extent to which CAPT assisted five Iranian EFL learners to improve their pronunciation skills in suprasegmental level. A detailed analysis revealed that individuals made almost the same gains in their use of suprasegmentals with different amounts. Rating and analysis of participants' speech samples showed that all of them made a noticeable degree of improvement in using suprasegmentals. For example, Daniel made an evident improvement in this area and showed major progress in the use of stress, rising and falling intonation specifically.

Based on the results of the present study, all the participants had serious problems with timing that is a crucial and common issue among Persian speakers of the English language. Persians speak a syllable-timed language (Farsi) in which the length of an utterance depends on the number of syllables, while English is a stress-timed language and moves from stress to stress and it does not matter how many unstressed syllables fall in between. Therefore, the learners were supposed to reduce the time assigned to unstressed syllables; it would be achieved by spending more time practicing rhythmical utterances with the help of CAPT.

Scrutinizing speech samples also showed that Negar gained more progress compared to other participants during the project. She tried to speak fluently with evidence of self-correction. After a while of training sessions, she mentioned in her reflective notes that she became much self-conscious of her pronunciation that she tried to use more suprasegmental features in speaking. CAPT may have raised her self-awareness of using suprasegmental features, which led to self-monitoring. The enhancement of Negar's self-awareness and self-monitoring skills in this study was consistent with Kim (2012), stating that CAPT increases these two skills in learners.

The results of the present study also are consistent with Calvo-Benzies' (2017) study. She declared that after a computer-based training period, the participants tended to use similar tools again for learning other language skills because they are motivating and engaging. One of the participants of the present study, Arad, was relatively a less-motivated learner and would spend less time and pay less attention to the program

initially. Low motivated students like Arad can benefit from more attractive applications with interesting features, more fun, and pleasant visual effects provided by CAPT. In his interview, he stated that observing his progress with scores, the waveforms that were getting close to standard form, and variety in lessons and tasks that served his needs in communication with his classmates and friends, made him more motivated and interested in using CAPT.

The literature emphasized that suprasegmental features play a more important role in listeners' evaluation of non-native speakers' speech than segmental features. Researchers such as Saito, Trofimovich, and Issacs (2016) maintained that prosody deficiencies might affect comprehension more than segmental deviation. The present study affirmed previous studies; for example, Tara became easier to make herself understood in English and when the program proceeded, the pace of her improvement became faster. However, the researchers in their observations noticed that her segmentals were not accurate, her speech was perceived to be comprehensible due to the proper use of prosody.

The findings of this study corroborate studies such as Hardison's (2004), Levis and Pickering's (2004), Abuseileek's (2007), and Tanner and Landon's (2009). All of them approved the positive effects of CAPT on suprasegmental features like the present study did. The key similarity of this study with previous ones was focusing on the effects of CAPT on suprasegmental features of pronunciation. This study used a qualitative approach to investigate individual differences, along with calculated data in three different aspects of suprasegmentals i.e. timing, intonation, and stress.

## **Conclusion and implications**

Pronunciation has been neglected and isolated in many EFL contexts. Many tasks and drills presented in modern textbooks look like traditional forms and cannot serve language learners' needs in the communication approach. Therefore, a new approach to pronunciation teaching might compensate for this deficiency. This alternative should motivate learners to enhance their English pronunciation skills both inside and outside the classrooms. It is believed that programs, Apps, and Websites allow students to practice pronunciation wherever and whenever they want while having fun at the same time. In this study, the results of the analysis of learners' speech by MyET, showed that the participants' use of suprasegmentals performance significantly improved after the training. Therefore, using CAPT as a powerful and efficient tool is strongly recommended in teaching and learning suprasegmental features of English pronunciation.

CAPT can assist learners in learning suprasegmental features in all levels of proficiency from beginning to advanced level. CAPT also allows the learners to repeat words, phrases, and sentences as much as they want and need. Furthermore, it can aid learners to save time and money. It is also an effective method for slow learners; thus, it offers them an engaging, encouraging, and additional form of training. CAPT can reduce anxiety in shy learners and increase autonomy and self-confidence as well.

This study also showed that providing instant, private, visual, and individual feedback is the unique privilege of CAPT that human teachers lack it. This option is beneficial for the teachers who deal with a lot of students in crowded classrooms, or

students who want to observe their performance and progress directly. Traditional training is limited to the classrooms and predetermined time and it cannot meet the demands of a large number of students who have only a limited amount of time for face-to-face instruction. Computers remove classroom limitations and combine informal education and formal one. CAPT allows the learners to continue ubiquitous learning.

This study was conducted with only five volunteers and the results may not be easily generalized to other contexts. Future studies may use random sampling strategies to include more participants and conduct in a different context. Furthermore, this program placed a heavy time limitation on the study and lasted for two months. The researchers and the participants met once a week for approximately 60 to 70 minutes. A longitudinal study with longer and more extensive treatment might have a more significant impact on the learners and provide more discerning findings. This project focused on the learners whose level of proficiency estimated lower intermediate. It is suggested that other researchers shift their focus on other target groups such as intermediate or advanced learners.

## References

- Abuseileek, A.F (2007). Computer-assisted pronunciation instruction as an effective means of teaching stress. *JALT CALL SIG*, 3(1-2), 3-14.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, and syllable structure. *Language Learning*, 42, 529-555. <https://doi.org/10.1111/j.1467-1770.1992.tb01043.x>
- Breitkreutz, J., Derwing, T.M, & Rossiter, M. (2002). Pronunciation teaching and practices in Canada. *TESL Canada Journal*, 19, 51-61. <https://doi.org/10.18806/tesl.v19i1.919>
- Calvo Benzie, Y. (2017). Contribution of new technologies to the teaching of English pronunciation. *Language Value*, 9(1), 1-35. <https://doi.org/10.6035/LanguageV.2017.9.2>
- Celce-Murcia, M., Brinton, D.M, & Goodwin, J.M (2008). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.
- Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Amsterdam: Benjamins.
- Derwing, T.M., & Munro, M.J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Philadelphia: John Benjamins.
- Elimat, A.K., & Abuseileek, A.F. (2014). Automatic speech recognition technology as an effective means for teaching pronunciation. *The JALT CALL Journal*, 10(1), 21-47. <http://eric.ed.gov/?id=EJ1107929>
- Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing*, 27(3), 379-399. <https://doi.org/10.1177/0265532210364407>
- Hahn, C. (2008). *Doing qualitative research using your computer: A practical guide*. London: Sage publications.

- Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology*, 8, 34-52. <http://llt.msu.edu/vol8num1/hardison/>
- Hardison, D. M., & Sonchaeng, C. (2005). Theatre voice training and technology in teaching oral skills: Integrating the components of a speech event. *System*, 33(4), 593-608. <https://doi.org/10.1016/j.system.2005.02.001>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Huensch, A. (2018). Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Language Teaching Research*, 1-20. <https://doi.org/10.1177/1362168818767182>
- Kim, A. (2012). Investigating the effectiveness of computer-assisted language learning (CALL) in improving pronunciation: A case study. *Multimedia Language Learning*, 15(3), 11-33. <http://kmjournal.bada.cc/wp-content/uploads/2013/05/15-3-1Kim.pdf>
- Kuckartz, U. (2014). *Qualitative text analysis*. London: Sage publications.
- Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics* 27, 184-202. <https://doi.org/10.1017/S0267190508070098>
- Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32(4), 505-524. <https://doi.org/10.1016/j.system.2004.09.009>
- Merriam, S. (2007). *Qualitative research and case study application in education*. Auckland: Jossy-Bass.
- Munro, M.J., & Derwing, T.M. (2015). A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation* 1:1, 11-42. <https://doi.org/10.1075/jslp.1.1.01mun>
- Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-teaching interface in Computer-Assisted pronunciation training. *Computer Assisted Language Learning* 15(5), 441-467. <https://doi.org/10.1076/call.441.13473>
- Neri, A., Cucchiarini, C., & Strik, H. (2002). Feedback in computer-assisted pronunciation training: When technology meets pedagogy. *CALL professional and the future of CALL research* (pp. 179-188). Antwerp, Belgium: <https://repository.ubn.ru.nl/bitstream/handle/2066/76206/76206.pdf?sequence=1>
- Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393-408. <https://doi.org/10.1080/09588220802447651>
- Olson, D. J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning & Technology*, 18(3), 173-192. <http://llt.msu.edu/issues/october2014/olson.pdf>
- Pokrivcakova, S. (2015). CALL and teaching pronunciation. *CALL and Foreign Language Education*, 29-37. <https://doi.org/10.17846/CALL.2014.29-37>
- Reed, M., & Levis, J. (2015). *The Handbook of English Pronunciation*. Sussex: Wiley & Sons
- Roller, M. R., & Lavrakas, P.J. (2015). *Applied Qualitative Research Design: A Total Quality Framework Approach*. New York: The Guilford Press

- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different levels. *Applied Psycholinguistics*, 37(2), 217-240. <https://doi.org/10.1017/s0142716414000502>
- Tanner, M.W., & Landon, M.M (2009). The effects of computer-assisted pronunciation reading on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology* 13 (3), 51-65. [https://scholarspace.manoa.hawaii.edu/bitstream/10125/44191/13\\_03\\_tannerlandon.pdf](https://scholarspace.manoa.hawaii.edu/bitstream/10125/44191/13_03_tannerlandon.pdf)
- Tsai, P. (2019). Beyond self-directed computer-assisted pronunciation learning: A qualitative investigation of a collaborative approach. *Computer Assisted Language Learning* 32(7), 713-744. <https://doi.org/10.1080/09588221.2019.1614069>
- Warren, P., Elgort, I., & Crabbe, D. (2009). Comprehensibility and prosody ratings for pronunciation software development. *Language Learning and Technology*, 3(3), 87-102. <http://llt.msu.edu/vol13num3/warrenelgortcrabbe.pdf>
- Yoshida, M.T (2018). Choosing technology tools to meet pronunciation teaching and learning goals. *The CATESOL Journal*, 30(1), 195-212. <https://files.eric.ed.gov/fulltext/EJ1174226.pdf>

## Appendix A

Interview questions (adopted from Kim, 2012)

1. Did instant and visual feedback provided by CAPT, help you to learn suprasegmentals properly? If so, how?
2. Did the program help you to be an independent learner? If so, how?
3. What kind of teaching method did you find to be more useful (computer-assisted or traditional human-teacher instruction)? And why?
4. How do you feel about learning pronunciation with the help of a computer? Did you enjoy it? If not, why?
5. What kinds of difficulties and challenges did you encounter while using CAPT?
6. What were the deficiencies and weaknesses of the program?