IMPROVING CHINESE UNIVERSITY EFL STUDENTS' ENGLISH INTONATION THROUGH PITCH MODIFICATION: PRECISION LANGUAGE EDUCATION-INFORMED PEDAGOGY

Zhongmin Li

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in English Language Studies

Suranaree University of Technology

Academic Year 2019

การปรับปรุงทำนองเสียงภาษาอังกฤษให้กับนักศึกษาระดับมหาวิทยาลัยชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศโดยผ่านการปรับแปลงเสียงนำเข้า: การสอนโดยให้ข้อมูลการศึกษาภาษาที่แม่นยำ

นายจงหมิน หลี่

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2562

IMPROVING CHINESE UNIVERSITY EFL STUDENTS' ENGLISH INTONATION THROUGH PITCH MODIFICATION: PRECISION LANGUAGE EDUCATION-INFORMED PEDAGOGY

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Thesis Examining Committee

(Assoc. Prof. Dr. Punyathon Sangarun)
Chairperson

(Dr. Butsakorn Yodkamlue)
Member (Thesis Advisor)

(Asst. Prof. Dr. Harald Kraus)
Member

(Asst. Prof. Arjuna Peter Chaiyasena)
Member

(Dr. Nattaya Puakpong)
Member

(Prof. Dr. Andrew Lian)
Member

(Assoc. Prof. Flt. Lt. Dr. Kontorn Chamniprasart)
Vice Rector for Academic Affairs
and Internationalization

(Assoc. Prof. Dr. Weerapong Polnigongit)
Dean of Institute of Social Technology

จงหมิน หลี่ : การปรับปรุงทำนองเสียงภาษาอังกฤษให้กับนักศึกษาระดับมหาวิทยาลัยชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศโดยผ่านการปรับแปลงเสียงนำเข้า: การสอนโดยให้ข้อมูลการศึกษาภาษาที่แม่นยำ (IMPROVING CHINESE UNIVERSITY EFL STUDENTS' ENGLISH INTONATION THROUGH PITCH MODIFICATION: PRECISION LANGUAGE EDUCATION-INFORMED PEDAGOGY) อาจารย์ที่ปรึกษา : อาจารย์ ดร.บุษกร ยอดคำลือ, 300 หน้า

การศึกษานี้มุ่งพัฒนาการออกทำนองเสียงการพูดภาษาอังกฤษของนักศึกษาระดับมหาวิทยาลัยชาวจีนที่เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศโดยการปรับระดับเสียง (การปรับทำนองเสียงที่ไม่ถูกต้องของนักเรียนให้ถูกต้องเพื่อเป็นข้อมูลรับเข้าในการเรียนรู้ผ่านการฟังและออกเสียงตามภายหลัง) การศึกษานี้มีวัตถุประสงค์เพื่อ 1) ตรวจสอบประสิทธิภาพของการปรับระดับเสียงต่อการเรียนทำนองเสียง 2) ตรวจสอบทัศนคติของนักเรียนเกี่ยวกับการเรียนทำนองเสียงโดยการปรับระดับเสียง

ผู้เข้าร่วมการศึกษานี้ได้แก่ นักศึกษาวิชาเอกภาษาอังกฤษที่กำลังศึกษาอยู่ที่มหาวิทยาลัยประเทศจีนจำนวน 66 คน โดยแบ่งเป็นสองกลุ่ม กลุ่มทดลองฝึกออกเสียงภาษาอังกฤษโดยใช้ต้นแบบจากการปรับเสียงของตัวนักเรียนเองเป็นต้นแบบ ขณะที่กลุ่มควบคุมใช้เสียงของเจ้าของภาษาอังกฤษเป็นต้นแบบ ผลการศึกษาแสดงให้เห็นว่าทั้งสองกลุ่มพัฒนาความสามารถด้านการรับรู้ทำนองเสียงและการออกทำนองเสียงอย่างมีนัยสำคัญ อย่างไรก็ตามเมื่อเปรียบเทียบการพัฒนาระดับความสามารถด้านการรับรู้ทำนองเสียงระหว่างกลุ่มควบคุมและกลุ่มทดลอง พบว่าไม่มีความต่างอย่างมีนัยสำคัญ ขณะที่กลุ่มควบคุมพัฒนาความสามารถด้านการออกเสียงได้สูงกว่ากลุ่มควบคุม นอกจากนี้ยัง พบว่าการฝึกการออกเสียงโดยการปรับระดับเสียงพูดของนักเรียนด้วยการเพิ่มสมรรถนะตัวบ่งชี้นั้น กระตุ้นการรับรู้เกี่ยวกับตัวบ่งชี้ทางกลสัทศาสตร์ และทำให้นักเรียนสามารถพูดสภาพจริงทางเสียงระดับสัทศาสตร์ของรูปแบบทำนองเสียงที่ยาก (การขึ้นเสียงสูง และการลงเสียงต่ำและสูง) หรือซับซ้อน (เสียงผสม)ได้ถูกต้องมากขึ้น ผลของแบบสอบถามและการสัมภาษณ์กึ่งโครงสร้างยังแสดงให้เห็นว่านักเรียนกลุ่มทดลองมีทัศนคติที่ดีต่อการฝึกออกเสียงโดยการปรับระดับเสียงของตนเองอีกด้วย โดยเห็นว่าวิธีดังกล่าวน่าสนใจ มีประสิทธิผลและประสิทธิภาพในการพัฒนาปรับปรุงการออกเสียงของนักเรียน

การศึกษานี้นำแนวคิดเกี่ยวกับการศึกษาภาษาที่แม่นยำของการรับรู้การพูดมาใช้ในการสอนการออกเสียง จากหลักฐานเชิงประจักษ์แสดงให้เห็นว่าการสอนแบบแม่นยำที่มุ่งเน้นปัญหาเฉพาะของผู้เรียนรายบุคคล และการแก้ปัญหาดังกล่าวด้วยการแทรกแซงที่เหมาะสมกับแต่ละบุคคลทำให้การเรียนมีประสิทธิภาพมากขึ้น

สาขาวิชาภาษาต่างประเทศ
ปีการศึกษา 2562

ลายมือชื่อนักศึกษา Zhongmin Li
ลายมือชื่ออาจารย์ที่ปรึกษา
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

ZHONGMIN LI : IMPROVING CHINESE UNIVERSITY EFL STUDENTS' ENGLISH INTONATION THROUGH PITCH MODIFICATION: PRECISION LANGUAGE EDUCATION-INFORMED PEDAGOGY. THESIS ADVISOR : BUTSAKORN YODKAMLUE, Ph.D., 300 PP.

ENGLISH INTONATION/ PITCH MODIFICATION/ PERCEPTION/ PRODUCTION/ PRECISION LANGUAGE EDUCATION

The present study aimed at improving Chinese university EFL students' English intonation through pitch modification (modifying students' incorrect intonation and playing it back to them as the input for learning). The objectives of this study are twofold: 1) to examine the effectiveness of pitch modification for intonation learning; 2) to investigate the students' opinions of learning intonation through pitch modification.

The participants of this study were 66 first year English major students studying at a university in China. They were assigned into two groups: the experimental group and the control group. The treatment was a form of English intonation training wherein the experimental group received the training by using the students' own modified speech as the model for learning while the control group used a model produced by a native English speaker. The results indicated that both groups' ability in intonation perception and production were significantly improved after the training. Comparisons of the two groups' performances showed that in perception level, there were no significant differences between the two groups; while in production level, the experimental group outperformed the control group. It was found

that compared to the native speaker's model, pitch modification as a means of cue enhancement could arouse students' awareness on critical acoustic cues and enable students to produce more accurate phonetic realizations of the intonation patterns that were difficult (rising tone and falling-rising tone) or complex (compound tones) to them. Moreover, results from the questionnaire and semi-structured interviews with the students from the experimental group revealed that the students held positive attitudes towards pitch modification for intonation learning. They thought that it was interesting, efficient and effective for improving their pronunciation.

This study was an implementation of the concept of precision language education in the field of pronunciation instruction. It proposed a precision approach for pronunciation instruction and provided empirical evidence that this kind of instruction by targeting individual learner's specific problems and dealing with those problems with tailored interventions could make learning more effective.

School of Foreign Languages

Academic Year 2019

Student's Signature Zhongmin Li

Advisor's Signature

Co-Advisor's Signature

ACKNOWLEDGEMENTS

When I was about to complete this dissertation, the rainy season of Thailand has not yet come to the end. The big drought of this rainy season made me lose my sense to season. Or, maybe more than three years of stay in Thailand faded my memory of the rotation of four seasons in my hometown. Or, maybe just like what people say, when you are too concentrated on one thing, you will forget about the ambient world. I will always feel astonished when I look back on how long and how far did I go. But most importantly, I should keep in mind that I would not have gone this long and far without the help, support, cooperation, and encouragement from a number of people, to whom I would like to express my sincere gratitude here.

First and foremost, my special thanks are extended to my supervisor, Prof. Dr. Andrew Lian, who enlightened me about the research topic and gave me inspiring guidance, invaluable advice, and constant encouragement. My heartfelt thanks also go to my other supervisor, Dr. Butsakorn Yodkamlue, for her detailed comments, constructive suggestions, and academic supervision in the research and dissertation writing.

Deepest thanks are also expressed to the members of my thesis examination committee: Assoc. Prof. Dr. Pannathon Sangarun (who served as chair), Asst. Prof. Dr. Harald Kraus, Asst. Prof. Dr. Arjuna Chaiyasena and Dr. Nattaya Puakpong, who gave me valuable comments and insightful suggestions on my work.

I am greatly indebted to my teachers, Dr. Sirinthorn Seepho, Dr. Anchalee Wannaruk, Dr. Adcharawan Buripakdi, Dr. Jitpanat Suwanthep, Dr. Issra Pramoolsook, Dr. Suksan Suppasetseree, as well as many other faculty and staff members of the School of Foreign Languages, Suranaree University of Technology (SUT). They brought me to the field of scientific research and offered me great help during my study in Thailand.

I also owe great thanks to my home university, Hunan University of Science & Engineering (HUSE), for rendering me conveniences to study abroad. Great thanks also go to my colleagues and students in HUSE, who helped me collect data or participated in the experiment, as well as my Chinese and international friends in SUT, who contributed laughter, company, and enrichment to my PhD life.

Last but not the least, I wish to express my greatest gratitude to my family members and my girlfriend, Cui Gu, for their love, understanding, encouragement and support throughout the journey of my PhD study. They did not care much about how far I can go, but they did care about how tired I was in the journey. With their care, love and support, I can always feel relieved and rejuvenated to go further.

Zhongmin Li

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

AC	Air Conducted
BC	Bone Conducted
AM	Autosegmental Metrical
CALL	Computer Assisted Language Learning
CET-4	College English Test Band 4
CG	Control group
EG	Experimental group
EFL	English as a Foreign Language
ESL	English as a Second Language
HUSE	Hunan University of Science & Engineering
L1	First Language
MT	The Motor Theory of Speech Perception
NNESTA	Nonnative English-Speaking Teaching Assistant
IELTS	the International English Language Testing System
PE	Precision Education
PLE	Precision Language Education
SFL	School of Foreign Languages
TEM-4	Test for English Majors Band 4

CHAPTER 1

INTRODUCTION

The present study aims at improving Chinese university EFL students' English intonation through modifying the pitch of the students' speech productions and playing it back to them as the model for learning. This introductory chapter first presents the background of this study. Then, the statement of the problem, research significance, research objectives, and research questions of this study are detailed. Lastly, definitions of some key terms related to this study are presented.

1.1 Background of the study

The main aim of language teaching and learning is to enable students to communicate in the target language (Harmer, 2001). Being able to speak in a foreign language involves a number of sub-skills of which pronunciation is the most important (Fraser, 2000), and intelligible pronunciation is an essential component of communicative competence (Morley, 1991). Lacking an intelligible pronunciation can result to an abortive and/or meaningless communication (Nikbakht, 2010). Celce-Murcia et al. (1996) hold that there is a "threshold level" of pronunciation for nonnative speakers of English, and if their pronunciation falls below this level, it is likely that

their speech will be difficult to be understood despite that their English proficiency is claimed to be the high level. Previous studies (Hinofotis et al., 1981; Rubin & Smith,1989; Gallego, 1990) investigated student-reported communication breakdowns with their nonnative English-speaking teaching assistants (NNESTAs) and found that among the linguistic factors, pronunciation was the leading cause of unintelligibility in the NNESTAs' presentations. Therefore, "with good pronunciation, a speaker is intelligible despite other errors; with poor pronunciation, understanding a speaker will be very difficult, despite accuracy in other areas" (Fraser, 2000, p.7). Pronunciation ability also influences language learners' performance in other aspects of language proficiency. Leather (1983) claimed that the learning of pronunciation may interfere with the learning of grammar or vocabulary; Wong (1993) found that poor pronunciation could affect students' reading and spelling; Gilbert (1995) believes that pronunciation practice is an aid to listening comprehension.

As English has become a lingua franca, the past decades witnessed a population explosion in numbers of English learners. Gary & Charles (2017) reported that the total number of English speakers around the world was about 1.121 billion. However, only 378.2 million of them were native speakers, which was about a half of the number of non-native speakers. Celce-Murcia et al. (1996) pointed out that there are particular groups of non-native English learners whose pronunciation deficiencies may place them at a professional or social disadvantages, such as international business personnel,

immigrant residents, refugees, professors or research scholars in higher education, and international students studying abroad, etc.. Jenkins (2004) argued that pronunciation plays a critical role in preventing communication breakdowns in international context, especially interactions between non-native speakers from different L1s. Furthermore, as a surface structure phenomenon that is most noticeable, pronunciation is perhaps the linguistic feature most open to judgement (Canagarajah, 2005). Pronunciation deficiency can decrease language learners' confidence of speaking, restrict their social interactions, and negatively affect the listeners affective judgement on them (Gilakjani, 2012). All these contribute to the point that pronunciation is important, and it is as important as other areas of language learning like grammar or vocabulary. Therefore, the teaching of pronunciation should be emphasized or at least be treated equal with other areas of language education in EFL classroom.

However, in reality, pronunciation practice is not always emphasized (Dalton,1997). The position of pronunciation in language teaching has always been determined by ideology and intuition rather than research (Levis, 2005), and teaching pronunciation is viewed as meaningless non-communicative drill-and-exercise gambits (Morley, 1991). Furthermore, most language proficiency tests are in written form focusing on testing students' reading, writing, or listening, with little attention paid on speaking. Thus, the teaching of pronunciation is thought to have little contribution to the teachers' teaching achievements. As a result, pronunciation is always described as "the

Cinderella of language teaching", which means that often low level of emphasis was placed on this very important language skill (Dalton, 1997).

There is a gap between the unequal treatment of pronunciation and students' needs, given that speaking is one of the most required job skills for graduate students and pronunciation is one of the most difficult skills to acquire (Gilakjani, 2012). Alghazo (2015) found that EFL learners were usually empowered with motivation and willingness to improve their pronunciation and asked more time to be devoted to pronunciation. Considering that ignoring students' needs is an abrogation of professional responsibility (Morely, 1991), it is imperative that students' urgent need-intelligible pronunciation for communication-should be served with instruction.

Pronunciation consists of the segmental aspects and the suprasegmental aspects. Segmentals refer to individual sounds, including consonants and vowels. Suprasegmentals, also known as prosodic features, include intonation, rhythm, stress, intensity, duration, etc.. In pronunciation teaching practice, the segmentals have traditionally received more attention than the suprasegmentals (Golombek, 2012), and the majority of pronunciation handbooks or textbooks also spend most ink on the vowels and consonants while leaving the instruction of suprasegmentals as the luxuries for the end of a course (Leather, 1983). Some teachers are not aware of the importance of suprasegmentals and take it for granted that they can be learned naturally if students can pronounce the individual sounds well. Other teachers think that the

suprasegmentals are important but they do not have adequate knowledge and therefore lack the confidence to teach it (Clennell, 1997; Paunović & Savić, 2008; Lengeris, 2012).

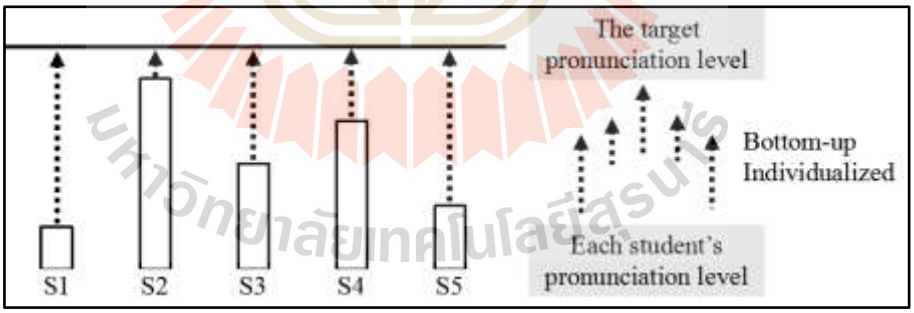
For several decades, researchers have appealed for a paradigm shift of pronunciation teaching-the teaching priority should be shifted from the segmentals to the suprasegmentals (Lantolf, 1976; Morley, 1991; Anderson et al., 1992; McNerney & Mendelsohn, 1992; Chun,1998; Derwing et al., 1998; Munro & Derwing, 1999; Pickering, 2001; Chela-Flores, 2003; Jenkins, 2004; Kang, 2010; Gilbert, 2014). Previous studies have shown that the suprasegmental aspects of pronunciation contributes to speech accentedness and intelligibility more than the segmental aspects do (Anderson et al., 1992; Wennerstrom, 1997; Munro & Derwing, 1999; Jilka, 2000; Pickering, 2001; Pickering, 2009; Kang, 2010; Kang et al., 2010). This means that a speaker is likely to have stronger accent and less likely to be understood by the listeners if s/he is weak in pronouncing suprasegmentals. Empirical evidence has also shown that pronunciation instruction by assigning priority to suprasegmentals had greater benefits, for not only can it facilitate the learning of the segmentals and enhance the learners' speech accuracy and intelligibility (Prator,1971; Derwing et al.,1998; Hardison, 2004; Sonia & Abdelkader, 2016), but also it can help to improve the listener's memory and increase the speakers' confidence (Hahn, 2004; Hardison, 2004). Therefore, a pronunciation classroom with priority assigned to and more importance

attached on the suprasegmentals will turn out to be more efficient and less frustrating for students.

The widespread of Internet and smart phones has greatly benefited language learning, especially for the learning of pronunciation. Learners can easily obtain all kinds of pronunciation learning materials from Internet and practice pronunciation at anytime in anywhere they want. However, facing with more choices does not mean that learners are able to choose the materials or methods that can best fit them. As learner's characteristics determine how much and how well they can learn, instruction should be designed to fit learner's characteristics, enabling learning occur in a "just in time, just enough, just for me" (Lian, 2014) fashion.

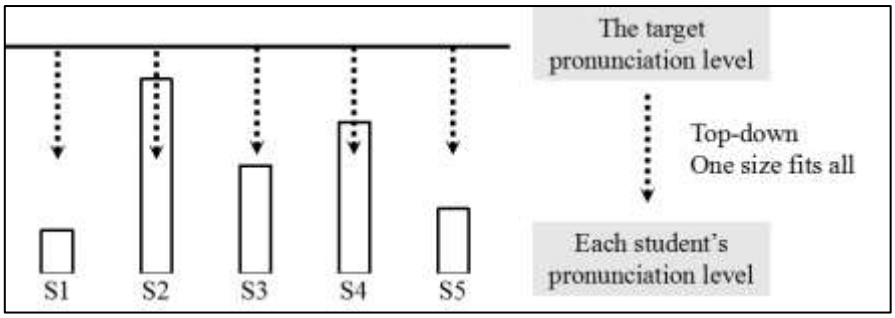
Cook et al. (2018) claimed that modern education should go beyond the question of "did an intervention work?" to "what intervention worked for whom and how did it work?". This involves the question of how precisely we can make use of language learners' information to tailor the instruction to each learner's needs, which is the core concept of precision education (Hart, 2016; Lian & Sangarun, 2017; Cook et al., 2018). Speech modification is a good example of the practice of precision education in the field of pronunciation instruction, which manages to modify the properties of the input speech signal to make it better fit for language learners' perception so as to improve their production. It touches upon the question of "what kind of input should we provide to learners and how should we provide it to them?".

Pronunciation learning involves the process of perceiving the incoming speech signal and producing it with own articulators. The way people perceive sounds and the way people use their articulators are different, which determine that pronunciation learning should be individualized. Pitch modification for intonation learning, informed by the concept of precision education, assumes that each learner's characteristics are different and their pronunciation levels and pronunciation problems are also different. It entails efforts to modify the input to best fit each individual learner's characteristics, i.e., "getting the right intervention in place for the right person for the right reason"(Cook et al., 2018, p.5). Therefore it is a bottom-up approach in nature (Figure 1.1), compared to the traditional standard protocol (top-down) approach (Figure 1.2) which views all students as a homogeneous group (Fuchs et al., 2003).



Note: The bottom-up approach assumes that each student's pronunciation level is different and each student's pronunciation problems are different. This approach starts from each student's problems and gradually approximate to the target level

Figure 1.1 Bottom-up approach for pronunciation teaching



Note: The top-down approach views the whole students as a homogeneous group and offers "one size fits all" instruction to students. As a result, some students' needs can not be met, while to other students, it is a waste of time.

Figure 1.2 Top-down approach for pronunciation teaching

China, with the largest number of EFL learners in the world, has been increasingly attaching importance to its English education. Since opened up to the world in 1978, China has begun to involve in international communications in all areas. More and more Chinese begin to realize the importance of English as a communication tool to engage in international communications. Since 2001, English has become a compulsory subject in primary school. For most Chinese, especially the younger generations, there has been an escalating demand for learning English as it is not only the prerequisite for stepping to higher education, finding a good job, and getting promotion, but also a necessary tool for engaging themselves in international political, economic, or cultural activities. And, to the Chinese students studying out of China, English is even one of the essential skills for survival.

However, despite the efforts, money and resources expended, the outcomes of English language education in China are unsatisfactory and discouraging. According to the statistics of IELTS (Academic) test taker performance in 2017 (IELTS, 2017), among the 40 most frequent test-taking countries or regions, the overall mean band score of mainland China is 5.76, ranking 34. This means that Chinese EFL learners' English proficiency is lagging far behind that of the EFL/ESL learners from other countries. In terms of performance in the four language skills, the mean score for Chinese test takers' speaking is 5.39, ranking 39, last but one, which reveals that Chinese EFL learners have more prominent problems in speaking, the weakest among the four skills (Wei & Su, 2012; He, 2015; Yang, 2016). Many students in China with more than 10 years' study of English are incapable of communicating in English (Han, 2013; Yang, 2016; Hu, 2017). Wei & Su (2012) explained that this is the result of a traditional teaching methodology with insufficient attention paid to listening and speaking and hence producing "deaf-and-dumb" English learners.

Hunan University of Science and Engineering (HUSE), a typical intermediate-level university in China, is located Yongzhou city. Built in 1941, it was originally a teacher training college and it has become a comprehensive university since 2002. HUSE has 15 academic departments offering 46 majors and 16 of which are in pedagogic specialties. School of Foreign Languages (SFL) is one of the academic departments which has two majors: English and Japanese, and they are also teacher

training majors at the same time. In September, 2018, altogether 272 full time students were enrolled to study in SFL, and 204 of them were English majors assigned to 6 classes with about 34 students in each class.

The majority of the English majors will work in primary schools, middle schools, or English training centers after graduation. This determines that their English learning objectives are different from those of other majors. They should have more adequate and accurate knowledge of English and know not only "what" but also "why". In terms of pronunciation, they should have a good mastery of the basic knowledge of phonetics and phonology of English and their English pronunciation level should be relatively high so that they can provide pronunciation model for their students in the future. Teacher's inadequate knowledge of English intonation has always been the excuse for intonation being neglected in the pronunciation classroom (McNerney & Mendelsohn, 1992; Clennell, 1997; Paunović & Savić, 2008; Lengeris, 2012; Reed & Michaud, 2015). It is now urgent to change this situation by teaching intonation to English majors in pre-service teacher training programs, so that they can be equipped with the required knowledge and informed with a better way to deal with pronunciation teaching, and therefore be confident to teach it to their students in the future.

1.2 Statement of the problem

The problems of pronunciation teaching and learning in Chinese universities, represented by HUSE, were analyzed from three sources of data: 1) previous studies

about Chinese university EFL students' pronunciation problems; 2) a questionnaire survey among HUSE English majors; 3) semi-structured interviews with HUSE teachers.

Firstly, previous studies have identified some common characteristics and problems regarding Chinese university EFL students' English pronunciation. Zhang (2015) investigated the international intelligibility of the English produced by Chinese students and found that it was largely intelligible to international listeners, while the intelligibility score was not high, with a mean score of 5.8 (out of 9), and their speech was frequently evaluated by the international listeners as "strong accented, fast, wrong, choppy, monotonous, truncated and hesitant"(p.51). These evaluations were related to suprasegmental features, which implied that the students' pronunciation in the suprasegmental aspect largely affected their overall pronunciation performance. Many researchers have identified Chinese EFL students' pronunciation problems from the aspect of suprasegmental features, such as incorrect stress placement at word level (Gao & Deng, 2009) and sentence level (Chen, 2006b), stress-timed rhythm patterns (Zhu, 2007), incorrect tone choices (Hong, 2012; Chen et al., 2008; Huo & Luo, 2017; Makarova & Zhou, 2006; Rui, 2007), inappropriate pauses (Yang & Mu, 2011), and incorrect assignment of intonation boundaries (Meng & Wang, 2009; Chen, 2006a; Yang, 2006; Xia & Mu, 2008).

Bi and Chen's (2013) cross-sectional study revealed that Chinese university EFL

students' problems in the suprasegmental aspect remained unchanged throughout four years of study, while the segmental aspect was improved to some extent. All these implied that Chinese university EFL students have prominent problems in pronunciation and their problems in suprasegmental aspect were more salient. However, facing with these unsolved problems, the current pronunciation teaching in China is ineffective and unsatisfactory, and it is urgent to change the condition.

In order to investigate the pronunciation teaching and learning problems in HUSE, the researcher conducted a questionnaire survey (Appendix A) among HUSE 204 first year English majors as well as semi-structured interviews (Appendix B) with 4 English pronunciation teachers. Results of the questionnaire and interviews were analyzed from two aspects: student-related pronunciation learning problems and teacher-related pronunciation teaching problems.

From the students' aspect:

(1) Most of the students had problems with their English pronunciation, especially in the suprasegmental aspect. In the questionnaire, the students were asked to respond to the question "How do you evaluate your pronunciation?" by ticking a 5-point Likert scale ranging from "very good" to "very bad". Results showed that most students were not satisfied with their pronunciation (Figure 1.3). Their pronunciation problems lied in various aspects such as vowels, consonants, stress, rhythm, and intonation, among which intonation ranked the first (Figure 1.4).

Results from the interviews with the 4 teachers were also in accordance with the students' responses. The teachers frequently reported that their pronunciation was "strong accented", "monotonous", or "flat". Even though some students can correctly pronounce the individual sounds but their pronunciation was still "strong Chinese accented",, making their pronunciation "hard to understand".

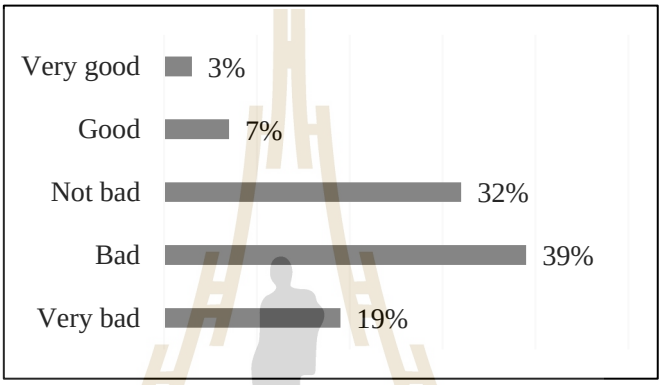


Figure 1.3 Students' self-evaluation of their pronunciation levels

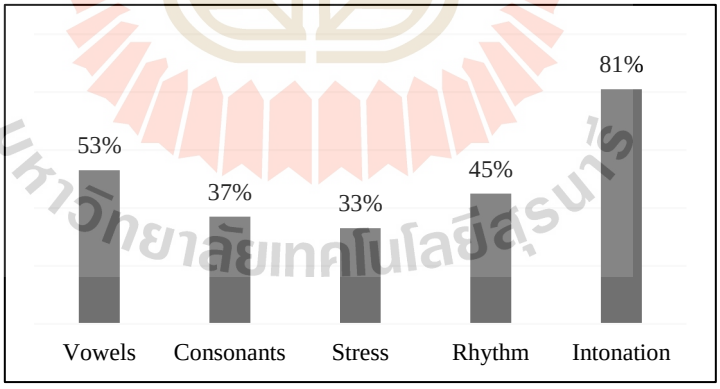


Figure 1.4 Students' self-reported pronunciation problems

(2) Most of the students had high motivation to learn pronunciation and desired to reduce their speech accent. The students' responses to the question "Do you think that you need to improve your English pronunciation?" showed that more than half

(57%) of the students wanted to improve their pronunciation urgently. About 92% of the students reported that they spent more than one hour per week on practicing pronunciation out of class. All of the 4 interviewed teachers also thought that their students had high level of motivation to learn pronunciation. The concept of World Englishes informs English learners that native speaker's model might be an unrealistic pronunciation goal. However, for university EFL learners who are in English major, they want to reduce their accent and approximate to the native speakers' model.

These results were not unexpected considering that these first year English majors were just emancipated from the shackles of all kinds of written tests-driven English teaching which gave no room to pronunciation before they entered into higher education. After they came to this university, especially knowing that they were English majors in the pre-service teacher training program, they began to realize that they must learn something that are practical for hunting jobs in the future, as results of the questionnaire showed that 84% of the students held that speaking is the most important job required skill among the four language skills. While having a good command of pronunciation is the first step to improve speaking. Therefore, it is essential and urgent for them to receive systematic phonetics course or pronunciation training at their first year of undergraduate study.

(3) The students' needs can not be fully met with the current pronunciation teaching in HUSE. Although more than 60% of the students indicated that they were

satisfied with the phonetics course, 47% of the students thought that the current pronunciation teaching was not efficient for solving their individual specific problems. Many students reported that their pronunciation went to a plateau or fossilization period, i.e., they had no idea how to make further improvement although there was still a large space for improvement. This means that the current pronunciation teaching in HUSE is targeting at most students' common problems while leaving the students' individual specific problems unsolved.

From the teachers' aspect:

(1) Teacher-centered, one-size-fits-all approach for pronunciation teaching. Information from the interviews with the 4 teachers indicated that most of them conducted pronunciation teaching by first introducing the knowledge of phonetics and then providing models for students to imitate. The large classroom size, with an average of 34 students in each class, made it impossible for the teacher to monitor each student's performance in class. As a results, the whole students were treated as a homogeneous group, and the teacher can only conduct pronunciation teaching by predicting students' problems according to intuitions or teaching traditions. Therefore, the teacher-centered, one-size-fits-all approach for pronunciation teaching in HUSE made the teacher and students exhausted while the outcomes were unsatisfactory.

(2) Problems about the pronunciation textbook. The textbook used for the HUSE students' phonetics course has inappropriate treatment of the pronunciation

components, which assigns priority to and lays emphasis on the segmentals while gives little attention to the suprasegmentals. For the intonation part, it just introduces the definitions, meaning and functions in a very general way and provides some decontexualized examples or patterned drills which neglect the communicative value of intonation. As a result, the pronunciation teaching in HUSE usually followed this order: vowels, consonants, stress, rhythm, intonation. The teaching of vowels and consonants always took up at least three fourths of the total teaching time. With too much time spent on the segmentals, the students always found themselves in a quandary that they can perfectly produce the vowels and consonants while their pronunciation was still strongly-accented and sometimes unintelligible. Therefore, the current pronunciation textbook was not reasonably designed and can not meet the students' needs.

(3) The "read-after-me" pattern for pronunciation teaching. In SFL of HUSE, there are 4 multimedia classrooms equipped with computers and supported hardware and software which can meet the requirements of computer assisted pronunciation teaching (CAPT). However, these facilities have not been efficiently used and the teachers lacked necessary knowledge of using speech technology to conduct phonetic experiments or transfer the speech theories or technological efforts to pronunciation teaching. Pronunciation teaching in HUSE was usually conducted in a "chalk-talk" way, following the "read-after-me" pattern. The results were: firstly, the students had little

chance to practice and they did more “listening” than “pronouncing”; secondly, each student’s pronunciation would be influenced by other students’ when they were simultaneously imitating the teacher’s pronunciation; thirdly, the teacher had little chance to monitor each student’s performance and the students had little chance to know their pronunciation problems. Therefore, the rare integration of educational technology or CPAT tools made pronunciation teaching and learning in HUSE effortful while the results were unsatisfactory.

To sum up, the first year English majors of HUSE were weak in pronunciation especially in the suprasegmental aspect. They showed high motivation in learning pronunciation, however, the traditional teacher-centered, one-size-fits-all approach for pronunciation teaching could not solve their problems and meet their needs. Facing with these problems, this study aimed to improve HUSE first year English majors’ English intonation by modifying the pitch of the students’ productions and playing the modified version back to the students as the input. The pedagogic procedures was informed by the concept of precision language education which treated students individually and gave each student a tailored instruction according to their specific problems.

Instruction of this kind was more mechanical which focused on individual student’s specific pronunciation problems and provided corrective feedback to their problems through pitch modification. Specifically, this approach took students’

incorrect intonation, digitally altered it to the correct one and fed it back to the students as corrective input, so as to raise their awareness of their own pronunciation problems. Then, students could imitate the models of their own voices, the voices with which they were most familiar. This study assumed that this approach could better arouse students' awareness of their own pronunciation problems, facilitate their intonation learning, and improve their pronunciation ability.

1.3 Significance of the study

This study intended to improve Chinese university EFL students' English intonation through pitch modification. Theoretically, it could test the effectiveness of using modified speech for pronunciation learning where students listen to their own modified voices. It aimed also to enrich the theory of precision language education and explore new ways for its implementation. Methodologically, it explored a new way for using students' own voices for pronunciation teaching. Pedagogically, it could provide implications for the teaching of intonation as well as the teaching of pronunciation to Chinese university EFL students. To be more specific, the significant points of this study are as follows:

Firstly, this study aimed to expand the ways of using speech modification for teaching English intonation and explore a new perspective for dealing with students' pronunciation problems by providing corrective feedback to students' incorrect pronunciations. Speech modification had been employed by many previous studies for

pronunciation teaching, especially for the teaching of individual sounds (Strange & Dittmann, 1984; Jamieson & Morosan, 1989; Wang & Munro, 2004; Hardison, 2012). In the field of intonation instruction, most studies stayed at the stage of exploring the techniques for modifying pitch or testing its viability (Yoon, 2009; Lu et al., 2012; Tang et al., 2001; Felps et al., 2009). Few research had conducted empirical studies to investigate its effectiveness for teaching intonation (Sundström, 1998; Hirose, 2004; Bissiri & Pfitzinger, 2009; Pellegrino & Vigliano, 2015), and those limited studies were not on the English language or were featured by a very small sample size or quite limited teaching materials, and no studies taught English intonation systematically by modifying the pitch. The present study featured more precise techniques, larger sample size, more systematic teaching materials, and a consideration for acceptability from the students' perspective. By so doing, this study attempted to verify that pitch modification of this kind can enable pronunciation learning occur in a more effective and efficient way.

Secondly, previous studies employing speech modification for pronunciation teaching used students' recorded voices, rendering the results unsatisfactory or caused negative effects (Holzman et al., 1966; Gaviria, 1967;Weston & Rousey, 1970; Daryadar & Raghibi, 2015). This study instead used students' self-perceived voices for intonation teaching combining the bone-conducted voice and air-conducted voice based bone conduction theory (Békésy,1949; Tonndorf, 1976; Maurer & Landis,1990;

Hansen&Stinson, 1998; Pörschmann, 2000; Shuster&Durrant, 2003; Reinfeldt et al., 2010) so that students' voices sounded the most familiar to themselves (the detailed process will be explained in Chapter 2). By so doing, it overcame the deficiencies of using students' recorded voices for pronunciation learning and further showed the advantages of using one's own voice for learning.

Thirdly, this study was the first attempt of translating precision language education (PLE) into intonation teaching practice and would provide empirical evidence of applying PLE in the field of pronunciation teaching. PLE is a new conception of teaching emerged in very recent years. Applied to limited numbers of previous studies, it has been claimed to be a promising mindset for future education. However, PLE is still in its infancy and there is ample space for exploration. The pedagogic procedures for the present study was a precision approach based on the main concept of PLE, which would turn out to be more effective and efficient than the traditional standard-protocol approach.

1.4 Objectives of the study

The main objective of the present study was to improve HUSE English majors' intonation through pitch modification. To be more specific, this study targeted the following two objectives:

1) To examine the effectiveness of pitch modification for English intonation learning;

2) To investigate students' opinions of using pitch modification for intonation learning.

1.5 Research questions

In order to fulfil the above objectives of this study, the following research questions were proposed:

1) Is pitch modification effective for English intonation learning? If yes, in what ways?

2) What are the students' opinions of pitch modification for intonation learning?

1.6 Definitions of key terms

In order to avoid ambiguities and clarify the scope of the key terms used in this study, some key terms are defined and explained below:

English intonation Intonation roughly refers to "the way something was said", which includes tonality (chunking), tonicity (the information focus), and tone (the falls and rises of pitch). This study focused on the linguistic aspect of pitch variations which operate together with other prosodic features to realize intonation purposes. The intonation used in utterances is different from the tones used in tonal languages for signifying words. Therefore, the definition of English intonation for this study was "the linguistic use of pitch variations in utterances of English language".

Pitch modification Pitch modification refers to modifying the pitch contour (fundamental frequency, F0) of the speech signal. In the present study, pitch modification refers to modifying the students' incorrect intonation to the correct one and feeding it back to the students as the input model for intonation learning.

The motor theory of speech perception The motor theory of speech perception (MT) was proposed by Liberman and his colleagues (the revised version see Liberman & Mattingly, 1985). It has two main claims: 1) perceiving speech is perceiving the speaker's intended gestures; 2) the motor system is recruited in speech perception.

Bone conduction theory. Bone conduction theory (Békésy,1949) revealed that one's voice heard by oneself and other people are different. The latter was mainly conducted by air while the former was mainly conducted by a combination of air and bones (and tissues) in his/her head.

One's own voice One's own voice refers to both one's voice heard by others (i.e., the recorded voice) and one's voice heard by oneself (self-perceived voice). Based on the bone conduction theory, the present study used students' self-perceived voice for pitch modification which contained a mixture of the recorded voice and bone conducted voice.

Precision education Precision education is the tailoring of education to the specific characteristics of the individual student (Hart, 2016).

Precision language education Precision language education (PLE) specifies the

application of precision education in the field of language education. It heralds a new way of dealing with individual differences by effecting as precise a diagnosis as possible on each language learner, thus triggering specific interventions designed to target and respond to each person's specific language-learning problems (Lian & Sangarun, 2017).

1.7 Summary

This chapter presented the background of this study by addressing the importance and the priority of pronunciation teaching, English education in China, and English pronunciation teaching and learning in HUSE. Focusing on the pronunciation teaching and learning problems in HUSE, this chapter proposed the statement of the problem of this study. The significance and objectives of this study were then outlined. Based on the objectives, the research questions were proposed. Finally, some key terms related to this study were defined.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews the literature pertinent to the present study with six sections. It first introduces the phonetics and phonology of English intonation. Then, literature related to English intonation instruction is reviewed. The third section introduces the motor theory of speech perception. The fourth section discusses using one's own voice for learning and details the method of obtaining one's self-perceived voice based on bone conduction theory. The fifth section profiles precision language education and its implications for pronunciation instruction. The last section presents the theoretical framework of this study.

2.1 The phonetics and phonology of English intonation

This section introduces the knowledge of English intonation from the phonetic and phonological aspect. It starts with clarifying intonation and the scope of intonation for this study. Then, it presents the hierarchical structure and trio systems of intonation, followed by interpreting the meaning of intonation from "three biological codes" and tone choices in communication.

2.1.1 The definition of intonation

Intonation, in the most general way, can be defined as “the melody of speech”. Without intonation, our speech would be monotonous. More strictly speaking, it is impossible to say something without intonation since monotone can also be classified as a type of intonation. We often hear people say “it’s not what you said, but the way you said it!”. “The way you said it” involves the use of intonation in speech and roughly refers to what intonation is.

The widely accepted definition of intonation from previous researchers (Ladd,1996; Kelly, 2000; Wells, 2006; Nolan, 2008; Tench, 1996; Tench, 2011; Aslam & Kak, 2007; Levis & Wichmann, 2015) can be summarized as “the linguistic use of pitch variations in utterances”, where “linguistic use” specifies the purpose of intonation; “pitch variations” reflects the physical and acoustic properties of intonation; “in utterances” specifies the range of intonation use .

Firstly, the main reason for limiting the use of pitch within linguistic purposes is to avoid the reference to other purposes, especially the aesthetic evaluations of speech. For example, we could evaluate one’s speech as “nice”, “pleasant”, or “harsh”; we could also say that a male’s voice sounds feminine. These are subjective impressions that the speaker’s voice leaves on the listeners, which do not affect the speaker’s intended meaning. However, most English language learners hold wrong beliefs about the function of intonation and think that the use of intonation is simply

decorative (Gilbert, 2014:125). Therefore, in teaching English intonation, it is essential to inform students that the linguistic use of pitch variations can change the meaning of an utterance and that intonation is an important way to convey information rather than be simply decorative.

Secondly, pitch variations are realized by modulating the prominence of stress and duration, so we can not isolate pitch from other prosodic features of speech, as noted by Levis (1999): “the way something is said includes not only pitch movement but also length, intensity, and a host of other factors, such as voice quality”(p.38). Actually, the term “intonation” has often been used interchangeably in the literature with prosody (Hirst & Cristo, 1998). In researching intonation, most researchers integrated intonation with other prosodic features. Wells (2006) held that intonation is the study of how the interplay of accented, stressed and unstressed syllables functions as a framework onto which the intonation patterns are attached. Allen (1971) defined intonation as “a quality of language that includes both rhythm and melody and is produced by tonal height and depth along with stress, volume and varying lengths of pause”(p.74). Therefore, a broader sense of intonation includes not only the falls and rises of tone, but also other prosodic features like stress, pause, or intensity, etc.. In the present study, the word “intonation” was used to refer to the collaboration of all these dimensions.

The last point is that intonation is used at the utterance level. This point is

addressed to differentiate intonation from the tone used for signifying words by tonal languages, such as Chinese and Thai. Non-tonal languages also use tones on single word, but the difference is that the basic meaning of that word does not change. Here, we should be careful to avoid a dichotomy between tonal and non-tonal languages. Tonal languages use pitch at word level and at the same time at sentence level. All languages use pitch variations for intonation purposes. Tench (1996) noted that "a tonal language often has a simpler intonation system, using alternative linguistic devices to compensate to make sure the use of pitch variations is not overloaded" (p.6).

To conclude, the definition of intonation used for this study was "the linguistic use of pitch variations in utterances". This study focused on the linguistic aspect of pitch variations which operate hand in hand with other prosodic features to realize the intonation purposes, and the intonation used in utterances is different from the tones used in tonal languages for signifying words.

2.1.2 The internal structure of English intonation

Intonation, the occurrence of various tunes or melodies in utterances, is "the result of the operation of a set of prosodic systems" (Chamonikolasová, 2007, p.11). Intonation is not a nebulous phenomenon, and each language has a specific intonation system. Tench (1996) claimed that intonation is as systematic as other parts of the phonology, which can be taught and learned. "Although there is a personal, subjective, and emotional element to it, intonation is mainly conventional" (p.9), otherwise people

will not recognize the meaning that a certain intonation pattern carries.

In the field of English intonation research, the British school (Halliday,1967; Crystal,1969; O'Connor&Arnold,1973; Cruttenden,1997) proposed and further developed some of the earlier intonation systems. Crystal (1969) focused on the phonetic nature of discussion, while Halliday (1967) sought to present a more linguistic orientation and emphasized the phonological nature of intonation. In order to show how intonation belongs to phonology, Halliday (1967) proposed a hierarchy of phonological structures and units of intonation. He recognized four units which specify a set of phonological constituent types parallel to those of the grammar (Figure 2.1).

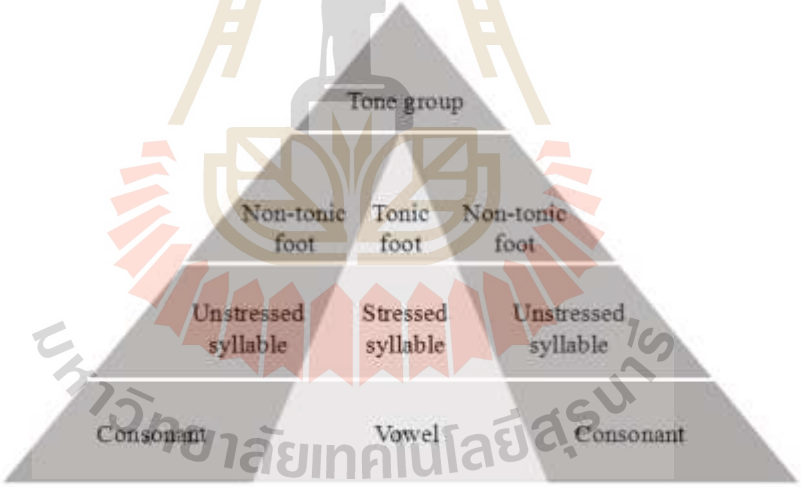


Figure 2.1 The hierarchical structure of English intonation

These four ranks of phonological structure are: at the highest level, the tone group, then the foot, syllable, and phoneme (Some researchers may name them differently, i.e., intonation group for tone group, rythmic unit for foot). These four units constitute a hierarchical structure in that "they are related taxonomically as the units of

the grammatical rank scale: each one consists of one or more of the one below it" (Halliday, 1967, p.12). More specifically, a tone group is made up of one or more feet; a foot contains one or more syllables; a syllable is composed by one or more phonemes.

The hierarchical structure of intonation implies that intonation has a phonological organization. This point is important because it shapes the nature of intonation and further determines its functions. However, many researchers, teachers, and students still feel puzzled about the nature and function of intonation, since intonation itself "sits uneasily with many ordinary linguistic assumptions" (Ladd, 1996, p.3). As intonation is closely linked to a paralinguistic vocal code functioning as a parallel communicative channel, to signal information about our gender, age, attitude, even social status, it is prone to be removed from the linguistic functions it carries. Only if we talk about the phonology of intonation as natural as the phonology of words, can we reach the point of "being able to describe in explicit and testable terms how intonation affects the meaning and function of utterances" (Ladd, 1996, p.4).

2.1.3 The trio-system of English intonation

The core concept of the trio-system of English intonation was proposed by Halliday (1967) and Crystal (1969). The trio-system operating in English intonation consists of tone, tonality, and tonicity (the three "Ts"). Wells (2006) integrated the three "Ts" into a mechanism of making decisions about intonation when people are speaking: how to break the material up into chunks (tonality), what is to be accented (tonicity),

and what tones are to be used (tone). According to Wells, the logical sequence of the three elements that the speaker has to decide firstly is tonality, then the tonicity, and tone goes the last. Tench (1996, p.8) defined the three terms in details as follows:

> *Tonality is the system by which a stretch of spoken text is segmented into a series of discrete units of intonation which correspond to the speaker's perception of pieces (or 'chunks') of information;*
> *Tonicity is the system by which an individual, discrete, unit of intonation is shown to have a prominent word which indicates the focus of information.*
> *Tone is the system of contrasting pitch movements in each unit of intonation, which identifies the status of the information, e.g. major, minor or incomplete.*

In intonation realization, tonality involves the use of pauses or intonation boundaries to break the utterances into pieces of information; tonicity refers to the focus of each piece of information; and tone constitutes the falls and rises of the pitch. These three "Ts" will be discussed in detail next.

Tonality. Listeners perceive the information conveyed by the speakers through capturing the meaning of chunks. This means that we receive information by taking the chunks as information units rather than word by word. These chunks are "the most readily perceivable, recurrent, maximal functional units to which linguistic meanings can be attached" (Crystal, 1969, p.204). As we all know, punctuation in written text can change the meaning. In the same vein, the spoken text with intonation as the punctuation will also convey different meanings when broken into different intonation groups. Tonality in speech plays a role like punctuation in writing (Wells, 2006; Nolan, 2008), so intonation has the function of disambiguating ambiguous syntactic structures.

Intonation groups are divided by intonation boundaries. Crystal (1969) found two boundary markers: a pitch change and a slight pause. At the beginning of a new intonation group, the pitch steps up after a falling tone or steps down after a rising tone, and gradually approaches to the speaker's neutral voice. However, in connected speech, when the speech is fast or drawled, the two markers may become ambiguous with the pause missing or pitch change hard to detect. In reality, speech transcribers do encounter ambiguities caused by boundary markers missing, which is evidence that intonation plays an essential role in disambiguating spoken languages.

Tonicity In oral communication, speakers use intonation to highlight some words to focus the hearers' attention on some particular pieces of information. These words are the focus of information or the burden of message (Halliday, 1967). Speakers highlight the information focus by accenting the stressed syllable in speech. This is where the nucleus locates and the pitch movement begins, and "the 'hooks' on which the intonation pattern is hung"(Wells, 2006, p.7). As discussed above, pitch variations are realized by the interplay of many factors on stressed syllables, like length, volume, voice quality, etc.. Therefore, stress gradient is the foundation of pitch variation, and pitch prominence is gained by shifting the levels of stress in an utterance. Stress gradient is meaningful only within one intonation group, because the level of the most prominent syllable in one intonation group might be the lowest level in another group. So, the degree of stress is infinite and relative (Liberman & Prince, 1977), and we are

not able to establish its value, or presence, unless there are other syllables with different degrees of stress to measure it (Gussenhoven, 2004).

Tonicity concerns the focus of information which is carried by the nucleus of the intonation group. Tench (1996) held that when people speak, some units can be abandoned if they lack a tonic syllable. The nucleus of an intonation group usually falls on the stressed syllable of the last content word. However, this is not always the case. The nucleus can fall on any word in an utterance depending on the speakers' intent. In general, the nucleus is likely to arise when the speaker wants to express a narrow focus-the particular information that the speaker wants to emphasize; a contrastive focus-the information that forms a contrast with the corresponding information that has been said; and the new information (Wells, 2006; Tench, 1996).

Tone Tone refers to pitch movement, the choice of going up or going down of the pitch contour. That there are choices means that there are contrastive pitch movements which convey different meanings. There are different approaches to describe pitch movements. The most influential and widely accepted approaches are the American school's autosegmental metrical (AM) approach (Pierrehumbert,1980; Beckman & Pierrehumbert,1986) and the British school's pitch contour approach. The former is dominant in the field of intonation research while the latter is dominant in the field of intonation teaching.

The British school's pitch contour analysis is based on tone pattern

identification and marks falling tone as "\" , and rising tone as "/". Crystal (1969) distinguishes three types of nucleus tone: simple, complex, and compound (see Table 2.1). O'Connor & Arnold (1973) counted the degrees of pitch movements and identified seven tone patterns: high fall, low fall, rise-fall, high rise, low rise, fall-rise, and mid-level. Tench (1996) viewed the degrees of pitch movements as secondary tones, such as high fall or low fall, compared to the primary tones which only focus on the overall falls or rises in pitch.

Table 2.1 Crystal's (1969) categorization of tone patterns

Simple	**Complex**	**Compound**
fall \	fall-rise \/	fall + rise \ + /
level >	rise-fall /\	rise+ fall / + \
rise /		

Tench (1996) and Wells (2006) recognized three types of primary tones that can lead to contrastive pitch movements: fall, rise, and fall-rise. The rising-falling tone was not included because it generally has the same function as the falling tone, while the falling-rising tone often "signals particular implications" (Wells, 2006, p.10). The primary tones have the function of organizing information while the role of secondary tones is in the expression of attitudes or the "degree of emotional involvement" (Wells, 2006, p.218). Therefore, interpretations of the identified three primary tones are more generalizable across different speakers or different variations of English, while the realization of the secondary tones is much more individual-dependent and complex.

2.1.4 The interpretation of intonation

Intonation is universal and language specific. The universal meaning of intonation is exercised in the phonetic implementation, while the language specific meaning is located in morphology and phonology (Gussenhoven, 2002). Ohala (1984) tried to interpret the universal meaning of intonation from the ethological perspective and proposed the notion of "frequency code". Gussenhoven (2002; 2004) further developed it into three biological codes-frequency code, effort code, and production code-based on the effects of physiological properties during speech production.

The frequency code The frequency code is related to the size of larynx (vocalizer). This can be dated back to animal behavior of which the high pitch is associated with small animals (small larynxes) while the low pitch is associated with big animals (big larynxes) (Wichmann, 2015). Therefore, the frequency code is exploited for the representation of power relations with the high frequency representing "subordinate, submissive, non threatening, desirous of the receiver's goodwill" (Ohala, 1984, p.1), and the low frequency representing "dominant, aggressive, threatening". Gussenhoven (2002) interprets the higher pitch as "friendliness or politeness" and lower pitch as "protectiveness or scathingness". As male's larynxes are about twice the size of female's, lower pitch is associated with masculinity and higher pitch with femininity. In addition, in Hadding-koch and Studdert-kennedy's (2009) experiment, participants were exposed to an artificial utterance with different intonation contours

and were asked to decide whether the utterance was a statement or a question. Results showed that a higher pitch was rated as question while lower pitch as statement. Therefore, there is also a connection between higher pitch and uncertainty, lower pitch and certainty, and hence “questioning vs. asserting”.

The effort code The effort code is associated with the energy expended in speech production. Gussenhoven (2002) explained that more energy will lead to greater articulatory precision and more canonical pitch movements, as well as wider pitch excursions. By exploiting the expenditure of effort, speakers allot the energy expense on each piece of information based on which part of information is more important or expected to highlight. Therefore, the interpretation of the effort code is closely connected to information focus-which part of information is to be emphasized. The affective interpretation of it can be “surprise” and “agitation”.

The production code The production code concerns the breathing process during speech production. The interpretation bases on the fact that speakers tend to spend more effort on the beginning of an utterance than on the end since the subglottal air pressure at the beginning will be higher than towards its end. The decrease of energy causes a drop in intensity and a fall in fundamental frequency (the pitch). Therefore the production code “associates high pitch with the utterance beginning and low pitch with its end”(Gussenhoven, 2002, p.5). Interpretation of the production code is connected with initiation and finality, and thus “high endings signal continuation, low endings

finality and end of turn"(p.5).

The trio-system of English intonation reflects the operation mechanisms of the three biological codes in speech production. The production code is linked to tonality which involves the use of pitch variations to signal the intonation boundaries; the production code relates to tonicity which concerns the allocation of energy to highlight or deaccentuate some pieces of information according to their degrees of importance; the frequency code interprets the specific meaning of tones in conveying information which is critically important for correct understanding of the speaker's intention. The operation of tone in communication is more complex than that of tonality and tonicity considering that there are various tone patterns and interpretations of each tone pattern is complex and highly context-dependent.

O'Connor & Arnold (1973) stated that "there is a limited number of pitch patterns in any one language and we use them to produce definite meaningful effects" (p.1). Only if there is a universally accepted agreement, recognition, or convention on the way we said something, can the intonation pattern used by one person mean the same thing when used by someone else. Therefore, it is possible "to describe frequently recurring patterns of pitch and to give rules for their use" (O'Connor & Arnold, 1973, p.1).

Büring (2016) proposed the notion of default accenting/intonation, or default tone by Wells (2006). By Wells' definition, the default tone means "the

unmarked or neutral tone that is used under no special circumstances"(p.15). Büring defined it as the out-of-the-blue intonation which is analogous for "accenting in particular, and prosody in general" (p.6). For example, the default tone for statements, commands, and wh-questions is a fall, but for yes-no questions it is a rise. However, in reality, the situation is more complex, and we should consider the the meaning and speaker's intention in specific context.

The present study followed the British school's norm of intonation analysis, focusing on the trio-system English intonation: tonality, tonicity, and tone. In dealing with tone patterns, this study only touched upon the primary tones that can result in contrastive pitch movements, i.e., fall, rise and fall-rise, and can be recognized by not only British or American English speakers but also other speakers of English. Agian, the main purpose of this study was not to teach students how to speak exactly the same as native speakers but to inform them the basic rules of intonation which function like grammar, and showing theme that violating the basic rules will lead to the change of meaning and cause misunderstanding or communication breakdowns.

2.2 English intonation instruction

This section reviews the literature related to English intonation instruction. It first states the reasons intonation should be taught, and then introduces corrective feedback in pronunciation teaching. The third part analyzes the current condition and existing problems of intonation instruction, and based on this, five principles for intonation

instruction are proposed in the last part.

2.2.1 Intonation as an essential component of pronunciation instruction

Gilbert's (2014) article, titled "intonation is hard to teach" in the book *Pronunciation Myth* edited by Grant and Brinton, highlighted the importance of intonation instruction while on the other hand elaborated the difficulties in teaching intonation. Most previous researchers agreed upon the point that intonation is notoriously difficult to teach (De Bot, 1983; Clennell,1997; Celce-Murcia et al.,1996; Dalton & Seidlhofer,1994; Wrembel, 2007; Lengeris, 2012). Léon & Martin (1972) even claimed that "of all the elements of a target language, intonation appears to be the most difficult to acquire"(p.141). Roach (1991) explained that "the complexity of the total set of sequential and prosodic components of intonation and of paralinguistic features makes it a very difficult thing to teach"(p.11).

As a result, intonation instruction has long been neglected in EFL classroom. Even now, many teachers and learners still hold that pronunciation and above all intonation cannot be taught (Lengeris, 2012) but only be acquired through long term exposure to the language (Tench, 1996). Others believe that intonation is such a thing that we are so sensitive to and students can acquire it naturally without being taught. Focusing on the importance of intonation instruction, the following part argues that intonation instruction is an indispensable component in EFL pronunciation classroom and intonation can be and should be taught.

Firstly, English is a special language that makes more elaborate use of intonation to signal meaning than do most other languages (Wells, 2004), and some functions of intonation appear to be specific to the English language, such as discourse and pragmatic functions (Clennell, 1997; Wells, 2008). Intonation is language-specific and different languages employ different intonation patterns to convey meaning and the extent to which their users rely on intonation to follow the meaning are also different. For example, the pitch of German is relatively low and level and Spanish has a noticeably narrow pitch range compared to English (Mennen, 2007), and Japanese females tend to use higher pitch to signal their social identity. However, EFL learners tend to assume that English intonation is the same as their native languages and thus transfer the intonation habits of their L1 to L2 (Wells, 2004). The positive transfer can facilitate their learning, however, the negative transfer is likely to cause communication breakdowns. Clennell (1997) claimed that "at the heart of many cross-cultural misunderstandings lie problems associated with intonation features of learner English"(p.117).

Secondly, like grammar, intonation itself also has a grammatical (phonological) structure (Beckman,1996). Braun et al. (2006) asked a group of subjects to mimic 100 randomly generated intonation contours and found that "the produced f0 contours gradually converge towards a limited set of distinct, previously recognized basic English intonation patterns (p.3)", which indicated that intonation in speech is

restricted to a small, limited sets of patterns. This implied that English intonation functions in systematic ways and the essential features of the systems can be taught and learned (Allen, 1971; Tench, 1996).

Thirdly, native English speakers do not make allowance for suprasegmental errors as they do for segmental errors (O'Connor & Arnold,1973; Tench, 1996; Wells, 2004; Wichmann, 2015). For one reason, native speakers' tolerance of learners' segmental errors may derive from their awareness of the phonology of individual sounds and they are quite familiar with learners' problems of pronouncing these sounds. However, most native speakers were not aware of the particularities of English intonation and did not realize that intonation can be erroneous. Pickering (2001) investigated the role of intonation in communication between the native-English-speaking teaching assistants and their international students, and found that the teaching assistants, as native English speakers, frequently exploited intonation to increase the accessibility of the lecture material and establish rapport with their students. However, their students' responses showed that the implicational function of intonation obfuscated the information structure and those teaching assistants were frequently characterized as unsympathetic and uninvolved by their students. This implied that native English speakers' unawareness of the role of intonation in communication will contribute to communication failures between native speakers and non-native speakers.

For another reason, an incorrect intonation still carries meaning, and native

speakers would misinterpret it as the speaker's intended meaning. For example, a Chinese EFL learner intended to show politeness when saying "Open the door, please" to a native speaker, so he highlighted the word "please" with a high fall and lengthened duration. The native speaker might feel that the learner was something like pushing or impatient other than polite. This kind of misinterpretation is prone to lead to cross-cultural misunderstanding. Furthermore, Mennen (2007) warned that speakers' inappropriate use of intonation may leave a negative impression on the listeners or lead to ill-founded stereotypes towards national or linguistic groups.

Fourthly, intonation plays an important role in conveying meaning and intonation itself has meaning. Frazier (1979) found that under difficult listening conditions (e.g., noisy environment) listeners capture information mainly relying on intonation, known as "cocktail party effect" (p.15). This is because that intonation is the most resistant parts of the speech signal to any form of distortion. Ohala & Gilbert (1981) found that listeners are capable of distinguishing different languages using the intonation alone. Therefore, intonation might play an more important role in conveying meaning than do individual sounds (De Bot, 1983), as we can recall that in daily communication we did not necessarily need to pronounce every single sound accurately or everyone would speak like a news reporter. In fact, the English language is weakening its pronunciation of the segmentals, especially the vowels, and gradually approaching the schwa, known as vowel neutralization (Bolinger, 1986). However, the

intonation is never weakened, retaining its importance. As claimed by Gilbert (2014), intonation signals are crucial “road signs” for the listeners to follow the meaning, if not taught, efforts at achieving pronunciation comprehensibility by drilling individual sounds will be proved frustrating.

Fifthly, intonation has many functions in spoken language and some particular functions provide meaning not coded in any other form (Chun, 1988). People will argue that some functions of intonation can be replaced by detailed words description. The fact is that, in daily communication, intonation is more often employed by native English speakers to convey meaning than mere words description (Tench, 1996). Apart from linguistic function, intonation plays an particularly important role in regulating discourse (Dalton & Seidlhofer, 1994) and reflecting the speaker’s age, gender, identity, physical and psychological state, as well as sociolinguistic membership (Mennen, 2007). In order for learners to develop communicative competence, they need to command the ability to interpret the implicational meanings and interactional conventions conveyed by intonation. Being unaware of intonation in communication will lead to the loss of propositional content of the message, the misunderstanding of the illocutionary force of utterances, and the failure to regulate the conversation and manage inter-speaker cooperation (Clennell,1997). Efforts to gain communicative competence without learning intonation but focusing on how to accurately produce individual sounds is just the “icing on the cake” (Chun, 2002).

Lastly, intonation learning can facilitate the development of other language abilities and is an essential component of language learning. Intonation is one of the first aspects of one's first language (L1) acquisition (Reed & Michaud, 2015). Observations from the behavior of infant L1 acquisition process indicated that prosody plays a central role in the initial "boot-strapping" of language, guiding the development of mental representations and producing the "root" processes of parsing (Warren, 1996). Previous studies have shown that prosodic aspects of pronunciation is closely connected to the perceived accentedness and speech intelligibility. Anderson-Hsieh et al.(1992) found that prosody has stronger effect on pronunciation ratings than segmentals; Kang (2010) claimed that measures of pitch range and word stress contributed most to accent ratings whereas comprehensibility scores were mostly associated with speaking rates; Munro & Derwing (1999) pointed out that prosodic errors appear to be a more potent force in the loss of intelligibility than phonetic errors; Kang et al. (2010) further found that suprasegmental factors account for about 50% of the variances in proficiency and comprehensibility ratings; Benrabah's (1997) study showed that incorrect placement in word stress reduces speech intelligibility; Wennerstrom (1997), Jilka (2000), Pickering (2001), and Pickering (2009) claimed that intonation contributes to the perception of foreign accent and speech intelligibility.

Previous studies have also shown that pronunciation instruction focusing on suprasegmental features can have greater overall benefits for learners' speech

intelligibility (Derwing et al.,1998), enhance students' overall pronunciation accuracy (Sonia & Lotfi, 2016; Hardison,2004), contribute to speech coherence (Wennerstrom, 1997), fluency (He, 2015), and interactional success (Pickering, 2009), improve student's phonological working memory (Yang, 2016) and listener's memory (Hahn, 2004; Hardison, 2004), and help to raise students' awareness of the various aspects of speech and increase their confidence during speaking (Hardison, 2004).

To sum up, the particularities of English intonation itself and its role in conveying information determine that intonation should be an indispensable component of pronunciation instruction. Intonation has phonological rules and is as systematic as other aspects of language which makes it teachable and can be commanded by learners. Previous studies showed that the learning of intonation not only benefits the learning of suprasegmental features but also improves the pronunciation of segmentals and contributes to overall pronunciation ability. The importance of intonation instruction calls for a paradigm shift for pronunciation instruction: giving priority to the teaching of suprasegmental features in EFL pronunciation classroom.

2.2.2 Corrective feedback in pronunciation instruction

One of the biggest controversial issues in second language teaching has been that of feedback (Hansen, 2006). The controversies on feedback originated from the complexity of feedback itself. Although researchers have identified several different

types of feedback, its exact notion is still far from clear. According to Neri et al. (2002b), "feedback is used as an umbrella term to refer to different types of information on the learner's performance on a given task " (p.1209). Lyster and Ranta (1997) investigated the feedback and learner uptake in four immersion classrooms and classified the feedback used by four teachers into six different types. They are:

> *Explicit correction*: the teacher provides the correct form.
> *Recasts*: the teacher reformulates the student's utterance, but removes the error.
> *Clarification requests*: the teacher asks the student to repeat or reformulate the utterance.
> *Metalinguistic feedback*: the teacher gives metalinguistic comments, information, or questions to arouse student's awareness of the error.
> *Elicitation*: teacher uses elicitation techniques to directly elicit the correct form from the student.
> *Repetition*: the teacher repeats the student's erroneous utterance.

Lyster and Ranta found that the teachers showed an overwhelming tendency to use recasts as the feedback strategy; however, recasts were found to be ineffective for eliciting students' repair. In comparison, the four other types of feedback except explicit correction, were more successful in initiating students' repair which lead to the students' uptake through negotiation of form. Havranek and Cesnik (2001) also investigated the effectiveness of different types of feedback and claimed that the most successful type of correction is feedback successfully eliciting self-correction in practice situations. Therefore, the key factors relate to the effectiveness of feedback lie in: 1) whether the feedback can arouse students' awareness of their errors; 2) whether the situations where the feedback is provided allow students to take remedial steps (Neri et al., 2002a).

Feedback focusing on learner's error is called corrective feedback. In the field of language education, according to Spada and Lightbown's (1999) definition, corrective feedback refers to "any indication to the learners that their use of the target language is incorrect" (p.171). Corrective feedback in pronunciation teaching is more crucial because the learner's L1 influence can be so overwhelming that the learner is not able to notice the discrepancies between the mispronunciations and the correct forms (Flege, 1995). In a pronunciation classroom, corrective feedback can be given in many different forms, either by teachers or computer assisted pronunciation training (CAPT) systems.

The teacher's instantaneous corrective feedback in pronunciation classroom allows students to notice their mispronunciations immediately. However, there is usually little time for teachers to focus on individual problems in class (Hansen, 2006) and little chance is left for students to practice the correct form. As compensation, teachers always choose to provide delayed-feedback for students out of class. Delayed-feedback can be provided in many forms, such as numeric scores, literal assessments, or the teacher's model. This kind of feedback, individualized though, is proved to be both time consuming and to go against the idea of instantaneous feedback (Hansen, 2006), leaving students to continue making the same mistakes.

The development of CAPT makes more possible ways available for providing individualized instantaneous feedback. Some CAPT applications can give

visual feedback on students' pronunciation, such as intonation curve or spectrum, and students can compare the visual information of their speech signals with the model speaker's. With the inclusion of automatic speech recognition (ASR), some applications can give instantaneous feedback on students' pronunciation by providing numeric scores, color-coding (highlighting students' mispronunciations in different colors), or evaluations on students' speech volume, fluency, rate or pitch values, based on a statistical comparison with that of native speakers.

However, feedback strategies, even with the inclusion of ASR, are still criticized as to be error-prone (Engwall & Bälter, 2007) and students usually do not get any information on how their pronunciation can be correct, i.e., the feedback is not comprehensible to the students. There is no doubt that some applications providing scores to students' pronunciation can inspire the students to keep on practicing to achieve a higher score, but students can not comprehend why he/she got that score, making the practice purely trial and error. Therefore, Engwall & Bälter (2007) claimed that "pedagogically more effective CAPT software should help the learner by indicating what the error is, and how the pronunciation could be improved" (p.236).

Given that unawareness of the differences between correct and incorrect pronunciations is one of the largest obstacles for mispronunciation correction (Zhao et al., 2013), the main aim of corrective feedback is thus to allow students to notice the discrepancies between their mispronunciations and the target desired ones. The key lies

in how to make the feedback more discriminatively perceptible to the students. Only this kind of awareness can lead to the acquisition of a specific linguistic item, according to Schmidt's (1990) "noticing hypothesis". Once this awareness has been raised, the student can take remedial steps (Neri et al., 2002a) through effective trials targeting on reducing deviations from the model.

Speech signal modification has become a frequently employed way of providing corrective feedback in pronunciation instruction. It entails efforts to modify the speech signals of student's mispronunciation and to play it back to the student, so the student can compare the modified version with his/her original incorrect version. Speech modification has greatly benefited from the development of speech signal processing and synthesis technology, making it possible for teachers or researchers to modify speech signals from various aspects, such as temporal features: duration, rate, volume, intensity; and spectral features like fundamental frequency, frequency components, spectral flux, spectral density, etc..

Pitch modification specifically refers to modifying the pitch of the speech signals. Pitch related acoustic factors are frequency, fundamental frequency (F0), duration, and intensity. In the past two decades, many studies have explored the techniques of modifying pitch or using pitch-modified stimuli for pronunciation teaching.

Yoon (2009) introduced the technique of exaggerating prosody by

manipulating either the fundamental frequency (F0), the segmental durations, or the intensity contour of an utterance, which supported the viability of using exaggerated prosody for foreign language pronunciation training. Lu et al. (2012) employed the automatic stress exaggeration method to enlarge the differences between stressed and unstressed syllables in order to help learners to perceive sentence stress. This approach was called hyper-pronunciation training. The duration, pitch, and intensity of the model speech were exaggerated to increase the students' awareness of the acoustic features. Results showed that the resynthesised stimuli with exaggerated features significantly facilitated students to better perceive English stress patterns.

Tang et al. (2001) invented a voice transformation technique which can flexibly manipulate various aspects of the input signal, e.g., pitch, duration, intensity, and formant. It can potentially be used to enhance English learners' pronunciation especially in terms of prosody by modifying the learners' speech as a model for imitation. Felps et al. (2009) invented another voice-transformation technique which can be used to transfer native speaker's accent to student's speech. Experimental results indicated that the technique can reduce foreign accentedness without significantly altering the voice quality properties of the foreign speaker, which provided the viability of using students' modified speech for learning prosody.

Sundström (1998) used the automatic prosody modification method for Bosnian students learning Swedish pronunciation. The students' speech were

resynthesized which allowed the students to hear the correct prosody of a foreign language as spoken with their own voice. However, the training was mostly on the word or phrase level, and the limited amount of training stimuli and small sample size rendered the results ambiguous.

Pellegrino & Vigliano (2015) also used modified speech to help Japanese learners of Italian. The suprasegmental features of native Italian speakers' speech were transferred to Japanese learners' speech. Then the Japanese learners were asked to imitate their own modified speech. Results showed that self-imitation improved the learners' performance significantly in terms of communicative effectiveness while the level of their accentedness remained unchanged. Similar methods of using students' prosodically corrected speech as training stimuli employed by Bissiri & Pfitzinger (2009) to teach Italian speakers learning German lexical stress and by Hirose (2004) to teach non-Japanese learners pronunciation of Japanese accents, also yielded significantly positive results.

Modified stimuli allow the instructors to make the critical acoustic cues more prominent so as to draw students' awareness on those critical cues (Hardison, 2012). In the present study, pitch modification was employed as a means of providing corrective feedback to the students. By comparing their original mispronunciations with the modified correct version, the students can easily perceive the discrepancies between their errors and the correct form. In this way, students can be well-informed what is

incorrect in their pronunciation, why it should be modified, and how could it be correct. Therefore, pitch modification can be an effective means of providing corrective feedback for intonation instruction.

2.2.3 The existing problems of intonation instruction

In 1999, Levis stated that "present intonational research is almost completely divorced from modern language teaching and is rarely reflected in teaching materials" (p.37). Nearly two decades have passed and conditions have not changed much. Language researchers and teachers are not so enthusiastic as they are in dealing with other aspects of language, and "intonation remains a challenge for teachers and students alike, at both the metacognitive and skill levels (Reed& Michaud, 2015, p.454)".

The teaching of intonation has been traditionally neglected in language classrooms (Clennell, 1997; Chela-Flores, 2003; Wrembel, 2007; Paunović & Savić, 2008; Lengeris,2012) and relegated to secondary status in pronunciation teaching (Lantolf, 1976; McNerney & Mendelsohn, 1992; Chun, 2002; Jenkins,2004; Gilbert,2014). For one thing, the characteristics of intonation itself make it difficult to teach and teachers tend to avoid teaching it explicitly in the classroom (Chun, 2002). Furthermore, students hold wrong beliefs about the functions of intonation. Moreover, intonational research findings were rarely transferred to practical use in language classrooms. Chun (1988) commented that the relatively extensive literature on intonation has had minimal impact on the pedagogy in language syllabuses, and

research findings are usually disseminated only to academic audiences and do not reach the classroom (Lengeris, 2012). These unsolved problems leave intonation the "problem child" of pronunciation teaching (Dalton & Seidlhofer, 1994).

From the teachers' perspective, most non-native teachers do not have the adequate knowledge of English intonation (Clennell, 1997; Paunović & Savić, 2008), hence they may lack the confidence or the ability to reproduce the prosodic patterns in a native-like manner (Lengeris, 2012). According to Reed & Michaud (2015), less than 30% of non-native teachers have ever received training in how to teach intonation. Some teachers even hold that intonation can not be taught especially after the students have past the so-called critical period but are best learned by listening to spoken language (Chun, 2002). Some teachers also overlook the importance of intonation and treated it as the "peripheral frill" rather than central to conveying of meaning (McNerney & Mendelsohn, 1992). Others may realize the importance of intonation but they lack the confidence to teach it. Actually, Reves & Medgyes (1994) claimed that non-native teachers do not necessarily feel unconfident. To the contrary, they may have advantages over native teachers in that non-native teachers can have better knowledge of their students' needs and problems, make appropriate use of the students' L1, and teach the learning strategies more effectively. Therefore, it is essential to equip teachers with the knowledge of English intonation and pass the knowledge to the students now, so that the teachers of the next generation can cut off the vicious circle: "my teachers

have never taught me, so I am not able to teach you".

From the students' perspective, given the particularities of English intonation, most students without a linguistic background do not have a clear idea what is meant by intonation (De Bot & Mailfert,1982) and they may not actually believe that intonation affects meaning (Gilbert, 2014), and therefore lack the motivation to master it (Paunović & Savić, 2008). Influenced by the traditional pronunciation teaching paradigm with focus on individual sounds, students spent most of their efforts on improving the pronunciation of individual sounds. However, they may usually find that their overall pronunciation is still strongly accented or largely unintelligible though their pronunciation on individual sounds are accurate, because they did not realize that intonation is an important factor influencing accentedness and intelligibility.

Students' inadequate knowledge and wrong beliefs about intonation also cause the problem that they may produce intonation as well as the teacher or coursebook required, while "they may walk out of the class without having accepted the system at all, or they may think intonation is simply decorative" (Gilbert, 2014, p.125). In other words, they are uncertain about the real-life applications of intonation and express ambivalence about adopting the intonation patterns in their own speech outside the classroom (Reed & Michaud, 2015).

In terms of the materials for teaching intonation, Levis (1999) pointed out that the treatment of intonation in most textbooks or teacher's reference books for

pronunciation are startlingly similar to each other and differing mostly in typesetting and page layout but not in content and philosophy. Even till now, the majority of published materials for pronunciation teaching or learning resemble each other in the order of content arrangement, descriptions of the form and functions of intonation, and the exercises offered. The major problems of current materials for pronunciation teaching can be summarized as the following points:

1) Focusing on the teaching of segmental features while giving little attention to suprasegmental features;

2) Inadequate descriptions of the functions of intonation and the overemphasized attitudinal function of intonation;

3) Decontextualized examples or exercises, where many of the exercises or examples in the textbooks occur in single, isolated sentences. This assumes that learners can fill in the blanks with native speaker intuitions (Levis, & Pickering, 2004) and puts students in the danger of "practicing the forms without learning their meanings or uses and without perceiving the real shape of pitch patterns in discourse"(p.506);

4) Inauthentic and impractical content, in that, much of what the textbooks transpired tended to set up a one-sided production or a stimulus-response structure, not a true social interaction (Chun, 2002). Consequently, students usually performed well in classroom without using it in real life communications;

5) Missing the communicative focus: most of the current pronunciation

textbooks teach intonation in much the same way as they teach the individual sounds. Listening exercises come first, followed by description of the common patterns, and then the patterned drills with those patterns. Intonation is a fundamental component of the communicative process (Chun, 2002), however, its communicative power in current textbooks was shackled by outdated descriptions and teaching practices with lack of communicative purpose. As claimed by Levis (1999), “Intonation can take its rightful place as part of a communicative approach to language teaching only if its communicative value is made obvious and if it is used for communication rather than pattern practice”(p.60).

In conclusion, although researchers have addressed the importance of intonation instruction for decades, the teaching of intonation is still often relegated to secondary status in pronunciation classrooms. Teachers usually have inadequate knowledge of intonation and lack the confidence to teach it. Students hold wrong beliefs about intonation. Problems also lie in the current materials for intonation teaching. Therefore, intonation instruction lags behind the intonational research and is far from the status of what the researchers have expected.

2.2.4 Principles for teaching intonation

Intonation instruction, as well as pronunciation instruction, has gone through ups and downs in the past century. Previous research and empirical studies on intonation and intonation teaching have informed us of the directions where we should

go and of what we should do. Based on the existing problems of intonation teaching as well as the light gleaned from previous research, the researcher of this study proposed the five principles for teaching intonation:

Principle 1: Giving priority to and attaching more importance on intonation in pronunciation classroom

The appeal for a paradigm shift of pronunciation teaching has been made for decades, while we are still facing with the quandary that on the one hand, there are many researchers and teachers realizing the significance of intonation teaching; on the other, the priority and focus of pronunciation teaching in EFL/ESL classroom are still on the segmental features. Intonation should be assigned foremost priority because it conveys meaning that cannot be coded in any other form, influences the intelligibility of segmentals, and facilitates the learning of segmentals (Prator,1971). In addition, McNerney & Mendelsohn (1992) argued that if students start learning pronunciation with intonation, they will feel less frustrated, because greater changes can be effected in a short time. Therefore, students should stop worrying too much about the tricky individual sounds but turn to the more important and less frustrating aspect.

Principle 2: Teaching intonation in context

There are consistent, systematic differences between the way intonation functions in discourse and in an isolated sentence (Levis, & Pickering, 2004). Actually all linguistic signs are indexical and interpretations of their meanings are embedded in

contexts. In the same vein, the communicative value of a particular intonation contour often changes dramatically (Levis, 1999), and differences and disagreements have been reported between decontextualized interpretations and the speaker's intended meaning (Reed & Michaud, 2015). Therefore, the teaching of intonation should be conducted by using coherent and contextualized materials which can help to make the learning process as a holistic experience and the exercises more relevant to students' real life and thus add to motivation (Tench, 2005).

Principle 3: Teaching learnable and practical content

Within the English-speaking world, the intonation used show a great diversity from one dialect to another. Even two speakers of the same dialect may have differences in phonetic implementations of intonation, but the underlying system-the number of basic tones and the principles of stressing and destressing- remains much the same (Thompson, 1981). Therefore, faced with the bewildering range of tone choices, what teachers need to do is to "narrow down the number of available options and locate the relevant tone patterns into specific context" (Kelly, 2000, p.105).

Allen (1971) suggested that teachers need to refocus on the data provided by linguists, in order to discriminate between the essential and the merely interesting, because the data used for intonational research may not be appropriate for practical use; Thompson (1981) chose what appear to be the most generally serviceable tones, maintaining their distinctness as far as is natural; Tench (2005) suggested to use the

more normal intonation forms in initial practice; Gilbert (2014) proposed that instruction should be focused on the main and most teachable functions of intonation and start with "listener-friendly" intonation-intonation that helps the listener "follow" what the speaker is saying; Nolan (2008) argued that the examples used for intonation teaching should be accessible not only to the large number of speakers of American or British English but also to the much larger population of ESL or EFL speakers. In a word, the teaching of intonation in classroom should chose the materials that are learnable and not so frustrating at the beginning, and the examples should be representative and ready for practical use.

Principle 4: Addressing the communicative purpose

Language is a collaborative social process rather than a product of grammatical and lexical rules (Chun, 1988). Missing the communicative value of intonation, students may perform well in classroom while they may discard what they have learned outside the classroom because they do not believe what they have learned, or, when students apply what they have learned in communication, they may fall short of interpreting intended or implied meanings, i.e., "understanding the words but not the message" (Vandergrift and Goh, 2012, p. 22). Levis (1999) argued that the abundance of exercises in textbooks has little power against students' entrenched beliefs about intonation, and "a change in the approach to teaching intonation will come not with better exercises alone but with a change in thinking (p.60)". If the communicative value

of intonation is not made clear, students may decide learning suprasegmental features is not worth the effort (Grant ,2014).

Principle 5: Using appropriate techniques to ensure quality imitation

Imitation is the basic technique in all pronunciation practice (Tench, 2005). By listening to and repeating a model, students can perform a set of input and output practice in a short time. Quality imitation, or quality repetition (Gilbert, 2014), offers students an opportunity to absorb into their personal long term memory banks an accurate sample of the target models, and helps student to internalize the prosodic flow of an utterance before they understand the underlying mechanisms of it. The properties of the input material determine the quality of imitation, i.e., what to imitate and how well the imitation.

To improve the effectiveness of input, teachers can use the technique of "noticing" by asking leading-in questions like "which word does the speaker stress?" or "why does the speaker say it in this way?" to raise students' awareness on the target point. Models with minimal pair examples can help draw students' focus on the intonational contrasts. Seeded conversation (Tench, 2005), a constructed dialogue deliberately devised for a particular phonological feature, can ensure rigorous practice on that feature. Teachers can also use speech technology to modify the input to the wanted form, such as using exaggeration techniques to make one prosodic feature more prominent or using filtering techniques to make the intonation pattern more perceivable and imitable.

The development of speech technology provides more possibilities to facilitate intonation instruction. Relying on the assistants of technology for intonation instruction does not mean that we can bypass these basic principles of intonation instruction. Teaching pronunciation without intonation is like teaching ballroom dancing without music (Gilbert, 2014). Only with the first step of changing teachers and students' thinking on intonation, can teachers lay emphasis on intonation teaching and students believe what have been taught and apply it to practical use. At the same time, learnable practices should be designed in context and ready for practical use, thus to realize the communicative value of intonation.

2.3 The motor theory of speech perception

The motor theory of speech perception (MT) was initially proposed by Liberman and his colleagues in Haskins Laboratories in 1960s (see Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), and since then has become one of the most cited theories in the field of speech. The theory has gone through refinements along its development and the revised version was published in 1985 (see Liberman & Mattingly, 1985). The discovery of mirror neurons in 1990s brought the theory back to life and aroused renewed interest inside and outside the field of speech, promoting its further development.

MT has two major claims. The first claim (the revised version, see Liberman & Mattingly, 1985) is that "the objects of speech perception are the intended phonetic

gestures of the speaker, represented in the brain as invariant motor commands that call for movements of the articulators through certain linguistically significant configurations" (p.2). To put it simply, perceiving speech is perceiving gestures, or in other words, perception tracks articulation (Liberman, 1957). Liberman held that there is an intimate link between speech perception and production. The internal process of phonation is to make the conversion from acoustic signals to gestures automatically. Thus speech perception and speech production share the same set of invariants (Liberman & Mattingly, 1985), which means "what people hear when they listen to speech is what they do when they speak"(p.3). The link between perception and production leads to the second claim of the motor theory: the motor system is involved in speech perception.

2.3.1 Motor involvement in speech perception

The motor system is a part of the central nervous system involved in the generation and control of voluntary and reflex movements. An innovative discovery in the field of motor control is that the motor system is also recruited in the process of perceiving an action, i.e., perception and action share the same coding system. This added new insights in understanding the relationship between speech perception and production.

The discovery of mirror neurons in the early 1990s provided the neural evidence for motor involvement in perception and evoked a renaissance for MT (Lotto,

Hickok & Holt, 2009). Mirror neurons were initially discovered in an experiment on macaque monkey (Rizzolatti et al.,1988). The experiment found that some neurons in the monkey's premotor cortex (area F5) were activated when the monkey performed a grasping activity (such as grasping a banana), and these neurons were also activated when the monkey saw another monkey or the experimenter (Di Pellegrino et al.,1992) performing the same grasping activity. These neurons, called mirror neurons, were fired both during executing an action and perceiving the action.

The mirror neurons were also found when the same experiments were conducted on human (see Fadiga et al., 1995; Iacoboni et al., 1999). These findings revealed that the motor controls of an action were stored in memory, and when the action was being observed, the motor controls were automatically covertly retrieved. This is in accordance with the common coding theory (Prinz,1984) which claimed that there is a shared representation (a common code) for both perception (event codes) and action (action codes). The common coding theory inspired the arising of the action-effect principle: planned actions are represented in terms of the effects they produce in the world (Galantucci et al., 2006).

Fadiga and his colleagues (2002) extended the study of motor involvement in perception to the field of speech and demonstrated for the first time that speech perception activates the cortical areas involved in speech production. By using transcranial magnetic stimulation (TMS) technology, they found that during speech

listening, there is an increase of motor-evoked potentials recorded from the listeners' tongue muscles if the stimuli involve tongue movement. Similar TMS studies conducted by Watkins et al. (2003) have also detected the potentiation of motor cortex of representations of the lip muscles during listening tasks. One EPG (electropalatography) study (Yuen, 2010) and four functional magnetic resonance imaging (fMRI) studies (Hickok et al., 2003; Wilson et al., 2004; Skipper et al., 2005; Pulvermüller et al., 2006) have also reported that cortical areas active during passive listening to speech overlap with those active during speech production.

The neural evidences of motor involvement in speech perception suggest an internal link between speech perception and production, i.e., how a listener perceives speech is informed by how the listener, as a speaker, produces it (Galantucci et al., 2006, p.368), or vise versa, how a speaker produces speech depends on how s/he perceives it. However, the link between perception and production is not direct. There is a third factor mediating them. As stated by Bradlow et al.(1999), whether speech perceptual training can promote speech production depends on whether the training causes "reorganization of the auditory-acoustic phonetic space which is the underlying system used for both speech perception and production"(p.14). In other words, the key lies in sensory-motor mapping.

2.3.2 Sensory-motor integration in speech production

The two claims of MT indicate that speech perception and production are both inherently motoric (Liberman & Mattingly, 1985), and speech perception is carried out by analyzing the signal based on the innate knowledge of the articulatory production of the particular sound (Ananthakrishnan & Engwall, 2011). The discovery of mirror neurons provides neural confirmation of a perception-production link (though not direct) for MT and mirror neurons seem to accomplish the one to one mapping between perception and action (Möttönen, Dutton & Watkins, 2012). The mapping requires the transformation of acoustic information into motor programs to be executed, also called auditory-motor mapping or acoustic-articulatory mapping (Canevari, et al., 2013). In the present study, we used the term “sensory” to include a more general sense (such as auditory and visual stimulation), and hence we call it sensory-motor mapping.

Disconnection between sensory and motor maps can lead to conduction aphasia which is caused by the arcuate fasciculus damage (Canevari et al., 2013). The arcuate fasciculus may serve in language development by facilitating the repetition of phonological elements in speech, and therefore helping in learning language and monitoring speech (Bernal & Ardila, 2009). Speech repetition or imitation is crucial for language development, because it not only promotes production through monitoring (by overt imitation), but also help to form the perceptual abilities (by covert imitation).

Speech imitation occurs in shadowing, repetition, passive listening tasks

(Gambi & Pickering, 2013), and even when stimuli are presented only visually (Miller et al., 2010). Hickok et al.(2003) and Buchsbaum et al. (2001) used fMRI to identify participants' sensory and motor responses when listening to speech and covert rehearsal/humming of the stimuli. Results showed that the cortical areas responded during both the listening phase and the rehearsal/humming phase. Hickok et al. (2003) explained that the sensory-motor interaction circuit enables acoustic input to guide the acquisition of language-specific articulatory gestures, and supports some aspects of speech production. This network plays a critical role in helping language learners tune their articulatory gestures through continuously comparing the ambient speech sounds with their own speech output attempts.

Wilson and Knoblich (2005) pointed out that covert imitation after perception is not only used to promote overt imitation, but also used to predict what is going to happen next. The prediction is the result of sensory-motor mapping and the link between perception and production (See Figure 2.2). Möttönen& Watkins (2009) claimed that the process of phonemic categorization is to map acoustically highly variable speech sounds onto less variable motor representations. The motor representations of articulators are the predictions made based on sensory-motor mapping.

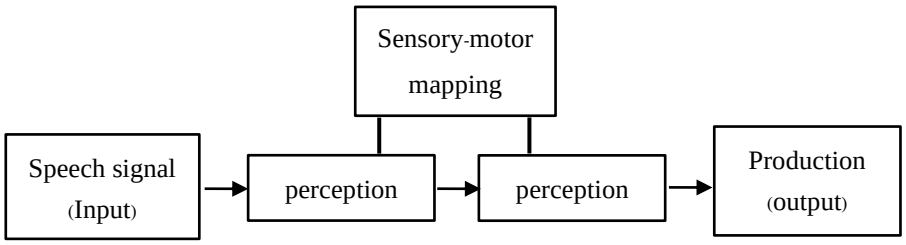


Figure 2.2 Sensory-motor mapping for speech prediction

Researchers used various terms for "prediction". For Möttönen& Watkins (2009), it was "motor representation"; Fairbanks (1954) called it "desired sensory outcome"; Skipper et al. (2005) and Wilson & Knoblich (2005) used the term "motor plan"; Grush (2004) named it "emulator", although this did not originally come from the speech field, but the whole motor control field. Nevertheless, Grush's emulation theory of representation, which claimed that an emulator incorporates a forward model of an external system that runs simulations of that system in real time, can be well applied to interpret the relationship between speech perception and production, whose framework was based on the internal forward model of motor control theory (Wolpert, Ghahramani, & Jordan, 1995).

An internal model, a system that mimics the behavior of a natural process, is the internal representation of the external world used to predict and adjust movements (Zuccaro, 2013). Wolpert et al. (1995) proposed two varieties of the internal model: the forward model and the inverse model. The former "mimics the causal flow of a process by predicting its next state given the current state and the the motor command"; the latter "inverts the causal flow by estimating the motor command that caused a particular

state transition" (p.1880). Therefore, the forward model enables the central nervous system to predict the consequences of motor commands while the inverse model determines the motor commands required to perform specific tasks (Kawato & Wolpert, 1998). Thus, prediction is the output of an forward model that maps the motor command to its desired consequence. The forward model is not fixed and can be trained and updated using prediction errors (Wolpert & Flanagan, 2001).

The internal model of motor control has also been applied to the field of speech. Heinks-Maldonado et al. (2009) used magnetoencephalography (MEG) to investigate the matching process of real time self-auditory feedback during speaking with its internal prediction and provided evidence for a precise forward model in speech production. Gambi & Pickering (2013) argued that during speech perception listeners covertly imitate speakers' speech by computing forward model predictions which are then compared to the incoming input, and phonetic imitation ability can be improved through sensory-motor adjustment, i.e., the correction of prediction errors.

Prediction is the result of sensory-motor mapping following perception, and the process of making predictions is the process of executing the internal forward model. While the forward model aims to predict the consequences of a motor command, the internal inverse model attempts to form the perceptual outcome to generate the motor command. Therefore, the inverse model implies that perception of the modified outcome tends to arouse changes on the motor command, hence the correction of

prediction errors.

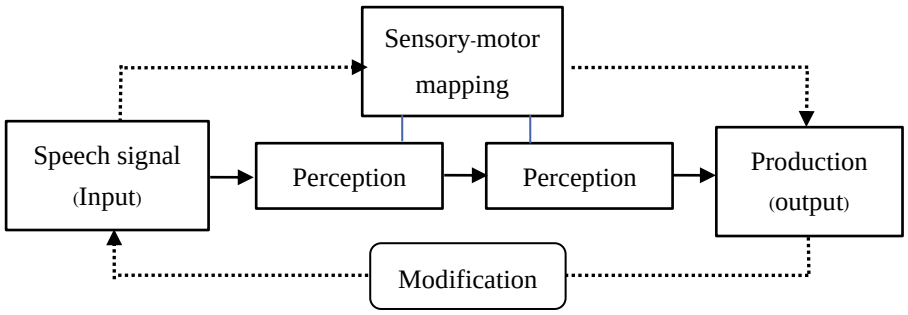


Figure 2.3 Modified output as input for improving speech production

The present study is based on the two claims of MT that perceiving speech is perceiving gestures and the motor system is recruited for speech perception. More importantly, given the empirical evidences of perception and production link and the fact that perception and production share the same internal coding system, this study followed the claim that sensory-motor mapping is the key factor that determines whether the desired production can be achieved. Based on this, this study will modify students' incorrect output to the correct one and play it back to the students as the input, i.e., the modified output as the input (see Figure 2.3 marked by the dotted line). By so doing, this study expects that the modified output can lead to students' sensory-motor remapping, so as to improve their speech production.

2.4 Learning from one's own voice

This section discusses using one's own voice for learning and how to obtain one's self-perceived voice based on bone conduction theory. It first states the benefits of

using one's own voice for learning and reviews the related studies on using students' own voice for pronunciation teaching. Then it argues that using students' recorded voice for learning might cause negative effects, and claims that students' self-perceived voices should be used instead of the recorded voice. Next, using bone conduction theory, the method and procedure is depicted for obtaining one's self-perceived voice.

2.4.1 The activating effect of listening to one's own voice

Pronunciation learning involves the process of perceiving others' speech and then producing it with one's own articulators. According to the motor theory of perception, this process entails sensory-motor mapping. Listeners receive stimulation from the speech signal and the stimuli activate the motor neurons linking those articulators involved in producing this speech. This evoked the researcher's interest in the present study whether or not using students' own voice for pronunciation learning can adjust their sensory-motor remapping and facilitate their learning.

As early as 1966, Holzman, Rousey, and Snyder conducted experiments to investigate the effects of listening to own voice on psychophysiological responses. Results indicated that the participants showed significantly greater psychophysiological activation responses upon listening to their own voices than to other people's voices. This kind of activation was detected no matter the participants consciously recognized their voice or not. The same results were also verified by Olivos' study (1967).

Neuroimaging studies have also provided evidences for the special activation

effect of listening to one's own voice. Nakamura et al. (2001) conducted a PET (positron emission tomography) study to investigate the neural substrates for recognition of familiar voices and found that special brain regions were activated to a greater extent during discrimination of familiar voices than that of unfamiliar voices. The fMRI study of Allen et al. (2005) showed that listening to self-generated words was associated with more activation in the left inferior frontal and right anterior cingulate cortex than words in others' voices. Another fMRI study (Kaplan et al., 2008) found that the right inferior frontal gyrus is involved in processing self-related stimuli, and listening to one's voice showed increased activity in this region which may contribute to an abstract self-representation.

Using electrophysiology technology, Graux et al. (2013) examined the neural processes underlying own voice discrimination and found that own voice discrimination involves neural processes that are different from those involved in discrimination of unknown voices. The former was associated with an early response involving a neuronal network including the left frontal region. In addition, Aruffo and Shore (2012) found that listening to one's own voice can reduce the McGurk effect, as fewer stimuli were perceived as illusory when one's own voice was used, which suggested that auditory self-speech conveys a processing advantage.

Yeager (1966) directly measured whether listening to one's own voice could facilitate memorizing verbal materials. Two groups of participants were required to

memorize the materials delivered to them under only listening conditions. One group received stimuli produced by their own voices while the other group's stimuli were from other people's voices. However, results did not show a significant advantage of learning from own voice over others' voices. Nevertheless, some of the discarded score sheets indicating better learning from one's own voice led Yeager to the hypothesis that "under normal listening conditions people learn as well from either voice, but that under difficult listening conditions, one's own delivery voice might facilitate more accurate perception and subsequent learning of the material"(p.577).

Crowder (1970) investigated the role of one's own voice in immediate memory by comparing the different effects of two kinds of auditory presentation-active vocalization (the listener reads the materials) and passive vocalization (the listener hears other person reads the materials)-on immediate memory. Results showed that active vocalization contributed more to immediate memory than passive vocalization, which supported the assumption that "there might be some special advantage when auditory input is received over a channel (one's own voice) which is highly familiar" (p.158). Similar results were also obtained by MacLeod et al. (2010) who discovered that words read aloud by oneself were more likely to be remembered than words read silently, or by another person. This was called the Production Effect, which claims that "produced items have the additional information that they were spoken aloud encoded in their representations, and this information is useful during retrieval in certifying prior

encoding" (MacLeod, 2011, p.1197).

These previous studies showed that listening to one's own voice and listening to others' voices employ different processing mechanisms, and listening to one's oven voice causes physiological activation automatically that is free of the control of consciousness. These evidences provide the possibility that listening to one's own voice may facilitate spoken language learning.

2.4.2 Using one's own voice for pronunciation learning

Pronunciation learning involves the process of receiving input and transferring it to output. To a large extent, the nature of input determines the nature of output. In other words, what you said is what you hear (Curio et al., 2000). The key factor for pronunciation learning is how the learner perceives incoming speech signals. Therefore, it is important to rethink what kind of input should be provided to learners to facilitate their perception. The activation effect of listening to one's own voice implied that using one's own voice as the speech model may have advantages over other persons' voices for pronunciation learning. Previous studies have explored this and the results were positive.

Probst et al. (2002) examined whose voice a language learner should imitate when learning pronunciation. Results indicated that students who imitated the model speakers' voices close to their own outperformed than those who imitated voices that were dissimilar. Probst et al. claimed that it would be advantageous to match students

with model speakers whose voices are similar to their own. The possible reason lay in that the model speakers' voices close to students' tended to be better perceived by those students, making it easier for the students to link the speech signals with their own articulatory gestures and thus improved their imitation quality.

Jilka & Möhler (1998) adjusted the F0 patterns (pitch contour) of the speech produced by American L2 speakers of German and asked native German speakers to evaluate the different versions of American L2 speakers' speech with modified F0 patterns. Results indicated that F0 generation in combination with resynthesis is very effective in identifying intonational foreign accents and the resynthesized speech with native-like contour was perceived as less foreign accented. As claimed by Jilka & Möhler, this resynthesis method enabled foreign language learners to hear their own, incorrect version of an utterance as compared to the correct, native−like version spoken by their own voice. Therefore, pitch modification is beneficial for intonation learning since "the direct comparison of both versions is helping the learner realize faster what is wrong about their intonation and how it could be improved" (p.118).

In order to help Italian speakers to learn German lexical stress, Bissiri & Pfitzinger (2009) used the prosodically corrected utterances in the learners' own voices as the training stimuli by transferring the pitch contour, local speech rate, and intensity from reference utterances of a German native speaker to the learners' speech signals. A perception test with German native speakers revealed that the learners' modified speech

significantly better reflected lexical stress than the corresponding original utterances. The training results indicated that the resynthesised stimuli were more effective than the natural native speakers' stimuli in learning stress.

Tang et al.'s (2001) voice transformation technique enabled the use of students' modified speech as the model for imitation. They also tested this idea by modifying the lexical tones of few Mandarin Chinese phrases spoken by a native English speaker. Results showed that the modified speech was significantly improved as judged by native Chines speakers, which provided the viability of using students' own voice for learning prosody. A similar technique was also used by Felps et al. (2009) which can provide learners with prosodically corrected versions of their own utterances. Felps et al. claimed that the student's own voice with a native speaker's accent is the ideal voice for a student to imitate. By so doing, the learner can "better perceive the differences and imitate the target speech, without the distractions from less relevant factors such as voice characteristics" (Tang et al.,2001, p.3). This means that using the student's own voice for pronunciation learning is more likely to draw the student's attention on the discrepancies between incorrect pronunciation and the target model.

Hirose (2004) developed a pronunciation training system for teaching Japanese accents to non-Japanese learners. This system could automatically recognize a learner's accent type and correct the prosodic features by referring to the teacher's features. The modified speech would be played back to the learner, so the learner could

be better informed where his/her problems were located as compared to the case where only the teacher's model was provided. Training results showed that students' learning process was facilitated, especially on sentence level.

Results from previous studies showed that using the students' own voices for pronunciation learning is beneficial and promising. However, negative effects of listening to one's own voice have also been encountered. Holzman et al. (1966) found that there is a trend of constriction of free associations after participants heard their recorded voices. The researchers explained that affective impact was responsible for this phenomenon by mobilizing defensive reactions following listening to one's recorded voice. The affective reaction was aroused by the discrepancy between listening to one's recordings and the voices they had expected to hear, as exemplified by their comments on listening to their voice recordings: "It's too high pitched", "It's too nasal"(p.432). Even many people failed to recognize the recorded voices as their own. In Holzman et al.'s experiment (1966), 13 out of the 24 participants did not recognize their own voices and so did with 32.5% of the participants in Gaviria's study (1967) and 43.6% in Daryadar & Raghibi's (2015) study.

Daryadar & Raghibi (2015) claimed that listening to a recording of one's own voice can cause attentional bias. Participants' reaction times in the listening task were greater when they listened to their recorded voices than that of the unfamiliar recorded voices of their own gender. Daryadar & Raghibi ascribed this phenomenon to

the negative affective reaction when they listened to the recordings of their voices as compared to the voices they had expected to hear.

This brought using learners' own (recorded) voices for pronunciation training into an "own voice fallacy"- the so called learners' own voices that could not be recognized by the learners themselves were nothing more than other persons' voices. In other words, the prerequisite of using learners' own voices as the model for pronunciation learning is that the voices should be the learners' self-perceived voices rather than recorded voices, otherwise the advantages of using one's own voice for learning might be very limited. The next part of this section will discuss how to obtain one's self-perceived voice based on bone conduction theory.

2.4.3 Obtaining one's self-perceived voice

According to bone conduction theory (Békésy, 1949), the transmission of one's self-perceived voice is a combination of air conduction and bone conduction, while one's recorded voice is mainly conducted by air. Therefore, to obtain one's self-perceived voice, it is needed to figure out the mixing ratio between the air conducted (AC) component and the bone conducted (BC) counterpart. The AC component is easy to measure through recording. However, the BC component is difficult to measure because its conduction "involves multiple pathways and there is no obvious way to distinguish between them" (Stenfelt, 2011, p.4). Moreover, "it is impossible to apply a direct method to measure the influence of all the bone conduction components on the

perception of one's own voice (Pörschmann, 2000, p.1040).

Notwithstanding, previous studies (Hansen&Stinson,1998; Shuster&Durrant, 2003; Pörschmann, 2000) found that BC sound can be obtained by low-pass filtering the air conducted sound. Therefore, one's self-perceived voice can be obtained through two steps: 1) low-pass filtering the air conducted voice (recorded voice) to get the rough bone conducted voice; 2) Mixing the low-pass filtered voice with the original recorded voice according to a certain ratio. The key is to figure out the value for the low-pass filtering and the ratio for the sound mixture.

As the structure of bones and tissues in each person's head is different, the way that one perceives his/her voice is highly individual dependent, so that the values for the low-pass filtering and the mixing ratio should be individualized and depend on individual's perception. The present study followed a delayed auditory feedback paradigm (Maurer & Landis, 1990; Shuster & Durrant, 2003) to investigate the participants' preferred values for the low-pass filtering and the mixing ratio. This method allowed the participants to make immediate comparisons between their real-time speech and the delayed feedback self-speech samples. The detailed procedures are as follows.

Firstly, in order to investigate the participants' preferred values for the low-pass filtering, the participants were asked to do the following steps by using Praat:

1) Speak a sentence into the microphone and record it. The sentence can be

anyone that the participants would like to speak.

2) Low-pass filter the recorded speech with any values (from 0-5000 Hz).

3) Speak that sentence again with covered ears (equals to bone conducted voice) and at the same time listen to the real-time speech and then, compare it with the low-pass filtered speech. If the low-pass filtered speech does not sound like the real-time speech, repeat step 2 and 3 till the selected low-pass filtering value can make the filtered speech sounded the most like the real-time speech. The participants can report their final selected values to the researcher.

Secondly, the participants were asked to select their preferred ratio for the mixing of the bone conducted voice (low-pass filtered speech) and the air conducted voice (recorded speech). They were required to:

1) Speak a sentence to the microphone and record it.

2) Low-pass filter the recorded speech with their preferred values.

3) Mix the recorded speech and the low-pass filtered speech with a certain ratio (from 0-100%), such as 30% for the recorded speech and 70% for the low-pass filtered counterpart. A sound mixture is thus generated.

Speak that sentence again with opened ears and at the same time listen to the real-time speech and then, compare it with the sound mixture. In the same way, step 3 and 4 can be repeated until the sound mixture sounded the most like the participant's self-perceived voice. Finally, the participants can report their preferred mixing ratios to the researcher.

This method of obtaining students' self-perceived voices was viable and reliable because the participants' self-perception was taken as the final criterion, i.e., the values for the low-pass filtering or mixing were subject to adjustment until the participants thought that the sound mixtures sounded natural or "right" to them. Using students' self-perceived voices rather than their recorded voices for pronunciation teaching has never been conducted before. One innovative point of the present study was that students' self-perceived voices were used as the training stimuli. By so doing, the negative effects of listening to the recorded voice can be eliminated and the advantages of listening to one's own voice for learning can be brought into full play.

2.5 Precision language education

This section introduces the concept of precision education and reviews some related studies which put precision education into practice. Considering the characteristics of language learning, this section argues that precision education is ideal to be applied to the field of language education, hence precision language education, especially for pronunciation instruction.

2.5.1 The concept of Precision education

Precision education (PE) is a new concept in the field of education which has just come into notice in recent years. It was inspired by the concept of precision medicine which is an innovative approach to personalizing healthcare delivery. The Precision Medicine Initiative, launched by president Obama (White House, 2015) was

based on the rationale that most medical treatments have been designed for the “average patient” and this “one-size-fits-all approach” to treatment can only cure the diseases of some patients, while with other people, it may be ineffective or even have adverse effects. Precision medicine, on the other hand, takes individual differences (such as patient’s genes, environments, and lifestyles) into consideration and enables medical professionals to tailor treatment to each patient’s unique needs.

The precision medicine approach to treat diseases can be ideally applied to the field of education in dealing with learning disabilities, because learning disabilities have remarkable similarities to biomedical diseases. Both of them are influenced by psychological, genetic, and environmental risk factors, and treatment is complicated by individual differences in etiology and response to treatment (Hart, 2016). The current educational system has been designed for the benefits of most students or the average student, with uniform instruction, broad assessment and fixed teaching methods. However, each student is unique and each student’s learning disabilities are unique, so students should be treated as a heterogeneous group with considerable individual differences (Cook et al., 2018). Mirroring the precision medicine approach of matching the right drugs to the right people, instructions could also be designed by tailoring to students’ characteristics.

In defining PE, Cook et al. (2018) pointed out that education should go beyond the question of “did an intervention work?” to “what intervention worked for

whom and how did it work?". The former question involves selecting the appropriate intervention for students in an overall sense, while the latter question reflects a more precise understanding of intervention which involves strategic efforts "to get the right intervention in place for the right person for the right reason" (p.5). Cook et al.'s (2018) definition of PE is as follows, together with the definitions made by Hart (2016) and Lian & Sangarun (2017).

> Precision Education is defined as an approach to research and practice that is concerned with tailoring preventive and intervention practices to individuals based on the best available evidence (Cook et al., 2018, p.4).
>
> Precision Education is the tailoring of education to the specific characteristics of the individual student, mirroring precision medicine (Hart, 2016, p.1).
>
> Precision education is the desire to access information that is as detailed and accurate as possible about learner characteristics and performance in order to initiate the most effective intervention in support of the students' learning efforts (Lian &Sangarun, 2017, p.3).

From the above definitions, it can be summarized that PE has the following characteristics.

(1) Encouraging individualized learning. As knowledge construction is based on individuals' past experience and influenced by psychological, genetic, and environmental factors, and those factors act not only independently but also in an interactive fashion, no single instruction can solve all students' learning problems and meet all their needs. PE objects to the "one-size-fits-all" approach of current educational system that treats students as a homogeneous group of which individual

differences are neglected. However, although individualization or personalization was frequently mentioned when PE is discussed, it should be noticed that PE is different from individualized education. The latter focuses on individual differences while neglecting how to actually deal with these differences. The former focuses on how accurate we can access individuals' information and how precisely we can solve individuals' specific problems based on this information.

(2) Targeting students' specific problems. Individual differences determine that each student's learning disorders are doomed to be different, so students' learning problems are also different from each other. From a horizontal view, each student manifests different problems; from a longitudinal view, each student's problems are different in different stages. PE encourages a problem-solving oriented instruction approach which contains an essential process of diagnosis of students' learning problems. Data on students' current performance, interests, multiple intelligence, learning methods, and aspirations are collected to match them with tailored instructions. Increased accuracy and details of the data increase the validity of the instruction.

(3) Tailored instruction. The process of accessing students' information gives teachers an extraordinary depth of knowledge about each student and allows them to create an individual customized learning experience for all students. The designing of tailored instruction is not a once-for-all effort, which means that the instruction should be tailored not only at the initial stage but also throughout the learning process with

continuous modification of the instruction to eliminate inappropriate factors while increase those are effective. Therefore, the tailored instruction for each student is elicited based on evidence of students' current existing problems and an analysis to their problems. An effective instruction signifies an instruction that can solve "my problem", not others' problems, and enable learning to occur in a "just in time, just enough and just for me" (Lian, 2014) fashion.

(4) Creating a student-centered, teacher-assisted learning environment. Since the appearance of communicative teaching methods, there has been a gradual shift from an emphasis on teaching and a teacher-centered classroom to an emphasis on learning and a learner-centered classroom. Teacher acts as an assisting role who does not "teach" but facilitates learning in a very special learner-centered way (Morley, 1991). The assisting role of teacher does not mean the weakening of teacher's work, but teachers do most work at "backstage". Before conducting the teaching, teacher collects information about the students and designed tailored instructions for each student. In classrooms, teacher's main work is to offer help to facilitate students' learning and monitor students' learning process, thus making the instruction more effective. Students can enjoy their learning as "free-operant conditioning" (Lindsley, 1990) which refers to a process of learning in which students are free to respond at their own pace without having restraints placed on them by the limits of the materials or the instructional procedures of the teachers.

(5) Flexible and adaptive interventions. PE views students' learning as a dynamic changing process. No single instructional strategy works all the time. The tailored instruction matched the initial needs of the students will often lose its effectiveness as the needs and abilities of the students change (White, 1986). Instruction approach based on PE is flexible and allows for adaptation. The tailored instruction matches to students' needs, preferences, and motivation to change, permitting the pruning of treatment components, so as to ensure the instruction tailors to each individual throughout the learning process.

In conclusion, proponents of PE claim that students are a heterogeneous group with individual differences. Effective instruction entails efforts to identify each students' learning disorders underlying their learning problems and employ tailored interventions to solve individuals student's specific problems. The learning process is flexible and adaptive to students' change. Students can control their learning process according to their needs. Teachers can monitor students' learning and conduct formative assessment in a more effective way. In a word, PE has the potential of promoting student-centered, individualized learning, allowing more precise problem-solving, and avoiding "educational waste, potential iatrogenic effects and counterproductive effects" (Cook et al., 2018, p.5).

2.5.2 Precision education in action

As a concept in the field of education, PE is still in its infancy, and there is much room for exploration and more research needs to be conducted to establish the evidentiary supports. As the theory develops, some educators have put it into practice. In the United States, National University has launched the Precision Education Initiative aiming at "creating a new paradigm for student success by exploring ways to leverage technology, data, and communications to create a truly customized learning experience for all students" (National University Precision Institute, 2017). The Precision Institute of National University has been created in support of this initiative and to conduct evidence-based research. Now the Institute is identifying research-based projects to explore multiple pathways for implementing PE, such as adaptive machine learning instruction, competency-based learning, online career and academic goal setting, individualized course navigation, etc..

In China, the Center for Language Acquisition & Precision Education (CLAPE, 2017) was established in the city of Xi'an through the cooperation of some leading universities (e.g.,Yale University, Harvard University, University of Toronto, etc.). The research focus of CLAPE is the implementation of PE in the field of language acquisition. For example, one of the ongoing research projects is *Research on early diagnosis and precision intervention of reading difficulties*. This research takes behavioral and genetic approaches to investigate the causal factors of children's

learning difficulties, thus to design effective intervention programs for those with learning problems from co-occurring conditions. To specify the application of PE in the field of language education, Lian & Sangarun (2017) proposed the concept of precision language education (PLE).

> Precision language education heralds a new way of dealing with individual differences by effecting as precise a diagnosis as possible on each language learner, thus triggering specific interventions designed to target and respond to each person's specific language-learning problems (Lian&Sangarun,2017, p1).

Lian's (1984) computer-based answer-evaluation and markup system is a good example of instruction activity based on precision language education, which provides precise feedback for listener's answers in a listening-transcription task. It uses the student's own input to identify whether students' answers are correct or not, and then provides specific feedback to help students repair the identified problems. By so doing, students are able to modify their perceptual and comprehension systems according to their specific problems, and get closer to the correct answers.

A series of empirical studies, focusing on a pre-intervention stage of investigation of students' needs and diagnosis of their problems so as to accurately match interventions to their needs, were published in a special issue of *Journal of School Psychology*: *Advancing the science and practice of precision education*. The following are some selected studies:

Miller, Cook, and Zhang (2018) developed and tested the treatment validity

of an intervention matching protocol-the Student Intervention Matching (SIM) Form.To examine the relative effectiveness of matched/mismatched interventions according to the SIM Form, eight elementary-age students were arranged in sets of four student dyads. Results showed that matched interventions using the SIM Form were more effective than the mismatched ones across a variety of dependent variables, which revealed that effective pre-intervention assessments are essential for matching students to specific evidence-based interventions.

O'Connor and Daly (2018) examined the effectiveness of consequence-based and antecedent-based treatments and their combinations on improving the math computation fluency of four elementary-aged students with escape-motivated problems and found that all students responded differentially to the treatments, implying the heterogeneity among students with escape-motivated behavior and there is a need to investigate and adapt interventions on a case-by-case basis.

Connor et al. (2018) examined the efficacy of Individualizing Student Instruction in mathematics by using algorithm-based intervention protocol to provide tailored instruction for second graders, as compared to the standard protocol approach. Results confirmed that assessment-informed personalized instruction, tailored to students' individual skills and abilities, is more effective than "one-size-fits-all" approaches. This study implied that precision education approaches can be implemented by regular classroom teachers in everyday school settings despite the

school-level poverty, students' gender and their initial scores.

Lemons et al. (2018) evaluated the efficacy of an early reading intervention for 6 children with Down syndrome. Tailored intervention was developed according to the participants' Down syndrome behavioral phenotype. Results indicated that the intervention was effective for most children (4 out of 6), with one child showing mixed results and no functional relation for the other. This study also reminded precision education practitioners that intervention can be tailored to subgroups of learners with similar characteristics rather than to each particular individual. In other words, precision education does not necessarily need to be individualized provided that homogeneity can be identified out of a heterogeneous group of students.

Hall and Burns (2018) meta-analyzed 26 studies concerning small-group reading interventions and found that interventions targeted to a specific skill were more effective than the comprehensive intervention programs that addressed multiple skills. These findings, as commented by Cook et al. (2018), "provide robust and generalizable support for the precision education approach, suggesting the importance of collecting problem analytic data to inform intervention tailoring efforts" (p.8).

These innovative studies have made some precise and tailored instructions possible by implementing the conception of PE. The problem-solving oriented process was conducted by strategic use of students' information data, analytics, algorithms, and CALL tools to inform the decision, as well as designing and modification of educational

services. However, although PE advocates student-centered, teacher-assisted learning environment in classrooms, it is actually a new challenge for teachers to put PE in practice in everyday classroom setting, since it requires more “backstage” work for teachers and they need to continuously monitor students’ learning and modify the teaching according to students’ needs. Therefore, more research is expected to explore the implementation of PE, to identify more possible pathways, and to investigate teachers’ and students’ as well as education decision makers’ opinions for the application of PE.

2.5.3 The necessity of applying PE in pronunciation instruction

Learning is the objective of teaching. However, teachers sometimes focus too much on teaching without investigating whether or not their teaching has had any effect on the students’ learning. West et al. (1990) claimed that an instruction will be more effective if teachers clearly identify students’ problems, specify what they want to teach, regularly monitor students’ performance, and adjust instruction according to their needs. This claim accords with the concept of PE, which advocates the precision approach for instruction in relative to the traditional standard-protocol approach.

A standard-protocol approach to instruction, which is deductive in nature, uses empirically validated treatment for all students with similar problems in a given domain (Fuchs et al., 2003). Teachers assume that students will have certain learning problems and the selected intervention can solve these problems, so the process of

instruction is a top-down approach.

Fuchs et al. (2003) identified two advantages of the standard-protocol approach. On one hand, it is easier to train practitioners to conduct one intervention coherently and correctly and to assess the accuracy of implementation. One the other hand, since it targets the common problems that a group of students share, large numbers of students can participate in a generally effective treatment protocol. Students can also work in a cooperative way and compare or monitor the learning process with each other. In this sense, the standard-protocol approach can be very effective at solving the homogeneous problems of a group, and it embraces more validity as it follows the mature models established and tested by previous research. However, in reality, the standard protocol approach is often criticized for failure in achieving desired outcomes, resulting in programs, practices, and interventions with modest or negligible effects (Cook et al., 2018). This may be because teachers often select interventions in a trial-and-error fashion based on their predictions or conventions, which is the same as "shooting in the dark and hitting targets indiscriminately" (p.5).

In comparison with the standard-protocol approach, the instruction approach that PE advocates is inductive in nature since "no student characteristic dictates *a priori* what intervention will work" (Fuchs et al., 2003, p.160), nor will a given intervention be effective for all students and all the time if the heterogeneity of the group of students is neglected, given the fact that students' learning is a dynamic changing process.

Proponents of PE hold that solutions to students' problems should be induced by the strategic use of students' information data, and interventions should be generated from a pre-intervention stage of problem identification and problem analysis, so the process is a bottom-up approach.

However, the comparison between the standard-protocol approach and the precision approach should not bring a dichotomy between them. Both of them have their advantages: standard-protocol approach seems more likely to capture the heterogeneity among students and to facilitate greater quality control (Fuchs et al., 2003); the precision approach focuses more on individual differences and how precisely teachers can locate students' problems and tailor the instruction to students' needs.

Several aforementioned studies have explored the ways to translate the concept of PE into precision instruction approaches. The researcher of the present study holds that precision approach is ideally for language education applications, especially for pronunciation instruction. As was stated earlier, pronunciation refers to the purposeful use of articulator(s) to generate the target sound. Everyone's articulators are different, such as vocal tract size, tongue, lips, mouth, etc., which leads to the fact that everyone's voice is different. As a result, each individual student will encounter different pronunciation problems when learning a second language, since even no two students' pronunciation problems are the same. In this sense, pronunciation is especially related to individual differences as compared to other areas of language.

Current pronunciation instruction approaches focus more on how to set or select good models for students to imitate rather than how to deal with students' specific personal problems. In other words, they are trying to address how teachers or the models can be correct rather than how the students can be correct. In contrast, the precision approach targets students' problems, providing tailored instructions for solving these problems. Therefore, the precision approach is suitable for pronunciation instruction.

The present study was a trial of implementing the concept of PE in intonation instruction. However, as White (1986) claimed, adopting precision teaching does not mean abandoning a "hard learned style" of teaching. Participants in the present study received intonation training by performing simple listen-repeat exercises. However, the focus of this study lies in how precise students' problems can be identified, what kind of input they receive, how the materials are arranged, and how effective the learning plan can be. Employing the precision approach does not reject any learning styles but focuses on how to become more effective and efficient in adapting those styles to meet the individual needs of each student.

2.6 Theoretical framework

The present study was based on four theories. The British school's approach to intonation analysis guided the development of the intonation training materials for this

study; MT provided the theoretical basis for employing pitch modification for intonation learning; bone conduction theory was referred to manipulate students' voices; and precision language education informed the pedagogic procedures for conducting the intonation training of this study.

Firstly, the designing of the training materials of this study was based on the British school's approach to intonation analysis which is also the dominant approach in the field of intonation instruction. Early contributors to this approach were Halliday (1967), Crystal (1969), Allen (1971), O'Connor&Arnold (1973), etc., and it was further developed by Cruttenden (1997), Tench (1996), and Wells (2009). The British school proposed the trio-system of English intonation: tonality, tonicity, and tone. This study did not touch upon the secondary tones and the tone patterns that could be interpreted individual-dependently, but focused on the universally accepted usage of the three primary tone patterns in conveying meaning. The rigorously selected intonation materials for this study were adapted from previous intonation related studies and were prudently designed in contexts that the target intonation should be interpreted correctly unless it will cause misunderstanding or communication breakdowns.

Secondly, the motor theory of speech perception (MT) provided the theoretical basis for using pitch modification for intonation learning. MT claimed that perceiving speech is perceiving gestures and the motor system is recruited in speech perception, which implies that there is an innate link between perception and production. Based on

MT, the process of pronunciation learning follows an internal forward model: receiving input-forming prediction-producing output. The internal inverse model inverts this process and claims that exposing learners to the desired output could adjust their sensory-motor mapping and facilitate their prediction error correction, thus to improve their output. Following this reversed causal flow, the present study employed pitch modification method by modifying students' undesired output to the desired one and playing it back to students as the input, expecting that it can better facilitate students' correct perception of the input so as to correct their prediction errors and improve their production.

Thirdly, manipulation of the students' own voices was based on bone conduction theory. Previous studies have shown that listening to one's own voice can arouse activating effects and has the potential to facilitate language learning, but using a recorded voice may result to negative affective reaction. Therefore, it is necessary to use students' self-perceived voices for intonation teaching in this study in order to eliminate the negative effects of the recorded voices. The method for obtaining students' self-perceived voices is based on the main claims of bone conduction theory and considers students' perception and evaluation of their manipulated voices as first priority, so as to ensure the quality of modified stimuli and make full use of students' own voices for intonation learning.

Finally, the pedagogic procedures for conducting the intonation training in this

study was guided by the concept of precision language education. PLE advocates educators to treat students as a homogeneous group, pay attention to each student's specific problems and design tailored instruction targeting each individual's problems. This study proposed that pitch modification for intonation learning can be best fitted in this approach since it will precisely diagnose each student's specific pronunciation problems and focus on solving these problems through modifying the mispronunciations and playing back to them. Therefore, the precision approach for intonation instruction in this study was flexible, effective, tailoring to students' needs, and targeting at students' individual specific problems.

2.7 Summary

This chapter provided a review of the literature related to this study. Firstly, it presented the knowledge of English intonation with a special focus on the British school's approach to English intonation analysis. Then, the literature related to English intonation instruction was reviewed , including the importance of teaching intonation, pitch modification in pronunciation instruction, the existing problems and principles of intonation instruction. Next, MT was introduced to provide the theoretical basis for using pitch modification in this study, followed by a literature review on learning from one's own voice and the discussion of bone conduction theory for obtaining one's self-perceived voice. Then, precision language education and its related literature were depicted and reviewed. Finally, the theoretical framework that informed the present study was presented in the last section.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter presents the methods and materials employed in the present study. It first discusses the research design, variables, and participants of this study. The research instruments used to collect data are then explicated. The third section details the procedures for data collection. Next, the procedures and methods for data analysis are described. Finally, the pilot study and its preliminary results are discussed.

3.1 Research design

The research objectives and research questions determined the research design of this study. In order to investigate the effectiveness of pitch modification for English intonation learning and the participants' opinions of this kind of learning, this study employed a mixed methods research design. A mixed methods research design combines or integrates both quantitative and qualitative research and data. Researchers choose mixed methods design to avoid the bias and make up the weakness of using one single method (Creswell, 2012). In the present study, the quantitative part aimed at examining the effectiveness of pitch modification for intonation learning; the qualitative part aimed at triangulating the quantitative findings so as to answer the first

research question and explore the participants' opinions to answer the second research question.

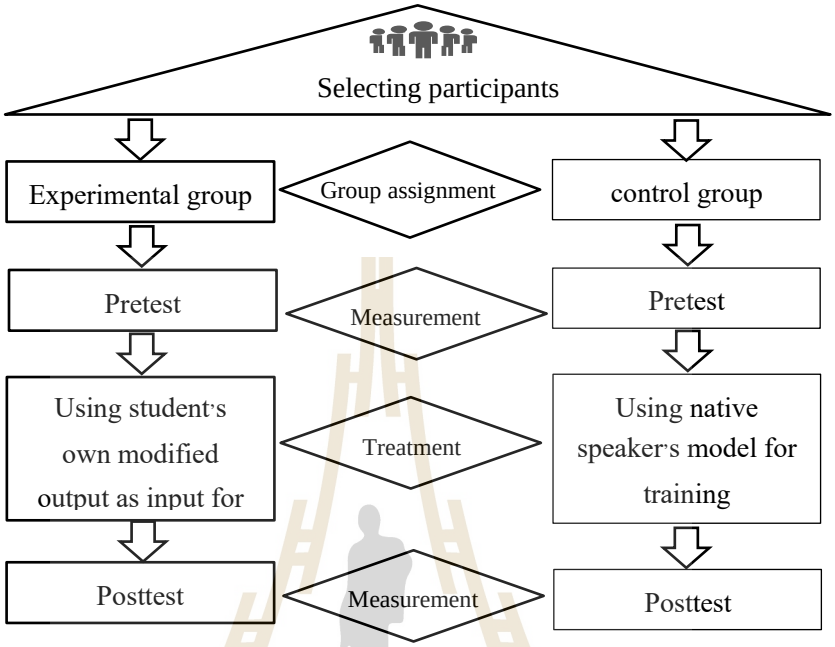


Figure 3.1 Research design of the present study

Specifically, this study used a pretest-posttest quasi-experimental design (Figure 3.1) to examine the effectiveness of pitch modification quantitatively, following Creswell (2012): experimental research seeks to determine if a specific treatment influences an outcome by providing the treatment to one group and withholding it from the other group and then determining how the two groups perform. The researcher of this study first selected participants from the population pool and then assigned them to the experimental group (EG) and the control group (CG). After a pretest, both groups received a treatment: an English intonation training. The only difference of the treatment between the two groups was that they used different pronunciation models for training-the CG used the native speaker's model while the EG used the students'

own modified output as the model. Lastly, the two groups' performance on the posttest were compared by using tests of statistical significance.

The qualitative data were from two sources: the participants' learning journals during the training and semi-structured interviews about the participants' opinions of the training. The qualitative findings were supposed to make up the weakness and bias of the quantitative findings and further explain the initial results, thus making this study more convincing.

3.1.1 Variables

The present study aimed at improving Chinese university EFL students' English intonation by intervention of students' intonation learning through modifying the pitch of the input speech signals. The independent variable was pitch modification, i.e., modifying students' incorrect production to the correct one and feeding it back to the students as the input model for learning. The dependent variables were students' English intonation perception and production performances.

3.1.2 Participants

The participants of this study were the Chinese first year undergraduate English majors studying at Hunan University of Science & Engineering (HUSE), China. In HUSE, there were altogether 204 first year English majors assigned to 6 academic classes (with about 34 students in each class). Students of two intact classes with the same teacher of English phonetics course were selected as the participants in order to

control the confounding variables that might arise from the phonetics course. One of the two classes was randomly selected as the experimental group (EG) and the other as the control group (CG).

Before the experiment, the participants were clearly informed of the information of this study, including the purposes, experimental procedures, potential benefits or risks, and confidentiality. They were given time to ask questions and to make decisions on whether to participate or not. A consent form was signed with the researcher and those who decided to participate in the study. Finally, altogether 66 students (33 in each group) were recruited in the experiment.

3.2 Instruments

In order to collect the quantitative and qualitative data, seven different types of instruments were employed. They were the English intonation learning materials, pretest and posttest, intonation training courseware, students' learning journals, questionnaire, and semi-structured interviews.

3.2.1 Learning materials

The learning materials used for English intonation training were developed by the researcher. The intonation exercises were adapted from previous studies, pronunciation handbooks and intonation textbooks (Thompson, 1981; Brazil, 1997; Kelly, 2000; Levis, 2002; Wang, 2005; Wells, 2006; Hewings, 2007; Nolan, 2008; Levis

& Wichmann, 2015). Design of the materials followed the British school's approach to intonation analysis. Thus, the contents for each training session were divided into three modules-tonality, tonicity and tone-according to the trio-system of English intonation. The materials altogether included 150 target utterances which were specially designed in particular contexts for practicing intonation patterns. (See Appendix C for an example of the training materials used in one training session). By referring to the principles for English intonation instruction proposed by previous researchers (Levis,1999; Tench,2005; Clennell, 1997), the materials were selected and designed according to the following principles:

(1) Teaching intonation in context. As the meaning of intonation is context-dependent, it is inappropriate to talk about the meaning of intonation without context. In this study, the intonation patterns were taught in dialogues rather than decontextualized sentence level. This enabled students to interpret the meaning of intonation patterns in discourse level and make the learning of intonation occur in contexts.

(2) Focusing on generalizable contents. In the field of intonation research, researchers have not reached universal agreement on the descriptions of English intonation. Even within the British school's contour approach, researchers have discrepancies on classifying intonation patterns. Considering the diversities of intonation among different varieties of English, this study did not touch upon those

discrepancies, but focused on the universally accepted intonation patterns among different varieties of English, so that students can generalize what they have learned to various contexts and stay “safe” in international communications.

(3) Providing learnable contents. Given the complexity of the interpretations of the secondary tones (the degrees of pitch movements), which make intonation difficult to teach, this study only focused on the more teachable primary tones that can lead to contrastive pitch movements, i.e., the falling tone, the rising tone, and the falling-rising tone, which have been adopted by some intonation handbooks or textbooks (Tench, 1996; Wang, 2005; Wells, 2006; Hewings, 2007).

(4) Addressing the communicative purpose. Using language for communication is the chief goal for most foreign language learners. It is essential to address the communicative value of intonation in teaching intonation. The intonation examples selected for the training are those utterances frequently used in daily communication. This can make it easier for the students to apply what they have learned to daily use outside the classroom. It can also contribute to the students’ beliefs toward intonation that what they have learned about intonation is practical and essential for communication.

3.2.2 Pretest and posttest

The pretest and posttest were used to test the students’ English pronunciation ability before and after the training. The pretest and posttest were identical, but none of

the testing items was directly taught during the training. Given the complexity of assessing pronunciation ability and the fact that no previous research had developed a complete test for assessing EFL students' pronunciation ability in terms of intonation, the tests used in this study were constructed by adapting and integrating intonation examples and exercises from previous studies, pronunciation handbooks and intonation textbooks (Thompson, 1981; Brazil, 1997; Kelly, 2000; Levis, 2002; Wang, 2005; Wells, 2006; Hewings, 2007; Nolan, 2008; Bradford, 1988).

As Bachman & Palmer (1996) claimed, a valid language test should cover the learning content and reflect the teaching objectives. The construction of the test was also rigorously designed to cover the training contents and measure students' learning outcomes guided by the teaching objectives, so as to enhance the construct and content validity of the test (Creswell, 2012).

The pretest and posttest included two sub-tests: the pronunciation perception test and the pronunciation production test. Previous studies have investigated the relationship between speech perception and production, and most of them found that there is a link between perception and production and perceptual training can improve language learners' production ability (e.g., Bradlow et al., 1997; Wang et al., 1999; Hardison, 2012). However, some researchers held that there is an imbalance between perception and production (Elliot, 1991; Sung & Flege, 2005) and "perceptual mastery is not necessarily a causative factor in the acquisition of productive skills" (Sheldon &

Strange, 1982, p.257). Therefore, it is necessary to test students' English intonation perception and production abilities separately.

The perception test (Appendix D) was composed of 38 multiple choice questions categorized into three modules, respectively testing students' intonation perception ability in tonicity, tonality and tone. The students were asked to perform listening comprehension tasks in order to test their perception and understanding of English intonation. The production test (Appendix F) had the same structure with the perception test, but in the production test, the students were asked to produce the target utterances with proper intonation according to the contexts provided.

The validity and reliability of the test were checked by three experts specializing in English language education. They were invited to rate the relevance of the content and the appropriateness of each testing item by using the Item-Objective Congruence Index (IOC). The IOC validation method used a 3-point scale evaluation (1 = relevant, 0 = uncertain, -1 = irrelevant). Items scored lower than 0.5 ($-1 \leq IOC \leq 1$) were revised and reevaluated under the experts' suggestions, so as to ensure that the mean IOC score of each item was higher than 0.5 (See Appendix E and Appendix G for the final results of the IOC rating).

3.2.3 English intonation training courseware

The students received the training by using an English intonation training courseware. The courseware was developed by the researcher with an e-learning

authoring tool-Lectora Inspire (version 17, commercially available online). Lectora Inspire, released by *Trivantis* Corporation, is one of the leading authoring tools used for responsive e-learning development (Rosevella, 2016). A powerful tool, at the same time it is easy to operate for average computer users and enables non-programmers to develop e-learning courseware (Wibawa, 2017). The development of the intonation training courseware for the present study followed 5 major steps:

Step 1: Create a new blank page and customize the master page by selecting desired template and adding necessary objects that can be inherited.

Step 2: Manage the structure of the content in the "Title Explorer" area by adding and editing necessary chapters, sub-chapters, or pages.

Step 3: Import the materials into each page and insert necessary medias such as texts, audios, recorders, and corresponding actions.

Step 4: Publish the courseware package to Web (HTML). Before publishing the courseware to web, it is essential to preview and check the its operation. Lectora Inspire has an automatic error checking procedure if the user clicks the "publish" button. It is still necessary for the user to run the courseware and check for potential errors again since some kinds of errors cannot be automatically detected by the system.

After the courseware was developed, three experts specializing in instructional design and e-learning development were invited to check its usability by using an evaluation form adapted from the *CALL Software Evaluation Form* (Davies,

2004) previously used for the training programme of ICT4LT (Information and Communications Technology for Language Teachers). The evaluation form (Appendix H) indicated that this courseware was feasible for English intonation training with high degree of usability.

3.2.4 Students' learning journals

A learning journal is a writing assignment, typically to be performed as a follow-up course-work activity (Hübner et al., 2010). The writing is about one's learning experiences and it is "a vehicle for reflection" (Moon, 2006, p.1). In the field of education, it is employed as a strategy to enhance learning (Miller, 2017) and monitor learning. The benefits of using learning journal lie in that it can:

1) record the students' learning experiences by tracking what they did, how they did it, and what they thought;

2) facilitate learning and deepen learning experiences (Miller, 2017);

3) increase the students' ability in self-reflection and critical thinking (Miller, 2017; Moon, 2006);

4) foster beneficial cognitive and metacognitive learning strategies (Hübner et al., 2010);

5) develop the students' writing skills;

6) be used as a means of assessment in formal education (Moon, 2006).

7) encourage self-expressions, give or improve “voice” (Moon, 2006);

In the present study, English intonation was not taught explicitly to students. Instead, students were exposed to systematic and representative intonation examples and they were expected to extract the intonation rules from those examples. The reasons are as follows: firstly, the main objective of this study was to compare the effectiveness of two different kinds of models for pronunciation learning-the native-speaker model and students’ own modified speech as the model. By employing implicit teaching, the researcher of this study could better investigate the effectiveness of the two kinds of input on students’ learning behaviors and learning outcomes, so as to answer the research questions of this study.

Secondly, the nature of language learning determines that pronunciation teaching should be implicit. Krashen (1981) believes that learners should learn languages just like children: with natural input without any explicit teaching, because learning can never be converted into acquisition and learned rules always decrease a learners’ performance. Criticisms to explicit teaching (Krashen & Terrel, 1983; Gilbert, 2014; Reed & Michaud, 2015; Atar, 2018) hold that teaching learned rules to students has the danger that they will very unlikely use them out of the class. In other words, they do not believe what they have learned. Implicit teaching, as a bottom-up approach, views students as knowledge explorers and enables learning to occur in a process of constructing meaning. The teacher’s job is to provide stimuli for students and monitor

their learning process through instruments like observations, dairies, or learning journals.

In the present study, the learning journal was used as a means to collect data about the students' learning experience. The learning journals were used to monitor the students' learning process and learning outcomes, as well as to find out whether the students in the EG could gain more knowledge than those in the CG. Therefore, the students in both groups were asked to keep learning journals during each training session. Considering the objectives of this study and the fact that "a journal is a matter for personal experiment and choice for the learner (Moon, 2006, p.95)" and there is no fixed format for journal writing, the students in this study were allowed to write the journals freely but focus on three points: 1) the intonation rules they extracted from the exercises; 2) their reflections on what they have learned; 3) the questions they encountered during learning.

3.2.5 Questionnaire

Questionnaire is an effective instrument for collecting survey information, providing structured, often numerical data, being able to be administered without the presence of the researcher, and often being comparatively straightforward to analyse (Cohen et al., 2007). It is also the most frequently used instrument to gather information which cannot be easily observed, such as opinions, attitudes, beliefs (Oppenheim, 2000). In the present study, a questionnaire survey was conducted to investigate the

students' opinions of learning intonation through pitch modification.

The questionnaire (Appendix I) was composed of 10 statements regarding the students' opinions of the training. The students can show their degrees of agreement on each statement by ticking a 5-point Likert scale ranging from "Strongly agree" to "Strongly disagree". To avoid misunderstanding and confusion, the questionnaire was written in both English and Chinese. Construction of the questionnaire was based on previous related studies (Kang, 2010; He, 2015; Yang, 2016) and the objectives of the present study. In addition, in order to validate the content of the questionnaire, three experts were asked to rate each item using IOC (Appendix J), and subsequent revisions were made based on their suggestions.

3.2.6 Semi-structured interviews

An interview, as remarked by Cohen et al. (2007), refers to an interchange of views, thus "inter-view", between two or more people on a topic. Therefore, interview is a frequently used data collection method to elicit information about participants' views, attitudes, or opinions. Especially in qualitative research, it is called the "gold standard of qualitative research" (Silverman, 2000, p.51). Interviews are most commonly conducted by face-to-face conversations between the interviewer and interviewees. There are three types of interviews: the unstructured (open) interview, the semi-structured interview, and the structured interview. A typical semi-structured interview allows the interviewer to ask the questions prepared in advance and at the

same time, enables the interviewer to ask extra questions initiated from the answers of the interviewee in order to obtain deeper information. Therefore, a semi-structured interview is more flexible than the structured interview and more effective for obtaining the target information than the unstructured interview.

In the present study, semi-structured interviews were employed to triangulate the results obtained from the questionnaire and access more detailed information with regard to the students' opinions of the training. The interview was composed of 7 structured questions (Appendix K), most of which were based on the 10 statements of the questionnaire. The 7 questions were also validated by 3 experts using IOC (Appendix L) and tested in the pilot study.

3.3 Data collection

The general procedures for data collection were divided into 3 phases: in phase 1, the pretest was administered to all the participants; in phase 2, the intonation training was conducted; in phase 3, the posttest was administered, and immediately after the posttest, the questionnaire survey and interviews were conducted.

3.3.1 Procedures for conducting the pretest and posttest

The pretest and posttest were conducted in a computer lab in HUSE. The computers were equipped with earphones and microphones. The first section of the pretest was intonation perception test, and the test takers were required to do the test

on the intonation training courseware as the following steps:

Step 1: Sit before a computer, test the microphone, and adjust the sound volume to a comfortable level.

Step 2: Run the software and enter the student name and ID on the welcome page.

Step 3: Do the testing items according to the instructions (Figure 3.2). The students were allowed to complete the test within 90 minutes.

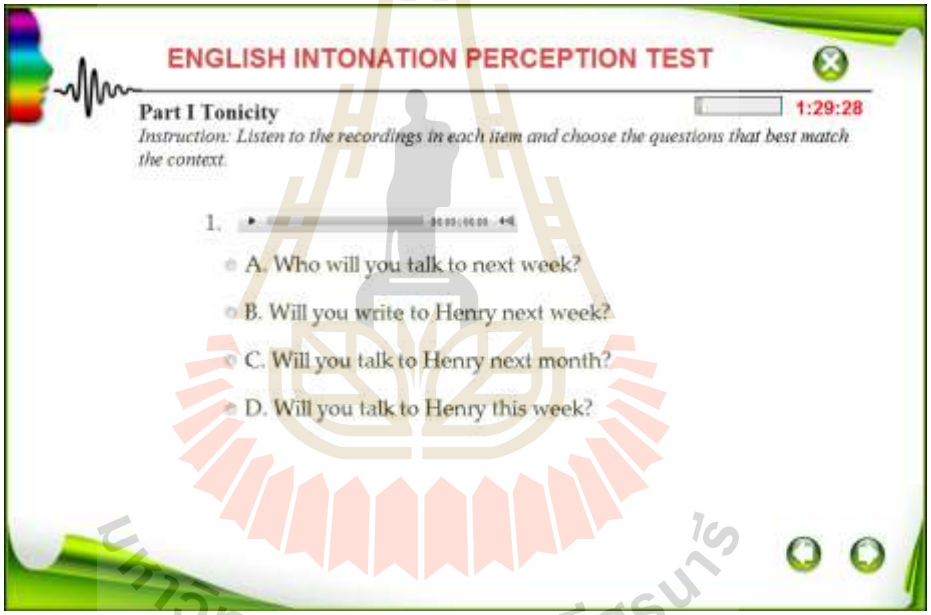


Figure 3.2 Screenshot of doing the pronunciation perception test

Step 4: Submit the answers. After the students finished the test, they can click the “Submit” button to send their answers to the server.

For the second section of the pretest-the intonation production test, the students took the test one by one in order to avoid disturbing each other. The specific procedures was the same as the procedures for the perception test except for step 3. For

the production test, in step 3, the students were asked to produce the underlined sentences in each dialogue with proper intonation according to the contexts (See Figure 3.3). The students can record their speech productions and upload them to the server by using the recorder on the page. The procedures for conducting the posttest were identical to the pretest.

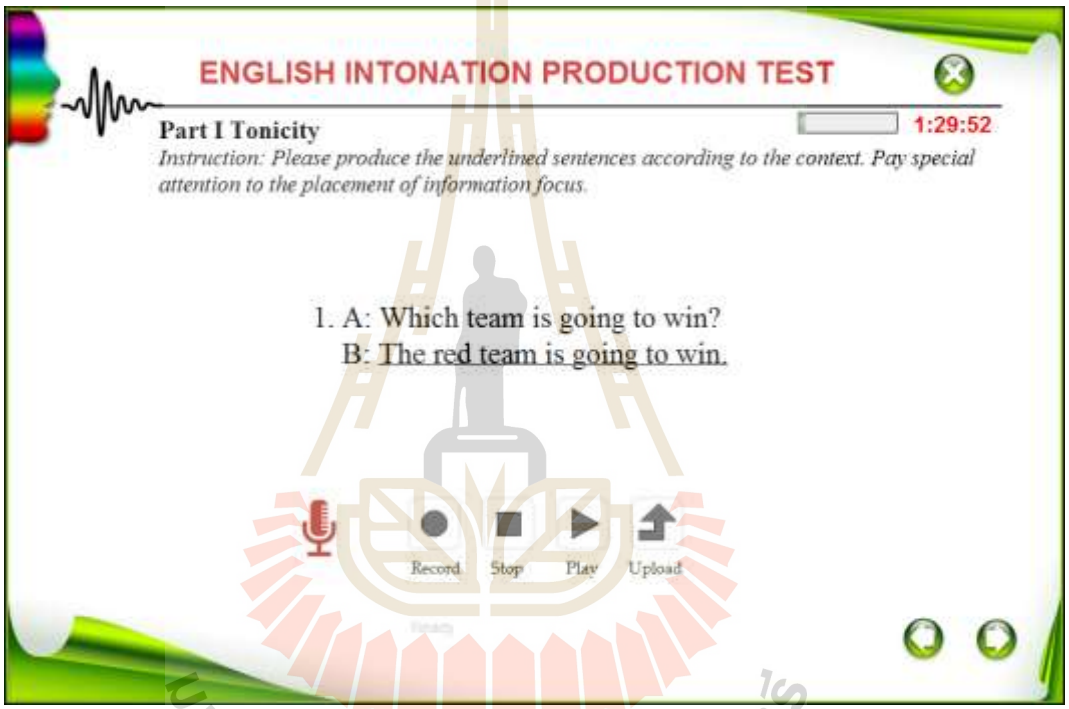


Figure 3.3 Screenshot of doing the production test

3.3.2 Procedures for intonation training

The English intonation training included 6 sessions, each session lasting for 3 hours. Considering that modifying the students' initial speech productions was quite time-consuming, a one week interval was placed between two consecutive sessions. Therefore, the 6 training sessions spanned 12 weeks. The training was conducted in computer laboratories, and the EG and CG took the training at the same time but in

different rooms. To take the training, the students were asked to do as follows:

Step 1: Run the intonation training courseware, test the microphones, and adjust the sound volume to a comfortable level.

Step 2: Log into the system with student name and ID number and click the "Start" button to start the training.

Step 3: Practice pronunciation through simple listen-compare-repeat exercises. The students can first try to produce the target utterances by themselves using the recorder on the page which can record and replay their productions. Then, they can listen to the pronunciation model and compare it with their own productions. Specifically, for students in the CG, they can click the underlined sentences to listen to the native speaker's model (see Figure 3.4).

Figure 3.4 Screenshot of the courseware page for the CG

For students in the EG, they can listen to their original productions and the modified version (see Figure 3.5). After a certain number of listen, compare, and repeat sequences, they were asked to upload a recording of each target utterance. Their uploaded recordings would be sent to the researcher's server, and the researcher could monitor their progress by reviewing these productions.

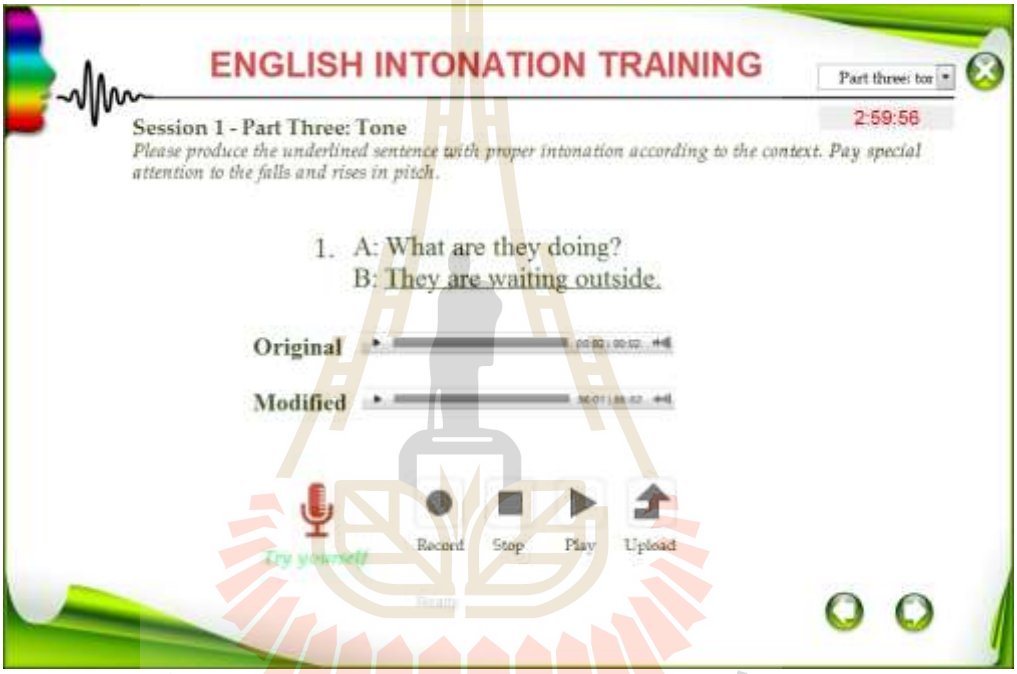


Figure 3.5 Screenshot of the courseware page for the EG

Step 4: Write the learning journal. After each training session, the students were required to keep learning journals. They could write the journal on papers in order to make it convenient for them to take notes or to make intonation marks. In addition, in order to ensure that they could clearly express what they wanted to say, they were allowed to write the journal in Chinese. Their journals were collected and transcribed verbatim for further analysis.

As the students in the EG used their modified initial productions as the model for learning, one week before each training session, the students were asked to produce target utterances. Then, the research would spend one week to modify their productions to be used as the pronunciation model for learning. The modifications of their initial productions focused on their incorrect intonation, and those productions with correct intonation were kept unchanged. The following is an example of modifying the students' production. First, see the following dialogue:

A: Jenny won a big prize!
B: What did she do?
A: I said she won a big prize.

Here, the students were expected to produce the underlined sentence "What did she do?" with a rising tone to indicate checking, as can be judged from the context. However, if one student produced it with a falling tone (see Table 3.1), his/her production would be modified by changing the falling tone to a rising tone.

Table 3.1 Student's undesirable production

Waveform	
Pitch (75-300Hz)	
Text annotation	what did she do

The modification was performed manually by using the phonetics software Praat (Boersma & Weenink, 2018). It followed five major steps. First, import the sound

file into Praat and generate a manipulation object; second, manipulate the pitch contour to the desired one (Figure 3.6); third, adjust the intensity or duration on necessary places to make it natural, especially for modifying sentence stresses and intonation boundaries; four, get the resynthesized speech. Table 3.2 shows the acoustic description of the resynthesized speech.

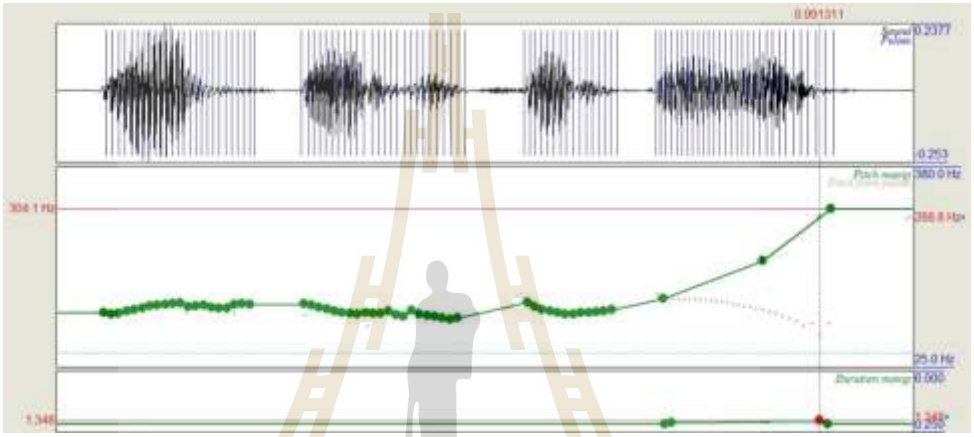


Figure 3.6 Manipulating pitch in Praat

Table 3.2 Student's speech after modification

Waveform	
Pitch (75-300Hz)	
Text annotation	what did she do

The last step involved transferring students' recorded voices to their self-perceived voices. As has been discussed in Chapter 2.4.3, to get one's self-perceived

voice, the recorded voice should be first low-pass filtered, and then mixed with the low-pass filtered voice. The generated sound mixture constituted student's self-perceived voice and was used as the stimuli for training. (See Figure 3.7 for student's recorded voice; Figure 3.8 for student's self-perceived voice)

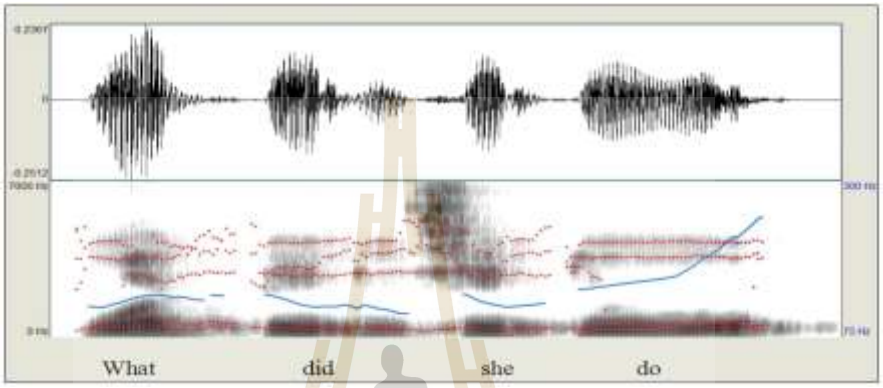


Figure 3.7 Student's recorded voice

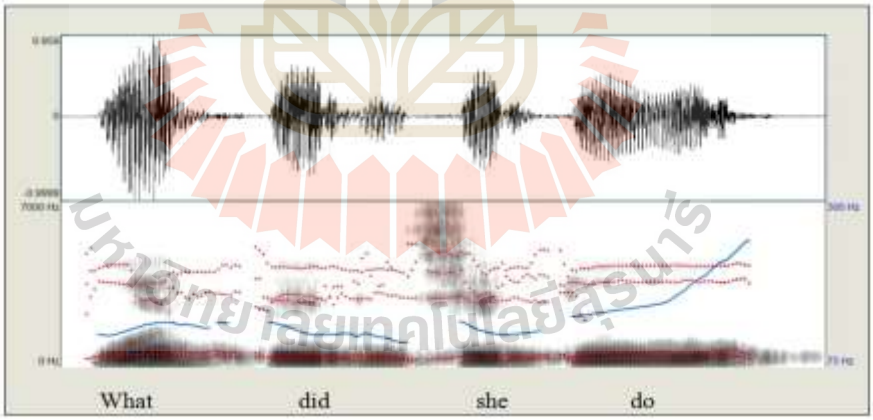


Figure 3.8 Student's self-perceived voice

For the purposes of this study, the following points needed to be addressed on modifying students' initial productions:

Firstly, the native speaker's speech, the model for the CG, was used as a reference point for modifying the students' speech in the EG. This study did not intend

to teach students to follow the native speaker's accent, but used the native speaker's speech as a reference to obtain the overall pitch contour trend, the values for F0, duration, and intensity, etc..

Secondly, the modification only focused on students' pronunciation problems, such as incorrect tone patterns, stress placements, and divisions of intonation boundaries. If there were no problems about their pronunciation, their original productions would be kept unchanged as long as the original speech was comfortably intelligible and comprehensible.

Thirdly, in order to ensure the usability of the modified speech and avoid the negative effects of resynthesized stimuli, three listeners were invited to crosscheck all of the students' modified speech and to make judgement about whether they were natural and intelligible, and revisions have been made under the listeners' suggestions.

3.3.3 Procedures for conducting the questionnaire and interviews

The questionnaire survey and semi-structured interviews for investigating the students' opinions of pitch modification for intonation learning were conducted immediately after the posttest. The questionnaire was created via Google Form and sent to all of the participants in the EG. The semi-structured interviews were conducted with 18 students based on their performance in the posttest. They were 6 students with the highest scores, 6 with the lowest scores, and 6 with the medium scores. The interviews were conducted in Chinese in order to elicit more detailed information and create more

relaxed atmosphere. Each participant was interviewed by the researcher for 7-10 minutes and their talk was audio recorded with the participants' consent.

3.3.4 Rating

The students' performance in the perception test (multiple choice questions) was automatically marked by the testing system. Their performance in the production test was rated by three teachers of English phonetics course in HUSE. The pronunciation rating involved two phases: 1) the rating for intonation choice, i.e., whether the student chose the correct intonation pattern for the target utterance; 2) the rating for the phonetic realizations of the intonation pattern, i.e., how well did the student realize the chosen intonation pattern in his/her pronunciation.

The first rating for intonation choice was carried out through perception tasks, which has been used in Bradlow et al.'s (1999) study. To be specific, the students' productions were designed into perception tasks and the three raters as listeners were asked to perform the perception tasks so as to make judgement about whether or not the students have employed the correct intonation patterns. So in this phase of rating the students' productions were scored dichotomously: "0" for incorrect intonation choice, and "5" for correct intonation choice.

The rating for phonetic realization focused on the degree to which the student's intonation deviated from the native speaker's intonation. This concept was proposed by Kennedy & Trofimovich (2008): "how closely the pronunciation of an

utterance approaches that of a native speaker" (p. 461), and by Isaacs & Thomson (2013): "how different the speaker sounds from a Native speaker" (p. 141). The rating employed a 5-point Likert scale, where "5" represents "near native speaker's intonation" and "1" represents "extremely different from the native speaker's intonation".

Therefore, for the 38 items of the intonation production test, the total score was 380, with 10 marks for each item. Of the 10 marks, 5 marks were for intonation choice and the other 5 marks were for the phonetic realizations of the chosen intonation. The rating employed a double-blind procedure. In order to ensure the inter-rater reliability, the three raters received rating training and participated in a pilot rating. Finally, after the rating, Person's Correlation Coefficient was employed to check the inter-rater reliability. The results showed that the rating in each phase were highly reliable ($r>0.80$).

3.4 Data analysis

Quantitative and qualitative analysis were employed for the data analysis in the present study (Table 3.3). Data obtained from the students' pretest and posttest scores together with the students' responses to the questionnaire were analyzed quantitatively, while data collected from the learning journals and the semi-structured interviews were applied for qualitative analysis.

Table 3.3 Data analysis for this study

Data Analysis	Quantitative	Qualitative	Purpose
Pretest & posttest	Descriptive statistics; t-tests		To answer RQ 1
Questionnaire	Descriptive statistics		To answer RQ 2
Learning journals		content analysis	To answer RQ 1
Interviews		content analysis	To answer RQ 2

3.4.1 Quantitative data analysis

Quantitative data analysis was presented in statistics by using SPSS (Version 20), which included descriptive statistics and inferential statistics. Descriptive statistics is applied to describe the data or summarize information about the data (Creswell, 2012). In the present study, students' scores on the pretest and posttest, as well as their responses to the questionnaire were first presented in descriptive statistics, including mean, frequency, and SD (standard deviation).

Inferential statistics is used to make inferences and predictions about a population based on the data (Creswell, 2012). In this study, a paired-samples t-test was applied to compare each group students' pretest scores with their posttest scores, thus, to examine whether the intonation training can improve the students' pronunciation performance. An independent-samples t-test was used to examine whether there were significant differences between the two groups of students' scores on the posttest, so as

to determine whether using students' modified productions as model was more effective than the native speaker's model for English intonation learning.

3.4.2 Qualitative data analysis

Qualitative analysis was applied to analyze the data obtained from the students' learning journals and the semi-structured interviews. The procedures for analysis followed Creswell's (2012, p.247) six-step qualitative data analysis:

Step 1: Organizing and preparing data for analysis. Students' learning journals were imported into Word files and coded with "E1- E33" for the EG and "C1-C33" for the CG, generating 66 word files. Audio recordings of the semi-structured interviews were transcribed verbatim into Word files and coded with "EI1-EI18", yielding 18 Word files.

Step 2: Reading through all data, to get a quick review of the raw data in order to get familiar with the data.

Step 3: Coding the data, referring to the initial coding for the data involving carefully reading the data and coding them by looking for units of meaning.

Step 4: Developing themes or categories. In this step, the coding process was reviewed to generate themes or categories by examining the relationships or patterns across the initial coding.

Step 5: Interrelating the themes or categories. Here, the themes or categories generated from the last step were summarized and connected by making meaning of

their interrelations.

Step 6: Interpreting the meaning of themes or categories. The researcher went beyond the descriptive data to extract meaning and insights from the data.

In order to ensure the dependability of the qualitative data analysis process, two experts specializing in qualitative research were invited to cross-check the codes and the coding process. The data analysis procedures have been further refined under the experts' suggestions.

3.5 Piloting

A pilot study was conducted to test the feasibility of the instruments and procedures. The pilot study spanned two weeks and the participants were 10 first year English majors of HUSE, with 5 students in the EG and 5 students in the CG. The participants received an intonation training of 2 sessions (one session per week). Problems encountered in the pilot study and some preliminary findings are described below.

3.5.1 The viability of using modified stimuli

As the students in the EG received the training by using their modified initial productions as the input model for learning, before each training session, the 5 students in the EG were asked to produce the target utterances and their productions were audio recorded. Then, their productions were modified through two steps: 1) transferring the students' recorded voices to their self-perceived voices; 2) modifying the incorrect

intonations of the students' productions to the correct ones.

As has been discussed in Chapter 2.4.3, the voice transformation involves two sub-steps: 1) low-pass filtering the students' recorded voices; 2) mixing the low-pass filtered voices with the original recorded voices. The students were asked to choose their preferred values for the low-pass filtering and the mixing ratio through delayed feedback trials which allowed them to select the values according to their own real-time perception. Figure 3.9 presents the results of the students' choices for the low-pass filtering.

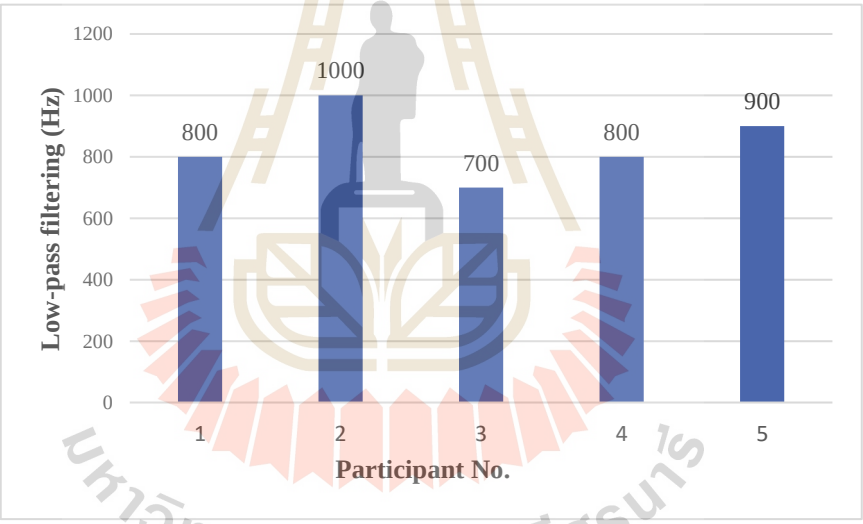


Figure 3.9 The students' preferred values for the low-pass filtering

As can be seen from the figure, the students' preferred values for the low-pass filtering range from 700 Hz to 1000 Hz, which was in accordance with Pörschmann's (2000) claim that bone conducted sound dominates the perception of one's own voice for frequencies between 700 Hz and 1200 Hz. Figure 3.10 shows the results of the students' preferences for the mixing ratio.

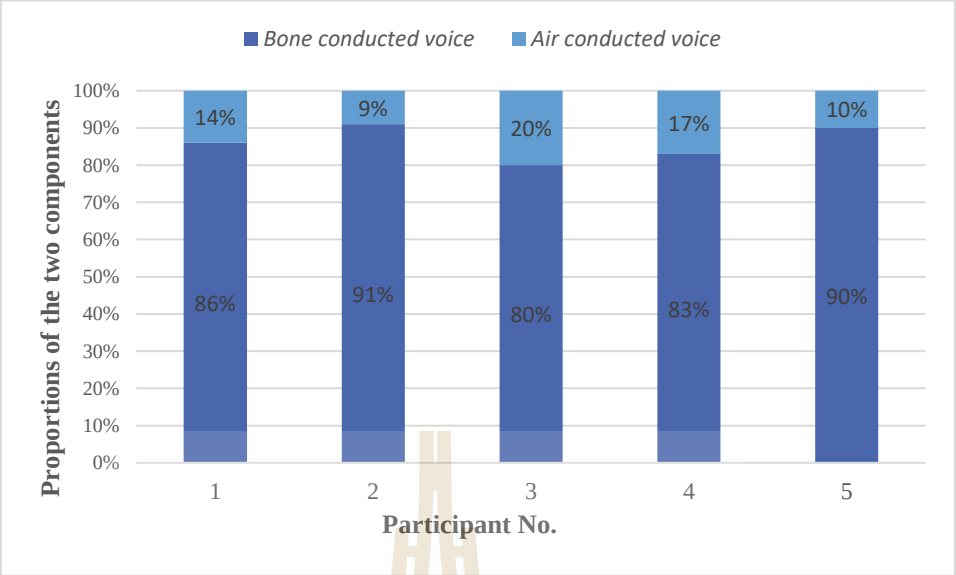


Figure 3.10 The students' preferred ratios of the sound mixture

Considering the students' choices for the low-pass filtering and mixing ratio, there was a tendency that the participants who chose a higher value for the low-pass filtering tended to choose a lower proportion for the air conducted component in the sound mixture. This means that the participants consciously adjusted the ratio between low frequencies and high frequencies to make the sound mixture sounded the most similar to their self-perceived voices.

These results revealed that the method for transferring students' recorded voice to their self-perceived voice was viable and reliable. The semi-structured interviews with the 5 students also showed that all of them immediately recognized their voices and none of them have reported negative evaluations on the manipulated voices. This implied that using students' self-perceived voices for pronunciation teaching can avoid the negative effects of using recorded voices as reported in former

studies (Holzman et al., 1966; Gaviria, 1967; Weston & Rousey, 1970; Daryadar & Raghibi, 2015).

For the modifications of the students' incorrect productions to the correct ones, the native speaker's productions were used as the reference points, and then the students' modified productions were checked by three experts. The experts frequently evaluated the modified speech as "natural", "no differences from the naturally produced speech", or "natural speech without modification". However, during the training, some students reported that sometimes they could not perceive the differences between their original productions and the modified version, especially when the speech was only slightly modified. Therefore, the researcher decided to add necessary exaggerations to the modified speech to make the modification more salient in the main experiment. By so doing, the students could better perceive what have been modified, and the modified speech can better arouse the students' awareness of their incorrect pronunciations and facilitate their meaning making in the process of learning.

3.5.2 Problems about the instruments

The first problem about the instruments involved the perception test and the production test. The two tests respectively included 38 testing items organized in the order of modules or sub-modules. For example, in the production test, module 3, items Nos.5-8 aimed to test students' pronunciation on wh-questions. However, when the testing items were organized in this order and administered to the students, most of the

students' performance on the items tended to be influenced by the neighboring items. For example, in the perception test, module 1, items Nos.3-4 belong to the sub-module of *Contrastive focus*. If one student's answer on item No.3 was incorrect, his/her answer for No.4 would also tend to be wrong. To avoid this phenomenon, the testing items of the perception test and production test were administered in random order to the participants in the main experiment.

The second problem was about preparing the training stimuli for the EG. As the students in this group received the training by using their own voices as the stimuli, the students' initial productions on the training materials should be collected immediately after the pretest and before the training. During this period, their initial productions should be modified and uploaded to each student's individualized intonation training courseware. This process turned out to be very time consuming in the pilot study. Therefore, the researcher decided to hire two assistants to help to collect data, manipulate students' speech, and develop the courseware in the main experiment.

The third problem involves the learning journal. The students in both groups were asked to keep learning journals during each training session, focusing on three points: the knowledge they learned, their reflections, and the questions they encountered. However, the pilot study showed that most students focused on summarizing what they learned and few reflections have been found in their journals. Therefore, in the main experiment, the researcher decided to give the students more

detailed instructions on how to write learning journals by showing them a good model selected from the pilot study.

3.5.3 Preliminary findings

The pilot study yielded some preliminary findings which could be outlined as follows:

Firstly, pitch modification had the strength of drawing students' attention on their pronunciation problems, made pronunciation learning more interesting, and motivated students to devote more time and energy to enhancing their pronunciation. Especially in the first training session, the students in the EG showed great interests in their modified productions. In the interviews, they reported that this was an interesting pronunciation learning experience that they had never had before. Their responses to the questionnaire also showed that most of them preferred to listen to their modified productions as the pronunciation model than the native-speaker model (5-point scale, mean=4.75).

Secondly, pitch modification set a more realistic pronunciation goal for students to pursue. The concept of World Englishes has influenced most pronunciation instructors to accept the fact that attaining a native-like pronunciation is an unrealistic goal (Morley, 1991) for most EFL learners. However, even though the goal of "comfortable intelligibility" (Kenworthy, 1987) is well-accepted, most instructors are confused about what a comfortably intelligible speech is like. From the students'

perspective, on the one hand they began to realize that the native-like goal is unrealistic, while on the other hand, they have no idea how to make their speech intelligible and to what extent their pronunciation can be accepted, leaving students' pronunciation learning in the dark.

Pitch modification focuses on the individual student's specific pronunciation problems and targets on how to solve their problems and how they can be correct. Students' modified output as input, with the voices with which they were most familiar, can be used as a more realistic model for students to imitate, and students can directly perceive the differences between their undesirable pronunciations and the desired ones and gradually approximate to the target goal. Therefore, pitch modification can inform students how their pronunciation can be correct and to what extent their pronunciation can be accepted, and thus had the potential of enhancing students' confidence of their own pronunciation and speaking English.

Thirdly, the students' performance on the perception test and production test showed that their problems in English intonation were caused by either phonological representation or phonetic realization, or both of the two dimensions. Hence, their problems can be classified into three types. Type 1 problem: the students have not formed the phonological representation of the target intonation pattern and they also had difficulties in pronouncing the pattern; type 2 problem: the students had no problems in pronouncing the intonation pattern but they were unable to choose that

pattern in the target contexts; type 3 problem: the students were able to choose the correct pattern but they failed to pronounce it accurately.

The results of the students' performance showed that both the native speakers' model (for the CG) and the students' own modified speech as the model (for the EG) was effective for solving type 2 problem, while the latter was more effective for solving type 1 and type 3 problems. In a word, imitating the pronunciation model of the students' own voices has the potential of facilitating the students' more accurate phonetic realizations of the target intonation.

In the pilot study, some students also raised some questions and pointed out some weaknesses of this training. Two students complained that sometimes they could not identify the modifications of their original speech and sometimes they did not know why some places should be modified. The reason might be that these two students' pronunciation levels were relatively high and little was changed to their original productions or the modifications were not so obvious as to arouse their awareness as mentioned above. There were also two students who reported that they also wanted to listen to the native speaker's model although they were quite in favor of their modified input. In other words, they were not so confident of using their modified input as the model and they wanted to know what the native speaker's model is like. In order to meet their demand and avoid confounding variables to the experiment, a feasible way to solve this problem was to provide the native speaker's model to them after the

posttest, so that they could either compare their original productions with the modified versions or with the native speaker's model.

3.6 Summary

This chapter presented the methods and materials used in the present study. Firstly the research design of this study was described. Then, the research instruments were detailed, including the learning materials, pretest and posttest, intonation training courseware, students' learning journals, questionnaire, and semi-structured interviews. Next, the detailed procedures for data collection by using these instruments were depicted, followed by the description of the methods and procedures for data analysis. Finally, the pilot study was discussed.

CHAPTER 4

RESULTS

This chapter aims at presenting the results yielded from the data in response to the two research questions. It first shows the results of the students' performance on the intonation perception test, followed by the production test. Then, the students' learning journals are analyzed. Next, results of the questionnaire survey and semi-structured interviews are presented. The last section wraps up the results to answer the research questions.

4.1 Results of the intonation perception test

As has been detailed in Section 3.2.2, the intonation perception test was specifically designed to test the students' ability of perceiving and understanding the meaning of English intonation. The test (Appendix D) included 38 multiple choice questions which were designed in the form of listening comprehension task. The students were required to listen to the recordings and make judgement according to the contexts and their understanding of intonation. The total score of the test was 38 (1 credit per item), and the students' answers were graded automatically by the testing system. The results of the students' performance are presented below.

Pretest

Firstly, descriptive statistics was employed to show the students' performance in the pretest, including the sample number, mean, standard deviation (SD), range, minimum, and maximum (Table 4.1).

Table 4.1 Descriptive statistics for the students' pretest scores (perception test)

Group	Number	Mean	SD	Min.	Max.
EG	33	21.79	3.81	15	30
CG	33	23.12	2.72	17	28

As can be seen from the table, both the experimental group's (EG) and the control group's (CG) performance were unsatisfactory, with their means approximating 60% of the total score (38*60% = 22.8). The highest score was 30 and the lowest was 15, both of which were from the EG, and the performance of the EG (SD=3.81) showed more variability than that of the CG (SD=2.72).

For the inferential statistics, the significance level set in the present study was 0.05, i.e., if p value<0.05, the null hypothesis will be rejected, and significant difference will be reached. The confidence interval was set at 95%. Shapiro-Wilk was used to test the distribution normality since the sample size of each group was smaller than 50. The tests of normality showed that both group's scores followed normal distribution with p=0.55>0.05 for the EG and p=0.17>0.05 for the CG. Thus, an independent-samples t-test was employed to compare the means of the two groups.

Table 4.2 Comparison between the two groups' pretest scores (perception test)

Levene's Test for Equality of Variances		t-test for Equality of Means					
F	Sig.	t	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
4.25	0.04	-1.64	0.11	-1.33	0.82	-2.96	0.29
		-1.64	0.11	-1.33	0.82	-2.96	0.30

Table 4.2 shows the results of the t-test. Firstly, Levene's test for equality of variances showed that the two groups' scores had unequal variances, which determined that the values of the t-test should be in the second line. Therefore, as can be known from the t-test, there was no statistically significant differences between the two groups' performances ($t=-1.64$, $p=0.11>0.05$). In other words, in terms of overall scores, the two groups' performances in the intonation perception test before the training were at about the same level. In order to examine whether there were differences between the two groups' scores in specific intonation patterns, independent-samples t-tests were carried out to compare the their performances across all of the patterns. Table 4.3 lists the two groups' scores as well as the results of the t-test in each one of the intonation patterns.

Table 4.3 Comparisons of the two groups' pretest scores in specific intonation patterns (perception test)

Intonation pattern	EG		CG		Indept-S. t test	
	Mean	SD	Mean	SD	t	Sig.(2-tailed)
Tonicity	3.45	1.18	3.61	1.20	-0.52	0.61
Tonality	3.21	0.93	3.61	1.14	-1.54	0.13
Statement (F)	1.7	0.47	1.76	0.44	-0.55	0.59
Statement (R)	1.76	0.44	1.70	0.47	0.55	0.59
Wh-question (F)	1.12	0.78	1.39	0.75	-1.45	0.15
Wh-question (R)	0.67	0.60	0.76	0.56	-0.64	0.52
Yes/no-question (F)	0.79	0.65	0.91	0.63	-0.77	0.45
Yes/no-question (R)	1.39	0.61	1.24	0.61	1.01	0.32
Tag-question (F)	0.67	0.54	0.88	0.55	-1.59	0.12
Tag-question (R)	1.09	0.63	1.21	0.55	-0.84	0.41
Command	1.73	0.45	1.82	0.39	-0.87	0.39
Exclamation	1.85	0.36	1.85	0.36	0.00	1.00
Implication	0.45	0.56	0.61	0.61	-1.05	0.30
Alternative question	1.15	0.62	0.94	0.61	1.40	0.17
Listing	0.76	0.56	0.85	0.57	-0.66	0.51

Note:

1. "F" in brackets stands for "Falling tone" and "R" stands for "rising tone". E.g., "Statement (F)" means the case of producing a statement with a falling tone.

2. "Indept-S. t test" is the abbreviation for "Independent-samples t test".

The results indicated that: 1) there were no significant differences between the two groups' performances across all of the intonation patterns (All of the p values were higher than 0.05); 2) the mean scores of both groups in tonicity and tonality were around 60% (60%*6=3.6) of the total score, which were not so satisfactory; 3) in terms of tone, both groups achieved high scores in *Statements (falling tone and rising tone)*, *Commands,* and *Exclamations* (mean>1.7=85% of the total), while performed very

poor in *Wh-questions (rising tone), Yes/no-questions (falling tone), Tag-questions (falling tone), Implications, and Listings* (M<1=50% of the total score).

Posttest

Table 4.4 shows the descriptive statistics for the students' performance in the posttest. It can be roughly seen that both groups performed better than the pretest, with their mean scores higher than 31 (about 82% of the total score). The lowest score was 26, still higher than their mean score in the pretest. One student in the EG achieved the full score 38. From these numbers, it can be estimated that students' performance in the posttest was better than that in the pretest.

Table 4.4 Descriptive statistics for the students' postttest scores (perception test)

Group	Number	Mean	SD	Range	Min.	Max.
EG	33	32.24	2.80	12	26	38
CG	33	31.36	2.54	10	26	36

Paired-samples t-tests were conducted to examine whether there were statistical significant differences between the students' performance in the pretest and posttest. Results (Table 4.5) showed that there were significant differences between their pretest scores and posttest scores, either for the EG (t=-13.24, p=0.00<0.05), or for the CG (t=-15.51, p=0.00<0.05). In other words, the students' (both groups) ability in perceiving or understanding the meaning of English intonation were significantly improved after the training.

Table 4.5 Comparisons between the students' pretest and posttest scores (perception test)

Pair (pretest-posttest)	Mean	SD	95% Confidence Interval of the Difference		t	Sig. (2-tailed)
			Lower	Upper		
EG	-10.46	4.54	-12.06	-8.85	-13.24	0.00
CG	-8.24	3.05	-9.33	-7.16	-15.51	0.00

To compare the two groups' performances in the posttest, an independent-samples t-test was conducted. Result (Table 4.6) showed that the EG performed a little bit better than the CG but the difference did not reach a significant level (t=1.34, p=0.19>0.05). That is to say, the two groups' performances were at about the same level in the posttest.

Table 4.6 Comparison between the two groups' posttest scores (perception test)

Levene's Test for Equality of Variances		t-test for Equality of Means					
F	Sig.	t	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
0.07	0.79	1.34	0.19	-1.33	0.88	-0.43	2.19
		1.34	0.19	-1.33	0.88	-0.43	2.19

Independent-samples t-tests were also employed to compare the two groups' performances in specific intonation patterns. Results revealed that there were also no significant differences between the two groups' mean scores across all of the intonation patterns. Furthermore, paired-samples t-tests were used to compare the students' pretest and posttest scores in specific intonation patterns. Results are shown in Table 4.7.

Table 4.7 Comparisons of the students' pretest and posttest scores in specific intonation patterns (perception test)

Pair: pretest - posttest	Experimental group			Control group		
	Mean differences	t	Sig. (2-tailed)	Mean differences	t	Sig. (2-tailed)
Tonicity	-1.97	-11.15	0.00	-1.76	-8.24	0.00
Tonality	-1.64	-9.17	0.00	-1.27	-5.94	0.00
Statement (F)	-0.15	-1.97	0.06	-0.12	-1.28	0.21
Statement (R)	-0.15	-1.54	0.13	-0.12	-1.68	0.10
Wh-question (F)	-0.73	-5.21	0.00	-0.39	-3.21	0.00
Wh-question (R)	-1.06	-9.25	0.00	-0.85	-6.84	0.00
Yes/no-question (F)	-0.76	-6.14	0.00	-0.49	-3.70	0.00
Yes/no-question (R)	-0.46	-3.67	0.00	-0.61	-4.94	0.00
Tag-question (F)	-0.79	-5.80	0.00	-0.46	-3.46	0.00
Tag-question (R)	-0.55	-4.71	0.00	-0.52	-4.44	0.00
Command	-0.15	-1.72	0.10	-0.09	-1.14	0.26
Exclamation	-0.06	-0.81	0.42	-0.03	-0.44	0.66
Implication	-0.88	-5.89	0.00	-0.61	-4.94	0.00
Alternative question	-0.49	-5.49	0.00	-0.46	-3.92	0.00
Listing	-0.61	-4.94	0.00	-0.49	-3.90	0.00

Note: This table shows the results of the paired-samples t-tests for comparing the students' pretest scores and posttest scores in specific intonation patterns for both groups. E,g., for the first pair tonicity, the students' pretest scores in tonicity minuses their posttest scores in tonicity. The results read: for the experimental group, mean differences=-1.97, t=-11.15, p=0.00; for the control group, mean differences=-1.76, t=-8.24, p=0.00.

As can be seen from the table, for both groups, there were statistically significant differences ($p<0.05$) between their pretest scores and posttest scores across all of the intonation patterns except for *Statements (falling tone), Statements (rising tone), Commands,* and Exclamations. The reason was that they had obtained high scores for these four cases in the pretest, which caused the ceiling effect.

To conclude the results of the intonation perception test, there were no significant differences between the experimental group's and the control group's performance, either in the pretest or the posttest, i.e., they performed equally poor in the pretest while equally satisfactory in the posttest. In terms of their performances in specific intonation patterns, both groups' performances in *Statements (both falling tone and rising tone), Commands,* and *Exclamations* were satisfactory in the pretest which left nearly no space for improvement, while their performances in the other patterns were far from satisfaction in the pretest but were significantly improved in the posttest. Therefore, one conclusion can be made that both the native-speaker model (for the control group) and the students' modified speech as the model (for the experimental group) for intonation training were equally effective for helping students perceive and understand the meaning of intonation. That is to say, at the perception level, the students' modified speech as the model for intonation learning showed no advantages over the native-speaker model.

4.2 Results of the intonation production test

The intonation production test had the same structure with the perception test. In the production test, the students were required to produce the target utterances with proper intonation according to the provided contexts. Their speech productions were audio recorded and coded for rating. The rating involves two phases. Phase 1 was the

rating for intonation choice, and phase 2 involved the rating for phonetic realizations of the chosen intonation using a 5-point Likert scale. The results were as follows.

Pretest

Descriptive statistics of the students' pretest scores (Table 4.8) showed that both two groups' mean scores approximated 56% of the total score (380), which was relatively low. None of the students' scores were higher than 300 (about 79% of the total score). The performance of the CG (SD=31.80) was more consistent than that of the EG (SD=36.41). Result of an independent-samples t-test showed that there was no significant differences between the two groups' performances (t=0.41, p=0.68>0.05), i.e., the EG and the CG performed equal in terms of the total score in the pretest.

Table 4.8 Descriptive statistics for the students' pretest scores (production test)

Group	Number	Mean	SD	Range	Min.	Max.
EG	33	213.48	36.41	131	157	288
CG	33	210.00	31.80	136	156	292

Table 4.9 lists the two groups' performances in specific intonation patterns and the comparisons (through independent-samples t-tests) of the two groups' mean scores in those patterns. As the results revealed, there were no significant differences between the two groups' performances across all of the intonation patterns. In terms of tonicity and tonality, their mean scores were around 50% the total score (60*50%=30). Their performances in *Statements (falling tone), Commands,* and *Exclamations* were

excellent, with the mean scores over 18 (90% of the total score). While their performances in *Wh-questions (falling tone), Yes/no-questions (falling tone), Tag-questions (falling tone), Implications, Alternative questions*, and *Listings* were far from satisfactory, with the mean scores lower than 8 (40% of the total score).

Table 4.9 Comparisons of the two groups' pretest scores in specific intonation patterns (production test)

Intonation pattern	EG		CG		Indept-S. t test	
	Mean	SD	Mean	SD	t	Sig.(2-tailed)
Tonicity	33.00	9.27	32.03	7.58	0.47	0.64
Tonality	26.24	8.22	24.24	8.27	0.99	0.32
Statement (F)	18.67	2.03	18.27	2.92	0.64	0.53
Statement (R)	13.67	5.76	13.42	5.79	0.17	0.87
Wh-question (F)	13.45	6.62	14.70	5.99	-0.80	0.43
Wh-question (R)	8.97	6.27	9.55	6.67	-0.36	0.72
Yes/no-question (F)	5.39	7.83	5.91	6.61	-0.29	0.77
Yes/no-question (R)	12.82	5.99	11.79	6.16	0.69	0.49
Tag-question (F)	8.79	6.96	9.21	6.96	-0.25	0.81
Tag-question (R)	13.91	6.95	13.27	6.86	0.37	0.71
Command	18.94	1.14	19.21	1.17	-0.96	0.34
Exclamation	19.00	1.23	19.06	1.14	-0.21	0.84
Implication	3.03	4.75	3.36	5.44	-0.27	0.79
Alternative question	8.79	6.55	8.42	6.05	0.23	0.82
Listing	8.82	6.20	7.55	5.33	0.90	0.37

Posttest

Descriptive statistics of the students' posttest scores are presented in Table 4.10. As can be seen from the table, both groups' mean scores were higher than 72% (273) of the total score. The lowest score was 223, which was still higher than the mean score

of the pretest (212). Paired-samples t tests were employed to compare the students' pretest and posttest scores. Results (Table 4.11) showed that there were significant differences between their pretest scores and posttest scores (EG, t=-17.51, p=0.00<0.05; CG, t=-14.13, p=0.00<0.05), which means that both groups' performance in intonation production were significantly improved after the training.

Table 4.10 Descriptive statistics for the students' posttest scores (production test)

Group	Number	Mean	SD	Range	Min.	Max.
EG	33	300.09	24.07	102	247	349
CG	33	275.52	26.92	106	223	329

Table 4.11 Comparisons between the students' pretest and posttest scores

Pair (pretest-posttest)	Mean differences	SD	95% Confidence Interval of the Difference		t	Sig. (2-tailed)
			Lower	Upper		
EG	-86.61	28.41	-96.68	-76.53	-17.51	0.00
CG	-65.52	26.64	-74.96	-56.07	-14.13	0.00

The students' detailed performance in the specific intonation patterns of the pretest and posttest were also compared by using paired-samples t tests. Results (Table 4.12) indicated that for both groups, their performances were significantly improved across all patterns except for *Statements (falling tone), Commands,* and *Exclamations*. The reason is that they had already achieved high scores for these three cases in the

pretest, resulting in the ceiling effect.

Table 4.12 Comparisons of the students' pretest and posttest scores in specific intonation patterns (production test)

Pair: pretest - posttest	Experimental group			Control group		
	Mean differences	t	Sig. (2-tailed)	Mean differences	t	Sig. (2-tailed)
Tonicity	-14.33	-10.97	0.00	-11.36	-8.28	0.00
Tonality	-14.52	-10.68	0.00	-11.58	-10.11	0.00
Statement (F)	-0.49	-1.15	0.26	-0.67	-1.01	0.32
Statement (R)	-4.21	-4.27	0.00	-2.55	-2.60	0.01
Wh-question (F)	-2.67	-2.56	0.02	-2.39	-2.7	0.01
Wh-question (R)	-6.21	-5.73	0.00	-4.46	-3.75	0.00
Yes/no-question (F)	-10.09	-6.98	0.00	-8.97	-7.24	0.00
Yes/no-question (R)	-4.15	-3.29	0.00	-3.18	-2.85	0.01
Tag-question (F)	-6.52	-5.57	0.00	-3.55	-3.20	0.00
Tag-question (R)	-3.58	-3.06	0.00	-3.42	-2.94	0.01
Command	-0.49	-1.97	0.06	-0.15	-0.65	0.52
Exclamation	-0.27	-1.06	0.30	-0.09	-0.39	0.70
Implication	-7.79	-8.35	0.00	-5.18	-5.90	0.00
Alternative question	-6.97	-5.82	0.00	-4.82	-3.64	0.00
Listing	-4.33	-4.02	0.00	-3.15	-4.44	0.00

Comparisons between the two groups' performances in the posttest were also conducted. Firstly, result of the independent-samples t-test showed that there was significant difference ($t=3.91$, $p=0.00<0.05$) between the EG's ($M=300.09$) and the CG's ($M=275.52$) mean total scores. That is to say, in terms of the total score, the EG

outperformed the CG in the posttest of the production test.

However, in terms of their detailed performances in specific intonation patterns, results of the independent-samples t-tests showed that there were no significant differences between the two groups' mean scores in *Tonality, Wh-questions (falling tone), Yes/no-questions (falling tone),* and *Tag-questions (rising tone)*. While the differences of the two groups' scores appeared in other patterns, *Tonicity*, *Statements (rising tone), Wh-questions (rising tone), Yes/no-questions (rising tone), tag-questions (falling tone), Implications, Alternative questions,* and *Listings,* reached significant level ($p<0.05$).

As the pronunciation rating was composed of two phases, it was necessary to examine the students' scores in each phase of the rating, so as to locate which phase of the rating caused the significant differences between the two groups' scores in the above mentioned eight intonation patterns. The first phase was the rating for intonation choice (whether the students used correct intonation) and the second phase involved the rating for the phonetic realizations of the chosen intonation (how well did the student pronunce the chosen intonation). Table 4.13 displays the results of the independent-samples t-tests for comparing the two groups' scores in the first phase of rating, and table 4.14 shows the results of the second phase of rating.

Table 4.13 Comparisons of the two groups' scores in intonation choice

Pattern	Exp. group (Mean)	Ctrl. group (Mean)	Independent-samples t test		
			Mean difference	t	Sig. (2-tailed)
Tonicity	25.30	24.24	1.06	1.15	0.26
Statement (R)	9.24	8.79	0.45	0.92	0.36
Wh-question (R)	7.88	7.58	0.30	0.49	0.63
Yes/no-question (R)	8.94	8.64	0.30	0.57	0.26
Tag-question (F)	8.03	7.42	0.61	0.98	0.33
Implication	5.76	5.30	0.46	0.69	0.49
Alternative question	8.33	7.73	0.61	1.00	0.32
Listing	6.97	6.36	0.61	1.04	0.30

Table 4.14 Comparisons of the two groups' scores in phonetic realization

Pattern	Exp. group (Mean)	Ctrl. group (Mean)	Independent-samples t test		
			Mean difference	t	Sig. (2-tailed)
Tonicity	24.24	19.15	2.88	3.32	0.00
Statement (R)	8.64	7.18	1.45	3.05	0.00
Wh-question (R)	7.30	5.94	1.36	2.24	0.03
Yes/no-question (R)	8.03	6.33	1.70	3.45	0.00
Tag-question (F)	7.27	5.33	1.94	3.35	0.00
Implication	5.06	3.24	1.82	3.74	0.00
Alternative question	7.42	5.52	1.91	3.52	0.00
Listing	6.18	4.33	1.85	3.37	0.00

As has been demonstrated above, for the eight intonation patterns shown in table 4.13 and table 4.14, there were significant differences between the two groups' performances. However, if one takes the scores separately, in terms of intonation choice (the first phase of rating), there was no significant difference between the two groups' performances across all of the eight patterns (Table 4.13); in terms of phonetic realization (the second phase of rating), there was a significant difference in each of the

8 patterns (Table 4.14). In other words, the EG and the CG showed equal performance in choosing intonation patterns, while in phonetic realizations of those patterns, the EG outperformed the CG. This means that it was the aspect of phonetic realization that caused the significant differences of the two group' performances.

To sum up the results of the intonation production test, there was no significant difference between the two groups' performances in the pretest. In the posttest, both groups' performances were significantly improved across all of the intonation patterns except for S*tatements (falling tone)*, *Commands*, and *Exclamations* due to a ceiling effect. In terms of the total score, the EG outperformed the CG; in terms of their scores in specific intonation patterns, the two groups showed no significant differences in *Tonality*, *Wh-questions (falling tone)*, *Yes/no-questions (falling tone)*, and *Tag-questions (rising tone)*, while for the other eight patterns, the differences reached significant levels. Comparisons of their scores in the two rating phases for the eight intonation patterns indicated that the differences did not lie in intonation choice but in the phonetic realizations of those patterns.

Here, conclusions can be drawn that both the native speakers' model (for the CG) and the students' modified speech as the model (for the EG) for intonation learning were equally effective for informing the students to choose correct intonation patterns for production. However, the latter was more effective than the former for enabling the students to produce more accurate phonetic realizations of intonation.

4.3 Students' learning journals

The students both in the EG and CG were required to keep learning journals during each training session. The learning journals were used by the students to record what they have learned, make reflections on the knowledge gained from the training, and write down the questions they encountered during the training. The students' learning journals were qualitatively analyzed by the researcher to monitor the students' learning process and access information about how well they learned, so as to examine whether or not the two types of pronunciation model had different influences on students' intonation learning.

Analysis of the students' learning journals showed that the students in both groups wrote their learning journals in great detail in the first training session. They took notes on what they have learned, gave deep and thorough reflections, and raised their questions during learning. In contrast, in the last training session, their journals were much simplified and the average length were about one third of the first entry. To some extent, this indicated that their motivation in learning attenuated from the first training session to the last one. This was also in accordance with the researcher's observations that most of the students were more devoted to the tasks in the former training sessions than the latter ones. However, the EG and the CG showed different degrees of motivation in different stages, as reflected by the learning journals and observations. It can be roughly estimated that the EG went to the languid period from the fourth training

session, while for the CG, the students' motivation for learning dropped abruptly from the second training session.

The learning journals also revealed that the students in both groups were able to extract most of the intonation rules from the exercises. For one thing, both types of pronunciation model were equally effective for helping students correctly recognize the intonation patterns. In the example below, the students who had produced the target utterance (the underlined sentence) with a falling tone in their first trials would immediately notice that the model speaker's intonation was different from theirs after they listened to the model. As one student wrote in the journal, "'*how much' here should be produced with a rising tone, not falling tone*" (CG 17). This means that by listening to the model, the students were able to perceive the differences between their intonation patterns and the model speaker's.

Example: Training session 1, part 3, No.4:

A: This new phone cost me 1,000 dollars.

B: <u>How much?</u> Are you kidding me?

For another thing, the students were also able to infer the meaning of intonation patterns by comparing the intonation contrasts in the examples or judging from the contexts. For the above example, as compared to their common knowledge of using falling tone to produce "How much", the students realized that using a rising tone would cause a change of the pragmatic meaning. Student EG 23 wrote in her journal "*According to speaker B's response 'Are you kidding me?', she intended to show*

surprise, so a rising tone should be used". Some students made further reflections and concluded that "*Wh-questions were normally produced with falling tones, but in special circumstances, like to show surprise or checking, it should be rising tones*" (EG 10), or "*for asking real questions, use falling tone; to show special emotions, use rising tone*" (CG 5). It was found that the EG wrote down more intonation rules and had more reflections in their journals than the CG. However, this was not reflected by their performance in the perception test as the two groups performed equal in the posttest.

A noteworthy difference between the two groups' journals was that the EG had more detailed descriptions of the pronunciation of the intonation patterns. Not only did they recognize the intonation patterns and decode the meaning, but also they described the skills of how to produce the intonation patterns, which was commonly found in their journals. For the above example, the students in the EG were able to recognize the rising tone pattern of "*How much*" in the pronunciation model and understand the meaning of the rising tone. Besides this, they also paid attention to the phonetic implementation of how to produce "How much" with a rising tone. For instance, one student noted that "*The stress should be placed on 'how', not 'much'*" (EG 10); four students mentioned that "*The time length of 'how' is longer than 'much'*" (EG 1, 6, 11, 29); several students described the starting point of the rising tone, i.e., "T*he rising should start from 'how' not 'much'*" (EG 3), "*The initial pitch should be lower so that the final pitch can climb higher*" (EG 24); a few students also described the movements

of articulators to better produce the target utterance, i.e., "*...opening mouth round and bigger on 'how'...sound nicer*" (EG 16), "*lower the tongue when pronounce the /ʌ/ in 'much'*" (EG 28).

In conclusion, the students' learning journals revealed that the students in both groups showed high motivation for learning at the beginning of the training, while their motivation declined with the proceedings of the training. The languid period of the CG was longer than that of the EG. For the former, their motivation dropped abruptly from the second training session; for the latter, their motivation was kept high till the fourth session. Their learning journals also showed that both groups were able to recognize the intonation patterns and decode the corresponding meaning by listening to the model. In the meantime, compared with the CG, the students in the EG paid more attention to the phonetic implementation of the intonation patterns, such as describing the stress placement, the starting point of tones, the variations of pitch, or even the movements of articulators. In other words, by listening to their modified speech as the model for learning, not only were they informed what intonation should be used but also they learned to how to pronounce the intonation in phonetic details.

4.4 Students' opinions of pitch modification for intonation learning

In order to investigate the students' opinions of pitch modification for intonation learning, a questionnaire survey and semi-structured interviews were offered to the students after the training. By using this two instruments, the researcher intended to

elicit information about the students' overall feelings towards the training, their evaluations on the effectiveness of pitch modification for intonation learning, as well as any problems they encountered during the training. The results are presented as follows.

4.4.1 Results from the questionnaire

The questionnaire (Appendix I) included 10 statements of which the former 5 statements were regarding to students' attitudes towards the training and the latter 5 were about their opinions of pitch modification for intonation learning. The participants were required to respond to each of the statement by ticking a 5-point Likert scale from "strongly agree" to "strongly disagree". The questionnaire survey was conducted to all of the participants (33 students) in the experimental group who used their own modified speech as the pronunciation model for intonation learning.

The results of the questionnaire were analyzed through Cronbach's Alpha Coefficient for reliability test, which indicated that the results were highly reliable (α=0.91). According to the the students' responses to the former 5 statements (Table 4.15), the majority of them showed a high degree of satisfaction (M=4.33) towards the training in terms of their overall feelings. They highly agreed that the contents of the training were what they wanted to learn (M=4.64), the courseware for the training was well-designed (M=4.33), and the training schedule was well-arranged (M=4.09). They also thought that this training was an enjoyable learning experience (M=4.45).

Table 4.15 Students' overall feelings towards the training

Statements	N	Mean	SD
1. I'm satisfied with this training.	33	4.33	0.54
2.The contents of this training are what I want to learn.	33	4.64	0.55
3.The courseware was well-designed.	33	4.33	0.74
4.The training schedule was well-arranged.	33	4.09	0.63
5.This is an enjoyable learning experience.	33	4.45	0.67

Results of the students' responses to the other 5 statements (Table 4.16) showed that the majority of the students highly agreed that pitch modification was effective for intonation learning. They thought that this kind of intonation learning effectively solved their individual specific pronunciation problems (M=4.73), enhanced their confidence in speaking English (M=4.36), and changed their cognition on pronunciation learning (M=4.12). The results also implied that most of the students held a positive attitude towards pitch modification for intonation learning. They thought that learning pronunciation by using their own modified productions was interesting (M=4.52), and they showed preference for their own modified speech as the model for learning than the native speaker's model (M=4.12).

Table 4.16 Students' opinions of pitch modification for pronunciation learning

Statements	N	Mean	SD
6. Pitch modification for pronunciation learning is effective for solving my own specific pronunciation problems.	33	4.73	0.45
7. Pitch modification for pronunciation learning enhanced my confidence of speaking English.	33	4.36	0.60
8. Pitch modification for pronunciation learning changed my cognition on pronunciation learning.	33	4.12	0.70
9. Pitch modification for pronunciation learning is interesting.	33	4.52	0.62
10. I prefer to use my modified speech for pronunciation learning than the native speaker's model.	33	4.12	0.82

4.4.2 Results from the semi-structured interviews

In order to obtain more detailed information about students' opinions of pitch modification for intonation learning, probe the reasons behind their choices in the questionnaire, and access information about their problems during the training, semi-structured interviews were conducted to 18 students in the experimental group immediately after the questionnaire survey. The interviews were guided by 7 open-ended questions (Appendix K) and the researcher also asked extra questions based on the participants' responses. Data obtained from the interviews were first coded, transcribed verbatim, and then analyzed qualitatively. Through a thematic analysis of the data, eight categories were developed from the students' responses, and two themes

were generated from these categories. The first theme and its categories as well as the evidence are presented in the following table:

Table 4.17 Theme and categories generated in the interviews (1)

Theme: overall feelings	Evidence (quotes from interviewees)
Category 1: interesting	EG 22: This was an enjoyable learning experience.
	EG 15: It's interesting to use my own voice for learning pronunciation.
	EG 7: I felt excited when I listened to my modified speech, especially in the first training session.
	EG 13: The intonation examples were vivid.
	EG 20: I'll never feel tired if I can learn pronunciation in this way.
Category 2: special	EG 2**:** I never thought that my own pronunciation can be my learning model**.**
	EG 25**:** I never received any kind of training on English intonation**.**
	EG 24**:** The courseware was concisely designed and easy to operate**.**
	EG 33**:** I always wanted to find a courseware for learning pronunciation that fits my particular needs**.**
Category 3: informative	EG 24**:** The training contents were what I wanted to learn**.**
	EG 5**:** This training gave me a chance to learn English intonation systematically
	EG 13**:** I didn't know how many intonation patterns are there in English, and I only knew falling tone and rising tone before this training**.**
	EG 20**:** Falling-rising tone was something that I never heard before, let alone its meaning**.**
	EG 11**:** The use of intonation can change the meaning of an utterance**.** This is what I never knew before
	EG 32**:** This training informed me of many of my pronunciation problems**.**

Table 4.17 Theme and categories generated in the interviews (1) (Continued)

Theme: overall feelings	Evidence (quotes from interviewees)
Category 4: relaxing	EG 7: The training task was simple.
	EG 8: This was a kind of autonomous learning.
	EG 24: I could learn on my own pace.
	EG 13: I could focus on my problems and I didn't need to waste time on things that I had already known.
	EG 27: The training was relax while efficient.

The evidence under the first theme indicated that the interviewees held positive attitudes towards the English intonation training. They thought that the training was interesting and the learning environment was student-friendly. Furthermore, they thought that they learned something from this training, and the knowledge was what they wanted to learn. The second theme was related to the students' opinions of the effectiveness of pitch modification for intonation learning. The details were shown in table 4.18.

Table 4.18 Theme and categories generated in the interviews (2)

Theme: opinions of the effectiveness	Evidence (quotes from interviewees)
Category 1: pronunciation improvement	EG 8: My pronunciation sounds nicer than before since I paid attention to intonation.
	EG 13: My Chinese accent was reduced after this training.
	EG 7: Intonation makes my speech more fluent.
	EG 13: I learned how to chunk when speaking.
	EG 22: I learned how to use intonation to signal particular meaning.

Table 4.18 Theme and categories generated in the interviews (2) (Continued)

Theme: opinions of the effectiveness	Evidence (quotes from interviewees)
Category 2: effective for solving individual specific problems	EG 11: This kind of learning gave me a chance to receive feedback on my own pronunciation.
	EG 32: It informed me where my problems locate.
	EG 15: Using my own voice for learning was more effective, because it targeted my problems.
	EG 14: It easier for me to perceive the differences.
Category 3: cognition change in pronunciation learning	EG 15: Now I think intonation is more important than vowels or consonants when learning pronunciation.
	EG 5: Intonation is a bigger picture for pronunciation.
	EG 13: Pronunciation learning is not about how long time I spend on practising, but about how effectively my problems can be solved.
	EG 25: Intonation is also very important for conveying meaning.
	EG 11: Intonation is not only decorative.
	EG 30: I gained more confidence in speaking.
Category 4: pronunciation confidence enhancement	EG 20: This kind of learning provided me a more realistic pronunciation goal.
	EG 8: Using my own voice as model made it easier for me to imitate.
	EG 25: I don't need to speak like a native speaker, as long as I know how I can be correct.
	EG 14: I'm not so nervous as before when speaking.
	EG 27: I can speak freely with less Chinese accent.

The evidence under the second theme indicated that the interviewees also held positive attitudes towards the effectiveness of pitch modification for intonation learning. The majority of them (91%) thought that their pronunciation was improved through this training, because this kind of intonation learning was effective for solving

their individual specific pronunciation problems. Pitch modification could also enhance their confidence of pronunciation and change their opinions of pronunciation learning. Moreover, most of the interviewees (87%) showed their preference for using their modified speech as the model for pronunciation learning as compared to native-speaker model.

However, some problems about the training were also reported in the interviews. Firstly, some students recommended that it will be better if the native speaker's model can also be provided along with their own modified speech. In other words, they still wanted to listen to the native speaker's model as a reference, although there were no problems about their modified model. Secondly, few students questioned about the meaning of some intonation patterns, such as the implication of falling-rising tone and the pragmatic function of "definite fall". There was no wonder that they might raise these questions considering that intonation has seldom been taught in their English learning experience and some intonation patterns and their meanings were relatively novel to them. Thirdly, some students also proposed some recommendations for pronunciation instruction. For example, "*More about intonation should be taught in class*" (EG8, 25); "*Teachers should guide students to use technology or more software for pronunciation learning*" (EG5,13, 22, 25, 27); "*We want to learn something more practical*" (EG22, 33); "*Teachers should teach us more techniques or tips on how to practice pronunciation*" (EG15); "*The pronunciation model provided in class should*

be more representative, or more various" (EG24, 33).

From the above analysis, it can be seen that the results of the semi-structured interviews were consistent with the results of the questionnaire. The two sources of data revealed that the students were satisfied with the training to a large extent, including its contents, training tasks, schedule, as well as courseware. They highly agreed that pitch modification for intonation learning was effective and efficient for improving their pronunciation. In addition, the students also thought that learning pronunciation by using their own voices was interesting and they preferred to use their modified speech as the model for learning rather than the native speaker's model.

4.5 Answers to research questions

This research posed two research questions:1) Is pitch modification effective for English intonation learning? If yes, in what ways? 2) What are the students' opinions of pitch modification for intonation learning? Results yielded from the intonation perception test, intonation production test, students' learning journals, questionnaire survey, and semi-structured interviews were collaborated to answer the two research questions. Detailed answers to each research question are as follows.

4.5.1 Answers to research question one

The results of the experiment showed that pitch modification was effective for English intonation learning. Firstly, student's own modified speech as the model for

intonation learning was as effective as the native speaker's model for improving the students' ability of perceiving and understanding the meaning of English intonation. Both of the two types of model could help students recognize the intonation patterns and understand the meaning of those patterns.

Secondly, pitch modification for intonation learning was effective for improving students' ability of pronouncing English intonation. Compared with the native speaker's model for intonation learning, pitch modification was more effective for enhancing students' pronunciation in eight intonation patterns: *Statements (rising tone), Wh-questions (rising tone), Yes/no-questions (rising tone), tag-questions (falling tone), Implications, Alternative questions,* and *Listings*. Listening to the native speaker's model could help students to form correct phonological representations of the intonation patterns, just as pitch modification did, however, the former was not so effective as the latter for facilitating the students' phonetic realizations of those intonation patterns.

Thirdly, pitch modification for intonation learning was effective for solving the students' individual specific problems in pronunciation. This kind of pronunciation learning targeted at students' mispronunciations and provided corrective feedback to the students with their own voices, the voices with which they are most familiar. In this way, the students could directly perceive the differences between their correct pronunciations and the incorrect ones. In other words, it facilitated students to precisely

locate their problems and how they can be corrected. Therefore, pitch modification for intonation learning provided tailored instruction to each student targeting on their own individual problems with the least time consumption.

Fourthly, pitch modification for intonation learning can enhance students' confidence in speaking English. By modifying students' mispronunciations to the correct ones, it informed students where their problems were and to what extent their pronunciations could be accepted. Students could listen to and imitate the model of their own voices, better facilitating the link between the speech signals with their own articulators, i.e., the sensory-motor mapping. In this sense, pitch modification for intonation learning provided a more realistic pronunciation goal for students to pursue, as compared to other models (e.g., native-speaker model, teacher's model) which focused on how the model was correct. This had the potential of reducing students' anxiety of speaking English as long as they realized where their problems were and how they could be corrected, and thus increase their confidence when speaking.

4.5.2 Answers to research question two

Results of the questionnaire survey and semi-structured interviews indicated that the students held positive attitudes towards pitch modification for intonation learning. Firstly, they thought that this kind of learning was interesting and the intonation training was an enjoyable learning experience. Most of the students felt excited when listening to their own modified productions as they had never been

exposed to their modified speech: they had never thought that their own voices could be manipulated and used as the model for pronunciation learning. They showed interest in their modified speech, and their motivation for learning was greatly enhanced especially in the early training sessions.

Secondly, the students showed their agreements on the effectiveness of pitch modification for intonation learning. They thought that this kind of learning was tailored to their needs and could effectively solve their individual specific problems in pronunciation. They also reported that the training changed their cognition in pronunciation learning, and once they paid attention to intonation, they felt that their overall pronunciation was significantly improved, which increased their confidence of speaking English. Compared to the native speaker's model, they preferred to their own modified productions as the model for pronunciation learning. In addition, some students showed that they also wanted to listen to the native speaker's model along with their modified model, so that they could have another reference to compare. Perhaps the best way is to employ a combination of the two types of model for pronunciation instruction.

4.6 Summary

This chapter presented the results of the present study. Firstly, it showed the results of the intonation perception test and found that there were no significant differences

between the experimental group's and the control group's performance, either in the pretest or the posttest; secondly, it presented the results of the production test and found that the two groups had equal performance in the pretest, while in the posttest, the experimental group outperformed the control group in eight intonation patterns; thirdly, it detailed the results of the questionnaire and the semi-structured interviews and presented the students' opinions of pitch modification for intonation learning; lastly, it summarized the results to answer the two research questions of this study.

CHAPTER 5

DISCUSSION

This chapter presents a discussion on the findings of this study. It first indicates the reasons behind the students' performance in the tests; then, considering the students' performance in the perception test and production test, it attempts to disentangle speech perception and production. Next, based on the relationship between perception and production, the effectiveness of pitch modification for intonation learning is elaborated. Lastly, by reviewing the pedagogic procedures of this study, a precision approach for pronunciation instruction is proposed.

5.1 Chinese students' problems in English intonation

There is no doubt that Chinese EFL students have various problems in English pronunciation, and their problems in suprasegmental aspects are more salient. Although previous studies claimed that Chinese university EFL students' overall pronunciation was intelligible to a large extent (Zhang, 2015), we should find ways to deal with the problems that hamper the intelligibility of their pronunciation before we are prematurely positive about their performance, especially in the case of those students who are pre-service English teachers, good pronunciation as well as good command of

the knowledge of pronunciation is an essential need for their future careers.

In the present study, the students' problems in English intonation were investigated through an intonation perception test and production test. The results of the pretest revealed that their overall performance were far from satisfactory, with their mean scores approximating 60% of the total score for the perception test and 50% for the production test. However, in terms of specific intonation patterns, their performance in *Statements (falling tone)*, *Commands* and *Exclamations* were satisfactory. This was in accordance with the results of previous studies (Hong, 2012; Jiang, 2012; Huo & Luo, 2017; Rui, 2007) which found that Chinese EFL students were good at producing falling tones, and Wells' (2006) claim that exclamations are the simplest kind of utterance for EFL students. The reason might be that falling tone is the most frequently used intonation patterns across all languages, and the default tone for these three sentence types is a fall.

For other intonation patterns, some were significantly improved after the training and some were more resistant to change. Generally speaking, the students' problems in intonation can be classified into two types: problems caused by phonological representation and problems caused by phonetic implementation. The first type is cognitive, i.e., the students lacked the intonation knowledge to form the correct phonological representations to perceive or understand the meaning of certain intonation patterns, thus failed to produce them correctly. The second type involves

how well did the students produce the intonation patterns in phonetic details.

5.1.1 Problems caused by phonological representation

The results of the perception test and production test showed that the students' problems in *Tonality*, *wh-questions (falling tone)*, *Yes/no-questions (falling tone)*, *Tag-questions (rising tone)* were mainly caused by phonological representation. The results revealed that Chinese EFL students had very poor knowledge of English intonation. Most of them did not realize that English intonation belongs to phonology and changing the intonation of an utterance could change the meaning. Specifically, their cognitive errors of intonation were manifested in the following aspects.

1) Choosing intonation arbitrarily or according to their first language intonation. Quite a few students had little knowledge about English intonation and paid no attention to intonation when speaking. They had no idea that a falling-rising tone has the meaning of implication, let alone the function of open and close contour in discourse regulation. As a result, they tended to transfer the intonation of Chinese language to produce English intonation. Although there are similarities between Chinese and English intonation system, the differences are more salient (Zhu, 2007). English is a non-tonal language while Chinese is a tonal language. Chinese uses tone at lexical level to signify words and at the same time uses intonation at sentence level, however, the tones used at sentence level are restricted by the tones used at lexical level in order to keep the basic meaning of each word in a sentence (Zhang, 2000). Therefore,

the pitch at sentence level of Chinese language can not be as various as non-tonal languages do, like English, and hence the Chinese-accented English always sounds flat or monotonous (Makarova & Zhou, 2006; Huo & Luo, 2017).

In the present study, acoustic analysis of the Chinese students' speech and the native speaker's speech showed that the English spoken by Chinese students was featured by less pitch variations and narrower pitch range. The phenomenon of overusing falling tones was very common among them, such as using falling tones to produce a normal Yes/no-question or a declarative question. In addition, the rhythm pattern of Chinese is syllable-timed, while for English, it is stress-timed (Zhu, 2007； Huo & Luo, 2017; Makarova & Zhou, 2006; Zhang, 2000). For Chinese, each word contains only one syllable and each syllable in a sentence takes approximately equal amounts of time to pronounce; for English, one word may contain several syllables and the stressed syllables in a sentence takes approximately equal amounts of time to pronounce. This leads to the students' problems in tonicity and tonality. For tonicity, they neglected the placement of nuclei and tended to stress every word in a sentence; for tonality, they over-relied on pauses to signal intonation boundaries, making their speech contain more chunks than necessary.

2) Using default tones inflexibly. Considering that intonation had little room in Chinese EFL pronunciation class, the limited things that have been taught were the default tones. Even in some pronunciation textbooks, the default tones are rigidly linked

to their corresponding sentence types. Most Chinese students took it for granted that a question should be produced with a rising tone and a statement with a falling tone. They were not aware that default tone is an "unmarked or neutral tone that is used under no special circumstances" (Wells, 2006, p.15) and intonation choices should be based on the speaker's intentions. As a result, most of the students in the present study tended to make incorrect intonation choices in cases where a sentence type should not be produced with its default tone, such as in *Statements (rising tone)*, *Wh-questions (rising tone)*, *Yes/no-questions (falling tone)*, and *Tag-questions (falling tone)*. As Jiang (2012) claimed, the way that Chinese students learn intonation was very rigid and they applied the rules of intonation without considering the contexts.

3) Linking punctuation with intonation. Intonation in speech sometimes can function like punctuation in written text (Nolan, 2008), however, punctuation does not always correspond to intonation (Well, 2006). For example, a sentence ending with an exclamation mark always signals an exclamation, which should be produced with a falling tone, and a comma in a sentence always signals a intonation boundary. These are positive links between punctuation and intonation. Negative links could be linking a question mark with rising tone, or the lack of divisions of intonation groups if there are no commas in a sentence. In the production test of the present study, several students chose rising tones to produce sentences ending with question marks without considering the contexts, such as in *Wh-questions (falling tone)*, *Yes/no-questions*

(falling tone), and *Tag-questions (falling tone)*. In tonality, most of the students viewed one sentence as one intonation group if there were no comma to signify the intonation boundary.

4) Choosing intonation according to intuition. In the semi-structured interviews, some students reported that their intonation choices largely depended on intuition or the language sense developed from reading. However, the intuitions or sense might not be so reliable as rules. Furthermore, the intonation used in reading is different from that in conversations, so their sense to intonation developed from reading might not properly fit the situations in conversations. The students' transfer of the intonation from reading to conversations further made their speech sound monotonous, flat and sometimes even strange.

In conclusion, Chinese EFL university students had very limited knowledge of English intonation, and their limited knowledge led to their lack of the awareness to form the correct phonological representation of the intonation patterns. Holding the erroneous cognition of intonation, some students negatively transferred the intonation of their first language to English; some resorted to default tones; some incorrectly linked punctuation with intonation; and some applied the intonation of reading to conversations. In a few words, the students' problems in cognition of intonation caused various problems in their intonation production. Therefore, it is important to teach students the knowledge of intonation before addressing their production ability,

because “it seems problematic to say that they have actually learned anything about intonation other than the ability to mimic it” (Reed & Michaud, 2015, p.461).

5.1.2 Problems caused by phonetic implementation

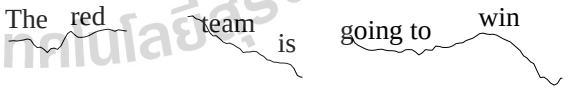
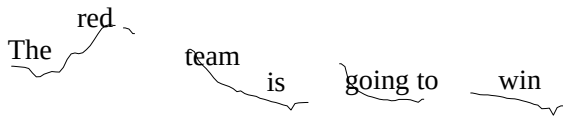
Intonation consists of a phonological and phonetic component (Pierrehumbert, 1980). Language learners’ phonological problems of L2 intonation may result from the intonational differences in the inventory of L1 and L2 phonological patterns (Mennen, 2007). For example, English speakers use rising tone (open contour) to signal nonfinality of the speech, while Chinese EFL learners would take it as a question. The phonetic problems of L2 intonation result from the differences between L1 and L2 in the phonetic realizations of the same phonological patterns, i.e., the same phonological category may be realized differently in different languages.

Mennen (2007) claimed that L2 intonation learning involves two stages. L2 learners may first acquire the phonological patterns before they acquire the correct phonetic implementation of these patterns. His former studies (1999, 2004) found that Dutch learners of Greek were perfectly able to produce the correct phonological tonal elements but implemented these structures by using L1 phonetic regularities. This means that learners may form the correct phonological patterns of L2 intonation, but fail to produce the phonetic details. Simply speaking, they know what intonation patterns should be chosen, but they do not know how to pronounce these patterns accurately. Phonetic realization was usually evaluated from the aspect of acoustic

properties, including the spectral features: the F0 value; and the temporal features: alignment of the segmental string. The Chinese students' problems in phonetic implementation of English intonation in the present study were manifested in the following aspects.

1) Problems in tonicity. The rhythm patterns of the English spoken by Chinese students reflect the typical features of syllable-timed language (Bi &Chen, 2013). As was shown in the present study, many students accented every word in a sentence, making their speech lack of prominence contrasts and strongly Chinese-accented. Even in the posttest of the production test, quite a few students still assigned more word prominence than necessary within one intonation group. Table 5.1 shows the pitch contour of the speech produced by a Chinese student and a native speaker in response to the question "Which team is going to win?".

Table 5.1 Example of student's problems in tonicity

Student's production	
Native speaker's production	

As can be known from the context, the nucleus of the target utterance should be placed on "red", indicating the new information, as the rest words constitute the background information. The native speaker's pitch contour shows that the sentence

stress (nucleus) was prominently placed on “red”, forming clear prominence contrast with other unstressed syllables, and the tail gradually falls to the end. For the student’s speech, it seems that he has noticed that the word “red” should be stressed, however when implementing the accent he placed a high prominence on “team. Then, following the nucleus, the tail went down then went up again on “win”, making the pitch prominence of “win” approximate that of “red”. It means that he also unconsciously accented the last word “win”. It is a very common phenomenon that Chinese EFL students tend to accent the last word of a sentence or phrase. This might due to that Chinese intonation system usually puts the nucleus at the sentence final (Pan, 2012).

The problem of accenting the last word even caused syllable shifting of the accented word. For example, in response to the question “When did Jack go to China?”, one student’s production of the target utterance “Jack went to China yesterday” accented the word “yesterday”, however the student misplaced the stress on “day”, making it [jestərˈdeɪ], which sounds rather strange. In addition, in strategies of pitch accent realizations, Chinese students tended to rely on pitch range amplification, featured by a sharper pitch contrast, stronger intensity, and bigger sound volume. In comparison, the native speaker’s strategies were more diverse. Except for using high falls, the native speaker preferred to use pitch variations or longer duration to signal the nucleus.

2) Problems in tonality. As has been discussed above, the students’

problems in tonality were mainly caused by cognitive errors. Before the training, most of the students were not aware of the function of tonality in distinguishing the alternative meaning of ambiguous structures. In the pretest, the students' productions in tonality went to two extremes: one was that some students' (48%) productions contained too much pauses which broke the speech into pieces and made it dysfluent or even hard to understand (Table 5.2, 1); the other extreme is that some students (35%) failed to mark the boundaries between intonation groups, leaving the speech ambiguous in meaning (Table 5.2, 2).

Table 5.2 Example of student's problems in tonality (1)

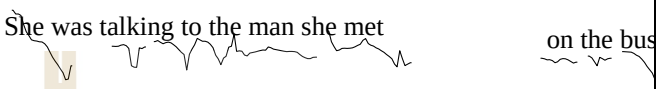
1	She was talking to the man she met on the bus
2	She was talking to the man she met on the bus

Note: Both of the above two productions were ambiguous in meaning as they failed to signal whether "She met the man on the bus" or "she and the man were talking on the bus".

Only about 17% of the students successfully employed tonality to disambiguate the target utterances. But this number rose significantly in the production test, as the students' awareness of tonality has been aroused after the training. However, there were still some problems in terms of implementing tonality, especially for the control group. They marked intonation boundaries by over-relying on pauses, and sometimes even exaggerated the pauses, making their speech unnatural and dysfluent.

In comparison, the native speaker tended to employ other strategies to mark the boundary except for pause, such as pitch reset or pitch variation. See the following example:

Table 5.3 Example of student's problems in tonality (2)

Student's production	
Native speaker's production	

Both of the two productions marked a boundary between the two intonation groups. As can be seen, the student used a long lengthened pause to mark the boundary, while the native speaker resorted to pitch reset to form a contrast between the two intonation groups. Although "the criterion most often mentioned in the demarcation of intonation-groups is that of pause" (Cruttenden, 1997, p.30), overusing or improperly using pauses would cause problems to the conveying of information. Chen (2006) claimed that too many pauses would chop up the speech, break down the fluency, and hinder the comprehensibility of the speech. In contrast, employing the strategy of pitch reset will make the speech more fluent and contain more pitch variations. Xue's (2016) study found that both native speakers and Chinese EFL learners significantly used variety of pitch range to judge foreign accent. Therefore, using pitch reset for marking intonation boundaries has the advantage of reducing foreign accent and improving the comprehensibility of the speech.

3) Problems in falling tone. Chinese students (even all EFL/ESL learners) are good at producing falling tones. This might because that falling tone is the most frequently used tone pattern across all human languages. However, in the present study, in phonetic realizations of falling tones, there were still some minor differences between the native speaker's speech and the students'. For example, several students added unnecessary pitch variations in their speech. As Table 5.4 shows, comparing to the native speaker's speech, the student's production used a falling-rising tone on "people" and a rising-falling + rising tone on "creative", making the overall falling tone sound rather confusing and strange. In the interviews, some students reported that they just produced it unconsciously, and some students thought that their speech would sound better if they add more pitch variations.

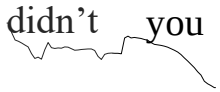
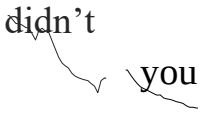
Table 5.4 Example of student's problems in falling tone (1)

Student's production	Most left handed people are creative
Native speaker's production	Most left handed people are creative

Another problem in falling tone involves the realization of falling tones in tag-questions. For example (Table 5.5), in pronouncing the falling tone of "didn't you", the native speaker's production used a high fall starting from the beginning of "didn't" and then the pitch drops sharply to the end. In comparison, although the student's

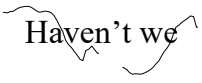
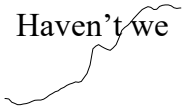
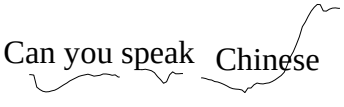
production was also a falling tone, he put the starting point of the fall on “you”, making his speech sound like stressing “you”, which is unnecessary. This also might be caused by Chinese students’ pronunciation habit of stressing the last word, which shifts the starting points of falling tones. Furthermore, if a target utterance was long, Chinese students would pronounce the former elements with a long flat intonation and put a fall on the tail of the utterance.

Table 5.5 Example of student’s problems in falling tone (2)

Student’s production	
Native speaker’s production	

4) Problems in rising tone. Comparing with falling tone, rising tone is relatively more difficult for Chinese EFL students to produce. In pronouncing rising tones, they showed two major problems. Firstly, most students misplaced the starting point of a rising. Mirroring their implementation of falling tones, they tended to start the rising from the last word/syllable of a sentence. See the following two examples:

Table 5.6 Example of student’s problems in rising tone

Student’s production		
Native speaker’s production		

The other problem was that some students set too high pitch at the beginning of the utterance, which made the pitch of the tail can not climb higher. In the following example, the student produced “can” with a high pitch, which was so high that the tail “Chinese” could not be higher. But she was still aware that the overall pitch contour of this sentence should be rising, so she turned to tone down the pitch at the beginning of “Chinese” and raised up the pitch at the last syllable to make it sound like a rising.

5) Problems in falling-rising tone. Among all of the intonation patterns, falling-rising tone is the most difficult for Chinese EFL students to produce. For one thing, the majority of the students have never learned this tone pattern, and thus they had no idea about the meaning of falling-rising tones. In the pretest, less than 30% of the students correctly perceived it or understood its meaning. For another thing, many students still could not correctly produce falling-rising tones even though they could correctly perceive them, as the production posttest showed that only about half of the students correctly produced this tone pattern.

Considering that falling-rising tone has no corresponding tone pattern in Chinese language, it might take longer time for Chinese students to form correct phonological representation of this pattern. Adding that falling-rising tone has been rarely taught, Chinese students might find this tone pattern was difficult to pronounce, or even to imitate. Intriguingly, most of the students who can not correctly produce

falling-rising tone tended to replace it with rising tones (Table 5.7). Or, they initially perceived it as a rising tone.

Table 5.7 Example of student's problems in rising tone

Student's production	she works very hard
Native speaker's production	she works very hard

6) Problems in compound tones. Compound tones are tone patterns that are composed of several single tone patterns, such as the tone pattern for alternative questions is composed of a rising tone for the first alternative and a falling tone for the second. In cognitive level, compound tones were relatively easier for students to command. They just need to keep in mind the rules of the composition of the tones. However, in production level, they showed many problems. The possible reason might be that the utterances for compound tones were relatively long and contained more than one intonation groups. See the following example:

I like apples oranges and pears

The student noticed that the intonation pattern for this utterance should be "rising + rising + falling", but in implementing this pattern, he seemed to forget the rising when the speech ran to "oranges", and went back to his "Chinese way" of

accenting the word. Therefore, it can be inferred that utterances with longer duration and more pitch variations would hinder students' phonological working memory and caused more problems in production.

In conclusion, the students' ability in perceiving English intonation does not guarantee their ability in producing it. In other words, even if they know it, they might not correctly pronounce it. In the present study, Chinese students showed various problems in producing English intonation. Influenced by the syllable-timed rhythm pattern of Chinese language, Chinese students had problems in placing the nuclei of intonation groups and tended to accent the last syllables of utterances. The results were: for tonality, the information focus were missed; for tonality, too much pauses broke the sense group; for rising tone, they misplaced the starting point of the rising. In addition, falling-rising tone and compound tones were found more difficult for them to acquire due to its novelty or complexity. All of these problems hindered Chinese students' production of English intonation and contributed to their Chinese accent.

5.2 Disentangling speech perception and production

The results of the experiment showed that the Chinese university EFL students had no problems in producing *Statements (falling tone)*, *Commands*, and *Exclamations*. Their problems in *Tonality*, *Wh-questions (falling tone)*, *Yes/no-questions (falling tone)*, and *Tag-questions (rising tone)* were mainly caused by phonological representation.

For the other intonation patterns, their problems lied in phonetic realization. The students' problems can be explained from the interplay of speech perception and production as well as the factor mediating the two modalities.

The relationship between perception and production has long been investigated in the field of second language phonological acquisition. In the segmental domain, many previous studies have shown that perception training on segments can facilitate the development of production of the segments, among which the most representative was a series of studies related to training Japanese learners to identify English /r/ and /l/ (e.g., Bradlow et al., 1997). In the suprasegmental domain, De Bot & Mailfert's (1982) study found that training in the perception of intonation resulted in a statistically significant improvement in the production of English intonation patterns. Wang et al's (2003) study revealed that after perception training, American learners' productions of Mandarin tones were significantly improved without any production training, and therefore "the facilitatory effect of perception training on production not only occurs for segmental learning, but also extends to suprasegmental learning" (p.1041).

However, though previous researchers agreed that there is a close relationship between perception and production, this relationship is still complex (Leather,1983). Hardison's (2012) study found that the facilitate effect of perception training was individual learner and specific phoneme dependent. Bradlow et al's (1997) study also indicated that the improvement in perception and production did not proceeded in

parallel. Hattori & Iverson (2010) claimed that perception and production processes and representations may be somewhat autonomous. Nevertheless, it is widely accepted among most researchers that production accuracy is constrained by perception accuracy. For example, Flege (1999) argued that the accuracy with which L2 segments are perceived limits how accurately they can be produced; De Bot & Mailfert (1982) claimed that the production of intonation depends heavily on the perception of intonation; Leather (1983) proposed that perception of a contrast is not just a necessary, but even a sufficient condition for adequate production.

In the present study, the relationship between the students' performances in the perception test and production test were examined by using Pearson correlation analysis. The results (Table 5.8) showed that in the pretest, the correlation for the experimental group did not reach a significant level ($r=0.33$, $p=0.07$), while for the control group, it showed significant positive correlation. In the posttest, both groups' perception and production scores showed significant positive correlation. Thus, the present study does not assert that perception and production is directly connected. Instead, it would rather go with the claim that L2 learners' ability in perception and production are not parallel. However, given the positive relationship between perception and production, the results of this study were in consistent with Lee and Lyster's (2017) finding that improvement in perception accuracy was a significant predictor of improvement in production accuracy. To disentangle perception and

production, it is critical to locate the mediating factor between them and reconsider the process of second language phonological acquisition.

Table 5.8 Correlation between students' perception and production performance

	Experimental group		Control group	
Correlation between perception & production	Pretest	Posttest	Pretest	Posttest
N	33	33	33	33
Pearson correlation	0.33	0.41*	0.43*	0.64*
Sig.	0.07	0.02	0.01	0.00

*. Correlation is significant at the 0.05 level (2-tailed).

Phonological acquisition involves the first step of the establishment of phonetic categories. The perception-first view (Lee & Lyster, 2017) has been supported by most of the influential speech acquisition theories, such as Speech Learning Model (SLM, Flege, 1995) and Perceptual Assimilation Model (PAM, Best, 1995). Both theories predict that the levels of difficulty in forming L2 categories depend on how L2 learners get rid of the restrictions from their L1 categories. As long as L1 categories subsume L2 categories, accurate perception (and hence production) of the target categories will be blocked. According to the category prototype concept (Hardison, 2012), a category is regarded as an aggregate of abstract knowledge derived from different exemplars. The detailed relevant information of an event provided by individual exemplars may be forgotten overtime, while the shared features of multiple traces (the prototype) will be retained long.

Thus, phonemic categorization is to map acoustically highly variable speech

sounds (exemplars) onto less variable motor representation (Möttönen & Watkins, 2009). Here motor representation refers to the prototype. For other researchers, it was called motor plan, desired sensory outcome, emulator, etc.. The present study used the phrase “phonological representation”. To put the process of phonemic categorization simple, a learner may be exposed to different realizations of a L2 category as input stimuli, and those stimuli share the same features of the prototype though with subtle differences. By perceiving different realizations of the same category, the learner can extract the common features shared by those stimuli and form the phonological representation of the category. Then, the learner’s production of that sound will be based on the representation and contain the learner’s own characteristics.

Based on this, it can be inferred that stimuli with high variability have the potential of facilitating the process of phonemic categorization. Many previous studies have supported this. For example, training techniques by using multitalkers (stimuli produced by different speakers, Lively et al, 1993; Wang et al, 2003), audiovisual stimuli (Hardison, 2003; Hazan et al, 2005), and modified stimuli (Jamieson & Morosan, 1989; Mcclelland et al, 2002) have yielded positive results. However, it should be noted that these studies were carried out in laboratory settings and the above described process of categorization is an ideal one. It is also noteworthy that phonemic categorization is a dynamic process and phonological representation is subject to adjustment/tuning, i.e., the learner will continuously adjust the representation as he/her

listens to new stimuli. The adjustment process is known as pruning or de-conditioning (Leather, 1983): continuously reducing learners' perceptual reliance on the established sound system of the mother tongue (de-conditioning).

Outside the laboratory setting, many factors may block the process of pruning, such as inadequate input, incorrect model from teachers or peers, and the learners' incorrect perception. These factors may hinder the learners to perceive the critical acoustic cues of the speech signal to form the correct phonological representation, as the development of L2 perceptual categories is dependent on the perceived differences between an L2 sound and the learner's closest L1 sound (Flege, 1995). If a learner is not aware of the critical issue, the perceived differences, the phonological categorization will go into a "plateau period" or fossilization. The results might be: 1) the learner can not correctly perceive the sound, and neither can he/she correctly produce it; 2) the learner can correctly perceive the sound but can not correctly produce it.

In a word, the key factor of L2 phonological acquisition lies in whether the input speech signal can arouse the learner's awareness of the perceived differences that lead to the adjustment of the representations. Adequate adjustment will lead to the establishment of new categorizations and the redundancy of input, while inadequate adjustment will result in perception and/or production failure (Figure 5.1). This might account for why in the present study some students could correctly perceive some

intonation patterns but could not produce them correctly.

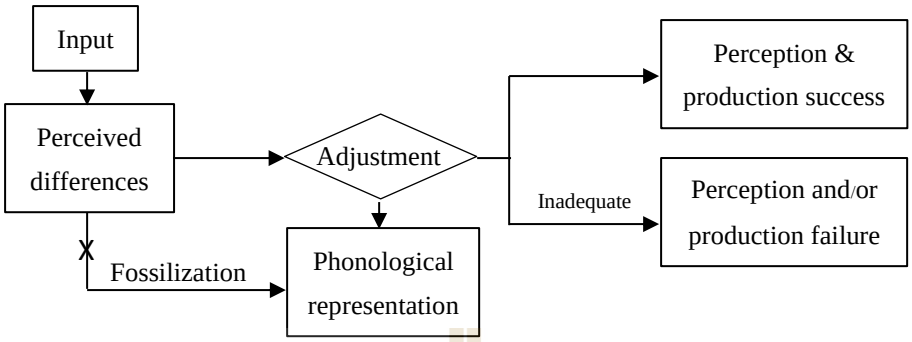


Figure 5.1 The process of phonological acquisition

5.3 Pitch modification as cue enhancement for intonation learning

The relationship between perception and production reveals that "well-formed phonological representation at the perception level is a *sine qua non* for target-like sensory motor skills and accurate L2 speech production" (Lee & Lyster, 2017, p.372). The motor theory of speech perception proposed that there is a specialized phonetic module that represents speech units in terms of the speaker's intended gestures, and this module mediates perception and production (Liberman & Mattingly, 1985). Thus, improvement in speech production as a consequence of perception training is due to a reorganization of the phonetic module, i.e., sensory-motor remapping. In other words, perceptual experience can guide sensory-motor learning (Kuhl, 2000).

The motor theory views a particular acoustic event as the product of a single vocal tract configuration or neuromuscular program. A learner attempting to reproduce a target sound must adjust the configurations and movements of his or her speech organs

according to the intended gestures that the speech signal carries. Learning in production involves a mechanism by which articulatory commands are tuned to internal acoustic representations (Bradlow et al., 1997). Leather (1983) claimed that the accuracy with which a learner can match the sound with the target articulatory command is limited by sensory information feedback. This feedback is determined by the learner's awareness of the perceived differences between the input speech signal and the auditory-perceptual target s/he has adopted. In this vein, whether the learner's awareness can be aroused depends on whether the input speech signal can attract his/her attention on the critical acoustic cues that lead to the perceived differences. For example, Japanese learners have difficulty in learning English /r/ and /l/ because they are overly sensitive to acoustic cues that are not reliable for /r/-/l/ categorization (Iverson et al, 2003).

Techniques employed to arouse students' awareness on a specific acoustic event are known as acoustic cue enhancement. Cue enhancement is a kind of input enhancement which aims at enhancing the salience of target forms by "exposing learners to comprehensible input and positive evidence to draw their attention to specific linguistic properties of the target language" (Benati, 2016, p.68). In the field of speech acquisition, efforts have long been paid to cue enhancement by manipulating the input speech signals. Verbotonalism, the theory underlying Guberina's SUVAG system of audio filtering, claims that if the learner could be facilitated to perceive the salient features of the L2 sound, the establishment of accurate production targets would

be sped up (see Leather, 1983).

Pitch modification as an acoustic cue enhancement technique in the present study, attempted to provide the students corrective speech feedback by modifying their incorrect intonation to the correct one and playing both of the two versions back to them. The direct comparison of both versions can help the learner realize faster the differences between their incorrect intonation and the target desired one. Furthermore, Using students' own voices (self-perceived voices), the voices with which they are most familiar, could enable them to pay attention to the critical acoustic cues "without the distractions from less relevant factors such as voice characteristics" (Tang et al.,2001, p.3). Brown (1999) pointed out that native speakers' model of the same category are continually faced with variable realizations influenced by factors like coarticulation, sloppy articulation or interspeaker variability. Those variations do not contribute to differences in meaning while distracting learners' attention to the critical cues. If this irrelevant 'noise' in the acoustic signal can be filtered out, the memory load put on the auditory system will be greatly reduced and processing can proceed more quickly.

In a word, pitch modification as cue enhancement in the present study facilitated the students' critical listening, while at the same time preserved the students' own acoustic characteristics, which made it easier for reproduction. Specifically, the effectiveness of pitch modification lied in three major aspects: 1) helping the students to deaccentuate unnecessary information focuses; 2) informing the students more

strategies for chunking and increased their speech fluency; 3) enabling the students to retune their phonological representations of tone patterns and facilitate the students' more accurate phonetic realization of these patterns.

Deaccentuation

As has been discussed above, influenced by the syllable-timed rhythm pattern of Chinese language, Chinese students tended to accent every word in a sentence, making their speech lack of prominence contrasts. It was relatively easier for the students to notice the information focus of an utterance, provided that they devoted attention to the contexts. In this sense, the native speaker's model was sufficient to arouse their awareness. However, in phonetic realization of the information focus, Chinese students would unconsciously add more prominence within one intonation group, and they were not aware of this problem by listening to the native speaker's model. In comparison, when providing the students with their original productions together with the modified versions, they would immediately perceive the differences and notice that the unnecessary stresses should be removed.

In the following example (Table 5.9), before the training, a student's (EG) production of the utterance "Yes, I like going on foot" in response to the question "Shall we walk there?" stressed two words, as the target focus should have been on "like". As can be seen from the pitch contour, the student's production contained two most prominent stresses, "like" and "foot", and the pitch of the rest elements were also lifted

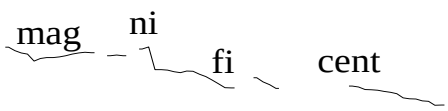
higher, making the overall pitch flat and lack of prominence contrasts. So, the modified version used for training raised the prominence of “like” by using a high rising and at the same time lowered the pitch of other elements. In the posttest, the student’s production successfully accented the nucleus and formed obvious prominence contrasts, which effectively reduced his Chinese accent.

Table 5.9 Example of deaccentuation (1)

Pretest:	
Modified version:	
Posttest:	

In the same vein, pitch modification was also effective for solving the stress shift problem in word level (Table 5.10), as stresses are realized by the variations of pitch prominence. It is noteworthy that in the control group, after the students listened to the native speaker’s model, rarely the problems of stress shift were solved. This might be due to: 1) the students did not pay attention to this problem; 2) even if they had paid attention, they still could not perceive the differences. In other words, the students failed to listen critically about the differences.

Table 5.10 Example of deaccentuation (2)

Pretest:	
Modified version:	
Posttest:	

Chunking

Chunking involves breaking the spoken materials into pieces of information. Speakers use chunking to distribute the energy of speaking in case the utterance is too long to produce within one intonation group, and more importantly, chunking is also used to disambiguate syntactic structures that are prone to cause misunderstanding. Though the use of pause is an important strategy of chunking, the overuse of pauses would break the overall information and decrease speech fluency.

In the present study, the students in the experiment group gained more strategies through the training to mark intonation boundaries than the students in the control group who merely relied on using pauses. By listening to the native speaker's model, the students could notice the intonation boundaries in the native speaker's speech, but they failed to extract the strategies used for signaling the boundaries. For example (Figure 5.2), if there was a pitch reset between two intonation groups (IP), students in the control group would tend to perceive it as a pause rather than notice the difference of the pitch variations between the two intonation groups.

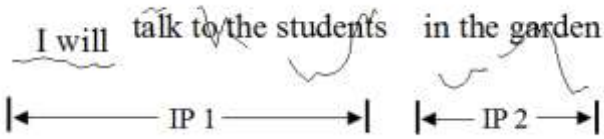
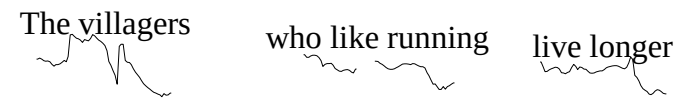


Figure 5.2 Example of using pitch reset to signal intonation boundary

In comparison, for the students in the experimental group, by comparing their original production with the modified version, they could be informed of the intonation boundary and become easily aware of the changes in their pitch. In other words, through the training, not only did they learned the importance of chunking in speaking, but they also gained the strategies for the realization of chunking. In the following example (table 5.11), the student relied on pause to mark the boundary before the training. After the training, the student learned to use pitch variations to signal the boundary. As can be seen from the pitch contour, the student's production in the posttest contained three intonation groups. The first and third group had a similar pitch contour and their pitch height were higher than that of the second group. It seems like the second part was inserted into the overall pitch contour of the utterance, which echoes the non-restrictive attributive clause of the syntactic structure.

Table 5.11 Example of chunking

Pretest:	
Modified version:	
Posttest:	

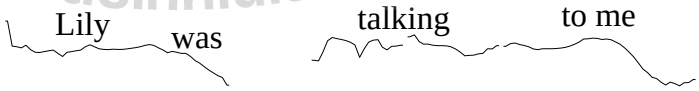
Retuning

Correct phonological representation is the prerequisite for accurate production. The establishment of correct phonological representation is a process of continuous adjustment, which is based on the perceived differences between the learners' existing representation and the incoming speech signal. In the present study it was found that listening to the native speaker's model could only inform students of the overall intonation patterns but could not facilitate their phonetic realization of those patterns. In contrast, pitch modification enabled students to listen critically to adjust their phonological representations, and thus to realize more accurate productions. In phonetic implementation of the falls and rises in pitch, pitch modification had the following effectiveness:

Firstly, it helped students to eliminate unnecessary pitch variations in speech. As has been discussed in the above section, some Chinese students were aware of the importance of intonation for pronunciation, and they thought that their speech would

sound “nicer” if contains more pitch variations, so they tended to employ more pitch variations when speaking. However, some of them did not realize that intonation serves for meaning and communication purposes and therefore they arbitrarily added unnecessary pitch variations to their speech in spite of intonation rules. As a result, this made their speech sound rather strange and sometimes even more prone to causing misunderstandings. By eliminating these unnecessary pitch variations (e.g., Table 5.12), their speech would be clearer, more fluent and comprehensible. In addition, this also informed the students that intonation serves for communicative purposes and the use of pitch variations is not arbitrary.

Table 5.12 Example of eliminating unnecessary pitch variations

Pretest:	
Modified version:	
Posttest:	

Secondly, it helped the students to produce more accurate phonetic realizations of rising tone patterns. Chinese students had more problems in producing rising tones than falling tones. They were commonly found to misplace the starting point of the rising and tend to rise the pitch on the last syllable of an utterance. For the students in the

experimental group, after listening to their modified productions, they noticed where their problems were and gained the strategies to implement risings by rearranging the pitch height and energy allocation for production (Table. 5.13). Actually, considering the limitations of technology and resynthesized speech, the students' productions in the posttest turned out to be better than the modified versions, which is more natural. This was due to the modified stimuli provided the students with critical acoustic cues for their acoustic-articulator remapping, which can yield more satisfactory productions.

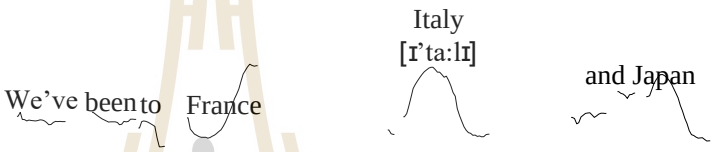
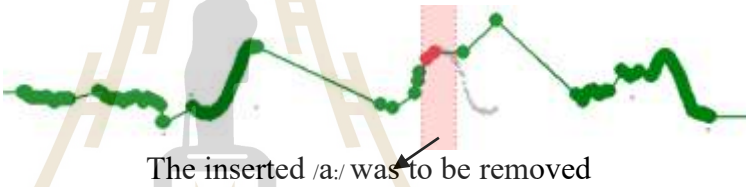
Table 5.13 Example of rising tone realization

Pretest:	
Modified version:	
Posttest:	

Thirdly, it enhanced the students' phonological memory in cases where the pitch variations were more complex. In the experiment, it was found that the students had more difficulties in producing the intonation patterns whose pitch variations were more complex, e.g., falling-rising tones, compound tones, or when the utterances were too long. In these cases, their attention was distracted by the segments, novel words, and sometimes long sentences caused their anxiety in production. Pitch modification as

corrective feedback for intonation learning targeted on the incorrect places based on the students' initial productions. By so doing, the students only needed to pay attention to their problems and their memory loads put on executing the motor commands would be greatly reduced. Therefore, in reproducing these intonation patterns, they might only need to make minor changes based on their initial productions (e.g., Table 5.14).

Table 5.14 Example of complex tone pattern realization

Pretest:	
Modified version:	
Posttest:	

In conclusion, the native speaker's model and the students' own pitch modified model for intonation learning were equally effective for arousing students' awareness of the placement of information focus, the divisions of intonation boundaries, and choosing correct intonation patterns according to communicative purposes. However, the pitch modified model was more effective for helping students listen critically and facilitated the students to produce a more accurate phonetic realization of intonation, including the deaccentuation of unnecessary prominence, more strategies for chunking,

and more accurate realizations of pitch variations.

5.4 The precision approach for pronunciation instruction

The results of the experiment showed that the Chinese university EFL students' problems in English intonation manifested in both perception and production. Their problems in perception were mainly caused by cognitive errors which hindered them to correctly perceive or understand the meaning of English intonation. The students' problems in production lay in the phonetic realization of intonation. Compared with perception, the students' problems in production were more resistant to change, especially for the control group. The native speaker's model was found ineffective for improving the students' accurate phonetic realization of intonation.

This might be due to that students' problems in production were much more individual-dependent than in perception. Strictly speaking, the acoustic characteristics of each individual student's phonetic realization of the same intonation pattern were different, which led to the fact that the students' problems in production were individually specific ones. By listening to the native speaker's model, the students were found hardly to perceive the subtle differences between their productions and the native speaker's, while those subtle differences might constitute the critical acoustic cues for accurate production. In this sense, the students' failure in listening to the native speaker's model or their own replays critically brought the model-oriented instruction

into a “bootstraps fallacy” (Sisson, 1970)-the phonetic criteria which the students needed for helpfully critical listening were precisely those which the instruction aimed to develop.

As has been mentioned, the model-oriented instruction focuses on how the model is correct, rather than how the students can be correct or how students’ problems can be solved. It fails to locate students’ individual specific problems and attempts to solve potential problems based on predictions. Specifically in pronunciation instruction, the “one-size fits-all” approach inevitably resulted in the students’ “plateau period”-when students’ pronunciation went to a certain level, they found that it was difficult to make further improvement while they felt that there was still much room for improvement-as quite a few students mentioned in the semi-structured interviews. Compared to the “one-size fits-all” approach, the pitch modification method for pronunciation learning in the present study, which targeted on each individual student’s specific problems and provided corrective feedback to their problems by modifying their incorrect intonation and playing back to them, was a precision approach informed by the concept of precision education.

Although precision language education has a myriad ways of implementation, a set of essential or common components can be identified. Hart (2016) suggested a near-term goal and a longer term goal to move toward. The near-term goal focuses on collecting data to obtain a better understanding of students’ learning disorders at

individual level. By using statistical analytics, each student's data sets can form a "knowledge network" to interpret the mechanisms underlying his/her learning problems. The longer term goal involves strategically using the knowledge gained from the near-term research and providing a personalized, or differentiated, intervention with the goal of remediating, or at least accommodating students' learning problems.

Cook et al. (2018) further identified four structural components for precision instruction approach: problem analysis, intervention design, performance monitoring, and collaboration. The purpose of the first component-problem analysis-is to identify students' learning problems. The second component pertains to the designing of tailored instructions based on students' problems, which are flexible and adaptive to student needs, preferences, and motivation to change. The third component is to monitor students' performance to determine whether the students have responded to the tailored instructions, so as to adjust the treatment to cater to students' needs. The last component refers to collaboration. To make timely and appropriate decisions, instructors or service providers meet to review monitoring data, derive consensus decisions, and plan for next steps.

Inspired d by the concepts of precision language education and based on previous educational evidences, the present study explored and validated a precision approach for pronunciation instruction. The pedagogic procedures of this approach can be divided into 5 cyclic steps: diagnosis, design, implementation, adjustment, and

assessment (Figure 5.3).

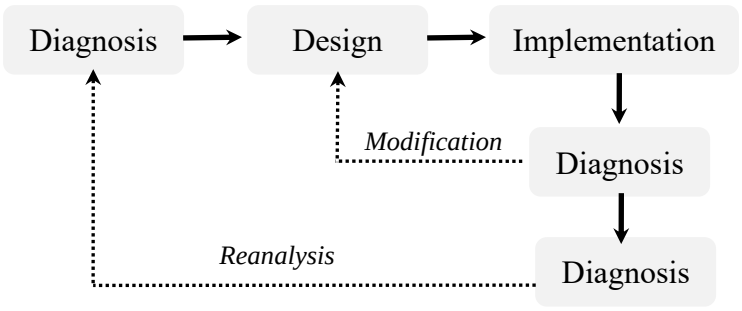


Figure 5.3 A precision approach for pronunciation instruction

Step 1: Diagnosis of students' problems

The first step involves collecting data about students' learning information so as to identify students' problems. Instruments like questionnaire, interviews, test, classroom observations can be used for collecting data. These data can form a knowledge network to access information about students' characteristics, needs, current performance, affective condition, etc., where their problems are rooted. This pre-intervention step provides a mechanism to better understand the nature of students' problem and engage in more systematic intervention planning efforts. Without this essential step, the intervention will delegate to one "based on hunches, best guesses, biases, or a trial-and-error procedure, which is likely to undermine intervention effectiveness" (Miller et al., 2018, p.12).

Specifically, in pronunciation instruction, a language background questionnaire, frequently used in previous studies, is an effective tool to elicit information about the current state of students' pronunciation learning. It is noteworthy that a teacher should

be cautious in dealing with students' needs and the pronunciation goals they made. Students' pronunciation goals are prone to be influenced by their teachers. As revealed by the present study, some students from the same class made the British accent as their pronunciation goal, as their pronunciation teacher was in favor of British accent. In another class, some students did not mind their pronunciation at all because their teacher was a World Englishes proponent, although the students might have mistaken their teacher's intentions. In a word, teachers should be capable of helping students make appropriate goals and thinking deeper about what they really need.

Pre-intervention pronunciation assessments are essential for identifying students' specific problems in pronunciation. Before this, it is of great importance to investigate students' attitudes towards and cognition in pronunciation. Attitude determines action. For example, if a student does not believe in the functions of intonation, any kinds of intervention will turn out to be in vain. Students' cognition in pronunciation determines whether they really gained the ability of production rather than mere mimicry (Reed & Michaud, 2015). In assessing pronunciation, it is necessary to respectively examine students' ability in perception and production, and further investigate whether their problems are caused by phonological representation or phonetic realization. The present study supported Mennen's (2007) finding that L2 pronunciation learners may first acquire the phonological patterns before they acquire the correct phonetic implementation of these patterns. Different students may be found in stage 1 or 2, or

even out of the two stages. This also accounts for why pronunciation instruction should be individualized and precise.

Step 2: Designing tailored intervention

As precision education highlights the heterogeneity among students with particular problems (Cook et al., 2018), interventions or instructions should be designed to cater to students' characteristics, target on their specific problems, and tailor to their needs. For pronunciation instruction, instructors should focus on the critical things and make it clear what factors are hindering students' improvement and what kind of input should be provided to students.

Considering the fact that individuals' responsiveness to intervention are different, the sequencing and intensity of intervention may differ depending on the individual. Previous educational evidences have shown the success of differentiating the time and intensity of the intervention (Fiorello et al., 2006; Fuchs et al., 2003). Though in the present study the students' time on task was strictly controlled, it was found that the students in the experimental group spent longer time on items that contained their modified speech. For those items that remained their original productions, they chose to gave another trial or just skipped them. In other words, pitch modification for intonation learning enabled them to save time on exercises with which they had no problems and spend more time focusing on their problems. In teaching practice, instructions can be designed in an adaptive way allowing students to learn in a free-

operant condition (Lindsley, 1990): students are free to respond at their own pace without having restraints (time or materials) placed on them.

Step 3: Implementing the tailored intervention

The third step is to implement the tailored intervention. As the interventions are specially designed according to each individual's needs, the precision approach embraces individualized learning. Under individual-centric interventions, students are allowed to learn at their own pace and even they are free to select the materials provided by the teacher. So the dominant role of the classroom shifts from the teacher to the students. The "precision teacher" performs like a coach, an advisor, and an on-line instructional designer (Lindsley, 1992) who does most of the work at "backstage" to assist and monitor students' learning. The teacher can use observation form to record students' learning behaviors or implant triggers in courseware to quantify their behaviors. In addition, the teacher can also ask the students to keep learning journals to take down what they have learned or have reflections on the learning process. Theses instruments can help the teacher to gather data for formative evaluation of student performance.

Step 4: Adjusting the intervention

As precision education advocates problem-solving bottom-up approach based instructions, the designed tailored interventions are subjective to adjustment. Through monitoring students' learning process, the teacher can gather information about

students' responsiveness to intervention. With the knowledge gained from the data, the teacher can modify the intervention by fading or pruning the treatments (if sufficient response) or increasing or changing the treatments (if insufficient response). For example, after the first training session of the present study, few students in the experimental group reported that they could not perceive the differences between their original production and the modified version. The researcher decided to increase the prominence of the modification by exaggeration, which turned out to better facilitate the students to perceive the critical cues in the following sessions.

Step 5: Assessing students' performance

In assessing students' performance, the routine teaching practice usually stops with scores, labeling students with these numbers. A final test cannot tell the whole story, especially with those tests developed from a single dimension. It is important for teachers to consider the criteria for evaluations. Lindsley (1992) proposed the concept of "fluency" which has a specific meaning about assessing students' performance: a true mastery or a level at which the skill becomes reliable and useful. Merbitz et al. (2004) further developed it into SARGE-the knowledge or skills gained is Stable, easily Applied, Retained over long periods, Generalized to new situations, and shows Endurance.

This criterion can also be adapted to assessing students' performance in pronunciation. Firstly, teachers need to examine whether the intervention has caused a

change in students' attitudes towards and cognition on pronunciation, and assessment should also go beyond the production level to cognitive level to ensure that the students have had a true mastery of the knowledge. Furthermore, teachers need to observe whether the students can apply what they have learned practically, as in pronunciation teaching practice students were usually found to perform well in class but discarded everything when they stepped out of the classroom. In addition, it is also necessary for teachers to investigate whether the students' pronunciation skills could be retained for a long time and the skills they gained from the training stimuli could be generalized to new stimuli.

It is inevitable that some of the students' problems are resistant to change or new problems will arise during the learning process, so after the assessment, teachers will need to reanalyze students' problems. A negotiation with students is recommended to better understand the mechanisms underlying their problems, so as to make the tailored interventions better fit their needs. In this way, the five steps cycle until the students' problems are solved. It can be seen that the precision approach poses a great challenge for teachers. They should be capable of using speech technology, spend time on analyzing data and designing instructions, be sensitive to students' changes, and even adapt to the change of their role. However, if teachers can overcome these challenges, instruction can avoid educational wastes, meet individual student's needs, and make learning less costly and more effective.

5.5 Summary

This chapter provided a discussion on the findings of this study. Firstly, by analyzing the students' performances in the pretest and posttest, it detailed the Chinese university EFL students' problems in English intonation, claiming that their problems were mainly caused by phonological representation and/or phonetic implementation. Secondly, it discussed the relationship between speech perception and production and the process and key factors of phonological acquisition. Thirdly, it illustrated the effectiveness of pitch modification for intonation learning. Lastly, it proposed the precision approach for pronunciation instruction which includes five cyclic step.

CHAPTER 6

CONCLUSION

This chapter concludes the present study. It is organized into 4 sections. The first section summarizes what the study did and the its major findings; the second section provides the pedagogical implications of the study; the third section presents the limitations of the study; and the last section proposes recommendations for future research.

6.1 Summary of the study

This study aimed at improving Chinese university EFL students' English intonation through pitch modification. Pitch modification is a method of acoustic cue enhancement for intonation learning which provides corrective feedback to students' productions by modifying students' incorrect intonation and feeding it back to the students as the model for leanring. This study assumed that by listening to the model of their own voices (self-perceived voices), the students can better perceive the differences between their incorrect productions with the target ones, focus on the critical acoustic cues, accurately locate their problems, so as to improve their performance in English intonation. To examine the effectiveness of this method for

intonation learning, this study proposed two research questions: 1) Is pitch modification effective for English intonation learning? If yes, in what ways? 2) What are the students' opinions of using pitch modification for intonation learning?

In order to answer the two research questions, a mixed methods research design was employed to collect both quantitative and qualitative data. The quantitative part was a quasi-experiment wherein an English intonation training was conducted to two groups of students: the experimental group and the control group. The training included 6 sessions and in each session the students were required to perform listen-compare-repeat exercises on a courseware designed by the researcher. The training materials contained 150 target utterances which were arranged into three modules: tonicity, tonality, and tone. The only difference between the experimental group and the control group was that: the control group used the native speaker's model while the experimental group used their own modified speech as the model for intonation training. An intonation perception test and production test were administered to the participants before and after the training to quantitatively examine the effectiveness of the treatment. In order to triangulate the quantitative findings and investigate the students' opinions of pitch modification for intonation learning, a questionnaire survey and semi-structured interviews were conducted immediately after the training.

The results of the experiment showed that both groups' performances in the perception test was significantly improved after the training, and there was no

significant difference between the two groups either in the pretest or posttest. That is to say, both the native speakers' model (for the control group) and the students' modified speech as model (for the experimental group) were equally effective for helping students perceive and understand the meaning of intonation. For the production aspect, the training also significantly improved both of the two groups' performances in the production test. There was no significant difference between the two groups' performance in the pretest, however, in the posttest, the experimental group outperformed the control group. This was due to the fact that the native speaker's model was as effective as the students' pitch modified model for informing the students to choose the correct intonation patterns for production, but it was not so effective as the students' pitch modified model for facilitating the students to produce more accurate phonetic realizations of those patterns.

In terms of the students' performance in specific intonation patterns, the results showed that the students had no problems in producing *Statements (falling tone), Commands,* and *Exclamations*. Their problems in *Tonality, Wh-questions (falling tone), Yes/no-questions (falling tone),* and *Tag-questions (rising tone)* were mainly caused by phonological representation which can be solved by both of the two types of model. For the other intonation patterns, their problems were caused by phonetic implementation which can be effectively solved by listening to the pitch modified model, while the native speaker's model was not so effective.

These results revealed that Chinese university EFL students' had prominent problems in English intonation before the training. They had very poor knowledge of English intonation and neglected the meaning and function of intonation. Thus, in producing speech, they chose intonation arbitrarily or according to default tones or intuitions. The results also implied that Chinese students were good at producing falling tones and sometimes even overused them. While they were not so good at producing rising tones and falling-rising tones, as well as compound tones. In addition, the results of this study support Mennen's (2007) finding that L2 intonation learning may go through different stages, and L2 learners may first acquire the phonological patterns before accurate phonetic implementation of these patterns.

This study indicated that pitch modification for intonation learning was effective (and more effective than the native speaker's model) for promoting students' development of intonation acquisition from phonological representation to phonetic realization. This was due to that pitch modification as a method of cue enhancement can make the critical acoustic cues more salient, and thus could help the students to listen to the model critically, i.e., they could easily perceive the differences between their incorrect productions with the correct ones. Furthermore, by listening to the model of their own voices, the students could directly link the speech signals with their own articulators without the distractions from uncritical acoustic elements, which reduced their memory load and sped up their processing.

The effectiveness of pitch modification for intonation learning was also validated by the data collected from the students' learning journals, questionnaire, and semi-structured interviews. Most of the students perceived that this kind of English intonation training was interesting, and their pronunciation as well as their confidence of speaking English was improved through the training. They also believed that the training effectively solved their individual specific problems in intonation.

Since pitch modification for intonation learning targeted the individual student's problems in intonation and provided corrective feedback to their problems, this kind of intervention reflects the main concept of precision education. By reviewing the pedagogic procedures of the intervention, this study proposes the precision approach for pronunciation instruction. This precision approach includes five cyclic steps: diagnosing students' problems, designing tailored interventions, implementing the interventions, adjusting the interventions, and assessing students' performance. This approach is not only confined to the instruction of the suprasegmental features of pronunciation, but also can be applied to the instruction of segmental features.

6.2 Pedagogical implications

The findings of this study bring potential enlightenment to the teaching practice of English pronunciation, especially for the teaching of English intonation. Hopefully, the implications shed light on the development of teaching materials, the diagnosis of

students' problems, and focusing on the critical things for solving students' problems in pronunciaiton teaching. To be more specific, the following implications can be drawn from the study:

Firstly, the findings of this study highlight the necessary of incorporating intonation as an essential compopnent in pronunciation teaching. Though the paradigm shift of pronunciation teaching-the teaching priority should be shifted from the segmental features to the suprasegmental features-has been claimed for decades, the teaching of suprasegmentals is still neglected in Chinese EFL pronunciation classroom. In the present study, Chinese university EFL students were found lack of basic knowledge of English intonation and perform very poor in either perception or production of intonation. This resulted to their strongly Chinese accented pronunciation, reduced their speech comprehensibility, and sometimes even caused communication break-downs. Quite a few students reported the "plateau period" phenomenon: their pronunciation was stuck at a certain level while they felt that there was still a great space for improvement. After the intonation training, many students showed that their overall pronunciation level has been significantly improved since they paid attention to intonation. Consequently, increasing the weighting of intonation in pronunciation teaching can make a breakthrough on students' "plateau period", reduce their speech accentedness, enhance their confidence of speaking English, and improve their overall pronunciation level.

Secondly, the content of intonation instruction should be more systematic and practical. As has been mentioned above, Chinese EFL pronunciation class room had little involvement of intonation and those limited instructions on intonation only focused on default tones and the examples given were on sentence level. As a result, Chinese EFL students had very limited knowledge of the meaning and function of intonation, and sometimes they incorrectly linked default tones with sentence types. This also contributed to the students' wrong beliefs about intonation. Many students held that intonation is simply decorative. They might perform well in class as the teacher required, but they would discard everything when they stepped out of the class. Therefore, the content of intonation teaching should be systematic to compensate their partial knowledge and be practical to encourage them to apply what they have learned to everyday use.

Thirdly, this study implied that pronunciation learners may go through different stages in phonological acquisition, and it is necessary to locate in which stage the student are and identify what types of problems they have. Pronunciation inevitably involves perception and production, and making clear the interplay between the two modalities is the key for understanding the mechanism underlying students' problems. This study supports the claim of previous researchers (De Bot & Mailfert, 1982; Flege, 1999; Lee& Lyster, 2017, etc..) that perception precedes production and correct phonological representation is the sine qua non for accurate phonetic implementation.

Hence, the core issue of phonological acquisition lies in the establishment of correct phonological representation of the target sound. So, pronunciation instruction should focus on how to arouse students' awareness on the critical acoustic cues so as to facilitate their establishment of the correct phonological representation, and thus to improve their phonetic realization of the target sound.

Fourthly, this study provides an example of using students' own productions for pronunciation learning. Pronunciation researchers and instructors have always been obsessed with the problem of what kind of stimuli should be provided to learners. They tended to find "standard" models and led students to focus on how the models are correct, rather than how the students can be correct. Some teachers may have asked students to listen to the replay of their productions, however, it turned out that students usually failed to listen critically, because there was no corrective feedback to facilitate their listening. The method of using students' own productions for learning in this study provided students a chance for critical listening. It might not be possible for teachers to employ this method for pronunciation teaching in every class since modifying students' productions is very time consuming. A practical way may be that teachers could incorporate speech modification tools in class, such as Praat, WinPitch, and guide students to use these tools, so that students can better know their productions, locate their problems, and retune their phonological representations through exposing to the variations of a target sound.

Lastly, the pedagogic procedures of this study shed light on implementing the concept of precision education in pronunciation instruction. The characteristics of pronunciation learning itself determine the uniqueness of each individual student's pronunciation problems and the one-size-fits-all approach can not meet each student's needs. The precision approach captures the heterogeneity of students' problems and its proponents believe that an effective intervention should start with a pre-intervention phase of problem diagnosis. Targeting on individual student's specific problems, interventions should be designed to tailor to each student's needs, and these interventions are subjective to adjustment according to new changes. The precision approach also embraces student-centered, self-paced learning and transfers the teacher's role from the "dominator" to "assistant", and thus leads to more efficient and effective instruction.

6.3 Limitations of the study

This study verified the effectiveness of using pitch modification as corrective feedback for improving Chinese university EFL students' English intonation. Through rigorous research design and experimental implementation, this study yielded significant findings which offers valuable insights and implications for pronunciation instruction. However, this study has some inevitable limitations.

First, the method of using pitch modification as corrective feedback for intonation

learning in this study was of a delayed-feedback paradigm: collecting students' productions-modifying the productions-replaying the modified version to students. Delayed-feedback has its inherent drawbacks. For example, a student's problem may be solved during the interval, or the problem may change to another problem. For this reason, immediate or simultaneous feedback could be more effective than the delayed one. However, considering the technological limitations, currently it is not possible to simultaneously modify students' speech when they are speaking.

Second, this study did not accurately record the students' learning behavior during training or track their mouse clicking on the courseware. So, this study can not point out what changes the modified speech brought to students' learning behavior. Furthermore, this study also can not specify what inner reflections (e.g., neuro activity in brain, muscle movements of articulators) did the students have when exposed to their own modified speech.

Third, the students' productions in this study were produced with scripts and they had time to consider the selection of intonation before production. In other words, their productions were not elicited by spontaneous speech. Hence, this study can not answer the question that whether their ability gained from the training can be generalized to real time conversations or be retained over a long period of time.

6.4 Recommendations for future research

Considering the limitations of the present study, there remains a large space for exploration in using modified speech for pronunciation teaching. In the meantime, the positive results of this study may also give enlightenment to future research. The detailed recommendations for future research are presented below:

Firstly, researchers who are specialized in speech technology are recommended to develop a real-time pitch modifier. The modifier has the function of recording students' speech and simultaneously modifying students' pitch when they are speaking. After students finished speaking, they can immediately listen to their original productions and the modified version. This means that with this kind of pitch modifier, students can get immediate corrective feedback (equals to recast) to their pronunciation, which can most efficiently facilitate students' uptake (Lyster & Ranta, 1997).

Secondly, it is recommended that neuroimaging experiments (such as fMRI) be carried out to investigate the neuron activity in the cortical area upon listening to modified self-produced speech. Focus could be placed on examining whether there are differences of the patterns of neuron activation in the process of perception and production between listening to other people's model and self-produced modified model. Hopefully, this kind of experiment can provide neuro-evidence that listening to one's own modified speech as the model for pronunciation learning can reduce the workload of the motor system and facilitate sensory-motor mapping. In addition,

experiments are also recommended to use transcranial magnetic stimulation (TMS) technology to detect the muscle activities of articulators to precisely specify the changes that motor control brings to articulators upon perceiving self-produced modified stimuli.

Thirdly, the method of using students' modified speech as corrective feedback for intonation learning can be applied to the learning of segmentals. This may entails more precise technology for manipulating individual sounds. Suggestions are given to start with vowels of which the modifications can be realized by manipulating the formants or only the second formant for vowel contrasts. Considering that L2 learners' problems in segmentals are more complex and individual-specific, one can look forward to the promising outcomes achieved here yielding positive results.

All in all, the present study validated a precision approach to intervention based on students' problems and coping with those problems with tailored interventions would turn out to expend the least effort and be the most effective. The significant improvement of the students' pronunciation verified Flege's (1995) claim that neural plasticity in terms of the ability to establish new perceptual categories exists throughout life. Intonation is learnable and can be learned by adults. The researcher of the present study further claims that intonation *must* be learned, taking account into the important role of intonation plays in pronunciation learning. This study was an attempt of implementing the concept of precision education in intonation instruction. However, as

Leather (1983) claimed that there is no single technique for facilitating the acquisition of all classes of L2 sounds under all conditions, future research needs to explore more means of its implementation and even expand the research of precision education to other fields of language education.

REFERENCES

Alghazo, S. M. (2015). Advanced EFL learners' beliefs about pronunciation teaching. *International Education Studies*, 8(11), 63-76.

Allen, V. F. (1971). Teaching intonation, from theory to practice. *Tesol Quarterly*, 5(1), 73-81.

Allen, P. P., Amaro, E., Fu, C. H., Williams, S. C., Brammer, M., Johns, L. C., & McGuire, P. K. (2005). Neural correlates of the misattribution of self-generated speech. *Human Brain Mapping*, 26(1), 44-53.

Ananthakrishnan, G., & Engwall, O. (2011). Mapping between acoustic and articulatory gestures. *Speech Communication*, 53(4), 567-589.

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

Aruffo, C., & Shore, D. I. (2012). Can you McGurk yourself? Self-face and self-voice in audiovisual speech. *Psychonomic Bulletin & Review*, 19(1), 66-72.

Aslam, M., & Kak, A. A. (2007). *Introduction to English phonetics and phonology*. Cambridge: Cambridge University Press.

Atar, C. (2018). Should We Teach Pronunciation Explicitly in L2/EFL Classrooms?. *International Journal of Contemporary Educational Research*, 5(2), 95-102.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*,

11(1-2), 17-68.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, 3, 255-309.

Békésy, G. V. (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America*, 21(3), 217-232.

Benati, A. (2016). Input manipulation, enhancement and processing: Theoretical views and empirical research. *Studies in Second Language Learning & Teaching*, 6(1), 65-88.

Benrabah, M. (1997). Word-stress—a source of unintelligibility in English. *International Review of Applied Linguistics in Language Teaching*, 35(3), 157-166.

Bernal, B., Ardila, A. (2009). The role of the arcuate fasciculus in conduction aphasia. *Brain*,132, 2309-2316.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore: York Press.

Bi, R., & Chen, H. (2013). Developmental changes of tone patterns in Chinese EFL students' read speech. *Foreign Languages and Their Teaching*, (1), 50-54.

Bissiri, M. P., & Pfitzinger, H. R. (2009). Italian speakers learn lexical stress of German morphologically complex words. *Speech Communication*, 51(10), 933-947.

Boersma, P., & Weenink, D. (2018). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.37, retrieved from http://www.praat.org/

Bolinger, D. (1986), *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press

Bradford, B. (1988). *Intonation in Context: Intonation Practice for Upper-intermediate and Advanced Learners of English (Student's book)*. Cambridge: Cambridge University Press.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61(5), 977-985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101(4), 2299-2310.

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

Braun, B., Kochanski, G., Grabe, E., & Rosner, B. S. (2006). Evidence for attractors in English intonation. *The Journal of the Acoustical Society of America*, 119(6), 4006-4015.

Brown, C. A. (1999). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp.4-63). Malden, MA: Blackwell.

Buchsbaum, B., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, 25, 663-678.

Büring, D. (2016). *Intonation and meaning*. Oxford: Oxford University Press.

Canagarajah, A. S. (2005). *Reclaiming the local in language policy and practice*. Mahwah, NJ: Lawrence Erlbaum.

Canevari, C., Badino, L., D' Ausilio, A., Fadiga, L., & Metta, G. (2013). Modeling speech imitation and ecological learning of auditory-motor maps. *Frontiers in Psychology*, 4, 364.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Center for Language Acquisition & Precision Education. (2017). [Center website]. Retrieved October 28, 2018, from http://www.clape.org/en/

Chamonikolasová, J. (2007). *Intonation in English and Czech dialogues*. Brno: Masarykova univerzita.

Chela-Flores, B. (2003). Optimizing the teaching of English suprasegmentals. *Barcelona English Language and Literature studies*, 12, 255-273.

Chen, H. (2006a). Intonation divisions in Chinese university students' read speech. *Foreign Language Teaching and Research*, 38(5), 272-278.

Chen, H. (2006b). EFL learners' placement of tonicity and information focus. *Journal of PLA University of Foreign Languages*, 29 (6), 32-38.

Chen, H., Sun, X. P., & Zhang, Y. (2008). On English intonation instruction. *Foreign Language Research*, (2), 135-139.

Chun, D. M. (1988). Teaching intonation as part of communicative competence: Suggestions for the classroom. *Die Unterrichtspraxis*, 21(1), 81-88.

Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Philadelphia, PA: John Benjamins.

Clennell, C. (1997). Raising the pedagogic status of discourse intonation teaching. *ELT Journal*, 51(2), 117-125.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th edition). London: Routledge.

Connor, C. M., Mazzocco, M. M., Kurz, T., Crowe, E. C., Tighe, E. L., Wood, T. S., & Morrison, F. J. (2018). Using assessment to individualize early mathematics instruction. *Journal of School Psychology*, 66, 97-113.

Cook, C. R., Kilgus, S. P., & Burns, M. K. (2018). Advancing the science and practice of precision education to enhance student outcomes. *Journal of School Psychology*, 66, 4-10.

Creswell, J. W. (2012). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th edition). London: Sage publications.

Crowder, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology*, 1(2), 157-178.

Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, 9(4), 183-191.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford University Press.

Dalton, D. F. (1997). Some techniques for teaching pronunciation. *The Internet TESL Journal*, 3(1). Retrieved from: http://www.aitech.ac.jp/iteslj/.

Daryadar, M., & Raghibi, M. (2015). The effect of listening to recordings of one's voice on attentional bias and auditory verbal learning. *International Journal of Psychological Studies*, 7(2), 155-163.

Davies, G. (2004). *ICT4LT Module 2.2: Introduction to multimedia CALL*. Retrieved from: http://www.ict4lt.org/en/en_mod2-2.htm

De Bot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior. *Language and Speech*, 26(4), 331-350.

De Bot, K., & Mailfert, K. (1982). The teaching of intonation: Fundamental research and classroom applications. *Tesol Quarterly*, 16(1), 71-77.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19(1), 1-16.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language teaching*, 42(4), 476-490.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2007). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29(3), 359-380.

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., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13(1), 1-17.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679.

Elliot, C. E. (1991). The relationship between the perception and production of Mandarin tones: An exploratory study. *University of Hawai'i Working Papers in ESL*. 10(2),177-204.

Engwall, O., & Bälter, O. (2007). Pronunciation feedback from real and virtual language teachers. *Computer Assisted Language Learning*, 20(3), 235-262.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15(2), 399-402.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73(6), 2608-2611.

Fairbanks, G. (1954). Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem. *Journal of Speech & Hearing Disorders*. 19, 133–139.

Felps, D., Bortfeld, H., & Gutierrez-Osuna, R. (2009). Foreign accent conversion in computer assisted pronunciation training. *Speech Communication*, 51(10), 920-932.

Fiorello, C. A., Hale, J. B., & Snyder, L. E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools*, 43, 835–853.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross- language research*, 92, 233-277.

Flege, J. E. (1999).*The relation between L2 production and perception*. In Proceedings of the XIVth International Congress of Phonetics Sciences (pp. 1273-1276), Berkeley.

Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a second language*. Canberra, Australia: Department of Education, Training and Youth Affairs.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Bloomington: Indiana University Linguistics Club.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, 18(3), 157-171.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13(3), 361-377.

Gallego, J. C. (1990). The intelligibility of three nonnative English-speaking teaching assistants: An analysis of student-reported communication breakdowns. *Issues in Applied Linguistics*, 1(2).

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*, 4, 340.

Gao, L., & Deng, Y. C. (2009). Study on English word stress misplacement of Chinese university EFL students—A corpus-based study. *Foreign Language World*, (3), 10-16.

Gary, F. S., & Charles, D. F. (2017). *Ethnologue: Languages of the world (20th edition)*. Dallas, TX: SIL International.

Gaviria, B. (1967). Autonomic reaction magnitude and habituation to different voices. *Psychosomatic Medicine*, 29(6), 598-605.

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, 2(3), 119-128.

Gilbert, J. (1995). Pronunciation practice as an aid to listening comprehension. In D. J. Mendelson & J. Rubin (Eds.), *A guide for the teaching of second language*

learning (pp. 97-111). San Diego: Dominic Press.

Gilbert, J. (2014). Myth 4: intonation is hard to teach. In Grant, L., & Brinton, D. (Eds.), *Pronunciation myths: Applying second language research to classroom teaching* (pp.107-137). Michigan: University of Michigan Press.

Golombek, P. R. (2012). Teaching pronunciation: a reference for teachers of English to speakers of other languages. *Tesol Quarterly*, 32(4), 783-784.

Grant, L. 2014. Prologue to the myths: what teachers need to know. In Grant, L., & Brinton, D. (Eds.), *Pronunciation myths: Applying second language research to classroom teaching* (pp.1-33). Michigan: University of Michigan Press.

Graux, J., Gomot, M., Roux, S., Bonnet-Brilhault, F., Camus, V., & Bruneau, N. (2013). My voice or yours? An electrophysiological study. *Brain Topography*, 26(1), 72-82.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3), 377-442.

Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 47–57.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours. *Phonetica*, 11(3-4), 175-185.

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

Hall, M. S., & Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology*, 66, 54-66.

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. Paris: Mouton.

Han, F. (2013). Pronunciation Problems of Chinese Learners of English. *ORTESOL Journal*, 30, 26-30.

Hansen, T. K. (2006). Computer assisted pronunciation training: The four 'K's of feedback. *Current Developments in Technology-Assisted Education*, 342-6.

Hansen, M. O., & Stinson, M. R. (1998). Air conducted and body conducted sound produced by own voice. *Canadian Acoustics*, 26(2), 11-19.

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

Hardison, D. M. (2012). Second language speech perception: A cross-disciplinary perspective on challenges and accomplishments. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 349-363). London: Routledge.

Hattori, K., and Iverson, P. (2010). *Examination of the relationship between L2 perception and production: An investigation of English /r/-/l/ perception and production by adult Japanese speakers*. In Interspeech Workshop on Second Language Studies: Acquisition, Learning, Education and Technology, ed. M. Nacano (Tokyo: Waseda University).

Harmer, J. (2001). *The practice of English language teaching*. London: Longman

Hart, S. A. (2016). Precision education initiative: moving toward personalized education. *Mind, Brain, and Education*, 10(4), 209-211.

Havranek, G., & Cesnik, H. (2001). Factors affecting the success of corrective feedback. *EUROSLA yearbook*, 1(1), 99-122.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese

learners of English. *Speech Communication*, 47(3), 360-378.

He, B., Sangarun, P., & Lian, A. (2015). Implementing autonomy: A model for improving pronunciation in Chinese EFL university students. *Rangsit Journal of Arts and Sciences*, 5(1), 1-12.

Heinks- Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *NeuroReport*, 17(13), 1375-1379.

Hewings, M. (2007). *English pronunciation in use (Advanced)*. Cambridge: Cambridge University Press.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, 15(5), 673-682.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3), 407-422.

Hinofotis, F. B., Bailey, K. M., & Stem, S. C. (1981). Assessing the oral proficiency of prospective foreign teaching assistants: Instrument development. In A. S. Palmer, P. J. M. Groot, & G. A. Trosper (Eds.), *The construct validation of tests of communicative competence* (pp. 106-126). Washington D.C.: TESOL.

Hirose, K. (2004). Accent type recognition of Japanese using perceived mora pitch values and its use for pronunciation training system. In *Proceedings of the International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages* (pp. 77-80), Beijing, China.

Hirst, D., & Cristo, A. D. (1998). A survey of intonation systems. *Physiological Reviews*, 6, 1-44.

Holzman, P. S., Rousey, C., & Snyder, C. (1966). On listening to one's own voice: Effects on psychophysiological responses and free associations. *Journal of Personality and Social Psychology*, 4(4), 432-441.

Hong, W. (2012). *Experimental study on Chinese EFL learners' English intonation patterns*. Doctoral dissertation, Nankai University, China.

Hu, F. (2017). A study on Chinese EFL learning of English pronunciation from the perspective of aesthetic linguistics. *Theory and Practice in Language Studies*, 7(7), 579-584.

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*, 20(1), 18-29.

Huo, S. Y., & Luo, Q. (2017). Misuses of English Intonation for Chinese Students in Cross-Cultural Communication. *Cross-Cultural Communication*, 13(1), 47-52.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449), 2526-2528.

IELTS. (2017). Test taker performance 2017. *International English Language Testing System*. Retrieved from: https://www.ielts.org

Isbell, D. R. (2017). Assessing pronunciation for research purposes with listener-based numerical scales. In Kang, O., & Ginther, A. (Eds.), *Assessment in Second Language Pronunciation* (pp. 101-124). London: Routledge.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y. I., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87(1), 47-57.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: a comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, 43(1), 88.

Jenkins, J. (2004). Research in teaching pronunciation and intonation. *Annual Review of Applied Linguistics*, 24, 109-125.

Jiang, H. L. (2012). An empirical study of the learning effect of college English majors in intonation. *Foreign Languages in China*, 9(2), 65-80.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Doctoral dissertation, University of Stuttgart, Germany.

Jilka, M., & Möhler, G. (1998). Intonational foreign accent: speech technology and foreign language teaching. In *Proceedings of ESCA Workshop on Speech Technology in Language Learning* (pp. 115-118). Marholmen, Sweden.

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

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.

Kaplan, J. T., Aziz-Zadeh, L., Uddin, L. Q., & Iacoboni, M. (2008). The self across the senses: an fMRI study of self-face and self-voice recognition. *Social Cognitive and Affective Neuroscience*, 3(3), 218-223.

Kawato, M., & Wolpert, D. (1998). Internal models for motor control. *Sensory Guidance of Movement*, 218, 291-307.

Kelly, G. (2000). *How to teach pronunciation.* Harlow: Pearson Education.

Kenworthy, J. (1987). *Teaching English Pronunciation.* London: Longman

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. D. & Terrel, T. D. (1983). *The natural approach*. Hayward, CA: Alemany Press.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, 97(22), 11850-11857.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Lantolf, J. P. (1976). On teaching intonation. *The Modern Language Journal*, 60(5-6), 267-274.

Leather, J. (1983). Second-language pronunciation learning and teaching. *Language Teaching*, *16*(3), 198-219.

Lee, A. H., & Lyster, R. (2017). Can corrective feedback on second language speech perception errors affect production accuracy?. *Applied Psycholinguistics*, 38, 1-23.

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Al Otaiba, S., & Fidler, D. J. (2018). Personalized reading intervention for children with Down syndrome. *Journal of School Psychology*, 66, 67-84.

Lengeris, A. (2012). Prosody and second language teaching: Lessons from L2 speech perception and production research. *Pragmatics and Prosody in English Language Teaching*, 15, 25-40.

Léon, P. R., & Martin, P. (1972). Applied linguistics and the teaching of intonation. *The Modern Language Journal*, 56(3), 139-144.

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

Levis, J. M. (2002). Reconsidering low-rising intonation in American English. *Applied Linguistics*, 23(1), 56-82.

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

Levis, J. M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, applied linguistics and TESOL: Challenges for theory and practice* (pp. 245-270). London: Palgrave Macmillan.

Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32(4), 505-524.

Levis, J., & Wichmann, A. (2015). English intonation-form and meaning. In Reed, M., & Levis, J. (Eds.), *The handbook of English pronunciation* (pp. 139-155). John Wiley & Sons.

Lian, A-P. (1984). Aspects of Answer-Evaluation in Traditional Computer-Assisted Language Learning, in Russell R. M. (ed.): *Proceedings of the 2nd CALITE Congress (pp. 150-160)*, Brisbane, University of Queensland.

Lian, A. P. (2014). On-Demand Generation of Individualized Language Learning Lessons. *Journal of Science*, 9(1), 25-38.

Lian, A. P., & Sangarun, P. (2017). Precision Language Education: A Glimpse Into a Possible Future. *GEMA Online® Journal of Language Studies*, 17(4),1-15.

Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117-123.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of speech code. *Psychological Review*, 74, 431-461

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.

Liberman, A. M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249-336.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children*, 22(3), 10-15.

Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis*, 25(1), 51-57.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94(3 Pt 1), 1242-1255.

Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13(3), 110-114.

Lu, J., Wang, R., & De Silva, L. C. (2012). Automatic stress exaggeration by prosody modification to assist language learners perceive sentence stress. *International Journal of Speech Technology*, 15(2), 87-98.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in second language acquisition*, 19(1), 37-66.

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review*, 18(6), 1197-1202.

Makarova, V., & Zhou, X. (2006). Prosodic characteristics in the Speech of Chinese EFL learners. *Proceedings of Speech Prosody*. Dresden, Germany.

Maurer, D., & Landis, T. (1990). Role of bone conduction in the self-perception of speech. *Folia Phoniatr*, 42(5), 226-229.

Mcclelland, J. L., Fiez, J. A., & Mccandliss, B. D. (2002). Teaching the /r/-/l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology & Behavior*, 77(4–5), 657-662.

McGurk H, MacDonald J. 1976. Hearing lips and seeing voices. *Nature*. 264:746--748.

McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 185–196). Oxford: Oxford University Press.

Meng, X. J, & Wang, H. M. (2009). Boundary tone patterns in Chinese English learners' read speech. *Foreign Language Teaching and Research (bimonthly)*, (6), 447-451.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of phonetics*, 32(4), 543-563.

Mennen, I. (1999, August). The realisation of nucleus placement in second language intonation. In Proceedings of the fourteenth international congress of phonetic sciences (pp. 555-558).

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. *Trends in Linguistics Studies and Monographs*, 186, 53-76.

Merbitz, C., Vieitez, D., Merbitz, N. H., & Binder, C. (2004). Precision teaching: Applications in education and beyond. In Moran, D. J. & Malott R. W. (Eds.), *Evidence-based educational methods*(pp. 63-78). Academic Press

Miller, F. G., Cook, C. R., & Zhang, Y. (2018). Initial development and evaluation of the student intervention matching (SIM) form. *Journal of School Psychology*, 66, 11-24.

Miller, L. B. (2017). Review of journaling as a teaching and learning strategy. *Teaching and Learning in Nursing*, 12(1), 39-42.

Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2010). Alignment to visual speech information. *Attention, Perception, & Psychophysics*, 72(6), 1614-1625.

Moon, J. A. (2006). *Learning journals: A handbook for reflective practice and professional development.* London: Routledge.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *Tesol Quarterly*, 25(3), 481-520.

Möttönen, R., Dutton, R., & Watkins, K. E. (2012). Auditory-motor processing of speech sounds. *Cerebral Cortex*, 23(5), 1190-1197.

Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *Journal of Neuroscience*, 29(31), 9819-9825.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49, 285-310.

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.

Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., ... & Kojima, S. (2001). Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia*, 39(10), 1047-1054.

National University Precision Institute. (2017). [University website]. Retrieved October 28, 2018, from https://www.nu.edu/precision/index.html

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

Neri, A., Cucchiarini, C., & Strik, H. (2002b) Feedback in Computer Assisted Pronunciation Training: Technology Push or Demand Pull?. *Proceedings of ICSLP 2002*: 1209–1212.

Nikbakht, H. (2010). EFL pronunciation teaching: A theoretical review. *Journal of English Language Pedagogy and Practice*, 4(8), 146-174.

Nolan, F. (2008). Intonation. In Aarts, B., & McMahon, A. (Eds.), *The handbook of English linguistics* (pp. 433-458). John Wiley & Sons.

O'Connor, J. D. & Arnold, G. F. (1973). *Intonation of colloquial English*. London: Longman

O'Connor, M. A., & Daly III, E. J. (2018). Selecting effective intervention strategies for escape-maintained academic-performance problems: Consider giving'em a break!. *Journal of School Psychology*, 66, 41-53.

Ohala, J. J. (1984) An ethological perspective on common crosslanguage utilization of F0 in voice. *Phonetica*, 41, 1-16.

Ohala, J. J., & Gilbert, J. B. (1981). Listeners' ability to identify languages by their prosody. In P. R. Léon & M. Rossi (Eds.), *Problèmes de Prosodie, II: Expérimentations, modèles et fonctions* (pp. 123–131). Ottawa: Didier.

Olivos, G. (1967). Response delay, psychophysiologic activation, and recognition of one's own voice. *Psychosomatic Medicine*, 29(5), 433-440.

Oppenheim, A. N. (2000). *Questionnaire design, interviewing and attitude measurement*. London: Bloomsbury Publishing.

Warren, P. (1996). Prosody and parsing: An introduction. *Language and Cognitive Processes*, 11(1-2), 1-16.

Paunović, T., & Savić, M. (2008). Discourse intonation-Making it work. *ELOPE: English Language Overseas Perspectives and Enquiries*, 5(1-2), 57-75.

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91, 176-180.

Pellegrino, E., Vigliano, D. (2015). Self imitation in prosody training: A study on Japanese learners of Italian. In Steidl, S., Batliner, A. & Jokisch, O. (Eds.), *Workshop on Speech and Language Technology in Education* (pp. 53-57), Leipzig, Germany.

Pierrehumbert, J. B. (1980). *The Phonetics and phonology of English intonation.* Doctoral dissertation, Massachusetts Institute of Technology, USA.

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

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

Pörschmann, C. (2000). Influences of bone conduction and air conduction on the sound of one's own voice. *Acta Acustica United with Acustica*, 86(6), 1038-1045.

Prator, C. H. (1971). Phonetics vs. phonemics in the ESL classroom: When is allophonic accuracy important?. *Tesol Quarterly*, 5(1), 61-72.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129-154

Probst, K., Ke, Y., & Eskenazi, M. (2002). Enhancing foreign language tutors–in search of the golden speaker. *Speech Communication*, 37(3-4), 161-173.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20), 7865-7870.

Reed, M., & Michaud, C. (2015). Intonation in research and practice: The importance of metacognition. In Reed, M., & Levis (Eds.), J. M. *The handbook of English pronunciation* (pp.454-470). John Wiley & Sons.

Reves, T., & Medgyes, P. (1994). The non-native English speaking EFL/ESL teacher's self-image: An international survey. *System*, 22(3), 353-367.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, 71(3), 491-507.

Roach, P. (1991). *English phonetics and phonology: A practical course.* Cambridge: Cambridge University Press.

Rosevella, D. N. (2016). Lectora Inspire as media in delivering TEFL materials for English department of Uniska. *RETAIN*, 4(2).

Rubin, D. & Smith, K. (1989). Linguistic and non-linguistic factors affecting undergraduates' attitudes toward NNSTAs. *23rd Aimual TESOL Convention*, San Antonio, Texas.

Rui, T. (2007). The English intonation of Chinese EFL learners: A comparative study. *CELEA Journal (Bimonthly)*, 12(6), 34-45.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied linguistics*, 11(2), 129-158.

Silverman, D. (2000). *Doing qualitative research: A practical guide*. London: Sage.

Sisson, C. R. (1970). The effect of delayed comparison in the language laboratory on phoneme discrimination and pronunciation accuracy. *Language Learning*, 20, 69-88.

Sonia, B., & Lotfi, B. A. (2016). The importance of prosody in a proper English pronunciation for EFL learners. *Arab World English Journal*, 7(2), 316-327.

Spada, N., & Lightbown, P. M. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal*, 83(1), 1-22.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3(3), 243-261.

Shuster, L. I., & Durrant, J. D. (2003). Toward a better understanding of the perception of self-produced speech. *Journal of Communication Disorders*, 36(1), 1-11.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, 25(1), 76-89.

Stenfelt, S. (2011). Acoustic and physiologic aspects of bone conduction hearing. In Kompis M. & Caversaccio M (Eds.), *Implantable Bone Conduction Hearing Aids* (pp.10-21). Basel: Karger Publishers.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, 36(2), 131-145.

Sundström, A. (1998). Automatic prosody modification as a means for foreign language pronunciation training. In *Proceedings of ISCA Workshop on Speech Technology in Language Learning* (pp. 49-52), Marholmen, Sweden.

Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, 33(3), 263-290.

Tang, M., C. Wang, and S. Seneff, (2001). Voice transformations: from speech synthesis to mammalian vocalizations. In *Proceedings of the 7th European*

Conference on Speech Communication and Technology (pp.357-360), Aalborg, Denmark.

Tench, P. (1996). *The Intonation Systems of English*. New York: Cassell.

Tench, P. (2005). Teaching intonation. *Speak Out*, 34, 45-53.

Tench, P. (2011). *Transcribing the sound of English: A phonetics workbook for words and discourse*. Cambridge: Cambridge University Press.

Thompson, I. (1981). *Intonation practice*. Oxford: Oxford University Press.

Thomson, R. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In Kang, O., & Ginther, A. (Eds.), *Assessment in second language pronunciation* (pp. 11-29). London: Routledge.

Tonndorf, J. (1976). Bone conduction. In Keidel, W. D. & Neff, W. D. (Eds.), *Handbook of sensory physiology* (pp. 37–84), New York: Springer-Verlag

Vandergrift, L. and Goh, C. 2012. *Teaching and learning second language listening: Metacognition in action*, New York: Routledge.

Wang, G. Z. (2005). *English Pronunciation Intonation for Communication*. Beijing: Higher Education Press.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, 32(4), 539-552.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive mandarin tones. *Journal of the Acoustical Society of America*, 106(6), 3649-3658.

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.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8), 989-994.

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

Wei, R., & Su, J. (2012). The statistics of English in China: An analysis of the best available data from government sources. *English Today*, 28(3), 10-14.

Wennerstrom, A. K. (1997). *Discourse intonation and second language acquisition: Three genre-based studies*. Doctoral dissertation, University of Washington, Seattle, USA.

West, R. P., Young, K. R., & Spooner, F. (1990). Precision teaching: An introduction. *Teaching Exceptional Children*, 22(3), 4-9.

Weston, A. J., & Rousey, C. (1970). Voice confrontation in individuals with normal and defective speech patterns. *Perceptual and Motor Skills*, 30, 187-190.

White, O. R. (1986). Precision teaching—precision learning. *Exceptional Children*, 52(6), 522-534.

White House (2015). *FACT SHEET: President Obama's Precision Medicine Initiative*. Retrieved from https://obamawhitehouse.archives.gov/precision-medicine

Wibawa, S. C. (2017). The design and implementation of an educational multimedia interactive operation system using Lectora Inspire. *Elinvo (Electronics, Informatics, and Vocational Education)*, 2(1), 74-79.

Wichmann, A. (2015). Functions of intonation in discourse. In Reed, M., & Levis, J. (Eds.), *The handbook of English pronunciation* (pp.175-189). John Wiley & Sons.

Wilson, M.,& Knoblich, G. (2005) The case for motor involvement in perceiving

conspecifics. *Psychol. Bull*. 131, 460–473.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, 11(18), 729-732.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880-1882.

Wong, R. (1993). *Teaching pronunciation: Focus on English rhythm and intonation*. Englewood Cliffs, NJ: Prentice, HallRegents.

Wrembel, M. (2007). Metacompetence-based approach to the teaching of L2 prosody: Practical implications. In J. Trouvain & U. Gut (Eds.), *Non-native prosody: Phonetic description & Teaching Practice* (pp. 189-210). Berlin: Mouton de Gruyter.

Xia, Z. H., & Mu, F. Y. (2008). A study of boundary tone use at turn-taking in Chinese EFL learners' conversations. *Journal of PLA University of Foreign Languages*, 31(1), 48-54.

Xue, X.J. (2016). Judging foreign accent of Chinese-accented English speech by Chinese EFL learners and native speakers of English: A comparative study. *Modern Foreign Languages*.39(05):647-657+729-730

Yang, J. (2006). Inappropriate divisions of intonation groups in Chinese university students' read speech. *Modern Foreign Languages (Quarterly)*, 29(4), 409-417.

Yang, Y. (2016). *Improving the phonological working memory and english speaking skills of Chinese primary EFL learners with a verbotonal-based approach*. Doctoral dissertation, Suranaree University of Technology, Thailand.

Yang, M., & Mu, F. Y. (2011). Study on fluency-related IP-internal pauses in Chinese EFL learners' spontaneous speech. *Foreign Languages and Their Teaching*, (6), 16-21.

Yeager, J. (1966). Learning from own vs other voice by air or bone conduction. *Auditory Perception*, 23(2), 575-578.

Yoon, K. C. (2009). Synthesis and evaluation of prosodically exaggerated utterances. *Phonetics and Speech Sciences*, 1(3), 73-85.

Yuen, I., Davis, M. H., Brysbaert, M., & Rastle, K. (2010). Activation of articulatory information in speech perception. *Proceedings of the National Academy of Sciences*, 107(2), 592-597.

Zhao, J., Yuan, H., Leung, W. K., Meng, H., Liu, J., & Xia, S. (2013). Audiovisual synthesis of exaggerated speech for corrective feedback in computer-assisted pronunciation training. *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 8218-8222).Vancouver, Canada.

Zhang, L. (2015). An empirical study on the intelligibility of English Spoken by Chinese university students. *Chinese Journal of Applied Linguistics*, 38(1), 36-54.

Zhang, P. (2000). A Comparative Study on English and Chinese Intonation and the Meaning and Function of English Intonation. *Foreign Language Research* (4), 33-35.

Zhu, L. (2007) *The rhythm patterns of English spoken by Chinese and its pedagogical implications*. Doctoral dissertation, Minzu University of China, China

Zuccaro, L. (2013). Internal models of motor control: reach adaptation and task autism. *Meducator*,1(23), 12-14

APPENDIX A

A Questionnaire Survey on HUSE English Majors' Pronunciation Learning

Name _____ Age _____ Gender: M___ F___

1. Do you have a hearing problem?

   A. Yes B. No

2. I have studied English for _____ years.

3. Have you ever been abroad? Which country? How long did you stay there?

4. Have you ever received any English pronunciation training before?

5. How do you evaluate your English pronunciation?

   A. perfect B. good C. not bad D. bad E. very bad.

6. Do you think you need to improve your English pronunciation?

   A. I need to improve my pronunciation urgently.

   B. I need to improve my pronunciation.

   C. I think it's both OK for me to improve or not.

   D. I don't need to improve my pronunciation much.

   E. I totally don't need to improve my pronunciation.

7. How many hours do you usually spend on practicing English pronunciation out of class every week?

8. Do you have any problems in English pronunciation? (You can choose more than one)

   A. I don't have problems in English pronunciation.

   B. I have problems with vowels.

   C. I have problems with consonants.

   D. I have problems with stress.

   E. I have problems with rhythm.

   F. I have problems with intonation.

9. How do you usually practice your English pronunciation?

10. For the four English language skills, which one do you think is the most important?

A. Reading B. Writing C. Speaking D. Listening

E. They are of equal importance.

11. Which one of the following language learning aspects do you want to improve the most?

A. Pronunciation B. Vocabulary C. Speaking D. Translation

E. Grammar F. Listening G. Reading H. Writing

I. They are of equal importance.

12. Are you satisfied with the current English pronunciation teaching in class?

A. Strongly satisfied.

B. Satisfied.

C. Neutral

D. Dissatisfied.

E. Strongly dissatisfied.

13. What are your suggestions for improving the English pronunciation teaching in class?

APPENDIX B

Structured Questions for Interviews with HUSE English Pronunciation Teachers

Name _____ Age _____ Gender: M___ F___

1. You have been teaching English for__years.
2. You have been teaching English pronunciation for_years.
3. Are you confident in teaching English pronunciation to your students?
   A. Very confident
   B. Confident
   C. Neutral
   D. Unconfident
   E. Very unconfident
4. How do you teach pronunciation? (theories, teaching methods, teaching objectives, teaching plan, teaching materials, etc..)
5. From your perspective, what are the students' English pronunciation problems? (their characteristics, motivation, level, problems, etc..)
6. Do you have any suggestions for improving the teaching of English pronunciation in HUSE (textbooks, facilities, curriculum, teachers, etc..)?

APPENDIX C

A Sample of the English Intonation Training Materials

TRAINING SESSION ONE

Module I Tonicity: the placement of information focus

Instruction: Please read the following dialogues and try to produce the underlined sentences with proper intonation according to the context. Then listen to the models and compare with your own productions. Pay special attention to the placement of information focus.

Part 1: Narrow focus

(1) A: Where did Jack go yesterday?

B: Jack went to China yesterday.

(2) A: When did Jack go to China?

B: Jack went to China yesterday.

Part 2: Contrastive focus

(1) A: This donation is for him?

B: From him, not for him.

(2) A: You bought it before Christmas?

B: After Christmas, not before Christmas.

Part 3: Old and new information

(1) A: Can you give me a cigarette?

B: I thought you have quit smoking.

(2) A: What soup do you want?

B: I prefer beef soup.

Module II Tonality: chunking

Instruction: Please read the following dialogues and try to produce the underlined sentences (the punctuation has been removed) with proper intonation according to the

context. Then listen to the models and compare the models with your own productions. Pay special attention to the intonation boundaries.

Part 1: Attributive clauses

(1) A: Jane is my sister who lives in Canada

B: Where's your other sister Ella?

(2) A: Who is Jane? Is she your only sister?

B: Yes. Jane is my sister who lives in Canada

Part 2: Adverbials

(1) A: I will talk to the students in the garden

B: OK. I'm going to take them to the garden.

(2) A: I will talk to the students in the garden

B: OK. I'm going to take them to your office.

Part 3: Parallel structures

(1) A: Has she washed the dishes?

B: She washed and ironed her blouse

(2) A: What did she do to her blouse?

B: She washed and ironed her blouse

Module III Tone: the falls and rises in pitch

Instruction: Please produce the underlined sentence with proper intonation according to the context. Pay special attention to the falls and rises in pitch.

(1) Statement: falling tone

A: What are they doing?

B: They are waiting outside.

(2) Statement: rising tone

A: They are waiting outside?

B: No, they aren't. I think inside.

(3) Wh-question: falling tone

A: I bought a new phone.

B: How much?

A: 1000 dollars.

(4) Wh-question: rising tone

A: This new phone cost me 1000 dollars.

B: How much?

(5) Yes/no-question: falling tone

A: It would be nice to have a new kitchen.

B: Can we afford one? You know we can't even afford the food.

(6) Yes/no-question: rising tone

A: Will you be at the meeting?

B: I'm not sure now.

(7) Tag-question: falling tone

Well it's not very good, is it?

(Note: the speaker is sure that the hearer will agree)

(8) Tag-question: rising tone

It's snowing, isn't it?

(Note: the speaker is not sure.)

(9) Command: falling tone

A: What should I do next?

B: Add the seasoning.

(Note: the speaker intends to give an normal instruction)

(10) Exclamation: falling tone

A: I just got a promotion.

B: What good news!

(11) Implication: falling-rising tone

A: Can we set up an appointment?

B: I could see you on Tuesday. (but that might not suit you)

(12) Alternative question: falling tone + rising tone

A: Is Mary ready or does she need some more time?

B: She is ready now.

(13) Listing: falling tone+ falling tone +...+ rising tone

A: What fruits do you like?

B: I like apples, oranges and pears.

APPENDIX D

Intonation Perception Test

Note: The "#" mark before a sentence represents that this sentence is the word transcription of the recording. In the transcriptions, the stressed words were written in bold; intonation boundaries were marked with "//"; falling tones were marked with "\", rising tones with "/", and falling-rising tones with "\/".

Part I Tonicity: the placement of information focus

Instruction: Listen to the recordings in each item and choose the questions that best match the context.

(1) # I will talk to Henry next **week**.

A. Who will you talk to next week?

B. Will you write to Henry next week?

C. Will you talk to Henry next month?

D. Will you talk to Henry this week?

(2) # My cat is sleeping **on** the table.

A. Is your cat sleeping under the table?

B. Is your cat eating on the table?

C. Whose cat is sleeping on the table?

D. Is your dog sleeping on the table?

(3) # We are **going to** win.

A. Who will win the match?

B. Have you won the match?

C. It seems that you will fail.

D. I think you are not going to win.

(4) # She bought a **blue** skirt.

A. Did she buy a red skirt?

B. Did she steal a blue skirt.

C. Did she buy a blue blouse?

D. Did Linda buy a blue skirt?

(5) # I **never** drink soft-drinks.

A. Do you want cola?

B. Sam never drinks beverages?

C. You never sell beverages?

D. You never drink soup?

(6) # We heard **Mary** singing upstairs.

A. Did you hear anyone singing upstairs?

B. Who heard Mary singing upstairs?

C. What was Mary doing upstairs?

D. Where was Mary singing?

Part II Tonality: chunking

Instruction: Listen to the recordings and choose the best answer for the questions in each item.

(1) # The villagers, who were rich, gave up their rights

Question: Who gave up their rights?

A. All the villagers.

B. Only the rich villagers. The poor villagers didn't.

(2) # "The president", said that reporter "is lying".

Question: Who is lying?

A. That reporter.

B. The president.

(3) # I saw the man // with the binoculars.

Question: Who was taking the binoculars?

A. Me.

B. The man.

(4) # The man who read my works frequently came to visit me.

Question: Which of the following statements is true?

A. The man read my works frequently.

B. The man visited me frequently.

(5) # I'm going to clean // and repaint the bathroom

Question: What am I going to do?

A. Clean the bathroom and repaint the bathroom.

B. Do the cleaning, and then repaint the bathroom.

(6) # Competent women and men hold all of the good jobs in the firm.

Question: Who hold the good jobs in the firm?

A. Competent women and competent men.

B. All men and competent women.

Part III Tone: the falls and rises in pitch

Instruction: Listen to the recordings, and choose the best answer for the question in each item.

(1) # He works in a supermarket \

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: Where does he work?

Speaker 2:

B. Speaker 1:

Speaker 2: No, He has quit his job already.

(2) # She will come with us /

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: All of us will go to the party? What about Linda?

Speaker 2:

B. Speaker 1:

Speaker 2: I'm not quite sure. You'd better ask her.

(3) # I have been here for three years \

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: How long have you been here?

Speaker 2:

B. Speaker 1:

Speaker 2: No, almost four years already.

(4) # She is a student /

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: Is she a teacher or student?

Speaker 2:

B. Speaker 1:

Speaker 2: No, I think she is a teacher.

(5) # Which one? \

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: One of the dogs is a pitbull.

Speaker 2:

B. Speaker 1: The big dog is a pitbull.

Speaker 2: _____. I thought the smaller one is.

(6) # How many? \

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: I will invite some friends to come.

Speaker 2:

B. Speaker 1: Ten of my friends will join our dinner.

Speaker 2: _____. Our table is not that big enough.

(7) # How much? /

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: I bought a new phone yesterday.

Speaker 2:

B. Speaker 1: My new phone cost me 1000 dollars.

Speaker 2:

(8) Speaker 1: My girlfriend said goodbye to me.

Speaker 2: # What did you say?

Question: Which of the following statements about speaker 2 is true?

A. She didn't catch what speaker 1 had said, so she is asking for repetition.

B. She is asking speaker 1 what said to his girlfriend.

(9) # Did you tell her my secret? \

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: Linda also feels sorry for you.

Speaker 2:

B. Speaker 1:

Speaker 2: No, I didn't.

(10) Speaker 1: Will Jim and Mary be coming for the awards banquet?

Speaker 2: # Are they invited? \

Question: Which of the following statements is true?

A. Speaker 2 knows that Jim and Mary were not invited.

B. Speaker 2 doesn't know whether Jim and Mary are invited or not.

(11) # Are you hungry? /

Question: This utterance could occur in which of the following contexts?

A. Speaker 1:

Speaker 2: No, I'm still very full.

B. Speaker 1: Can I have something to eat?

Speaker 2: _____ Here you are. Have some chocolate first.

(12) # Did I tell you that? /

Question: This utterance could occur in which of the following contexts?

A. Speaker 1: I can't believe it! It happened, as you said.

Speaker 2:

B. Speaker 1:

Speaker 2: No, I think It's Tom.

(13) # That won't be big enough, will it? \

Question: Which of the following statements is true?

A. The speaker knows that it will not be big enough.

B. The speaker is not sure whether it will be big enough.

(14) #He teaches physics, doesn't he? \

Question: Which of the following statements is true?

A. The speaker knows that he teaches physics.

B. The speaker is not sure whether he teaches physics or not.

(15) # That's correct, isn't it? /

Question: Which of the following statements is true?

A. She is sure that it it correct.

B. She is not sure whether it is correct or not.

(16) #We could start with the kitchen, couldn't we? /

Question: Which of the following statements is true?

A. The speaker just makes a suggestion that they can start with the kitchen.

B. The speaker strongly believes that they should start with the kitchen.

(17) # Don't move! Raise your hands! \

Question: Which of the following statements about speaker 2's response is true ?

A. It's an encouragement.

B. It's a gentle command.

C. It's a question.

D. It's a strong command.

(18) Speaker 1: Where should I go next?

Speaker 2: # Turn left. \

Question: Which of the following statements about speaker 2's response is true ?

A. It's an encouragement.

B. It's a question.

C. It's a normal instruction.

D. It's a warning.

(19) Speaker 1: Do you remember the days we spent on that island?

Speaker 2: # What a marvellous holiday we had! \

Question: Which of the following statements about speaker 2 is true?

A. It's a question.

B. It's an exclamation.

C. It's an instruction.

D. It's a command.

(20) Speaker 1: She finished her study within three years.

Speaker 2: # How could that be! \

Question: Which of the following statements about speaker 2 is true?

A. It's a question.

B. It's an exclamation.

C. It's an instruction.

D. It's a command.

(21) Speaker 1: Do you believe his words?

Speaker 2: # Yes. \/

Question: What is speaker 2 going to say next?

A. I trust him.

B. But only partly.

(22) Speaker 1: What is the food like?

Speaker 2: # It looks nice. \/

Question: What is speaker 2 going to say next?

A. And the taste is good.

B. But it tastes bad.

(23) Speaker 1: # Would you prefer coffee or tea?

Speaker 2: I'd like coffee.

Question: What is the tone pattern of "coffee or tea" in speaker 1's speech?

A. Falling tone + rising tone

B. Rising tone + falling tone

(24) Speaker 1: This drawing is excellent.

Speaker 2: # Do you really mean that, or are you just being polite?

Question: What is the tone pattern of speaker 2's speech?

A. Falling tone + rising tone

B. Rising tone + falling tone

(25) Speaker 1:Which day are you free?

Speaker 2: # Friday/, Saturday/, Sunday\.

Question: Has speaker 2 finished the listing?

A. Yes.

B. No.

(26) Speaker 1: What colors can be used?

Speaker 2: # Blue/, white/, red/

Question: Has speaker 2 finished the listing?

A. Yes.

B. No.

APPENDIX E

IOC Analysis for the Intonation Perception Test

Part	Item	Experts			Mean	Result
		1	2	3		
Part I Tonicity	1	+1	+1	+1	1	□
	2	+1	0	+1	0.67	□
	3	0	+1	+1	0.67	□
	4	+1	+1	+1	1	□
	5	0	+1	+1	0.67	□
	6	+1	+1	+1	1	□
Part II Tonicity	1	+1	+1	+1	1	□
	2	0	+1	+1	0.67	□
	3	+1	0	+1	0.67	□
	4	+1	+1	0	0.67	□
	5	+1	+1	0	0.67	□
	6	+1	+1	+1	1	□
Part III Tone	1	0	+1	+1	0.67	□
	2	+1	+1	+1	1	□
	3	+1	+1	+1	1	□
	4	0	+1	+1	0.67	□
	5	+1	+1	+1	1	□
	6	+1	+1	0	0.67	□
Part III Tone	7	+1	+1	+1	1	□
	8	+1	+1	0	0.67	□
	9	+1	0	+1	0.67	□
	10	+1	+1	+1	1	□
	11	+1	+1	+1	1	□
	12	+1	+1	+1	1	□
	13	+1	+1	0	0.67	□
	14	+1	+1	+1	1	□
	15	+1	0	+1	0.67	□
	16	+1	0	+1	0.67	□
	17	+1	+1	+1	1	□
	18	+1	+1	0	0.67	□
	19	+1	+1	+1	1	□
	20	+1	+1	+1	1	□
	21	+1	+1	+1	1	□
	22	+1	+1	+1	1	□
	23	+1	+1	+1	1	□
	24	+1	+1	+1	1	□
	25	0	+1	+1	0.67	□
	26	0	+1	+1	0.67	□
Total	42	31	33	32	0.84	□

Notes:

1. +1= the item is congruent with the objective

2. -1= the item is not congruent with the objective

3. 0=uncertain about this item

4. Items mean scored higher than 0.5 (-1□IOC□1) can be accepted

Result of IOC:

(IOC=□R/N)

Item number: 42

R=31+33+32=96 (Scores from experts)

N=3 (Number of experts)

IOC=96/3= 32

Percentage: 32/38×100%=84.20%

The table above shows that the analysis result of IOC is 32, and the percentage is 84.2% which is higher than 80%, and the mean score of each item is higher than 0.5. Therefore, all of the items in this test are suitable for adoption to test students' intonation perception ability in this study.

APPENDIX F

Intonation Production Test

Part I Tonicity: the placement of information focus

Instruction: Please produce the underlined sentences according to the context. Pay special attention to the placement of information focus.

(1) A: Which team is going to win?

B: The red team is going to win.

(2) A: That mobile looks familiar.

B: It's your phone.

(3) A: Do you remember what he said?

B: I only care how he said it.

(4) A: Does she write books?

B: No, but she used to write books.

(5) A: Shall we walk there?

B: Yes. I like going on foot.

(6) A: Do you like winter?

B: No, I can't stand cold weather.

Part II Tonality: chunking

Instruction: Please produce the underlined sentences according to the context. The punctuation has been removed. Pay special attention to the boundaries between intonation groups.

(1) A: The villagers who like running live longer

B: Yes, I can tell. All the people in this village, old or young, like running very much.

(2) A: The defendant said the accuser should be punished

B: I agree. Obviously, it's the defendant's fault.

(3) A: She was talking to the man she met on the bus

B: She told me already. They first met at a party.

(4) A: Those who spoke quickly got an angry response

B: He always requires students to keep quiet in class.

(5) A: Imported apples and oranges are expensive

B: The price of apples is reasonable. But the oranges are domestic.

(6) A: Who will clean the table?

B: I'm going to clean and repair the bathroom

A: You don't need to repair the bathroom. Just clean the table.

Part III Tone: the falls and rises in pitch

Instruction: Please produce the underlined sentences with proper intonation according to the context or instructions.

(1) A: Most left handed people are creative.

B: I agree. I'm left handed and obviously I'm creative.

(2) A: The sun rises from the east and sets in the west.

B: But to me, I feel it rises from the west and sets in the east.

(3) A: We're going to have to let you go.

B: You are firing me?

B: Yes. You disappointed all of us.

(4) A: Anybody home?

B: Oh, Tony. It's you! Come on, in.

(5) A: Where are the kids staying?

B: They are staying with their grandmother.

(6) A: How long did it take to get there?

B: Fifty minutes drive.

(7) A: Ten of my friends will join our dinner.

B: How many? Our table is not that big.

(8) A: Jenny won a big prize.

B: What did she do?

C: I said she won a big prize.A: Answer me!

(9) Will you marry me ?

B: I will, but please give me some time.

(10) A: I'm sorry. I really didn't want to hurt her. It was not on purpose.

B: But did you hurt her?

(11) A: Was she pleased to see you?

B: Yes, sure she was.

(12) A: Can you speak Chinese?

B: No, I can't. It's so difficult.

(13) A: Why did I only get a C?

B: Because you made a lot of mistakes, didn't you.

(14) A: She is pretty smart, isn't she? *(Note: The speaker is asking for agreement)*

B: Yes, she's always smart.

(15) A: You are Japanese, aren't you? *(Note: The speaker is quite sure)*

B: Yes, I am. How did you know?

(16) A: We have met before, haven't we? *(Note: The speaker is not sure)*

B: No, I think we haven't.

(17) A: Move out of my way!

B: Why are you shouting at me!

(18) A: Stop! I told you many times. Don't feed the dog from the table!

B: Alright. Don't be angry. I'll never do that again.

(19) A: I just won the lottery!

B: Why are you yelling at me?

A: I'm sorry. I'm just too excited.

(20) A: He has donated all of his properties.

B: He's such a kind soul!

(21) A: What time should I come in tomorrow?

B: Can you come in at 3?

A: I can...

B: So what?

A: But should I? The meeting starts at 1.

(22) A: Will their parents be coming to the dinner?

B: They're invited.

A: But?

B: They refused.

(23) A: Is something up?

B: Was that a knock at the door, or am I imaging things?

(24) A: He was very rude, wasn't he?

B: Is he always like that, or had something upset him?

(25) A: Do you have something to recommend?

B: We have fried chicken, hamburger, French fries...

A: OK. A hamburger please.

(26) A: We can paint it in red, white, blue...

B: Red and blue.

A: I haven't finished yet. We can also choose brown, purple, and green.

APPENDIX G

IOC Analysis for the Intonation Production Test

Part	Item	Experts			Mean	Result
		1	2	3		
Part I Tonicity	1	+1	+1	+1	1	
	2	+1	+1	+1	1	
	3	+1	+1	+1	1	
	4	0	+1	+1	0.67	
	5	+1	+1	+1	1	
	6	+1	+1	+1	1	
Part II Tonicity	1	+1	+1	+1	1	
	2	+1	0	+1	0.67	
	3	+1	+1	+1	1	
	4	+1	+1	+1	1	
	5	+1	+1	+1	1	
	6	0	+1	+1	0.67	
Part III Tone	1	+1	+1	+1	1	
	2	+1	+1	+1	1	
	3	+1	+1	+1	1	
	4	+1	0	+1	0.67	
	5	+1	+1	+1	1	
	6	+1	+1	+1	1	
	7	+1	+1	+1	1	
Part III Tone	8	+1	+1	+1	1	
	9	+1	0	+1	0.67	
	10	+1	+1	+1	1	
	11	+1	+1	+1	1	
	12	+1	+1	+1	1	
	13	+1	+1	+1	1	
	14	0	+1	+1	0.67	
	15	+1	+1	+1	1	
	16	+1	+1	0	0.67	
	17	+1	+1	+1	1	
	18	+1	+1	+1	1	
	19	0	+1	+1	0.67	
	20	0	+1	+1	0.67	
	21	+1	+1	+1	1	
	22	+1	+1	+1	1	
	23	+1	+1	+1	1	
	24	+1	+1	+1	1	
	25	+1	+1	0	0.67	
	26	+1	+1	+1	1	
Total	38	33	35	36	0.91	

Notes:

1. +1= the item is congruent with the objective

2. -1= the item is not congruent with the objective

3. 0=uncertain about this item

4. Items mean scored higher than 0.5 (-1□IOC□1) can be accepted

Result of IOC:

(IOC=□R/N)

Item number: 38

R=33+35+36=104 (Scores from experts)

N=3 (Number of experts)

IOC=104/3= 34.7

Percentage: 34.7/38×100%=91.3%

The table above shows that the analysis result of IOC is 34.7, and the percentage is 91.3% which is higher than 80%, and the mean score of each item is higher than 0.5. Therefore, all of the items in this test are suitable for adoption to test students' intonation production ability in this study.

APPENDIX H

CALL Software Evaluation Form

Instruction for raters: Please review the courseware and evaluate its feasibility according to the criteria below by using the 5-point Likert scale where "1= very bad; 2=bad; 3= average; 4= good; 5= very good".

Title of software package / program: English Intonation Training Courseware					
Criteria		**Experts**			**Mean**
		1	**2**	**3**	
Functionality	Ease of getting started	5	5	5	5
	Ease of navigation	5	5	5	5
	Flexibility of use	5	4	5	4.7
	Range of functions appropriate to purpose and content	5	5	5	5
	Level of student interaction	4	4	3	3.7
Media Content	Appropriate mix of text, images, sound, video	5	4	4	4.3
Quality of linguistic/ cultural content	Grammar	5	5	4	4.7
	Thematic vocabulary	4	4	4	4
	Cultural insights	5	4	5	4.7
	Functions	4	5	4	4.3
	Strategy training	4	4	4	4
	Accuracy of language used	4	5	5	4.7
	Insights into language learning skills	4	5	5	4.7
Relevance	Relevance to scheme of work	5	5	5	5
Relevance	Relevance to national / regional / departmental programmes of study	4	4	5	4.3
	Relevance to National Curriculum Attainment Targets	5	5	4	4.7
Outcomes	Clarity of the anticipated learning outcomes	4	4	3	3.7
	Ability of software to raise standard of student achievement beyond that expected from alternative resources	4	5	5	4.7
	Efficient use of student time	5	5	4	4.7
	Efficient use of teacher preparation time	3	3	3	3
Total		89	90	87	88.7

As can be seen from the table, the total mean score of the courseware as awarded by the three experts is 88.7 (with the total score of 100). The percentage is 88.7%, which is higher than the 80% standard. The average score for each criterion is higher than 3.5 except the last one. The result shows that this courseware is feasible for English intonation training with high degree of accessibility.

APPENDIX I

A Questionnaire on Students' Perceptions of Learning English Intonation Through Pitch Modification (English Version)

This questionnaire is designed to gather information about your perceptions of teaching English intonation through pitch modification. Please choose the degree of your agreement on each of the following statements by ticking the corresponding box with "□".

5 = Strongly agree

4 = Agree

3 = Neutral

2 = Disagree

1 = Strongly disagree

No.	Statements	Rating scales				
		1	2	3	4	5
1	I'm satisfied with this training.					
2	The contents of this training are what I want to learn.					
3	The courseware was well-designed.					
4	The training schedule was well-arranged.					
5	I like this kind of pronunciation training.					
6	Pitch modification is effective for solving my own specific pronunciation problems.					
7	Pitch modification for pronunciation learning enhanced my confidence of speaking English.					
8	Pitch modification for pronunciation learning changed my cognition on pronunciation learning.					
9	Pitch modification for pronunciation learning is interesting.					
10	I prefer to use my modified speech for pronunciation learning than the native speaker's model.					

A Questionnaire on Students' Perceptions of Learning English Intonation Through Pitch Modification

(Chinese Version)

基于语调调整的英语语调培训满意度调查问卷

感谢你参加此次英语语调培训。本调查问卷旨在收集你对本次培训的满意度情况。请根据你的实际感受填写本问卷调查表, 按照你对表中每一条陈述的同意度，在后边的数字表中打"□"， 其中：

1 =非常不同意

2 =不同意

3 =不确定

4 =同意

5 =非常同意

题号	内容	同意程度				
		1	2	3	4	5
1	我对本次的语调培训很满意					
2	本次培训的内容是我想学习的					
3	本次培训的课件设计的很合理					
4	本次培训的时间安排很恰当					
5	我喜欢这种语调培训方式					
6	语调调整法学习发音能更有针对性的解决我自己的发音问题					
7	语调调整法学习发音提高了我讲英语的自信					
8	语调调整法学习发音改变了我对发音学习的认知					
9	语调调整法学习发音很有趣					
10	相比母语者的示范我更愿意选择听我调整过的语音来学习发音					

APPENDIX J

IOC Analysis for a Questionnaire on Students' Perceptions of Learning English Intonation Through Pitch Modification

Item	Experts			Mean	Results
	1	2	3		
Q1	0	+1	+1	0.67	√
Q2	+1	+1	0	0.67	√
Q3	+1	0	+1	0.67	√
Q4	+1	+1	+1	1	√
Q5	+1	+1	+1	1	√
Q6	+1	+1	+1	1	√
Q7	0	+1	+1	0.67	√
Q8	+1	+1	+1	1	√
Q9	+1	+1	+1	1	√
Q10	0	+1	+1	0.67	√
Total	7	9	9	0.83	√

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item
5. Items mean scored higher than 0.5 (-1 □IOC□ 1) can be accepted

Result of IOC:

(IOC=□R/N)

Item number: 10

R=7+9+9=25 (Scores from experts)

N=3 (Number of experts)

IOC=25/3= 8.3

Percentage: 8.3/10×100%=83.3%

The table above shows that the analysis result of IOC is 8.3, and the percentage is 83.3% which is higher than 80%, and the mean score of each item is higher than 0.5. Therefore, the items are suitable for adoption to investigate students' opinions of the English intonation teaching in this study.

APPENDIX K

Semi-structured Interview Questions on Students' Perceptions of Learning English Intonation Through Pitch Modification (English Version)

1. Are you satisfied with training? Why or why not?
2. Do you think your English pronunciation was improved after this training? If yes, in what aspects? If no, why?
3. Did this training change your cognition in English pronunciation learning? If yes, how? If no, why?
4. Do you think pitch modification for intonation learning is effective for solving your own specific problems in pronunciation?
5. Do you think pitch modification for pronunciation learning is interesting?
6. Which model do you prefer for pronunciation learning? Your modified speech as model or native-speaker model?
7. What else do you want to say about this training?

Semi-structured Interview Questions on Students' Perceptions of Learning English Intonation Through Pitch Modification (Chinese Version)

基于语音调整的英语语调培训访谈问题

1. 你对这次培训满意吗？为什么？
2. 你认为本次培训是否提高了你的英语发音？如果有，在哪些方面？如果没有，为什么？
3. 这次培训是否改变了你对语调和英语发音学习的看法？
4. 你认为这种语调学习方法能更有针对性的解决你自己的发音问题吗？
5. 你认为这种语调学习方法有趣吗？
6. 相比其他语调学习方法，你是否更愿意用这种方法学习语调？
7. 关于本次培训，你还有其他要补充的吗？

APPENDIX L

IOC Analysis for the Semi-structured Interview Questions on Students' Perceptions of Teaching English Intonation Through Pitch Modification

Item	Experts			Mean	Results
	1	2	3		
Q1	+1	0	+1	0.67	□
Q2	0	+1	+1	0.67	□
Q3	+1	+1	+1	1	□
Q4	+1	0	+1	0.67	□
Q5	+1	+1	+1	1	□
Q6	+1	+1	+1	1	□
Q7	+1	+1	0	0.67	□
Total	6	5	6	5.7	□

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item
4. Items mean scored higher than 0.5 (-1□IOC□1) can be accepted

Result of IOC:

(IOC=□R/N)

Item number: 7

R=6+5+6=17 (Scores from experts)

N=3 (Number of experts)

IOC=7/3= 5.7

Percentage: 5.7/7×100%=81.4%

The table above shows that the analysis result of IOC is 5.7, and the percentage is 81.4% which is higher than 80%, and the mean score of each item is higher than 0.5. Therefore, the items are suitable for adoption to investigate students' opinions of the English intonation teaching in this study.

CURRICULUM VITAE

Zhongmin Li was born in 1988 in Linyi City, Shandong Province, China. He obtained his Bachelor of Arts degree in English Language and Literature from Jilin Agricultural Science and Technology University, China, in 2011. In 2013, he graduated from Northeast Normal University, China, with Master of Arts degree in English translation and interpreting.

Since July, 2013, he has been working as a teacher at School of Foreign Languages, Hunan University of Science and Engineering, China. From 2015 to 2016, he worked as a volunteer Chinese teacher in Buriram Rajabhat University, Thailand. In 2016, he was enrolled in the Ph.D. program of English language studies in the School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology, Thailand. His research interests are experimental phonetics, translation and interpreting, and computer assisted language learning.