A VERBOTONAL-BASED APPROACH TO PHONETIC CORRECTION OF A SELECTION OF ENGLISH VOWELS FOR CHINESE EFL LEARNERS

Fengwei Wen

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

หลักการเวอร์เบอร์โทนอลเพื่อแก้ไขการออกเสียงสระภาษาอังกฤษสำหรับผู้เรียนชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ

นายเฟิงเหว่ย เวิน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2562

A VERBOTONAL-BASED APPROACH TO PHONETIC CORRECTION OF A SELECTION OF ENGLISH VOWELS FOR CHINESE EFL LEARNERS

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

(Asst. Prof. Dr. Adcharawan Buripakdi)
Chairperson

(Assoc. Prof. Dr. Punyathon Sangarun)
Member (Thesis Advisor)

(Asst. Prof. Dr. Harald Kraus)
Member

(Asst. Prof. Dr. Arjuna Chaiyasena)
Member

(Dr. Supanit Kulsiri)
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

เฟิงเหว่ย เวิน : หลักการเวอร์เบอร์โทนอลเพื่อแก้ไขการออกเสียงสระภาษาอังกฤษสำหรับผู้เรียนชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ (A VERBOTONAL-BASED APPROACH TO PHONETIC CORRECTION OF A SELECTION OF ENGLISH VOWELS FOR CHINESE EFL LEARNERS) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.ปัณญาธร แสงอรุณ, 273 หน้า

งานวิจัยนี้มีจุดประสงค์สองประการ (ก) เพื่อช่วยให้ผู้เรียนชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ โดยไม่ได้เรียนภาษาอังกฤษเป็นวิชาเอก สามารถออกเสียงสระภาษาอังกฤษต่อไปนี้ /ɪ/, /i:/, /e/, /æ/, /ʊ/ และ /u:/ ได้ถูกต้องโดยใช้หลักการของทฤษฎีเวอร์เบอร์โทนอล (ข) เพื่อศึกษาการประยุกต์ใช้ หลักการของคอเล็คทีฟออฟติมอล (corrective optimal) ของทฤษฎีเวอร์เบอร์โทนอล ในงานวิจัยนี้กลวิธีที่ใช้ในพัฒนาการออกเสียง คือ การเรียนรู้การออกเสียงตามหลักการของทฤษฎีเวอร์เบอร์โทนอล (verbotonal-based pronunciation learning, VTPL) นั่นคือ งานวิจัยนี้ศึกษาผลของการใช้หลักการคอเล็คทีฟออฟติมอล ที่มีต่อการออกเสียงสระภาษาอังกฤษของผู้เรียนชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ และศึกษาว่าผู้เรียนดังกล่าวสามารถใช้หลักการคอเล็คทีฟออฟติมอลด้วยตนเองได้หรือไม่

ผู้เข้าร่วมการวิจัยครั้งนี้คือนักศึกษาชั้นปีที่ 1 ที่ไม่ได้เรียนภาษาอังกฤษเป็นวิชาเอกของมหาวิทยาลัยแห่งหนึ่งในประเทศจีนจำนวน 76 คน ผู้เข้าร่วมงานวิจัยทั้งหมดได้รับการสุ่มเพื่อแยกเป็นกลุ่มทดลองและกลุ่มควบคุม ผู้เข้าร่วมงานวิจัยในกลุ่มทดลองได้รับการฝึกแบบดิจิตอลฟิลเทอร์รังค์ตามหลักการของทฤษฎีเวอร์เบอร์โทนอล ในขณะที่ผู้เข้าร่วมงานวิจัยในกลุ่มควบคุมได้รับการฝึกแบบดั้งเดิม โดยผู้เข้าร่วมงานวิจัยทั้งสองกลุ่มใช้เนื้อหาในการฝึกที่เหมือนกัน ผู้วิจัยใช้วิธีการวิจัยแบบผสานวิธี (mixed-method) ดังนั้นจึงมีการเก็บข้อมูลในเชิงปริมาณและข้อมูลในเชิงคุณภาพ

ผลการการวิจัยพบว่า 1) ผู้เข้าร่วมการวิจัยสามารถออกเสียงสระได้ดีที่สุดเมื่อใช้ตัวกรองที่มีความถี่ต่างกันแบบไม่ต่อเนื่องตามหลักการคอเล็คทีฟออฟติมอล (corrective optimals) (2) คอเล็คทีพออฟติมอล (corrective optimals) มีความแคบ มีความละเอียด มีความหลากหลาย และมีรูปแบบที่แตกต่างกันมากกว่า เนทีฟออฟติมอล (native speaker optimals) (3) ผลจากการทดสอบก่อนการเรียนและการทดสอบหลังการเรียน พบว่าผู้เข้าร่วมงานวิจัยในกลุ่มทดลองมีการพัฒนาความสามารถอย่างมีนัยยะสำคัญ ทั้งในแง่ของความสามารถในด้านการรับรู้เสียงและการออกเสียง นอกจากนี้ยังพบว่าทักษะการพูดโดยรวมของผู้เข้าร่วมงานวิจัยในกลุ่มทดลอง ในด้านที่ทำให้ผู้ฟังสามารถเข้าใจความหมายของเสียงที่ผู้เข้าร่วมการวิจัยพูด และความคล่องแคล่วในการออกเสียงของ

ผู้เข้าร่วมการวิจัยยังเพิ่มขึ้นอีกด้วย ซึ่งประเด็นนี้ผู้วิจัยไม่ได้กำหนดไว้ว่าจะศึกษาในตอนแรก ผลการวิจัยในประเด็นนี้แสดงให้เห็นว่าคอเล็คทีพออฟติมอลเพียงอย่างเดียว สามารถพัฒนาการออกเสียงสระภาษาอังกฤษที่ได้รับเลือกเพื่อศึกษาในงานวิจัยนี้ของผู้เข้าร่วมการวิจัย (4) ผู้เข้าร่วมการวิจัยในกลุ่มทดลองและกลุ่มควบคุมมีความสามารถในการรับรู้เสียงและการออกเสียงไม่แตกต่างกันในการทดสอบก่อนเรียน แต่หลังจากการฝึกด้วยวิธีคอเล็คทีพออฟติมอล ผู้เข้าร่วมการวิจัยในกลุ่มทดลอง มีความสามารถเหนือกว่าผู้เข้าร่วมการวิจัยในกลุ่มควบคุม ในการรับรู้เสียงและการออกเสียง ของสระ 5 สระจากสระที่ได้รับเลือก 6 สระ โดยวัดจากการอ่านออกเสียงประโยคและการเล่าเรื่อง (5) ผู้เข้าร่วมการวิจัยส่วนใหญ่มีทัศคติเชิงบวกต่อการใช้ VTPL

โดยสรุป ผลงานวิจัยที่ค้นพบแสดงให้เห็นว่า หลักการคอเล็คทีฟออฟติมอลของทฤษฎีเวอร์เบอร์โทรนอล มีประสิทธิภาพมากกว่าวิธีการสอนแบบดั้งเดิม ในการส่งเสริมความสามารถในการออกเสียงสระภาษาอังกฤษที่ได้รับเลือกเพื่อศึกษาในงานวิจัยนี้ ผลงานวิจัยนี้มีความสำคัญต่อทฤษฎีเวอร์เบอร์โทรนอลและงานวิจัยทางด้านการจัดการเรียนการสอนภาษาต่างประเทศ

สาขาวิชาภาษาต่างประเทศ	ลายมือชื่อนักศึกษา Fengwei Wen
ปีการศึกษา 2562	ลายมือชื่ออาจารย์ที่ปรึกษา
	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

FENGWEI WEN: A VERBOTONAL-BASED APPROACH TO PHONETIC CORRECTION OF A SELECTION OF ENGLISH VOWELS FOR CHINESE EFL LEARNERS. THESIS ADVISOR : ASSOC. PROF. PUNYATHON SANGARUN, Ph.D., 273 PP.

OPTIMALS/CORRECTIVE OPTIMALS/VERBOTONAL THEORY/ /CHINESE EFL LEARNERS/ ENGLISH VOWELS

This research study has a two-fold objective: (a) to facilitate the acceptable production of a selection of English vowels, i.e. /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/ for Chinese non-English major EFL learners through the application of the verbotonal theory and (b) in a more technical perspective, to use the facilitation of acceptable pronunciation of the above-mentioned vowels to investigate the nature and application of the concept of corrective optimals in verbotonal theory. The approach used to facilitate pronunciation is referred to as the VTPL approach. Specifically, this study investigates (a) the nature of the corrective optimals of the target English vowels in the participants and (b) whether using corrective optimals alone suffices to bring about acceptable production of the target vowels using a self-managed approach.

Participants were 76 first-year non-English major undergraduates enrolled in a local university in China. They were randomly assigned to the experimental and control groups. Participants in the experimental group received digital filtering training using the VTPL approach while participants in the control group received no such training but followed a traditional approach using exactly the same materials. Both quantitative and qualitative data were gathered using a mixed methods approach.

The findings indicate that: (1) The target English vowels were best produced when filtered through corrective optimals consisting of discontinuous multiband filters containing a low-frequency component. (2) The corrective optimals proved to be narrower, finer and more diverse than the classical native speaker optimals. (3) The experimental group showed statistically significant improvements in terms of both perception and production of the target vowels from the pretests to the posttests. Surprisingly, after treatment, the experimental group also improved significantly in the areas of comprehensibility, fluency and pronunciation, although these areas were never deliberately focused upon. (4) Initially at the same level as the control group, the experimental group outperformed the control group, after training, in five out of six vowel sounds in both perception and production tasks as well as in comprehensibility, fluency and pronunciation in sentence-reading and storytelling tasks. (5) A significant majority of participants had positive attitudes toward the use of the VTPL approach.

In summary, findings demonstrate that the VTPL approach, focusing on corrective optimals only, was effective in itself as well as being more effective in comparison to the traditional approach in enabling the acceptable production of the target English vowels. These findings have important implications for verbotonal theory and research in foreign language learning and teaching.

School of Foreign Languages — Student's Signature Fengwei Wen

Academic Year 2019 — Advisor's Signature ____________

Co-Advisor's Signature ____________

ACKNOWLEDGEMENTS

I would never have been able to complete my thesis without help, support and encouragement from many people who deserve my sincere thanks and gratitude.

First and foremost, I would like to express my deepest gratitude to my advisors, Prof. Dr. Andrew Lian and Assoc. Prof. Dr. Punyathon Sangarun, for their expert guidance, invaluable inspiration, timely feedback and eternal optimism. Without their unrelenting guidance and encouragement, this thesis would not have materialized.

I would like to show my highest appreciation to all the members of my thesis examining committee for their constructive comments and insightful suggestions: Asst. Prof. Dr. Adcharawan Buripakdi, chair of the committee, Asst. Prof. Dr. Harald Kraus, Asst. Prof. Dr. Arjuna Chaiyasena and Dr. Supanit Kulsiri.

I want to express my deep and sincere gratitude to the faculty members of the School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology, Thailand, for giving me a solid research foundation, which enabled me to conduct my research project. My warm thanks go as well to the staff of the School of Foreign Languages for their help and patience.

I would like to offer my special thanks to all the participating experts for their invaluable assistance in checking the research instruments and rating the production performances. Special thanks are extended to the language lab staff for taking care of the language laboratories during the training process. My thanks also go to the participants from Gannan Normal University who were very cooperative since the inception of the training.

My sincere thanks and appreciation go to my employer, Gannan Normal University, for giving me the opportunity to pursue my research interests. Special thanks also to my colleagues in the School of Foreign Languages, for their help and support during my doctoral studies.

I wish to express my profound gratitude to my parents, my parents-in-law, and my sister, for their constant love, encouragement and moral support. Last but not least, my very special thanks go, of course, to my wife and my son, for always being on my side, giving me the courage and strength to move forward. Without their unwavering love and support, my years of studies in Thailand would not have been possible, and I thank them all from the bottom of my heart.

Fengwei Wen

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

LIST OF TABLES

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF ABBREVIATIONS

CALL	Computer-Assisted Language Learning
CECR	College English Curriculum Requirements
CET	College English Test
EFL	English as a Foreign Language
ESL	English as a Second Language
FL	Foreign Language
GNU	Gannan Normal University
IOC	Index of Item-Objective Congruence
JASP	Jeffreys's Amazing Statistics Program
L1	First Language
L2	Second Language
MOE	Ministry of Education
TESOL	Teaching English to Speakers of Other Languages
VTPL	Verbotonal-based Pronunciation Learning

CHAPTER 1

INTRODUCTION

The current study focuses on the development and testing of a self-managed approach based on the verbotonal theory designed to assist with the pronunciation of English language sounds by foreign language learners. To this end, and for the purposes of experimentation, this approach will be applied and tested in the context of a study that seeks to assist with the acceptable production of a selection of English vowel sounds by Chinese EFL learners. Specifically, the vowels to be studied are: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/. This chapter begins with a background section that discusses the importance of English, pronunciation problems of English vowels and the verbotonal theory of phonetic correction. After that, it focuses on the statement of the problem, purposes of the study, research questions, significance of the study, definitions of key terms and summary.

1.1 Background of the study

In the era of globalization, as an important medium for international communication in a wide range of contexts, English has been paid remarkably more attention than ever before. According to Graddol (2006), never has a language been as widely used as English is today and furthermore, English and globalization have helped acceleration of each other throughout the world. As a result, English is becoming the global working language for the entire community of nations, whereas English is actually the mother tongue of none. L2 users of English now greatly outnumber the total number of L1 users (Baker, 2015), and the number of users of English is still on the rise (Gowans, 2012). Moreover, communication among L2 speakers of English has exceeded the number of those that involve interactions in English between L1 and L2 speakers (Baumgarten & House, 2010). To serve the purposes of different cultures, English is in its new status of English as a foreign language (EFL), English as a second language (ESL), English as a lingua franca (ELF), English as an international language (EIL), English as a world language (EWL) and English as a global language (EGL) (Galloway, 2017; House, 2009, 2010, 2014; Jenkins, 2006, 2009, 2012, 2015; Matsuda, 2012; McArthur, 2002; Northrup, 2013; Pan, 2015; Rose & Galloway, 2019; Seidlhofer, 2011; Walker, 2010; Zacharias & Manara, 2013). Therefore, a good command of the English language is highly desirable, and as a consequence, the learning of English is sweeping the world (Godwin-Jones, 2018).

Language has traditionally been considered to consist of four skills: listening, reading, speaking and writing. Among the four language skills, speaking occupies a leading position in language learning (Pleuger, 2001) and is highlighted as an essential element in thinking and achieving academic success (Goh & Burns, 2012). At the same time, speaking is viewed as the most challenging skill for learners of English to master (Pawlak, Waniek-Klimczak, & Majer, 2011). Pronunciation is deemed as a sub-skill of speaking, which plays a vital role in determining the clarity of expression in communication. According to Trofimovich and Isaacs (2012), pronunciation is the key factor which seriously affects the communication process in ESL/EFL learners. Research has suggested that intelligible pronunciation is important for effective communication and that the majority of communication breakdowns are due to pronunciation errors (Boyer, 2002; Gilakjani, 2017; Jenkins, 2000; Thir, 2016; Trofimovich & Isaacs, 2012). However, for most L2 learners, pronunciation is frequently recognized as a seriously difficult aspect of language learning when considering the significant influence of learner's L1, age, motivation, pedagogic instruction, target language exposure, etc. (Gilakjani & Ahmadi, 2011).

There are two main aspects of pronunciation: the segmentals (vowels and consonants) and the suprasegmentals (stress, rhythm, intonation, etc.). According to a number of studies which have been conducted to investigate ESL/EFL learners' problems of pronunciation in English, both segmental and suprasegmental issues are documented (Caspers, 2010; Derwing & Munro, 2015; No, 2003; Ohata, 2004; Reed & Levis, 2019; Swan & Smith, 2001). In the field of second language acquisition (SLA), the importance of vowels in English pronunciation cannot be overlooked (Bohn & Munro, 2007). However, within the range of potential segmental errors, a majority of

L2 studies have shown that L2 vowels that are not present in L1 cause considerable difficulties which make L2 learners strive to produce these vowel sounds and fail to do so acceptably (Chan, 2011; Chan & Li, 2000; Colantoni, Steele, & Escudero, 2015; Flege, Bohn, & Jang, 1997; Lai, 2010; Ohata, 2004; Wong, 2015).

This phenomenon is also true in the case of Chinese learners using English as a foreign language. In general, Chinese speakers find English difficult to pronounce (Chang, 2001), and they are particularly weak in pronunciation (Qiang & Wolff, 2011). Previous research on segmental errors has reported that some English vowels and vowel contrasts remain difficult and still pose problems for Chinese EFL learners, e.g. /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/ (Chang, 2001; Jia, Strange, Wu, Collado, & Guan, 2006; Wang & Heuven, 2004; Zhang & Yin, 2009; Zheng & Liu, 2018). Another reason for choosing vowels as the target sounds is that vowels result from smooth airflow (Ball & Rahilly, 1999), and therefore, they are more regular in structure and more manageable digitally as compared to consonants.

The classroom has always been regarded as the privileged place for learning and teaching and as a place for synchronous pre-determined activity, where tight control still tends to be exercised by authorities and academic institutions such as universities, colleges and schools and usually in the name of quality control and the setting of standards, thus subjecting learners to one-size-fits-all pre-determined and pre-organized procedures (Lian, 2014; Lian & Pineda, 2014; Sangarun, 2014). However, language learners have different purposes and will need to be able to have ways of responding to these purposes. In other words, individuals are faced with the inability to perceive something which presents no difficulties to someone else. The one-size-fits-all approach to the learning/teaching of language skills is, necessarily, insufficient. It is further understood, increasingly, that more attention needs to be given to the existence of individual differences and that learners must be supported differently for optimal outcomes (Dörnyei, 2005; Ellis, 2008).

We as teachers, friends, or other human beings, simply do not know the details of what is going on inside someone's brain. According to Gattegno (1987), only awareness is educable. Systems for learning are primarily centered on changing our operational histories with awareness-raising as the key enabling learners to make meanings as well as create new personal knowledge (Lian, 2004, 2014; Lian & Lian, 1997; Lian & Sussex, 2018). If we are unaware of a phenomenon in some way, then it might be the evidence that it does not exist in our operational histories (Lian, 2000; Mason, 1998). The most important aspect of awareness-raising seems to be modification or manipulation of input for students to bypass learning difficulties caused by their operational histories in the process of meaning-making and to create new personal knowledge for the field of awareness later (Lian, 1987, 2013; Schmidt, 2010). Thus, at least part of the job of teaching is to find ways of manipulating the input so as to optimize the learning outcomes that they are trying to achieve.

Various phonetic treatment paradigms have been used by linguists and speech pathologists as ways of raising awareness so as to improve perception and production of L2/FL pronunciation. One paradigm that has been demonstrated to be worth trying in phonetic correction is the verbotonal approach, which helps improve learners' pronunciation concerning both segmental and suprasegmental features by exposing them to the materials using digital filtering (Asp, 1972, 2006; Asp & Kline, 2012; Hang, 2012; He, 2014; Lian, 1980; Mildner & Bakran,2001; Mildner & Tomić, 2007; Wu, 2013; Zhang, 2006). Moreover, several studies have reported good results, even after a small number of training sessions (Mildner & Bakran, 2001; Mildner & Tomić, 2007; Tomić, Kiš, & Mildner, 2011). From the above studies, it is suggested that the verbotonal approach might be a promising approach for phonetic correction. However, up to now, it has been insufficiently explored.

The verbotonal approach is an auditory rather than an articulatory approach, developed by Professor Petar Guberina in the 1950s, a researcher who has been specifically keen on speech perception. The approach was first developed for the training of the hearing impaired or deaf people and later used for improving the hearing of all speech sounds in the foreign languages for normal hearing people. The ultimate goal of this approach is the rewiring of the brain's neural connectivity and the development of self-correction skills when we speak (Asp, 2006; Asp, Kline, & Koike, 2012; Guberina, 1989).

The basic idea in Guberina's theory is the idea of optimality. The traditional idea of optimality consists of multimodal techniques, including optimal frequency bands, intonation, sound environment, placement of the sound in optimal and non-optimal contexts, gross motoric movement, body position etc. Based on the research of Guberina and other classical verbotonal practitioners, they do not talk about the optimal frequency bands alone, and instead, the verbotonal theory of phonetic correction has always been implemented in combination with multidimensional techniques for providing optimal learning condition (Asp, 2006; Asp & Kline, 2012). However, in order to figure out whether the corrective optimal frequency bands alone will be sufficient to bring about the necessary improvements, the present study will focus on only one aspect of optimality, i.e. corrective optimals.

According to the verbotonal theory, each language has its own set of optimal frequency bands for its sound repertoire. It proposes that segmental articulation causes comparatively little trouble if the optimal quality of the sound has been perceived. This optimal frequency band is described as the optimal octave with a one-octave bandwidth yielding the best identification score for a specific speech sound (phoneme) when the filtered phoneme sounds similar to the same unfiltered phoneme (Asp, 2006). In Guberina's (1972) experiment, he used one-octave-band filters to identify the possible optimal frequency range for the perception of a particular English sound (a vowel or a consonant) for native speakers of English. According to him, "all phonemes are contained in one phoneme" because the perception of different phonemes (e.g. the vowels /ɪ/, /e/, /ə/, /ɒ/, /ʊ/) was possible from the same phoneme (e.g. the vowel /ɪ/) depending on which octave bandwidth the phoneme was filtered through.

The following lists Guberina's optimal octaves for the six target vowels ("fcenter" indicates the center frequency): the optimal octave for /ɪ/ was 1600-3200 Hz (fcenter = 2263 Hz); the optimal octave for /i:/ was 3200-6400 Hz (fcenter = 4525 Hz); the optimal octave for /e/ was 1600-3200 Hz (fcenter = 2263 Hz); the optimal octave for /æ/ was 1200-2400 Hz (fcenter = 1697 Hz); the optimal octaves for /ʊ/ were 200-400 Hz (fcenter = 283 Hz), 300-600 Hz (fcenter = 424 Hz) and 400-800 Hz (fcenter = 566 Hz); the optimal octave for /u:/ was 200-400 Hz (fcenter = 283 Hz).

These optimal octaves (optimals) proposed by Petar Guberina have been used experimentally with good results for correcting the individual English sounds of native speakers of English. However, according to Guberina (1972), our brain does not always correctly perceive the frequencies sent through the ear but makes a selection, assuming that our mother tongue works as a natural filter when we perceive sounds. That is, based on the differences between the two languages, English and Chinese in our case, exposing the Chinese speakers of English to the native-speaker optimals might still be insufficient. Therefore, a new set of what has been called "corrective" optimals may need to be explored and refined to cater to each learner's individual perceptions and learning needs. Until now, no related research on corrective optimals for Chinese EFL learners has ever been carried out. Therefore, it would be valuable to investigate the corrective optimals for Chinese EFL learners using the native speaker optimals as the starting point. In addition to investigating the nature of the corrective optimals, the study will also investigate whether the use of the corrective optimals alone, with none of the other verbotonal techniques, will be sufficient to bring about acceptable production of the target vowels in Chinese EFL learners. This investigation will form the primary focus of our study but will necessarily be reflected in the other focus of our study which necessarily involves the actual impact of the approach on Chinese EFL learners of English as outlined below.

1.2 Statement of the problem

In the People's Republic of China, the first real surge of learning English began in the early 1980s. As the opportunities increased for China's interactions with the west, success in English has become an ideal for career and academic success as well as a better life. English has hence been assigned as a compulsory school subject from primary school Grade 3 onwards since 2001. By the time these students go to college or university, they might have received an average of nine or ten years of English instruction. If they pass the college entrance examination and get admitted, they will continue to learn English as English majors or non-English majors for another two or more years. Unfortunately, learning English in the classroom setting for a decade does not guarantee that learners can speak fluent English (Félix-Brasdefer & Koike, 2012). Chinese learners of English are comparatively weak in speaking and listening, particularly in pronunciation (Thomson, 1996). This unacceptable phenomenon has long existed and has been referred to as Mute English (a.k.a. Deaf and Mute English, Dumb English, Dumb and Deaf English) (Pan, 2015; Zhang, 2012). As a result, English syllabi at all levels have witnessed an increased emphasis on the communicative skills of listening and speaking (MOE, 2000; MOE, 2003; MOE, 2007; MOE, 2011).

The latest version of the national College English Curriculum Requirements (CECR) for Chinese non-English major undergraduates was issued by the Ministry of Education in 2007 (MOE, 2007). It sets up a nationally unified standard for present-day college English teaching in terms of character and objectives, teaching requirements, course design, teaching models, evaluation and teaching administration. The general objective of the CECR is to develop students' ability to use English in an all-round way with more emphasis on listening, speaking, learner autonomy and cultural awareness. As a requirement, all Chinese non-English major EFL learners need to reach the basic pronunciation proficiency level before graduation. This is described as "clear articulation and basically correct pronunciation and intonation" (MOE, 2007). However, in reality, many non-English major graduates fail to meet the requirements of the CECR. Moreover, a great majority of Chinese non-English major EFL learners have pronunciation problems at both segmental and suprasegmental levels (Cai, 2011; Liu &

Hu, 2010; Zhou & Song, 2015).

According to previous research studies, the problems of English pronunciation learning and teaching have been mainly discussed as follows: (1) Because of the negative transfer effect of Chinese in pronunciation learning, Chinese non-English major EFL learners tend to replace some English sounds by what competent English speakers perceive as "different" sounds. (2) Chinese non-English major EFL learners take their English teachers as pronunciation models. However, not all English teachers are good at pronunciation, and most of them often feel that they are not ready yet to teach pronunciation. (3) Chinese non-English major EFL learners, at most colleges and universities, cannot register for any pronunciation courses. Moreover, College English is commonly conducted in large groups. Students hardly have any individual treatment of pronunciation problems on account of the large class size. (4) Because of the nature of the current examination system, before graduation, the most important thing for Chinese non-English major EFL learners to do is to pass the English proficiency tests, such as the CET-4 (College English Test Band 4) and CET-6 (College English Test Band 6). Thus, students pay more attention to developing skills in other areas such as vocabulary and grammar to ensure good results (Cong, 2013; Shi, 2014; Wang & Shi, 2015; Zhang, 2013).

Having been teaching College English for more than six years at Gannan Normal University (GNU), the researcher discovers that many Chinese non-English major EFL learners are not proficient in English. Most of them have problems with English learning, especially in pronunciation learning. One particular pronunciation problem is that they have difficulty producing English sounds and sound contrasts, such as /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/. Moreover, according to the students' responses to questions about their pronunciation learning (see Appendix A), a vast majority of them reported that they had inadequate pronunciation training and were not confident in their pronunciation and of speaking English. However, they wanted to improve their pronunciation performance.

Given the overwhelmingly large numbers of non-English major students at the tertiary level in China, any effort targeting this population is likely to have a huge impact on learning as well as teaching. In this sense, it will be highly beneficial if we can develop an effective learning approach for Chinese non-English major EFL learners at GNU concerning the phonetic correction of these English vowels.

1.3 Purposes of the study

The purposes of the study were as follows:

(1) To investigate the corrective optimals for Chinese non-English major EFL learners for the following English vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/;

(2) To examine the differences between the native speaker optimals and the corrective optimals for Chinese non-English major EFL learners;

(3) To develop a verbotonal-based approach (henceforth referred to as the VTPL

approach which stands for the VerboTonal-based Pronunciation Learning approach) for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners;

(4) To determine the effectiveness of the VTPL approach to phonetic correction of the target English vowels for Chinese non-English major EFL learners and, as part of this, to determine whether simple exposure to the corrective optimals alone is sufficient to bring about the perceptual and articulatory changes necessary for the acceptable production of the vowels in question;

(5) To make a comparison between the VTPL approach and the traditional approach and discover which of the two approaches is more effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners;

(6) To examine the Chinese non-English major EFL learners' opinions of the VTPL approach to phonetic correction.

1.4 Research questions

The research questions for this study were:

(1) What are the corrective optimals of Chinese non-English major EFL learners for the following English vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/?

(2) Are there any differences between the native speaker optimals and the corrective optimals for Chinese non-English major EFL learners? If yes, what are these differences?

(3) Is the VTPL approach effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners? In particular, is simple exposure to the corrective optimals alone sufficient to bring about acceptable production of the target vowels?

(4) Which approach, VTPL or traditional, is more effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners?

(5) What are the Chinese non-English major EFL learners' opinions of the VTPL approach to phonetic correction?

1.5 Significance of the study

First and foremost, the present research attempts to help solve the pronunciation problems of a selection of English vowels for Chinese non-English major EFL learners through the concept of corrective optimals. Until now, no such empirical studies have been conducted for Chinese non-English major EFL learners. The current study can fill this research gap.

Second, the notion of optimal octaves is modified to include both single bandpass filters and discontinuous multiband filters. In this sense, the verbotonal theory of phonetic correction may be refined to be more effective.

Third, the study attempts to fill another research gap which is to determine whether simple exposure to corrective optimals alone is sufficient to bring about acceptable production of the target vowels. Typically, optimals are used in conjunction with a number of other corrective techniques. Here, they will form the sole approach to correction.

Fourth, the construction of the individualized perceptual training for Chinese non-English major EFL learners will help to restate the importance of individual differences in language learning and teaching.

Fifth, the implementation of different load-lightening and load-increasing activities for Chinese non-English major EFL learners will shed light on the issues of meaning-making and awareness-raising.

Sixth, the findings of the research may provide Chinese non-English major EFL learners with optimal vowel profiles for the target English vowels. Thus, it may provide an alternative way for effective pronunciation learning and teaching.

Finally, the findings of the research may have pedagogic implications for language learning and teaching of Chinese EFL learners and possibly speakers of other languages. Moreover, the findings may also have an impact on education in general.

1.6 Definitions of key terms

(1) Traditional approach

The traditional approach refers to the standard way of using training materials to learn the target English vowel sounds. In the traditional approach, the students are exposed to entirely natural training materials and no digital filtering is involved. It often involves the separation of segmental and suprasegmental elements of speech, emphasis on articulation (e.g. through the use of vowel quadrilaterals) rather than perception and intellectualization of the processes involved in pronunciation with copious teacher intervention and descriptions of linguistic phenomena.

(2) VTPL approach

The VTPL approach refers to the verbotonal-based pronunciation learning approach inspired by the verbotonal theory of phonetic correction. In the VTPL approach, the students are exposed to the training materials with the target English vowel sounds optimally filtered through their corrective optimals. The VTPL approach focuses only on filtered corrective optimals and does not employ the full range of techniques normally associated with the verbotonal approach to pronunciation correction.

(3) Lowpass filter

Lowpass filter means a filter that passes (or keeps) frequencies below a specific cutoff frequency and attenuates all frequencies higher than that cutoff frequency. In the current study, to follow the tradition, the cutoff frequency for the lowpass filter is set at 320 Hz.

(4) Optimals

The term "optimals" (sometimes optimal octaves) refers to a set of 13 overlapping octave bands (filters) ranging from 100 Hz to 12800 Hz for the English sound repertoire (vowels and consonants) proposed by Petar Guberina (Asp, 1972). According to Guberina (1972), when a particular sound is filtered through its optimal octave (1-octave band), the sound is best perceived. An octave refers to a frequency band whose upper cutoff frequency is twice as great as that of its lower cutoff frequency (e.g. 300-600 Hz is an octave, 1600-3200 Hz is an octave).

(5) Corrective optimals

The term "corrective optimals" refers to a set of filters identified for Chinese non-English major EFL learners at GNU to help them produce acceptable target English vowels. They use the native speaker optimals as the point of reference and are personal in nature.

(6) Single bandpass filters

The term "single bandpass filters" refers to three frequency bands, i.e. 1 octave, $\frac{1}{2}$ octave and $\frac{1}{3}$ octave. The concept of single bandpass filters is derived from Guberina's notion of optimal octaves as described above.

(7) Discontinuous multiband filters

The term "discontinuous multiband filters" refers to the single bandpass filters enriched by the addition of a lowpass filter (below 320 Hz) to the existing filters. They include 0-320 Hz + 1 octave, 0-320 Hz + $\frac{1}{2}$ octave and 0-320 Hz + $\frac{1}{3}$ octave.

(8) Filtering

Filtering means the manipulation of the target vowel sounds in logatomes, words and sentences using the students' corrective optimals. For instance, in the sentence "Is it pet or is it pat? Pet or pat?", only the vowel sounds /e/ and /æ/ in the words "pet" and "pat" will be filtered. This gives rise to the following sentence (bold underlined parts represent filtering): "Is it p**<u>e</u>**t or is it p**<u>a</u>**t? P**<u>e</u>**t or p**<u>a</u>**t?". Therefore, the sentence becomes a combination of filtered and natural language.

(9) Logatomes

The term "logatomes" refers to nonsense syllables in the pattern of CVCV (consonant-vowel-consonant-vowel), such as /dɪdɪ/, /hɪhɪ/, /lɪlɪ/, /sɪsɪ/, /zɪzɪ/ etc. in terms of the vowel sound /ɪ/. Logatomes are designed to minimize the students' cognitive load and maximize their exposure to the characteristics of the vowel sounds under study.

(10) Optimal sentences

The term "optimal sentences" refers to sentences that allow students to practice the pronunciation of the target English vowels under the best conditions. They are designed to help students produce the target vowel sounds easily and to give them practice in producing sounds in the most favorable surroundings before moving on to less favorable contexts. Take, for example, the optimal sentence for /iː/ (e.g. I see these bees.).

(11) Contrast-embedded sentences

The term "contrast-embedded sentences" refers to sentences that allow students to practice the pronunciation of the target English vowels in syntagmatic opposition to one another (i.e. /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/). They are designed to enhance the students' discrimination abilities of the contrasting vowel sounds using a sentence carrier "Is it __ or is it __? __or __?". Take, for example, the contrast-embedded sentence for /ɪ/-/iː/ (e.g. Is it ship or is it sheep? Ship or sheep?).

(12) Non-optimal sentences

The term "non-optimal sentences" refers to sentences that allow students to practice the pronunciation of the target English vowels under the worst conditions. They are designed to create difficulties for the students in producing the target English vowel sounds and to give them practice in producing these sounds in the least favorable surroundings. Take, for example, the non-optimal sentence for /ɪ/ (e.g. It is a ship which sits on the ripple heading for the city.).

(13) Chinese non-English major EFL learners

The term "Chinese non-English major EFL learners" refers to the Chinese undergraduate students at GNU who are not studying English as a major. Specifically, the students are first-year undergraduates enrolled by GNU in the first semester of the academic year 2016-2017.

1.7 Summary

In this chapter, a brief introduction to the study of the pronunciation problems of a selection of English vowels for Chinese non-English major EFL learners was provided. It first described the background of the study. This was followed by the statement of the problem, purposes of the study, research questions, significance of the study and definitions of key terms. In the next chapter, a review of the literature on pronunciation teaching and research, phonetic correction, verbotonal theory, understanding learning and awareness-raising will be covered, leading to the development of a theoretical framework upon which the present study is based.

CHAPTER 2

LITERATURE REVIEW

This chapter covers a review of the literature related to the present research. It consists of seven sections. The first section presents an overview of pronunciation teaching and research, giving particular attention to the history of pronunciation teaching, goals of pronunciation teaching, pronunciation teaching methods and pronunciation research in China. The second section discusses speech perception and production and training paradigms in terms of phonetic correction. The third section focuses on the verbotonal theory, which provides a discussion of fundamental principles, verbotonal procedures and techniques as well as verbotonal research in China. The fourth section is concerned with general remarks on learning and meaning-making. The fifth section is devoted to awareness-raising. The sixth section is the discussion of the theoretical framework underpinning this study. Finally, a summary of the chapter is given.

2.1 Pronunciation teaching and research

2.1.1 History of pronunciation teaching

Over the decades, in general, good pronunciation has always been acknowledged as a significant objective in L2/FL learning (Derwing & Munro, 2015;

Derwing, Munro, & Wiebe, 1998; Jarosz, 2019; Murphy, 1991; Stern, 1992; Wahiduzzaman, 2017). Many people, including scholars and non-scholars, equate non-native speakers' accents with their ability to communicate effectively (Trofimovich & Isaacs, 2012). As Jones (2011) stated, good pronunciation was a key part of confidence in speaking a language and making ourselves more easily understood. However, learning the pronunciation of English has been notoriously difficult for many non-native speakers of English.

Pronunciation involves not only the application of phonological rules but also perception and production (Fouz-González, 2019). As speech perception is closely connected to speech production, the perceptual treatment of sounds turns out to be an important part of language learning, especially for good pronunciation. However, in the history of English teaching, pronunciation has not always been considered in terms of perception and production.

In the 1960s, guided by the concept of transformational-generative grammar and the view of cognitive psychology, the cognitive approach found its way into pronunciation instruction. It deemphasized the role of pronunciation in favor of grammar and vocabulary (Morley, 1991; Seidlhofer, 2001), and at that time, people thought the nativeness of pronunciation was unachievable (Scovel, 1969). In the twentieth century, the grammar-translation method and the reading-based approach, which emphasized teaching the comprehension of texts and focused mainly on helping students read and appreciate foreign literature, had always considered pronunciation to be irrelevant.

The rest of the language teaching methods and approaches, nevertheless, placed more emphasis on verbal communication. For example, in the late 1800s and early 1900s, the direct method stated that pronunciation was of great importance and practiced it through teacher modeling or recording modeling. The method assumed that sounds presented through imitation and repetition would be internalized and correctly produced. In the 1940s and 1950s, with the influence of the reform movement, the audiolingual method proposed that pronunciation was a crucial component of language teaching and needed to be taught explicitly from day one. Influenced by behaviorist psychology, pronunciation training primarily focused on imitation and repetition modeled by teachers or recordings, using minimal pair drills both at the word and sentence levels (Celce-Murcia, Brinton, & Goodwin, 1996).

However, giving students the same set of minimal pair drills to help improve their pronunciation suggests that all learners have the same degree of difficulty in producing problematic sounds. For example, if learners make errors with the vowel contrast "/ɪ/-/iː/", they will then be exposed to the same minimal pair drills with the same auditory input, such as "ship-sheep" at the word level and "It's a cheap computer chip." at the sentence level. This presupposes that students have the same degree of difficulty in speech perception regarding L2/FL learning, which not only ignores their individuality but also denies many other possible causes that may lead to particular errors. What's more, most of the training materials (sound stimuli) discussed so far are presented in their natural phonetic environments (natural language), which indicates that no efforts have been made in terms of creating other possible phonetic learning environments. However, in reality, natural phonetic contexts alone may not be enough to raise L2/FL learners' phonological awareness. Therefore, they should be exposed to more phonetic environments to help sensitize and reinforce their perceptions of English speech sounds.

Even in the teaching methods that put emphasis on verbal communication mentioned above, comparatively, little attention has been paid to the complex nature of one's auditory perception in L2/FL learning. In fact, correcting people's wrong perceptions and teaching them to perceive L2/FL in an effortless manner is viewed as so difficult that most teaching methods just focus on teaching productions alone without taking into consideration the fact that wrong perceptions need to be reorganized.

Therefore, it will be desirable to explore an approach that acts on learners' perceptions, highlights different phonetic environments and capitalizes on learners' individuality, in order to generate carryover to conversational speech.

2.1.2 Goals of pronunciation teaching

In the field of L2/FL pronunciation learning and teaching, the goals of pronunciation teaching have always been the primary concern of practitioners and researchers. Up to now, a number of teaching goals have been discussed and implemented, such as nativeness, intelligibility, comprehensibility, accentedness and fluency (Derwing & Munro, 2009; Derwing et al., 1998; Hahn, 2004; Kang, 2010; Levis, 2005; Munro & Derwing, 1995a; Pennington & Rogerson-Revell, 2019; Reed & Levis, 2019; Saito, Trofimovich, & Isaacs, 2015; Suzuki & Kormos, 2020; Trofimovich & Isaacs, 2012).

Over the years, the idea of nativeness has been a very trendy topic among scholars as well as language learners. This instructional orientation is largely based on the distinctiveness between the two groups of speakers, namely, native and non-native speakers, with clear expectations for both: native speakers set examples for non-native speakers and non-native speakers perform as native speakers (Kramsch, 1998). On the contrary, some researchers argued that native-like pronunciation is an unrealistic and unachievable goal for most L2/FL learners and only a small number of them could speak native-like English (Celce-Murcia et al., 1996; González-Bueno, 1997; Jenkins, 2002; Levis, 2005; Moyer, 1999; Rokita-Jaskow, 2008; Scovel, 1969; Seidlhofer, 2001; Szpyra-Kozłowska, 2015). Moreover, some research studies have shown that there exists an age constraint (around age 6) for the acquisition of native-like pronunciation (Asher & García, 1969; Lenneberg, 1967; Long, 1990, 2007; Oyama, 1976). However, research studies also indicated that, in general, ESL/EFL learners have quite positive attitudes toward native-like pronunciation and many learners show their eagerness to pronounce English like native speakers or at least hope to approximate a native-like accent (Derwing, 2003; Golombek & Jordan, 2005; Lippi-Green, 1997; Low, 2015; Moyer, 1999; Sung, 2016; Timmis, 2002). Taking into consideration the fact that native-like pronunciation could be difficult to achieve in practice, a more realistic and attainable goal, known as intelligibility, is set for learners.

As emphasized by many scholars and phoneticians, the primary goal of pronunciation teaching for the majority of the language learners should be intelligibility (Celce-Murcia, Brinton, Goodwin, & Griner, 2010; Derwing & Munro, 2005; Levis, 2016; Levis & Sonsaat, 2019; Munro & Derwing, 2015). The concept of intelligibility can be traced back to the early 20th century, and can be broadly defined as "the extent to which a speaker's message is actually understood by a listener" (Munro & Derwing, 1995a, P. 76). As Kenworthy (1987) put it, "the more words a listener is able to identify accurately when said by a particular speaker, the more intelligible that speaker is" (P. 13). Increasingly, intelligibility is seen as a more desirable and attainable goal for many learners.

The notion of comprehensibility, concerning the degree of difficulty, is described as "the listener's judgment of how difficult it is to understand an L2 speech production" (Derwing et al., 1998). In terms of easiness, it refers to "listeners' subjective perception of how much or how easily they understand L2 speech" (Saito et al., 2015). The difference between comprehensibility and intelligibility, as pointed out by Levis

(2006), depends on whether the concept of intelligibility is defined in its narrow or broad sense. In its narrow sense, intelligibility is always assessed through the inspection of listeners' accuracy of written transcriptions of L2 learners' utterances. In its broad sense, intelligibility is often used interchangeably with comprehensibility and measured by listeners' scalar ratings of how easily they understand speech (Munro & Derwing, 1995a). Although several L2/FL assessment instruments adopt the term "intelligibility" in their speaking band descriptors (e.g. TOEFL, IELTS, CEFR), the use of scalar ratings implies that the targeted construct is in fact comprehensibility (Trofimovich & Isaacs, 2012). Therefore, as a teaching goal, comprehensibility should be given careful consideration. Up till now, linguistic variables such as vowel/consonant errors, word stress, fluency, lexis and grammar have been reported to be linked to comprehensibility (Saito et al., 2015; Trofimovich & Isaacs, 2012).

Accentedness refers to linguistic nativelikeness, which is defined as "the extent to which a listener judges second language (L2) speech to differ from NS norms" (Derwing et al., 1998, p. 396). The concept is "typically measured through listeners' perception of how closely speakers can approximate speech patterns of the target-language community" (Saito et al., 2015, p. 1). Accent has long been the focus of L2 speech research (Edwards, Zampini, & Cunningham, 2019; Flege, 1987; Kuhl, 2000, 2004; Meador, Flege, & MacKay, 2000; Munro & Derwing, 1995a, 1995b; Shintani, Saito, & Koizumi, 2019; Trofimovich & Isaacs, 2012). As Lev-Ari and Keysar (2010) declared, "most non-native speakers have an accent, and it could cause them to seem less credible for two main reasons: (1) The accent serves as a signal and (2) the accent makes the speech harder to process" (p. 1093). Similarly, Munro and Derwing (1995a) stated that the accented speech was harder to process. Previous studies have shown that pronunciation variables such as vowel/consonant errors, word stress and rhythm are closely associated with accentedness (Saito et al., 2015; Trofimovich & Isaacs, 2012).

Fluency refers to an automatic procedural skill in speech production (Derwing, Rossiter, Munro, & Thomson, 2004; Kovač & Vicko, 2019; Schmidt, 1992; Tavakoli, 2019), which is interpreted as "an impression on the listener's part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently" (Lennon, 1990, p. 391). The assessment of fluency is most often based on listeners' scalar ratings of how fluently learners speak. As a crucial aspect of successful communication, a group of L2/FL speech researchers have consistently shown their interest in exploring the concept as well as its relationship with accentedness and comprehensibility (Derwing, Munro, & Thomson, 2008; Derwing et al., 2004; Derwing, Thomson, & Munro, 2006; Edwards et al., 2019; Rossiter, 2009; Saito et al., 2015; Trofimovich & Baker, 2006; Trofimovich & Isaacs, 2012). The related literature has indicated that fluency problems such as inappropriate pausing, a slow articulation rate and false starts could all affect the listeners negatively (Derwing, 2017; Derwing et al., 2004; Derwing & Munro, 2001; Trofimovich & Baker, 2006).

According to the previous studies of L2/FL speech, one of the trends in research has been to investigate the influence of segmentals and suprasegmentals on L2/FL learners' speech concerning comprehensibility, accentedness and fluency. Many researchers have tried to determine which has more influence over the other concerning different aspects of L2/FL speech production. Up till now, there have been three different points of view toward the importance of segmentals and suprasegmentals, which are as follows: First, segmental features play a more important role than suprasegmentals. For example, Collins and Mees (2013) advocated that the segmental features had the greatest influence on intelligibility, and therefore, the highest priority in pronunciation instruction needed to be given to segmentals. This rests on the very view that intelligibility is impeded by segmental errors rather than by suprasegmental errors. Second, suprasegmental features are more important than segmental ones. For instance, Brown (1992) suggested that "suprasegmental features are, if anything, more important than the segmentals in terms of intelligibility and the acquiring of a quasi-native accent" (p. 11). Third, both segmental and suprasegmental features are important. Derwing et al. (1998), for example, investigated the influence of segmental and suprasegmental instructions on accent, comprehensibility and fluency, the results showed that both segmental and suprasegmental features benefited L2 learners.

With respect to the fact that English has now become a global language and worldwide lingua franca (Jenkins, 2015; Rose & Galloway, 2019), the current study will partially adopt the well-established pronunciation evaluation tradition proposed by Trofimovich and Isaacs (2012) as a starting point, focusing on both specific measures (i.e. segmentals and suprasegmentals) and general measures (i.e. comprehensibility and fluency). Moreover, to figure out how well all students are doing, a general assessment of the production of the sounds of English (vowels and consonants) will be conducted. Therefore, pronunciation will be another focus of the study. Five segmental and suprasegmental aspects of production will be adopted as the rating criteria to make the evaluation process more manageable, including sounds (vowels and consonants), rhythm, word stress, intonation and speech rate. That is, in the present study, both students' productions of the target vowels and their overall speaking proficiency concerning comprehensibility, fluency and pronunciation will be evaluated using the above-mentioned rating parameters.

2.1.3 Pronunciation teaching methods

As an integral part of developing communicative competence, the importance of English pronunciation teaching could not be underemphasized (Darcy, 2018; Gilakjani, 2016). Over the decades, three methods for pronunciation teaching have been developed to help improve L2/FL learners' pronunciation, including the articulatory method, the audiolingual method and the computer-assisted method. Of these teaching methods, the first two are more frequently used in present-day pronunciation instruction. They are discussed as follows.

The articulatory method

This method is based on the assumption that L2/FL learners will be able to produce L2 sounds correctly after having a good command of how L2/FL articulatory system functions. So, it is necessary that students be introduced to the physiology of pronunciation and getting familiarized with different parts of speech organs, such as lips, tongue, teeth, hard and soft palate, alveolar ridge, etc. They are then told how to articulate vowel and consonant sounds, particularly through the concept of "vowel space", "place of articulation" and "manner of articulation".

However, having something under control like that is not an easy job. For example, the vowel sound /ʊ/ is described as "the part of the tongue just behind the center is raised, just above the half-close position; the lips are rounded, but loosely so; the tongue is relatively relaxed" (Kelly, 2000, p. 31). Concerning the complex and vague description that cannot be easily understood, the phonation of /ʊ/ becomes much more difficult for L2/FL learners. Similarly, the instruction of pronunciation exercises is given as "pop your lips forward, using the musculature at the center of both the upper and lower lips...make a popping p sound in time with the ticking of the second hand of a clock, for one minute" (Cameron, 2018, p. 9). Once again, L2/FL learners may have problems understanding the unclear instruction, and most likely, they may be unable to pronounce it correctly without being helped.

In this method, the L2/FL speech is regarded as easy as putting a string of isolated sounds together. Although separate sounds can be isolated, the characteristics they show in isolation will not be the same as the characteristics they have in the connected speech. The oversimplification of English pronunciation frequently leads to learners' sense of frustration. Having practiced all the phonemes with considerable efforts, the learners often feel frustrated to find that they fail to understand native speakers of English, let alone the face-to-face communication. Moreover, the idea of auditory perception is totally neglected in this teaching method, which is likely to cause unstable speech perception and may lead to further unintelligible speech production.

The audiolingual method

This method originates from audiolingualism, which is greatly influenced by the theory of behaviorism. In this method, under the influence of the notion of contrast in structural linguistics, pronunciation training is mainly devoted to learners' imitation and repetition of sounds modeled by the teacher or recordings, using minimal pair drills that use words that differ by a single sound in the same position both at the word level (e.g. word drills: "bean-bin") and the sentence level (e.g. syntagmatic drills: "It's a cheap computer chip." and paradigmatic drills: "Look out for that sheep." and "Look out for that ship.") (Celce-Murcia et al., 1996). These pairs are, most often, drilled chorally and individually, in order to give students plenty of opportunities to listen out for differences and practice saying them. Finally, individual students will be asked to read the lists without models.

As indicated, in this method, the entire instruction process is highly structured and teacher-centered. The student's role is to respond to teachers' stimuli. Because the continuous repetition for memorization is monotonous and tiring, many students find the classes boring, unsatisfying and frustrated. Due to their "mother tongue sieve" (Trubetzkoy, 1939/1969), "students appear to have difficulty in hearing a difference between the two contrasting sounds and also seem to produce neither sounds accurately, tending instead to produce a sound which seems to be halfway between the two" (Kelly, 2000, p. 27). Also, students are not able to transfer pronunciation skills to real communication outside the classroom (Kapurani, 2016; Richards & Rodgers, 2001). Similarly, according to Goldstein (2014), after being taught to imitate patterns, students are more likely to become passive learners. Accordingly, their auditory perceptions are often less well developed with less satisfactory results because of inadequate awareness-raising support.

The computer-assisted method

The implementation of modern technology in language learning and teaching has become a trend since the 1980s. It is widely accepted that computer-assisted language learning (CALL) allows L2/FL learners more freedom in language learning, provides a more individualized learning environment and thus boosts learner autonomy

(Fischer, 2007; He, 2014; Jenkins, 2004; Ozawa, 2019; Pennington & Rogerson-Revell, 2019; Schwienhorst, 2008; Swann, 1992; Wagener, 2006).

Until now, many positive results have been found after computer-assisted pronunciation training concerning segmental and suprasegmental features. For example, Wang and Munro (2004) investigated computer-based training for learning English vowel contrasts, compared with the control group, the experimental group showed improved perceptual performance, transferred their perception to new contexts and maintained their improvement three months after training. Similarly, Luo (2016) declared that computer-assisted pronunciation training technique was more effective in reducing students' pronunciation problems in comparison to traditional in-class pronunciation instruction. However, Thomson (2011) argued that certain CALL packages were simply showy programs to pronunciation learning, in fact, which might cause some difficulty for language learners instead of supporting them.

Some CALL programs present visual display regarding the phonetic features (segmental and suprasegmental features) of the learner's recorded speech and allow comparison to that of a native speaker model. However, as far as the visual acoustic features of segmentals (spectrograms) are concerned, there is little benefit in helping learners remedy their segmental errors.

Non-expert learners actually often find spectrograms uninterpretable, which makes it impossible to provide any information that can be easily used to improve the pronunciation of vowels and consonants. Although more individualized learning can be achieved with the help of CALL, very few considerations have been given to the individual differences concerning pronunciation training. So, no matter how much time learners spend on pronunciation learning, their pronunciation problems are still there, and consequently, problematic sounds become fossilized errors after being exposed to the same training stimuli over and over again.

As Silverstein, Silverstein and Nunn (2001) argued, “all the sounds around us would be meaningless noise if it were not for the brain” (p. 24). In other words, we are aware of the sounds that are meaningful to us right now and unaware of sounds which are not meaningful to us right now. Thus, learning is making the meaningless meaningful (Lian, 2000). As indicated in the previous research, in order to speak acceptably, we have to perceive sounds correctly (Escudero, 2007; Flege, 1995a; Flege et al., 1997; Lee & Lyster, 2017; Trazo & Abocejo; 2019). Hence, pronunciation training needs to be more perceptual-based to make meaningless sounds meaningful.

According to the aforementioned pronunciation teaching methods, it can be seen that perceptual-based methods that take into account the individual differences are currently very few available. In response to this problem, the present research will be more focused on individualized perceptual training through awareness-raising activities.

2.1.4 Pronunciation research in China

After reviewing the research studies retrieved from China's well-known academic online database CNKI (i.e. China National Knowledge Infrastructure), it's clear that there have been many research studies conducted by Chinese scholars in terms of pronunciation learning and teaching. However, it also shows that the research in China concerning pronunciation learning and teaching is far from enough when compared to the efforts devoted to other aspects of language skills.

A close inspection of these studies indicated that research interests covered a number of topics, including theoretical considerations (Cheng, Zhang, & Zhang, 2017; Liu & Niu, 2018; Wang, 2018; Xu & Zeng, 2017; Zhao, 2019), pedagogic considerations (Cai, 2018; Tang, 2020; Xiang, 2019; Zhang, 2019), teaching methods (Hu, 2016; Lin, 2019; Liu, 2016; Song, 2015; Zhang & Ma, 2015) and phonological comparison between Chinese and English (Lin, 2017; Qian, 2015; Wen & Chen, 2019). Although more and more researchers are becoming aware of the importance of suprasegmental features, many researchers think that segmental problems are more serious and should be settled first before suprasegmental ones (Hai, 2018; Hong, 2017; Xia, 2019). Thus, when it comes to pronunciation instruction, the discussion of segmental features has outnumbered the concern of suprasegmentals.

A review of the literature indicated that numerous studies have been carried out concerning the interference of mother tongue on their pronunciation learning. Many scholars focus on the issues of negative transfer of learners' Chinese dialects in their English pronunciation learning (Cai, Zhu, & Chen, 2015; Chen, 2015; Fu, 2018; Li, 2015; Ma, 2019; Wu, 2020). Other research efforts are intended to diagnose learners' pronunciation errors in terms of speech perception and speech production (Tang & Ge, 2019; Zhang, 2013; Zhang & Wang, 2017; Zhou & Song, 2015). For example, Zhou and Song (2015) explored Chinese students' English pronunciation ability based on their speech production. In their study, 88 Chinese college students were first required to read twice the words containing 44 target segmentals in the carrier sentences ("Speak _ twice" and "Say _ again") and after that, they continued reading a short passage with the same amount of segmentals embedded. The results showed that listeners' intelligibility was negatively influenced by the mispronunciation. As they stated, for Chinese students, there was still a long way to go before achieving clear, fluent, accurate and effective cross-cultural communications concerning both native and non-native speakers of English from all over the world.

According to the previous research, the investigation of learners' pronunciation learning strategies turns out to be another research interest (Chang, 2019; He, 2016; He, 2019; Peng, 2014; Peng & Wang, 2014). For instance, Peng (2014) examined the differences in the use of pronunciation learning strategies between 77 English and 105 non-English major EFL learners. The results indicated that the two groups were statistically significantly different concerning memory strategies, cognitive strategies, metacognitive strategies, affective strategies and social strategies. It also argued that EFL learners' pronunciation achievements could be predicted by cognitive strategies.

With the rapid development of modern technologies, the numbers of CALL programs for pronunciation learning have been greatly increased over the past few years. In line with this, an increased amount of attention has been given to the discussions of the feasibility of using technologies in pronunciation learning and teaching, which have been claimed to be hypothetically promising (Li, 2014; Shang, 2016; Zhi & Li, 2020). For example, Shang (2016) proposed that the Praat program made it possible to visualize segmental and suprasegmental features in pronunciation teaching, which could help learners self-correct pronunciation errors. The author advocated that Praat was a promising software program for helping pronunciation learning as well as teaching.

As presented above, the issue of English pronunciation has not been well investigated in China, particularly when it comes to the correction of pronunciation errors, such as the notoriously difficult English vowels. Many studies are simply theoretical or pedagogic considerations, which actually leave pronunciation problems untouched. Although much more research has been conducted concerning the correction of pronunciation errors, not all research studies produce the same satisfactory results. One speculation is that perceptual correction has not been given sufficient attention by those researchers who have discussed the question of phonetic correction. Another speculation is that the theory of learning, which must surely be at the heart of education, is not generally taken into consideration in terms of language pedagogy and teaching methodology.

2.2 Phonetic correction

2.2.1 Speech perception and production

We are all involved in listener-speaker interaction every day. As a part of our daily lives, we perceive and produce all kinds of speech sounds in our native language. Sometimes, to communicate with the people from another culture, we must perceive and produce non-native L2/FL sounds which are often quite different from the sounds in our mother tongue. According to Lapteva (2011), speech processing is a complex process involving "activity in the speaker's brain in order to create a linguistic form; movements of the vocal organs in order to produce a message…perception of the linguistic form and processing it in the brain" (p. 45). It is generally accepted that speech perception and production abilities of late L2 speakers differ from those of native speakers. In terms of speech perception, even after years of exposure, adult L2 speakers still have considerable difficulty in identifying and discriminating many contrasting sounds (vowels and consonant contrasts) that do not exist in their native language (Werker & Tees, 1999).

Previous research has shown that infants have the language-general mechanism in perceiving the speech sounds of all languages (Rvachew & Brosseau-Lapré, 2012). However, with the increased exposure to their native language, their perceptions become more language-specific and gradually their abilities to perceive non-native sounds decline (Burnham & Mattock, 2010).

Based on Lenneberg's (1967) critical period hypothesis (CPH), such inability to perceive the non-native speech contrasts (phonemes) is caused by the loss of neuroplasticity, suggesting that native-like production in an L2 becomes more difficult and perhaps impossible to acquire if the learning takes place after the critical period. Similarly, Trubetzkoy (1939/1969) termed it as "mother tongue sieve". According to him, L2 learners were only able to perceive the sounds they were already familiar with due to their extensive exposure in their mother tongue. Since the phonological sieve of the mother tongue does not work for the L2, speech sounds of L2 are always distorted and misinterpreted, causing lots of production mistakes and errors.

With regard to the influence of speech perception on speech production, some studies have explored the potential relationship between them. The accuracy with which L2/FL segmentals (phonemes) are produced has been found to be positively influenced by the accuracy of L2/FL to which they are perceived (Flege et al., 1997; Flege, MacKay, & Meador, 1999; Wong, 2015). Additionally, some studies indicated improved L2/FL speech production after perceptual training of L2/FL speech sounds even though no production training is provided (Rochet, 1995; Thomson, 2011; Wang, Jongman, & Sereno, 2003). All of these studies suggested that auditory perception of L2/FL sounds play an important role in speech production. It seems reasonable to conclude that L2/FL learners' inability to produce non-native sounds are caused by their auditory perception problems, and furthermore, perceptual training can help improve L2/FL learners' speech perception as well as production. Therefore, all instructional efforts must be devoted to bypassing the mother tongue sieve that triggered learning problems (e.g. auditory perception problems of L2/FL sounds) for language learners, otherwise, almost assuredly, the learners would remain "deaf" to the foreign language system (Lian, 1980).

In line with the ideas presented above, the present study will focus on using the auditory perceptual training to assist the acceptable production of English vowels in Chinese non-English major EFL learners. A review of the related paradigms is discussed below.

2.2.2 Training paradigms

Choosing an appropriate perceptual training paradigm is of great importance if the treatment is to be effective concerning phonetic correction of the problematic sounds. According to the previous phonetic training studies, various perceptual training paradigms have been used to raise L2/FL learners' phonological awareness and improve their perception and/or production of non-native sounds, such as perceptual fading technique, high variability phonetic training (HVPT) approach, auditory-visual training (AV) approach and auditory-only training (AO) approach. Among these perceptual training paradigms, the HVPT approach is the one that receives more attention from researchers when it comes to the phonetic correction of L2/FL contrasts.

According to Wong (2015), the high variability phonetic training approach has received special attention in the past two decades due to its efficacy in improving L2/FL learners' perception and production of non-native sounds (i.e. vowel and consonant contrasts). As far as the perceptual training paradigms are concerned, they are always implemented with different training tasks.

Up till now, discrimination tasks and/or identification tasks have been most frequently used to help improve language learners' L2/FL speech perception and/or production (Carlet & Cebrian, 2015; Shinohara & Iverson, 2018). Over the decades, many research studies using discrimination tasks and/or identification tasks in perceptual training under laboratory conditions have reported positive results in terms of pronunciation improvements (Carlet & Cebrian, 2015; Flege, 1995b; Iverson & Evans, 2007; Logan & Pruitt, 1995; Shinohara & Iverson, 2018).

With regard to discrimination tasks, learners are asked to listen to the stimuli in sequence and determine whether they are the same or different in the paradigm. Basically, there are two types of discrimination tasks, namely, AX discrimination and ABX/AXB/XAB discrimination (oddity task) (Colantoni et al., 2015).

For the AX task, pairs of sounds are involved, such as /ʃiːp/-/ʃɪp/ or /ʃɪp /-/ʃɪp/ and learners are asked to hear and decide whether sounds A and X are the same or different. In terms of XAB task, learners hear triplets of sounds in order such as /ʃiːp/-/ʃɪp/-/ʃiːp/ and decide whether sound X they hear is the same as sound A or B. If learners can discriminate A and B as different sounds, then it should be easy for them to tell which one matches sound X. In identification tasks (forced-choice identification), each time learners hear only one sound stimulus, such as /ʃiːp/, then they have to select the letter ("p" or "b") or word ("sheep" or "ship") from the choices provided.

Although perceptual training has shown positive research results, there are some aspects that need careful consideration. First, the aforementioned sound stimuli are merely restricted to phonemes or words; however, in real communication outside the classroom, learners should speak in a connected way. Therefore, it will be desirable if sentences containing the target sounds can be involved in training sessions. Second, all learners have been exposed to the same sound stimuli, irrespective of L2/FL learners' individuality and speech perception difficulties, which refer to the inability to hear the sounds as they are. In other words, learners interpret the sounds they hear in the wrong way, using personal filters. Moreover, the sounds provided in the natural environments alone will be far from enough to raise learners' phonological awareness. Thus, L2/FL learners' differences, their speech perception problems, as well as training environments, need to be taken into consideration at a very early stage. Third, the previous training paradigms have been carried out under laboratory conditions, which make it impossible for learners to have any perceptual training after class. Therefore, it will be better if a training package can be developed to help learners access learning resources regardless of whether they are in classrooms or at some other places.

As mentioned earlier, the verbotonal theory of phonetic correction, which is different from all other training paradigms, focuses on providing optimal listening conditions for language learners. According to the verbotonal point of view, each language has a set of optimal frequency bandwidths for its speech sounds. Therefore, the students who experience difficulty with a particular foreign language sound is described as not being able to identify the "optimal" frequency band of the sound. It further proves that learners perceive the speech sounds best if they are exposed to the optimal frequency bands of the sounds (Guberina & Asp, 1981).

The optimal frequency band refers to the optimal octave, which is defined as "the octave bandwidth that produces the highest identification score for a particular speech sound (phoneme) and this filtered phoneme sounds similar to the same phoneme when it is not filtered" (Asp, 2006, p. 204). Taking into account the optimal listening conditions, together with the experimentally good results of the verbotonal theory of phonetic correction conducted at the individual level, a verbotonal-based approach will be developed in the present study.

The traditional octave bandwidths (optimals) determined by Guberina for both British and American English vowels and consonants were filtered through 13 overlapping octave bands ranging from 100 Hz to 12800 Hz (i.e. 100-200 Hz, 150-300 Hz, 200-400 Hz, 300-600 Hz, 400-800 Hz, 600-1200 Hz, 800-1600 Hz, 1200-2400 Hz, 1600-3200 Hz, 2400-4800 Hz, 3200-6400 Hz, 4800-9600 Hz, 6400-12800 Hz) (Asp, 1972; Koike, 2012) (see Table 2.1 and Table 2.2).

More recent studies indicate that Guberina's notion of optimals has been extended to exploring L1 optimals of other languages, such as Japanese, French, German, Italian and Spanish (Koike, 2012) and learning Croatian as L2 (Mildner & Tomić, 2007; Tomić et al., 2011).

Up till now, no research has ever been conducted concerning Chinese learners of English, let alone Chinese non-English major EFL learners. Hence, it will be beneficial to determine the corrective optimals by using the native speaker optimals as a point of departure.

Based on the optimals listed for both British and American consonants and vowels, it seems that there are very few differences between the two types of English. However, taking into consideration the fact that the participants have been instructed using more British English, only the optimal octaves for British English will be adopted for the diagnosis of the corrective optimals.

Table 2.1 Optimal octaves of English (British) consonants and vowels

Frequency	Optimal: English (American)	
	Consonants	Vowels
100-200 Hz	-	-
150-300 Hz	-	-
200-400 Hz	p b	uː u
300-600 Hz	w	u ou
400-800 Hz	l n	əː u
600-1200 Hz	k g f v m	ɔː ɔ ʌ
800-1600 Hz	t d h l	ɑː ə ɔi
1200-2400 Hz	ð ʒ r dʒ	æ ɛə
1600-3200 Hz	ʃ tʃ	i e iə
2400-4800 Hz	j	ei ai
3200-6400 Hz	-	iː
4800-9600 Hz	z	-
6400-12800 Hz	s	-

Note. Reprinted from *The effectiveness of low-frequency amplification and filtered-speech testing for preschool deaf children*, by Carl W. Asp, 1972, retrieved from https://eric.ed.gov/?id=ED065977. In the public domain.

Table 2.2 Optimal octaves of English (American) consonants and vowels

Frequency	Optimal: English (American)	
	Consonants	Vowels
100-200 Hz	-	-
150-300 Hz	-	-
200-400 Hz	p b	u
300-600 Hz	w	ʊ o
400-800 Hz	l n	aʊ
600-1200 Hz	k g f v m	ɔ ʌ
800-1600 Hz	t d h l hw	ɑ
1200-2400 Hz	ð r dʒ m j	æ
1600-3200 Hz	ʃ tʃ n	ɪ ɛ e
2400-4800 Hz	θ	eɪ aɪ
3200-6400 Hz	-	i
4800-9600 Hz	z	-
6400-12800 Hz	s	-

Note. Adapted from *Optimal filter perception of speech sounds: Implications to hearing aid fitting through verbotonal rehabilitation*, by Kazunari J. Koike, 2012, retrieved from https://fdocuments.in /document/optimal-filter-perception-of-speech-sounds-phonemic-audiogram-adjustment.html. In the public domain.

2.3 Verbotonal theory

2.3.1 Fundamental principles

The verbotonal theory of phonetic correction, which has been focused on an auditory perception rather than an articulatory training, was developed by academician Petar Guberina in the 1950s (Guberina, 1972). It was essentially developed for rehabilitation of deaf people or the hearing impaired, aiming at the optimization of the residual hearing and enhancing their speech intelligibility through binaural listening, so that they could communicate effectively with the normal hearing people and mainstream into the regular school classrooms (Asp, 2006; Guberina & Asp, 1981). From the verbotonal point of view, the verbotonal theory is guided by certain fundamental principles. In order to have a better understanding of the system (i.e. verbotonal system), it is important to know these guiding principles.

First, the basic principle of the verbotonal theory is the idea of optimality. As Guberina (1989) stated, "in the field of learning foreign languages, it is also necessary to start from the optimal" (p. 11). According to him, the notion of optimal should not be only restricted to the hearing sense, but it also needed to be extended to include other aspects such as the sense of the whole body. In the verbotonal point of view, the optimal segmental features can be transmitted through speech by using filters, such as optimal octaves, whereas the optimal suprasegmental features can be transmitted through body by using low pass filters set around 320 Hz because the body is the most sensitive to low frequencies. Besides hearing in the optimal frequencies, moreover, corrective body movements and vocalizations are practiced simultaneously to enhance language learners' motor coordination and control required for intelligible speech. Step by step, through different combinations of frequencies and the help of body movements, the hearing and producing field will become broader and broader.

Second, the concept of neuroplasticity is considered as a very important principle in understanding the working mechanism of the verbotonal theory. As mentioned earlier, because of the brain's plasticity, very young infants can discriminate various segmental contrasts of their mother tongue as well as non-native contrasts. However, as they grow older and have more exposure to their native language, their ability to distinguish non-native sounds gradually declines. After reaching the age of puberty, the acquisition of a foreign phonological system becomes much more difficult, because the brain has reached a level where the mother tongue is guiding the perception. According to Trubetzkoy (1939/1969), this phenomenon was termed as "mother tongue sieve". Influenced by the loss of cerebral plasticity, the brain does not always get the correct frequencies sent through the ear but makes a selection. Naturally, the brain that reaches maturity is choosing among the rich possibilities of sounds of its mother tongue. As a result, when someone wrongly hears and articulates the sounds of a foreign language, he or she is actually working on the phonological system of his or her mother tongue. Therefore, it seems reasonable that the instructional efforts should be given to helping language learners bypass the mother tongue sieve. Otherwise, the learners will remain unable to perceive non-native sounds of the foreign phonological system.

Traditionally, most phoneticians and pathologists work from the ear or tongue point of view concerning the phonetic correction, which indicates that when they work, they work with what they see. Because of the invisibleness of the brain, it has been a frequently neglected factor in the field of language learning and teaching. However, as a part of the central nervous system, the brain is the place where the speech processing occurs, including monitoring, controlling, producing as well as understanding speech. Hence, the importance of the brain in the process of phonetic correction cannot be neglected. Taking this into consideration, Petar Guberina proposed the idea of stimulating and restructuring the perception of the learners' brains (Asp, 2006). Based on the verbotonal point of view, learners will be able to develop their auditory speech perception for non-native sounds if the optimal listening condition is provided to each brain. In other words, the brain could be rewired, "and with time and training, it would be prepared to respond to more difficult tasks" (Guberina & Asp, 1981, p. 2).

Third, the principle of listening through optimal octave filters plays an important role in rewiring learners' brains. According to the verbotonal theory, each language has a set of optimal bands of frequencies for its sound system and language learners' inability to perceive the non-native sounds is due to the wrong filters that they use under the influence of their native language. It also proposes that the articulation at the segmental level causes relatively little difficulty if the optimal listening condition of the sounds has been given. According to Guberina (1972), vowel and consonant sounds that passed through specific octave bands were more easily identified. Following this optimal octave concept, he determined optimals for both British and American English vowels and consonants filtered through 13 overlapping octave bands ranging from 100 Hz to 12800 Hz (Asp, 1972; Koike, 2012). As proposed in the verbotonal theory, the optimal octave filter of each sound does two things: it passes the optimal frequency for the perception of the particular phoneme in that language and it filters out or attenuates the frequencies which might prevent it from being perceived.

As stated by Asp (2006), the concept of the optimal octave is important for successful speech processing because it "passes only the optimum, or the essence, of the target phoneme" (p. 96). Therefore, after a learner hears the phoneme in its optimal frequency, his or her brain is trained to be aware of the optimal octave of that phoneme. As a result, when the learner hears the sounds in natural language, the brain is still attuned to the optimal frequency of the phoneme. For example, to correct the /i:/ sound, the octave filter is generally set at the frequency range 3200-6400 Hz. As we know, /i:/ sound is much tenser than the /ɪ/ sound. Therefore, if a learner substitutes an /i:/ phoneme for an /ɪ/ phoneme, the teacher could use the filters above the optimal octave 3200-6400 Hz to increase the chances of a correction to the vowel /i:/. On the other hand, the teacher should set the filters below this optimal octave to reduce the tension.

When the learner perceives and produces the /i:/ correctly, the optimal octave is removed to facilitate carryover to everyday communication.

Fourth, the principle of listening through rhythm and intonation is greatly stressed in the verbotonal theory of phonetic correction. As mentioned earlier, because of the complexity of non-native prosodic (suprasegmental) features, traditionally these features have often been neglected in L2/FL pronunciation instruction. This is based on the assumption that the segmentals should be taught first and later on, the suprasegmentals (e.g. rhythm and intonation) can be fixed in some way. However, the fact of babies learning their mother tongue shows that they have already learned the manipulation of rhythm and intonation to show different intentions before they could talk in full sentences (Guberina & Asp, 1981). Besides, according to Mehrabian (1968), suprasegmental components make up more than one-third of information in human communication. This proves that rhythm and intonation also play a critical role in adult spoken language and need to be emphasized. From the verbotonal point of view, "rhythm and intonation (suprasegmental or prosodic) is the foundation of both listening and spoken language" (Asp et al., 2012, p. 323).

Low frequencies, as proposed by Guberina, were optimal for processing speech rhythm and intonation patterns of speech (Guberina & Asp, 1981). According to the verbotonal theory, the brain's perception of segmental features can be enhanced by listening through low frequencies which are generally set around 320 Hz. In other words, the use of suprasegmental features can help rewire the brain for a better perception of the segmental sounds (phonemes), even though the segmental features should be filtered out of the sound signals. In line with the verbotonal viewpoint, previous research studies have shown the positive training effects of the prosodic features on the segmental sounds. According to the previous research, many of the phonemes would be corrected in their way and "fall into place" after prosodic treatment (Lian, 1980; Renard, 1975). However, for some problematic sounds such as vowels, the treatment at the suprasegmental level alone will be far from adequate. Hence, a more intensive verbotonal-based phonetic correction should be conducted at the segmental level.

Fifth, the principle of using body movements either at the suprasegmental level or the segmental level to develop listening skills and spoken language is emphasized in the verbotonal theory of phonetic correction (Asp et al., 2012). Body movements are indispensable parts of the human communication system. Through these movements, we can convey our feelings, thoughts and intentions. Research has indicated that up to 70% of daily communication takes place non-verbally and that when a verbal message is contradicted by a non-verbal one, it is the non-verbal message we trust (Blom & Chaplin, 1988). Before children can actually speak much, they talk by using body movements instead of language. Even after acquiring the more sophisticated verbal language skills, body movements remain our clearest and most reliable line of communication. According to Sheets-Johnstone (2011), "when we learned our mother tongue, we spontaneously learned the specific tactile-kinesthetic invariants peculiar to it" (p. 334). Similarly, Johnson (2001) described movements as the mother tongue. Given the importance of body movements in language development, it is not surprising that the use of body movements is frequently emphasized in the field of language learning and teaching. For example, as Asher (1993) stated, the understanding of the foreign language should be developed through the learners' body movements. According to Gassin (1990), one of the main problems facing learners is that when they speak a foreign language, unconsciously they use body movements that normally function in their native language.

In line with the above, based on Laban's theory of body movements for dance, Guberina proposed the notion of corrective body movements in the verbotonal theory of phonetic correction, which has been used to strengthen the link between body and phonation and to indirectly help them develop a natural voice quality as well as good rhythm and intonation patterns of the speech (Asp, 2006). For example, more tension is required while articulating the vowel /i:/, so a learner can stand up and move both of their arms upward, creating more tension; whereas for the vowel /ʊ/, because little tension is needed, the learner can stand relaxed and move both arms downward to create little tension. Moreover, based on the verbotonal point of view, the corrective body movements for segmental and suprasegmental features are not fixed. According to verbotonal point of view, learners are allowed to use any kinds of body movements according to individual differences to help establish a harmonious relationship between movements and speech. This is also understood as self-synchrony. Eventually, when the body and phonation are in harmony, together with the stabilized speech perception and production, the learner will be able to produce the sounds acceptably without the support of corrective body movements.

2.3.2 Verbotonal procedures and techniques

As mentioned earlier, because of the negative influence of the phonological sieve of one's mother tongue, L2/FL learners always have difficulties in phonating non-native sounds of a second or foreign language. In other words, because L2/FL learners have been previously trained to perceive and produce their mother tongue, hence, when L2/FL learners are faced with a wide range of information in L2/FL, they are unable to make the correct selection because they are "deaf" to the foreign phonological system.

According to Lian (1980), there are two necessary phases that need to be emphasized in pronunciation learning. In the first phase of phonetic correction, every effort must be given to defeating their "deafness" to the sounds of a foreign language through listening to the sounds in optimal conditions. In the second phase of treatment, once their awareness has been raised, learners need to practice more intensively the newly-learnt speech motor (articulatory) patterns to help them develop a "feel" of body and phonation, at the same time such repeated rehearsal can fix the sounds being corrected in their long-term memories and later on, generate carryover to daily communication.

Based on the above-mentioned ideas, together with the guidance of the verbotonal theory of phonetic correction, some training activities can be developed to help Chinese non-English major EFL learners defeat their "deafness", raise their phonological awareness as well as reorganize their perceptions concerning the target English vowel sounds which are problematic to them, using learning materials with the target English vowels digitally filtered through personal corrective optimals newly defined in the current study.

Specifically, students will go through four stages of learning, ranging from load-lightening activities to load-increasing activities to assist the acceptable production of the target English vowels (for details, see Section 3.5). This is an attempt to get students involved in sets of activities of varying degrees of difficulty, i.e. the easier ones first and the more difficult ones later. In other words, the students will move from optimal environments to non-optimal environments. The reason for doing this is that in real life, we always encounter a mix of sounds rather than the same sets of sounds.

2.3.3 Verbotonal research in China

After reviewing the previous studies, it clearly indicates that the verbotonal theory is still underdeveloped in China when compared to other theories and approaches. Over the years, only a few research studies have been carried out by Chinese researchers in terms of language learning and teaching for L2/FL learners in China. According to the literature, most studies focus on learning and teaching of Japanese pronunciation. For example, Hang (2012) conducted a research to examine the effectiveness of a verbotonal-based approach in teaching Japanese pronunciation for Chinese Japanese major FL learners. A total of 18 students were randomly chosen from the first-year and second-year Japanese majors. They were evenly divided into two groups, namely, Group A and Group B. The students in Group A (experimental group) were instructed with the verbotonal-based approach, while the students in Group B (control group) were instructed with the traditional approach. The results turned out to be quite positive, showing that the verbotonal-based approach was overall better than the traditional approach and particularly helpful for Japanese beginners. Similarly, Wu (2013) carried out a survey on Japanese learning as well as a Japanese pronunciation test among 114 Japanese major students at Hunan First Normal University. The results showed that the pronunciation problems of Japanese FL learners included both segmental and suprasegmental features. In response to these problems, she proposed the idea of using a verbotonal-based approach in Japanese learning and teaching and explained the possible procedures for phonetic correction. She, therefore, concluded that with the help of phonation and body movements, the verbotonal-based approach would be a beneficial approach for helping Japanese learners overcome their pronunciation difficulties.

Up until now, only two research studies have been conducted when it comes to the discussion of the verbotonal theory of phonetic correction for Chinese EFL learners. For example, He (2014) explored the teaching of English pronunciation to Chinese English major EFL learners enrolled in compulsory English phonetics classes at Xinyi Normal University for Nationalities. It focused on an approach consisting of verbotonal theory, body movements, rhizomatic theory, autonomous learning and CALL. The focus was not on theoretical discussions of phonemes or prosody or any study of individual phonemes, but the perception and production of English intonation. A total of 96 first-year English majors from two intact classes participated in this research. They were randomly assigned to the control and experimental groups. A mixed-methods research design was adopted: the quantitative part was devoted to the assessment of the students' pronunciation, perceptions and learner autonomy, while the qualitative part put emphasis on the students' and teacher's attitudes toward the verbotonal-based approach. The pronunciation pretests and posttests were assessed in a double-blind way in terms of nativeness, comprehensibility and fluency by both Chinese experts and naïve native English speakers with tight controls of variables, including time on task. Importantly, the control group was significantly better than the experimental group in the pretest, and the teacher (not the researcher) was strongly in favor of the traditional approach and did not believe in the new system. The results showed that both the control and experimental groups improved significantly on all aspects tested.

However, despite the fact that it began with a significant disadvantage concerning the control group, the experimental group exceeded and outperformed the control group on every aspect tested. Qualitative results also revealed a high level of satisfaction with the approach, which corroborated the results of quantitative analysis. Also, some additional surprising and counter-intuitive results emerged: (a) It indicated a positive influence of the suprasegmental features on the segmental ones; and (b) The experimental group's levels of nativeness, comprehensibility and fluency were greater than that of the control group. To conclude, verbotonal-based pronunciation teaching has proved very effective and highly successful as the experiment achieved the progress expected.

As presented above, the issue of the verbotonal-based phonetic correction is still not well investigated in China when it comes to Chinese EFL learners. Up till now, verbotonal research has only been conducted in terms of prosodic features for Chinese EFL learners. No research has ever been carried out concerning segmental features, let alone the exploration of the corrective optimals of English vowels for Chinese EFL learners. Therefore, the present study is designed to fill this research gap.

2.4 Understanding learning

2.4.1 General remarks

Although learning goes by quite unnoticed in many cases, it is no stranger to any of us. Instead, learning plays a very important role in human development, and it is not only something that happens quite naturally but also something that we all participate in. However, according to Wiltsher (2005), learning is difficult to define. In spite of the fact that the concept of learning has been discussed by many researchers, the definitions vary in wording and detail from researcher to researcher, and there is no consensus on the definition of learning (Bunge & Ardila, 1987). For instance, according to Domjan (1998), learning refers to "an enduring change in the mechanisms of behavior involving specific stimuli and/or responses that results from prior experience with similar stimuli and responses" (p. 14). Rao (2002) defined learning as "modification of behavior and experience which is of a lasting nature not brought about by biological or physiological factors" (p. 112). As Sandhaas (1989) pointed out, "learning is understood as an active, mainly conscious operation of a person interacting with his or her environment and learning always implies understanding" (p. 81). According to Lian (2004), learning implies "an act of comprehension which challenges the learner's personal representational and logical systems" (p. 3). In general, they put more emphasis on the idea of change, modification, active operation, understanding and comprehension. All these indicate that learning is a dynamic process of individual knowledge construction.

However, throughout the years, under the influence of behaviorism, learners' understandings have long been neglected, and the focus is mainly on what can be seen happening-behavior. That is, learning is reduced to forms of behavior. Moreover, errors are often viewed as the result of bad habits, which can be removed if only learners have enough rote learning and pattern drills with the help of the target language models. Hence, the primary interest in learning or training has been centered completely on stimulus-response associations (habit formation). As a result, classroom teaching is fully teacher-centered, subjecting learners to one-size-fits-all pre-determined and pre-organized procedures (Lian, 2014; Lian & Pineda, 2014; Sangarun, 2014). In other words, every student in the classroom is engaged in the same learning task(s) without taking into account the individual differences. However, language learners have different purposes and will need to be supported differently. Furthermore, a particular problem might be difficult for one student but not difficult for another. It is also accepted that people inevitably understand differently as a result of individual differences that emerge from their diverse backgrounds ranging from prior experience and internal representations to sociocultural practices and cultural discourses (Eskey, 2005). It is further understood, increasingly, that for optimal outcomes, students need to be supported optimally in terms of individual differences (Dörnyei, 2005; Ellis, 2008).

In response to the problem of neglecting personal understanding, individuality and optimality in the behaviorist view of learning, Petar Guberina proposed the verbotonal theory to understand better the concept of learning which regarded these features as indispensable for knowledge construction. In the verbotonal point of view, as Lian (2011) noted, “knowledge construction is understood increasingly as an act of individual meaning-making rather than as an act of information-passing or simple memorisation” (p. 7); and understandings are constructed by using “optimal learning condition” (a.k.a. optimal listening condition, optimal frequency zone and optimal filter) to go beyond the information given (Asp, 2006; Guberina, 1972). The concept indicates that learning is an active, ongoing process, and individuals construct their new knowledge based upon their optimal understandings, which differ from person to person and is, therefore, idiosyncratic. Such an idea is also reflected in the notion of “habitus” (Bourdieu, 1990). Hence, when seeking to optimize students’ learning, we need to change their filter and make the meaningless filtered out, thus making the meaningful stayed in the new habitus.

However, students’ prior knowledge is personal, complex and highly resistant to change (Jensen, 2005). Moreover, knowledge construction requires a high level of mental effort which cannot be seen in the ordinary sense. Awareness, as stated by Gattegno (1987), is the only thing that is educable, and therefore, the only way that we can do to rewire the students’ brains is to raise their awareness. Similarly, according to Lian and Pineda (2014), “awareness-raising is the first step in the reconstruction of personal operational histories as, without it, it would be essentially impossible to bring into the learner’s field of relevance what had, until now, been irrelevant, i.e. unknown” (p. 20). In other words, in order to learn new knowledge, we need to learn with awareness and be aware of what we are unaware of, i.e. making the meaningless meaningful (Lian, 2000). Therefore, learning is understood as a meaning-making process (Lian, 2004; Lian & Pineda, 2014). This indicates that mechanisms for learning are based on and rely centrally on changing learners' personal operational histories (prior knowledge) with awareness-raising as the key to enabling them to construct meanings (new knowledge). As a result, the primary focus must be on raising learners' awareness, enabling them to make sense of linguistic as well as non-linguistic signals concerning spoken and written texts.

As discussed above, the conceptual basis of the present research is primarily shaped by Guberina's verbotonal point of view of knowledge construction. Theoretically, the verbotonal theory of learning is in line with the notion that learning is a dynamic individual meaning-making process. As a part of the theoretical framework of the current study, a discussion of meaning-making is presented below.

2.4.2 Meaning-making

Over the decades, although a great number of researchers, such as Halliday (1978), Harris (2018), Kress (2010), Lian (2000, 2011, 2014), Lian and Pineda (2014), Mortimer and Scott (2003), to name just a few, have devoted themselves to the understanding and discussion of the concept of meaning-making, there is no simple consensus about how to define the term. In general, meaning-making is essentially viewed as a dynamic process rather than a static process (Chun, 2015), which plays a central role in all aspects of our lives (Chen, 2011).

From the meaning-making point of view, learning, then, is not one that involves ideas being transferred directly from teacher to student, parent to child, or friend to friend, but rather one that involves individual meaning-making process (Mortimer & Scott, 2003). Similarly, according to Lian (2004), learning is considered as a process of dynamic individual meaning-making, in order to learn, we need to make sense of what is happening to generate meanings. In other words, it focuses on the recognition and theoretical inclusion of the diversity of learners, putting emphasis on the significance of the individual learner in the learning process and the way in which the meanings are constructed by the learner.

In the present study, in order to correct the learners' pronunciation problems of a selection of English vowels, in line with the above-mentioned ideas, learners will be supported individually according to individual learning difficulties with the target vowel sounds. Hence, after two awareness-raising activities (i.e. in-class activities and out-of-class activities), learners may be able to construct their personal meanings, make sense of the target English vowels and correct the pronunciation problems on their own through active learning. As an important starting point for developing learning procedures of the verbotonal-based phonetic correction, the general pedagogic considerations guiding the current study cannot be neglected, and they are presented below.

First, learning is a special case of individual meaning-making, rather than knowledge transmission, relying on our internal logical and representational systems necessarily contain our operational histories (Lian, 2004; Lian & Pineda, 2014). In other words, learning as a knowledge construction and the meaning-making process happens at the individual level. More importantly, when we learn something, we make sense of it, construct a meaning actively and attribute this meaning to ourselves, the people around us and the world we live in based on each of our personal histories. As a result, if something does not make sense to the learners, they will never learn it, therefore, pedagogically, the primary focus must be on helping individual learners construct and make sense of the new personal knowledge.

Second, learning occurs when the learners' internal logical and representational systems take in the new personal knowledge which has been hitherto blocked unconsciously by them. Therefore, it is reasonable to work on the learners' personal operational histories if they fail to construct and make sense of new personal knowledge. In the present study, blocked by the "mother tongue sieve" (Trubetzkoy, 1939/1969), namely, the phonological system of Mandarin Chinese, Chinese non-English major EFL learners reject the new sounds (i.e. the target English vowels) in the English phonological system that they are trying to learn. As mentioned earlier, the L2/FL learners' unintelligible speech production is caused by their inability to perceive the non-native sounds correctly. In this sense, learners' perceptual systems need to be reeducated and changed to embrace the new sounds.

Third, under these conditions, mechanisms for learning are dependent centrally on changing the learners' personal histories (i.e. perceptual systems) with awareness-raising as the key to enabling learners to create meanings and construct new personal knowledge, namely, the English vowels under study. (Gattegno, 1987; Lian, 2004, 2014; Schmidt, 2010). There are many ways of raising awareness. However, the most effective way of awareness-raising turns out to be modification or manipulation of input concerning sound signals. Moreover, this works directly on the neuroplasticity. Therefore, it helps reshape the learner's perceptual systems. In the present study, this enables the learners to listen to the same sounds filtered through their corrective optimals, thus enriching their understandings of the features of sounds that have yet to be perceived by them.

As presented above, meaning-making provides a basis for the guiding pedagogic principles that account for the current study. Therefore, learners' personal operational histories (i.e. perceptual systems) need to be changed with the help of awareness-raising (i.e. input manipulation) techniques to enable their knowledge construction and meaning-making of the sounds hitherto blocked by them.

2.5 Awareness-raising

Over the past several decades, there have been a series of studies concerning awareness-raising and most of them related to language awareness (Ahn, 2016; Carter, 2003; Hawkins, 1999; Kennedy & Trofimovich, 2010; Little, 1997; Lucas & Yiakoumetti, 2019; Schmidt, 1993, 1995, 2012; Valeo, 2013; Van Lier, 1996; White & Horst, 2012; Zenotz, 2012). As Schmidt (2012) pointed out, "to many people, the idea that SLA is largely driven by what learners pay attention to and become aware of in target language input seems the essence of common sense" (p. 27). However, in reality, awareness should not be restricted to the language aspect alone, and it consists of a range of awarenesses of many different kinds that help learners develop the ability to function properly in a language (Lian, 1993).

As mentioned in the previous chapter, our perceptions are mediated by our personal filter (Guberina, 1972) and our personal habitus (Bourdieu, 1990). According to Lian and Pineda (2014), we are physiological beings and "we do not necessarily sense the world as it really is but as we perceive it, as our past experience dictates" (p. 12). If we are not aware of a certain phenomenon in some way, then it might be the evidence that it does not exist in our personal operational histories (Lian, 2000; Mason, 1998). This indicates that it is our habits that prevent us from seeing new things, which makes meaning-making difficult to achieve and therefore in order to create new personal knowledge, we need to defeat the habits and change the ways in which we make sense of things. The basic assumption behind awareness-raising is that learners have been habituated to ignore or reject signals that they are not familiar with. Now, the habits are our current meaning-making mechanisms (Lian, 2000). Moreover, based on our normal meaning-making mechanisms, we do not always perceive things in the way they really are. In this sense, awareness-raising basically comes from a kind of breaking away from the habitual and presenting things in novel ways and in novel conditions to bypass our normal meaning-making mechanisms.

As we grow older, we become more efficient in known contexts and less efficient in unknown contexts, and this happens as a result of maturation and the accompanying neurological pruning (Jordan, Carlile, & Stack, 2008). Therefore, it is reasonable that neurological pruning should be defeated to function properly in unknown contexts. A good example is the so-called rubber-hand experiment first reported by Botvinick and Cohen (1998). In the rubber hand experiment, the participants sat with their left arm resting on a table, which was hidden behind a screen. A rubber hand model of a left hand and arm was placed directly in front of them, and they were asked to fixate on a rubber hand in front of them as if it was their hand. To induce the rubber hand illusion, the experimenter synchronously stroked the participant's hidden hand and the artificial rubber hand with two brushes. After a short while, the participants reported that they tended to feel the touch on the rubber hand instead of the real hidden hand at the location where they saw the rubber hand being stroked. Participants also reported feeling ownership over the rubber hand, i.e. the rubber hand was their own hand (Tsakiris & Haggard, 2005). This is a noticeable illusion as it indicates that our perception is partly dependent on visual elements. Moreover, the experiment shows that proprioceptive inputs are also influenced by this visual illusion. In this experiment, we do something in a way that we do not normally go through. This is an illustration of the reality of neuroplasticity in real life, which indicates the ability of the brain to change. Therefore, we can defeat neurological pruning by creating new connections in people's brain through special techniques which creates new associations between signals and understanding of these signals.

The above-mentioned ideas are in line with the verbotonal theory of phonetic correction. People hear by making choices from the income signals. The natural signals are rich and give listeners the opportunity of finding what they are familiar with within specific contexts. When it comes to the filtered signals, however, they are not familiar to the listeners, and their ears have fewer choices to select and therefore fewer opportunities for them to make mistakes in terms of the sounds that they are looking for. In other words, when they hear the filtered sounds, they are hearing the optimal sounds that their perceptual mechanisms have not allowed them to hear before. Unfortunately, people's brains may still distort those signals in order to make them familiar because our brain has certain habits and preferences and therefore, even if we present them with optimal frequencies for native speakers. Hence, we have to modify them to compensate for the distortion by the non-native speakers' ears. As a result, awareness-raising is not just listening to something but compensating for the non-native speakers' meaning-making mechanisms.

Taking into consideration the purpose of raising each student's awareness of the pronunciation, in the present study, Chinese non-English major EFL learners will be optimally supported by listening to the sound stimuli with the target English vowel sounds filtered through personal corrective optimals.

2.6 Theoretical framework

As presented above, the literature review of the present study discussed so far turns out to be a coherent, well-linked whole. In this sense, it sets a good stage for the theoretical framework underpinning the present study. In other words, the literature review provides a theoretical foundation, based on which the theoretical framework of the current study can be developed. In order to have a better understanding of the study, the importance of the theoretical framework cannot be neglected, which is discussed as follows.

The present study aims at investigating the corrective optimals and examining whether the use of the corrective optimals alone will be sufficient to bring about acceptable production of the target vowels in Chinese EFL learners. Taking into account that the core learning concept of the study is guided by meaning-making, individuality and optimality, the whole study is conducted using the verbotonal theory of phonetic correction (i.e. optimal octaves) as the starting point.

There is no denying that we are all meaning-making beings rather than simply meaning-receiving beings. We make meanings based on our personal operational histories. From this point of view, learning is an individual knowledge construction and meaning-making process, which depends on the individual's internal logical and representational systems (Lian, 2000, 2004; Lian & Pineda, 2014). Therefore, to construct meanings of new personal knowledge, the learners' internal logical and representational systems must be changed to bring in the new knowledge which has long been excluded from the personal operational histories.

Up till now, the only way, as well as the most effective way to change the personal operational histories, is awareness-raising (Gattegno, 1987; Lian, 2014). There are many ways of awareness-raising, such as input manipulation. Experimentally, the results of the input manipulation have been proved to be positive in terms of the verbotonal-based phonetic correction (Mildner & Bakran, 2001; Mildner & Tomić, 2007; Tomić et al., 2011).

Moreover, from the verbotonal point of view, each language system has a set of optimal octave bands, and language learners' auditory speech perception for non-native sounds can be developed if each individual is supported by the personal optimal listening condition (Asp, 2006). Learners will correctly perceive and acceptably produce the speech sounds if they are exposed to the optimal frequency bands of the sounds, i.e. optimals (Guberina & Asp, 1981). Based on these ideas, the concept of corrective optimals is developed utilizing the native speaker optimals as our point of departure.

In conclusion, based on the aforementioned considerations, the theoretical framework of the present study consists of three components: (a) meaning-making, (b) awareness-raising and (c) verbotonal system. Within this theoretical framework, all the components work together in a consistent and harmonious way to raise the Chinese non-English major EFL learners' phonological awareness. As a result, their perceptions will likely be reorganized. Moreover, they will be most likely to self-correct their pronunciation problems concerning the target English vowels. The figure below presents how this framework is constructed (see Figure 2.1).

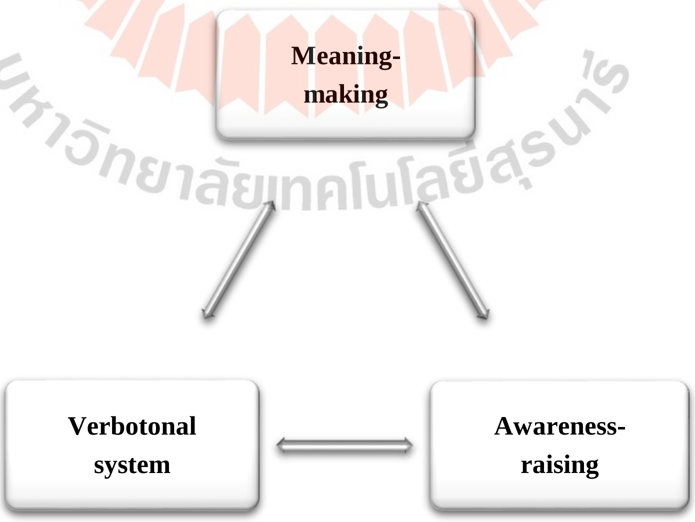

Figure 2.1 Theoretical framework of the present study

2.7 Summary

This chapter began with an emphasis on pronunciation teaching and research, followed by a discussion of phonetic correction, verbotonal theory, meaning-making and awareness-raising. Based on the theoretical foundation embedded in the literature review, it led to the construction of the theoretical framework of the current study. The review indicated that the verbotonal theory of phonetic correction was still not well investigated in China, and no research has ever been conducted concerning the corrective optimals for Chinese EFL learners. Therefore, the present study will fill this research gap by investigating the correctives optimals for Chinese non-English major EFL learners and further developing a verbotonal-based approach to assist their pronunciation learning. The next chapter explains the details of the research methodology for the current study.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter describes the methods, materials and procedures used in the present research. It consists of eight sections. The first section gives an account of the background and grouping details of the research participants. The second section discusses the research design. The third section provides an account of the variables. The fourth section gives a description of the research instruments. The fifth section focuses on pedagogic procedures. The sixth section is concerned with data collection procedures. The seventh section is devoted to data analysis. In section eight, a summary of the chapter is given.

As a reminder of what were explored in the present study, the research questions are restated below.

(1) What are the corrective optimals of Chinese non-English major EFL learners for the following English vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/?

(2) Are there any differences between the native speaker optimals and the corrective optimals for Chinese non-English major EFL learners? If yes, what are these differences?

(3) Is the VTPL approach effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners? In particular, is

simple exposure to the corrective optimals alone sufficient to bring about acceptable production of the target vowels?

(4) Which approach, VTPL or traditional, is more effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners?

(5) What are the Chinese non-English major EFL learners' opinions of the VTPL approach to phonetic correction?

3.1 Research participants

The participants in the present study were 76 first-year non-English major EFL learners enrolled in the full-time undergraduate degree course *College English 1* in the first semester of the academic year 2016-2017 at GNU, Ganzhou, Jiangxi, China. They ranged in age from 17 to 21, with a mean age of 19. The average length of time the participants had spent learning English was 9 years, ranging from a minimum of 6 years to a maximum of 13 years.

Ethics clearance was secured from GNU's Division of College English Teaching and Research. The procedures, potential benefits and main purposes of the study were carefully explained to all students who explicitly signified their assent to participate. This confirmed that the researcher was permitted to collect and analyze the experimental data for research purposes under strict rules of anonymity and security.

All participants reported having normal hearing as determined by a pure tone hearing test at the time of the college entrance physical examination. The 76 participants came from three majors, i.e. international economics and trade, financial management and human resource management. In addition, according to the paper-based college English placement test conducted yearly at GNU for the newly enrolled non-English major undergraduate students, they attained similar levels of English proficiency. They were randomly assigned to the experimental and control groups using a quasi-experimental design. Participants in the experimental group were 37 first-year non-English major EFL learners. Participants in the control group were another 39 first-year non-English major EFL learners. Two students were excluded from the experimental group because they provided incomplete data.

Experimental group

The experimental design of the present study was inspired by Guberina's (1972) concept of optimal octaves for English sounds (i.e. vowels and consonants) and was based on the verbotonal theory of phonetic correction. Under the guidance of these ideas, the notion of individualized perceptual training through a verbotonal-based approach (i.e. the VTPL approach) for Chinese non-English major EFL learners was proposed. This is the treatment that the experimental group received. Specifically, the experimental group was required to listen to and repeat recordings enhanced through the use of digital filtering and presenting the target vowels both in isolation and in contrast with one another. A diagnosis of the corrective optimals for each student was conducted prior to the first in-class training session. During the training sessions, the students listened to and repeated optimally-filtered vowels and vowel combinations. Thus, the students' perceptions were optimally supported at all times and therefore, were most likely to enable them to correct their wrong perceptions and produce acceptable target English vowels.

Control group

In the present study, the control group also went through exactly the same training sessions, but without the benefit of digital filtering (i.e. using the traditional approach). Specifically, the control group was required to listen to and repeat the same recordings as those of the experimental group, but these had not been enhanced with digital filtering. They listened to entirely natural language. Care was taken to ensure that the students in this group spent the same amount of time as the students in the experimental group practicing the pronunciation of the target vowels both in isolation and in contrast with one another.

The only differences between the two groups were that the experimental group went through a diagnostic phase for personal corrective optimals and the experimental group listened to optimally filtered materials while the control group did not. In this way, the impact of using digital filtering within the context of this experiment could be assessed.

3.2 Research design

It has been widely accepted that a mixed methods research design (i.e. both quantitative and qualitative methods) can help strengthen a study in a number of ways. According to Creswell (2014), the advantage of a mixed methods design is that limitations inherent in each method can be neutralized or offset. Besides, the ability to draw on both quantitative and qualitative data enables the triangulation of data (Tashakkori & Teddlie, 2003; Teddlie & Tashakkori, 2009). In other words, in order to corroborate results, the data collected from the quantitative approach is checked against and combined with the data collected from the qualitative approach.

In the present study, a mixed methods research design was applied, using both quantitative and qualitative approaches for collecting and analyzing data. Because the participants were not randomly selected, a truly experimental design was not possible. Therefore, a quasi-experimental design was developed and used instead.

In order not to affect the normal teaching of *College English I* at GNU, there was no formal classroom instruction on English vowels. In this study, the students from both the experimental and control groups participated in after-school supplementary classes for phonetic correction lasting 1 hour per week. In addition, they were told to study another 30 minutes a week out of the classroom to maximize the benefits from the classroom work. Prior to the experiment, the students from both groups were introduced to the training activities as well as procedures. This was to give them a good understanding of the whole training process (no content teaching).

The whole learning process lasted for an 8-week period, beginning in October 2016 and ending in December 2016 (see Table 3.1). Specifically, students in the experimental group listened to the sound stimuli with the vowel sounds under study filtered through their personal corrective optimals. Students in the control group were exposed to the same sound stimuli with no filtering.

Table 3.1 Learning procedures

Week	Learning activities	
	EG	**CG**
1	/ɪ/-/iː/ (Activity 1 & Activity 2)	/ɪ/-/iː/ (Activity 1 & Activity 2)
2	/ɪ/-/iː/ (Activity 3 & Activity 4)	/ɪ/-/iː/ (Activity 3 & Activity 4)
3	/e/-/æ/ (Activity 1 & Activity 2)	/e/-/æ/ (Activity 1 & Activity 2)
4	/e/-/æ/ (Activity 3 & Activity 4)	/e/-/æ/ (Activity 3 & Activity 4)
5	/ʊ/-/uː/ (Activity 1 & Activity 2)	/ʊ/-/uː/ (Activity 1 & Activity 2)
6	/ʊ/-/uː/ (Activity 3 & Activity 4)	/ʊ/-/uː/ (Activity 3 & Activity 4)
7	The first review of weeks 1-6	The first review of weeks 1-6
8	The second review of weeks 1-6	The second review of weeks 1-6

Note. EG = Experimental group; CG = Control group

Stimuli

Based on the fact that non-English major undergraduate students have been instructed using more British English, the students were more accustomed to British English. Thus, British English was adopted as the standard for sound stimuli in the current study.

The training stimuli for the experiment consisted of a number of logatomes (e.g. /sɪsɪ/), monosyllabic words (e.g. ship) and sentences (e.g. Is it ship or is it sheep? Ship or sheep?), involving the English vowels under study. Specifically, monosyllabic words were used both in the diagnosis of the students' corrective optimals and in the training sessions. In contrast, the logatomes and sentences were used only during the training process. They were recorded by a male English native speaker (with a British accent), who was not involved in any tests or training sessions of the study. The recordings were conducted in a soundproof booth at the Suranaree University of Technology Sound Studio with a Focusrite Scarlett 2i4 USB Audio Interface and Rode NT1-A microphone. Before recording, the speaker was given some time to read through the materials. Besides, the speaker was told to read at normal speed and loudness. The recordings were digitized at a sampling rate of 48000 Hz and stored as .WAV files for the purpose of sound editing. In the end, all stimuli were rechecked to ensure their quality and intelligibility.

Filtering

The vowel sounds and only the vowel sounds in recordings of the logatomes, words and sentences were filtered using the students' personal corrective optimals as described below. Taking the words "ship" and "sheep" as examples in the sentences "Is it ship or is it sheep? Ship or sheep?", only the vowel sounds /ɪ/ and /iː/in words were filtered, and the rest of the sounds remained the same. This resulted in the following sentence (underlined and enlarged portions represent filtering): "Is it ship or is it sheep? Ship or sheep?". The sentence, therefore, consisted of a mix of filtered and natural language, with optimal filters applied only to the vowel sounds under study. In the current study, the audio-editing program Audacity (Version 2.1.2; Audacity Team, 2016) was used for the digital filtering of the target English vowel sounds (see Figure 3.1, Figure 3.2 and Figure 3.3).

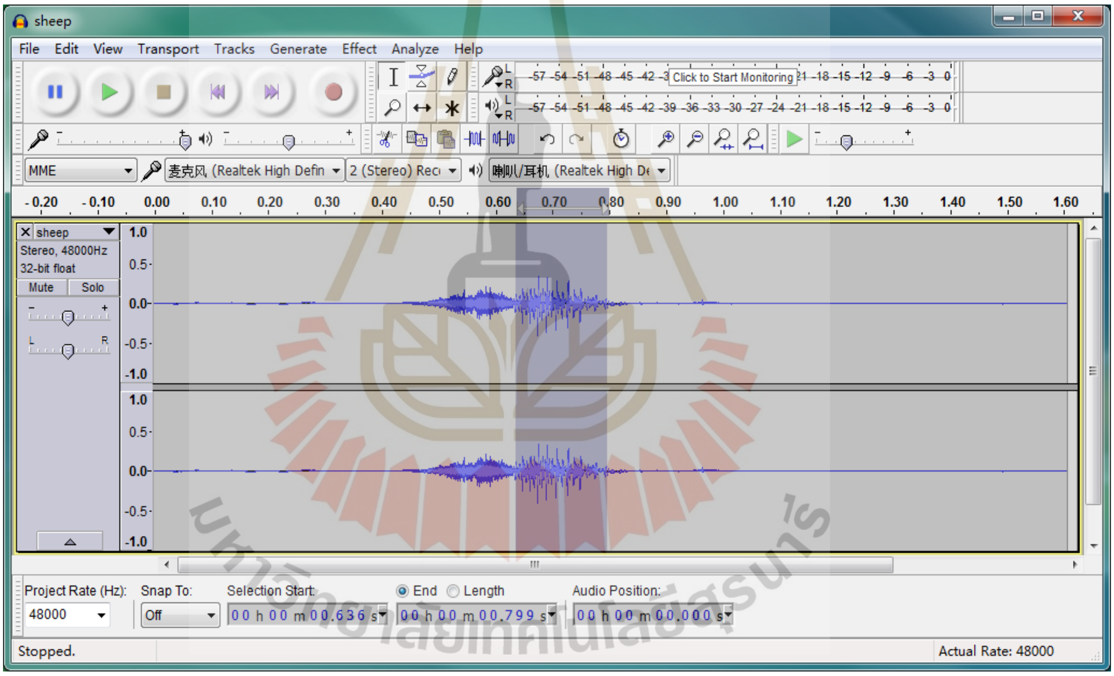

Figure 3.1 Sound selection for the vowel /i:/ in "sheep"– selection is unfiltered /i:/

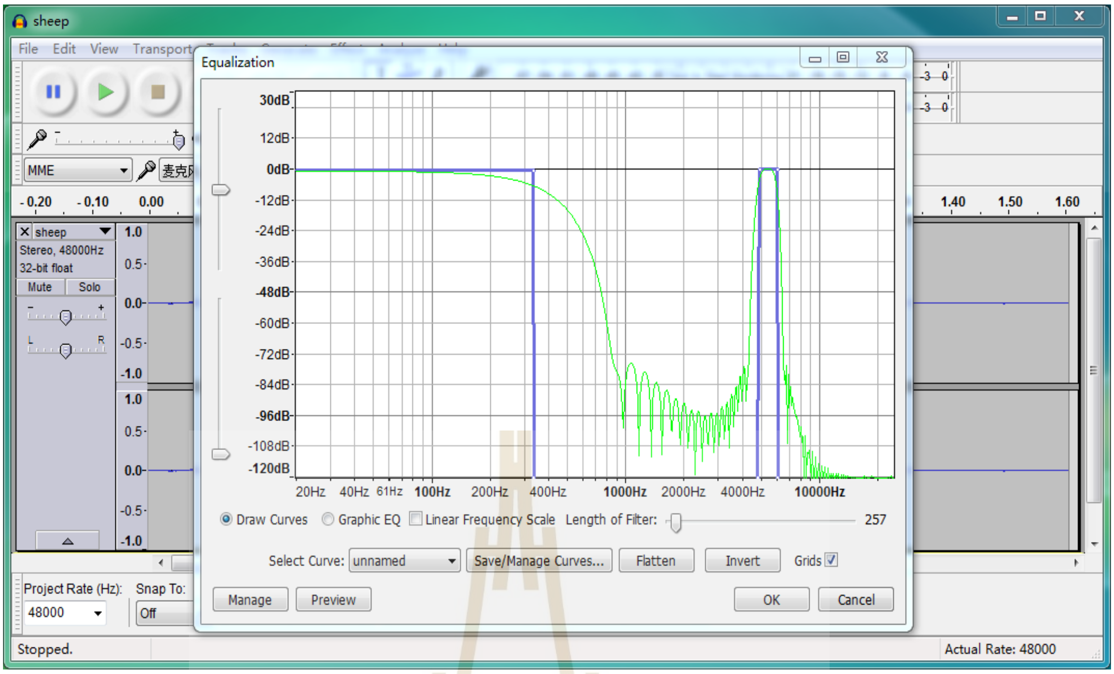

Figure 3.2 Filter details for the sound /i:/ in "sheep" (0-320 Hz + $\frac{1}{3}$ octave)

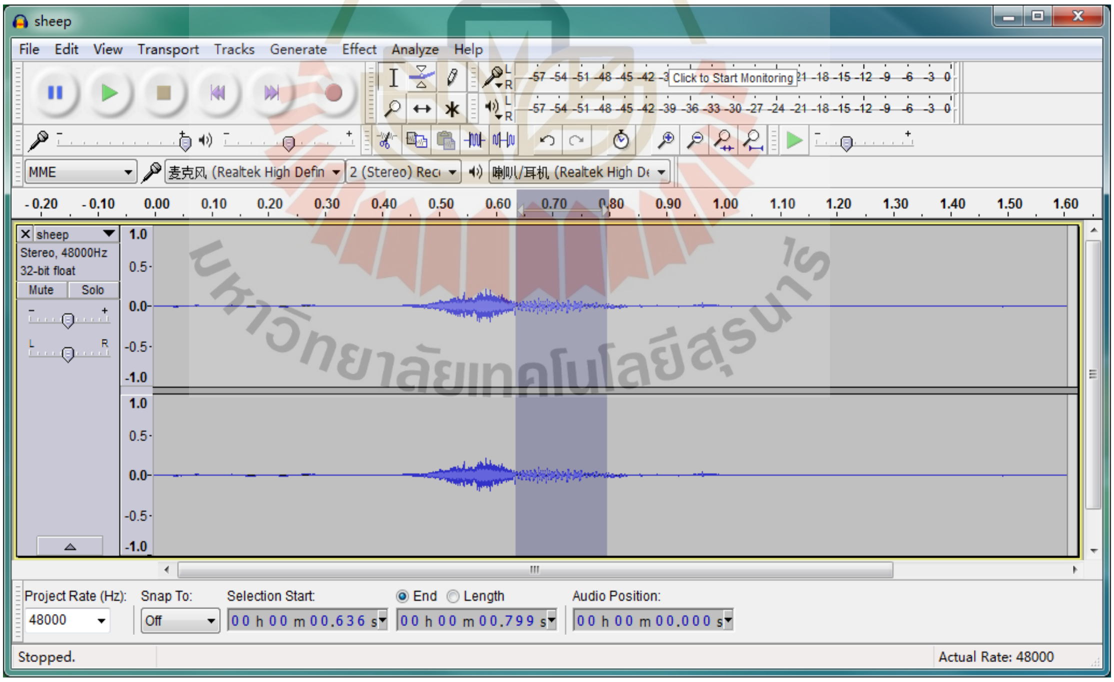

Figure 3.3 Sound filtering for the sound /i:/ in "sheep"

(0-320 Hz + 4838-6096 Hz) – selection is filtered /i:/

With regard to training, the researcher took care of both the experimental and control groups. The person who managed the lab was present to supervise the training process of the two groups during the in-class training activities to make sure that everything was working properly. However, the whole process was student-centered rather than teacher-centered. There was no teaching performed in any form by the researcher or anyone else. In other words, the whole process was teacherless. In the present study, all the participants were involved in self-managed learning. Therefore, the results were not biased in terms of approach.

As for the quantitative aspect, the data were collected from the diagnosis of corrective optimals, questionnaires, the perception pre- and posttests and the production pre- and posttests. When it came to the qualitative aspect, in order to corroborate results and achieve the purpose of data triangulation, students' questionnaires, self-reports as well as semi-structured interviews were applied.

3.3 Variables

With regard to the research questions of the present study, the independent and dependent variables were as follows:

The independent variables were the manipulated variables. They included a set of broad and modified corrective optimals and two different pronunciation learning materials (i.e. filtered and unfiltered) to the phonetic correction of the vowel sounds under study.

The dependent variables were the measured variables. For this, the dependent variables included the students' best production of the target vowels (i.e. identification of the corrective optimals), their scores on both perception and production pre- and posttests and their attitudes toward the approach used to learn the target English vowels.

3.4 Research instruments

The data collection instruments developed in the present study were the diagnosis of the corrective optimals, questionnaires, tests, self-reports and interviews. Specifically, questionnaires referred to the pronunciation learning questionnaire and the pronunciation learning satisfaction questionnaire. The tests referred to the perception pre/posttest and the production pre/posttest. The self-reports were students' written records of their pronunciation learning process. The interviews involved were semi-structured. They are described in more detail as follows.

3.4.1 Diagnosis of the corrective optimals

In this study, the concept of corrective optimals was inspired by Guberina's (1972) notion of optimals (i.e. optimal octaves) for English vowels and consonants. In other words, the determination of the corrective optimals of the target English vowels for the students was based on Guberina's native speaker optimals of the target vowel sounds.

To make the target English vowels more salient and act more effectively on students' perceptual systems, the notion of "optimals" was redefined to include both broad optimals (full octaves) as well as modified optimals. This was a significant departure from, as well as an important refinement of, the original dogma of Guberina optimals which work only with full octaves. Therefore, it was possible to get a much finer understanding of the frequencies that mattered to each student. As part of this study, it was assumed that is was possible that the corrective optimals might not be identical with the native speaker optimals. In other words, students' corrective optimals could be different from native speaker optimals and also from one another, i.e. that, unlike earlier studies, there was no single optimal for all ears.

To determine the Chinese non-English major EFL learners' corrective optimals for the target English vowels, students from the experimental group participated in the experiment. The sound stimuli used for diagnosis were a list of 6 monosyllabic words (i.e. ship, sheep, bed, bad, soot, suit). The diagnosis was individually carried out in a language laboratory at GNU. The students went through two steps as follows.

Step 1. Identify the fcenter for each corrective optimal

In step 1, as a starting point, in order to determine each individual student's preferred fcenter for a particular vowel sound, the procedure began with exposure to a particular fcenter (i.e. the fcenter for the native speaker optimal). The student listened to it 3 times. Then, the student was required to repeat the word he/she heard. It is worth noting that this did not necessarily lead to the acceptable production of the sound. When the student failed to produce the target sound acceptably, a different (higher or lower) fcenter was presented according to the student's performance. As a result, a set of frequency centers for each student was determined where the student was most likely to produce vowel sounds close to the target vowel sounds.

Step 2. Provide variations of the fcenter for each corrective optimal

In step 2, once the best fcenter has been determined, the target vowel of each word was filtered using both single bandpass filters and discontinuous multiband filters (i.e. modified filtering). Therefore, we were able to determine a battery of potential corrective optimals consisting of full octaves, partial octaves ($\frac{1}{2}$ octave and $\frac{1}{3}$ octave) and full and partial octaves with a lowpass component (0-320 Hz + 1 octave, 0-320 Hz + $\frac{1}{2}$ octave and 0-320 Hz + $\frac{1}{3}$ octave). Together with the 1-octave filter determined above, every vowel under study was enhanced using 6 filters (i.e. 1 octave, $\frac{1}{2}$ octave, $\frac{1}{3}$ octave, 0-320 Hz + 1 octave, 0-320 Hz + $\frac{1}{2}$ octave and 0-320 Hz + $\frac{1}{3}$ octave). Each time a student listened to a specific filter 3 times. The student was then required to try to say the word, and at some point in the process, the student was in a position to best produce the sound under study. When that happened, the particular filter was noted and used to give the student exercises tailored to their preferred perceptual profile.

The researcher and another experienced English teacher assessed the production quality of the target vowel sounds. After rechecking productions several times, the researcher determined each individual student's listening profile (the combinations of corrective optimals-6 filter settings for the 6 vowels studied) for the target English vowel sounds being studied.

3.4.2 Questionnaires

Broadly speaking, questionnaires can be used to collect three types of data from the respondents: factual, behavioral and attitudinal. Specifically, factual data typically cover demographic information as well as any other background information relevant to the study. Behavioral data focus on the respondent's actions, lifestyles, habits and personal history. Attitudinal data concern interests, attitudes, opinions, beliefs and values (Dörnyei & Taguchi, 2010).

Two questionnaires were constructed for the present study. Questions in the questionnaires were mainly closed-ended. The first one was designed to gather information about how Chinese non-English major EFL learners went about learning of English pronunciation. The second one was devoted to exploring the Chinese non-English major EFL learners' opinions after learning the target English vowels.

The first questionnaire for pronunciation learning consisted of two parts (see Appendix A): general information and pronunciation learning information. The first part included students' personal information, i.e. name, gender, major, age, place of birth, minority background, Chinese dialect background and college entrance exam score for English. The second part was made up of 7 questions focusing on the students' past pronunciation learning experience and their attitudes toward pronunciation learning. The second questionnaire consisted of 9 statements, where students declared their degree of agreement concerning pronunciation learning satisfaction (see Appendix D). In this questionnaire, students were told to rate their answers using a 5-point Likert scale. Values on the scale were "strongly agree", "agree", "undecided", "disagree" and "strongly disagree". Additional comments could be added at the end.

To avoid possible problems of ambiguity and misinterpretation, both questionnaires were then translated into Chinese. Four experts were invited to evaluate the content validity of the two questionnaires. These experts were academically qualified, and they all had a long and rich experience of teaching English. The experts rated each questionnaire item on content appropriateness and clarity using the item-objective congruence index (IOC) procedure as a content validation approach to determine whether each item was congruent with the objective of the questionnaire or not. The evaluation form used a 3-point scale (i.e. 1 = relevant, 0 = uncertain, -1 = irrelevant). An IOC value of more than 0.75 is considered to be acceptable (Rovinelli &

Hambleton, 1977). The IOC values of the two questionnaires were 0.96 and 0.94 respectively (see Appendix F and Appendix I). These indices indicated that both questionnaires were acceptable.

3.4.3 Self-report

In order to trace the students' progress as well as promote their responsibility, management and mastery for individual pronunciation learning, they were asked to self-report the general information about the pronunciation learning process over the 8 weeks period of the experiment (see Appendix L). Specifically, both the experimental group and the control group were required to report and record information, including name, date, place, starting time, ending time, materials, problems, progress and activities before and after learning. Based on the students' self-reporting data, the researcher was able to examine their pronunciation practice activities over time.

3.4.4 Tests

According to Phillips and Stawarski (2008), "testing is important for measuring learning in program evaluations, pre- and post- program comparisons using tests are common, an improvement in test scores shows the change in skills, knowledge or attitude attributed to the program" (p. 13). In the present study, two types of tests were constructed, i.e. perception test and production test. The perception test was adapted from the diagnostic test of sound discrimination (Baker, 2006). The production test was designed according to the speaking tests used for English major undergraduate students at GNU.

The perception pre/posttest was a combination of discrimination and identification tasks, which consisted of 90 test items (see Appendix B). In each test item, the students listened to two words. As for the discrimination task, the students should decide whether the sounds they heard were the same or different. When it came to the identification task, they were asked to complete a more challenging task, i.e. they should indicate which pair they were listening to and circled the relevant answer provided. The perception pre/posttest was a pencil and paper test. The students followed the instructions and listened to the 90 sound stimuli through headphones in a university language laboratory.

The production pre/posttest was made up of three parts (see Appendix C). Part I tested the students' discrimination abilities concerning the target English vowels in the form of word-reading (60 items). Part II evaluated their pronunciation of English in terms of the target vowel sounds, comprehensibility, fluency and pronunciation using a sentence-reading task (9 items). Part III measured the improvements of students' English pronunciation concerning the target vowels, comprehensibility, fluency and pronunciation through storytelling (1 item). At the time of the production test, the individual who took the test was given a piece of paper with all the test items printed on it. Their performances were recorded using a digital voice recorder in a university language laboratory. Before recording, the students were given some time to read through the materials. The students were asked to write their ideas (i.e. three stories) to elicit more information (i.e. the target vowel sounds), that is to say, the part of storytelling was designed as the prepared speech rather than the impromptu speech. In addition, the students were asked to speak as clearly as possible during the recording phase. In the end, each recording was assigned a 7-digit random number. All the recordings (i.e. word-reading, sentence-reading and storytelling) stored in the form of .WAV files, together with the scanned images of the students' written stories named after 7-digit random numbers used for recordings, were sent to three qualified experts for double-blind rating. Specifically, the raters were asked to give ratings in respect of the target vowels and for comprehensibility, fluency and pronunciation (see Appendix K).

The IOC value of the perception pre/posttest was 0.98 (see Appendix G). The IOC value of the production pre/posttest was 0.97 (see Appendix H). The indices of the two tests had the IOC values greater than 0.75, and therefore, both tests were acceptable.

Raters

In the present study, three experts were invited to score the recordings collected in the production pretest and posttest from the experimental and control groups. The three experts had a minimum of ten years' experience with English teaching and rich experience concerning the assessment of speech productions. In order to check the level of agreement among the raters, an inter-rater reliability evaluation (Pearson's correlation coefficient r) was performed. According to Muijs (2004), correlation is said to be reasonable if the absolute value of the correlation coefficient r is greater than 0.7. The results of inter-rater reliability for the assessment of the vowel sound /ɪ/ concerning word-reading in the production pretest for the experimental group indicated that correlation coefficients were higher than the threshold value of 0.70. Correlation coefficients (r) were 0.85 (rater 1 - rater 2), 0.83 (rater1 - rater 3) and 0.90 (rater 2 - rater 3) respectively (see Table 3.2). Therefore, the values of inter-rater reliability were acceptable for the current study.

Table 3.2 Results of inter-rater reliability analysis

	Pearson's r
Rater 1 - Rater 2	0.85
Rater 1 - Rater 3	0.83
Rater 2 - Rater 3	0.90

3.4.5 Semi-structured interview

The interview has been regarded as one of the most widely used and basic methods for collecting qualitative data concerning interviewee's opinions, beliefs and feelings (Ary, Jacobs, Sorensen, & Walker, 2014). Fundamentally, there are three main types of interviews: structured, unstructured and semi-structured. In the semi-structured interview, the interviewer has a list of prepared questions but may also modify the format or questions during the interview process (Ary et al., 2014). According to Barkhuizen, Benson and Chik (2014), "the semi-structured interview is the most commonly used format in language teaching and learning research" (p. 17). Similarly, as stated by Nunan (1992), the semi-structured interview has been well known for its flexibility and has attained a high degree of popularity among researchers. Besides the flexibility it gives to the interviewer, the semi-structured interview also somewhat empowers the interviewees and gives them control over the process of interview.

In the present study, a semi-structured interview was used (see Appendix E). Sixteen percent of students (the researcher expected a random selection of about 15%-20%) in the experimental group were interviewed for in-depth information about their attitudes and opinions toward learning a selection of English vowels. To avoid the possible problems of ambiguity and to elicit more information, the interview was conducted in Mandarin Chinese (Putonghua). The IOC value of the interview was 0.97 (see Appendix J). The index of the interview indicated that the IOC value was more than 0.75. Therefore, the interview was acceptable. The face-to-face interview was conducted in a university language laboratory with the help of another English teacher. All the interview data were recorded using the digital voice recorder, and the transcripts of these recordings were made and analyzed.

3.5 Pedagogic procedures

As for the pedagogic considerations, two sets of activities were involved: in-class and out-of-class activities. In other words, the students first participated in the classroom activities and then they took part in self-managed after-class actives. The purpose of designing these two activities was to raise the students' awareness of the target English vowels and help them further develop intelligible pronunciation of these vowels.

3.5.1 In-class activities

The classroom practice consisted of four training activities, i.e. logatomes and monosyllabic words, optimal sentences, contrast-embedded sentences and non-optimal sentences. The first two activities were designed for practicing the target English vowels. The third activity was used for training the target vowels in vowel contrasts. The last activity was aimed at practicing the target vowels in non-optimal environments. The idea here was to get students involved in activities of progressively increasing difficulty, i.e. the easier ones first and the more difficult ones later. They were conducted in a university language laboratory. Being exposed to these activities, the students worked intensively on their personal corrective optimals for the target English vowels. More detailed information is provided as follows.

Activity 1. Logatomes and monosyllabic words

The first activity aimed to raise the students' phonological awareness and minimize their cognitive load of the target English vowels. At this stage, logatomes and monosyllabic words were used to provide the students with a variety of linguistic contexts for the target vowel sounds. Taking the vowel sound /ɪ/ as an example, the logatomes used were, for example, /dɪdɪ/, /hɪhɪ/, /lɪlɪ/ and /sɪsɪ/ presented in the pattern of CVCV (consonant-vowel-consonant-vowel). The monosyllabic words used were, for example, “bin”, “chip”, “kin” and “ship”. The reason for using logatomes and monosyllabic words was that there was no need for the students to make any sense of these materials. The students were also told not to try and make sense of them (not entirely possible because of the way the brain works). This was designed to enable them to focus on listening rather than sense-making, resulting in a likely reduction in the processing load and a re-allocation of processing resources to perception. Under these circumstances, it was surmised that the optimally filtered target sound could be better processed and act more effectively on their perceptual systems.

Before practicing the exercises here, each student was provided with his/her optimal vowel profile as diagnosed prior to the first training session. The students were asked to listen to each optimally-filtered logatome/word (using personal corrective optimals) 10 times without repetition. They then heard the logatome/word again and repeated the logatome/word after each playback. After performing the exercises in this activity, the students were able to focus on the characteristics of each target English vowel and therefore drew similarities of each vowel sound among different contexts. As a result, the articulation of the vowels under study would pose relatively little difficulty for the students.

Activity 2. Optimal sentences

The second activity also aimed to raise the students’ phonological awareness and to lower their cognitive load of the target English vowels. In this activity, the optimal sentences were designed to allow students to practice the pronunciation of the target English vowels under the best conditions for them. Take, for example, the optimal sentence for /iː/ (e.g. Lee, do you see green beads in the sea?). The reason for using optimal sentences was to assist students in perceiving and producing the target vowel sounds easily. Hence, the students' perceptions of these vowels could be reorganized after being exposed to the optimally filtered sentences in the optimal learning environments.

At this stage, the students continued using their optimal vowel profiles, as mentioned above. In other words, the students were provided with individualized learning materials. Specifically, the students listened to each optimal sentence enhanced by personal digital filtering 10 times and then repeated the sentence on completion of the repetitions. Such procedures could effectively remove irrelevant frequencies of the target vowel sound, which might prevent it from being perceived. After doing exercises in the optimal learning environments, it was expected that the students would be well sensitized and likely to acceptably produce the vowel sounds under study.

Activity 3. Contrast-embedded sentences

The third activity aimed at further raising the students' phonological awareness, while at the same time increasing and bringing the cognitive load back to normal as well as practicing their discrimination abilities by exposing students to the practice of the target English vowels in vowel contrasts (i.e. /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/). After learning the individual vowel sounds for some time, the students moved forward to the practice of these vowel contrasts. In this activity, the contrast-embedded sentences (using the sentence carrier "Is it __ or is it __? __ or __?") were designed for practicing students' discrimination abilities concerning the contrasting sounds. Take, for example, the contrast-embedded sentence for /ɪ/-/iː/ (e.g. Is it ship or is it sheep? Ship or sheep?).

At this stage, the students continued to practice the target English vowels using their optimal vowel profiles as described above. They worked intensively on the contrast-embedded sentences, emphasizing the contrasting sounds. Specifically, the students listened to each contrast-embedded sentence enhanced by personal corrective optimals 10 times and, on completion of this activity, repeated the sentence. After doing exercises in this activity, they were more likely to produce the target vowel sounds intelligibly.

Activity 4. Non-optimal sentences

The last activity was devoted to further raising students' awareness and increasing their cognitive load concerning the target English vowels. At this stage, the students moved from optimal environments to non-optimal environments, which were used to create difficulties for the students in perceiving and producing the target English vowels. Specifically, these sentences were designed using repetition of sounds (the target vowels), contrasting sounds (the target vowels in contrasts) as well as confusing sounds (any sounds that will create difficulty) to establish non-optimal environments which could confuse their perceptions. Take, for example, the non-optimal sentences for /ɪ/ (e.g. It is a ship which sits on the ripple heading for the city.) and for /ɪ/-/iː/ (e.g. It is a ship with sheep on a sea of ripples near the sea.). Hence, the students' correct perceptions established so far concerning the target vowels could be reinforced after being exposed to the optimally filtered sentences in the non-optimal learning environments.

In this activity, the students continued practicing the target vowels using their optimal vowel profiles. They worked intensively on the non-optimal sentences, highlighting the target English vowels. Specifically, the students listened to each non-optimal sentence using personal corrective optimals 10 times and then follow the recording. Similar to the above-mentioned activities, such procedures could effectively act on the students' perceptions of the target vowels. As a result, the students' understanding of these vowels could be enhanced, and they were most likely to articulate them in an intelligible way.

During the training process, the students learned at their own speed and were optimally-supported at all times through the use of their personal corrective optimals. It was also worth pointing out that the students were free to listen to the filtered materials as many times as they wanted to. Through personally optimized pronunciation learning, it was surmised that the students' perceptions could be reorganized. In the end, they could be well sensitized and reinforced to correctly perceive and acceptably produce the target English vowels.

It should be noted that the progression from optimal to non-optimal was also provided to the control group, thus equalizing its impact across both groups. It was therefore unlikely to be of any significance in accounting for the differences in performance.

3.5.2 Out-of-class activities

In this training program, learning was not limited to in-class activities. As an indispensable part of learning, the out-of-class actives should not be neglected. In the present study, the students in both groups were told to practice for another 30 minutes per week out of the classroom so as to help students improve their pronunciation through personal reinforcement activities. More importantly, with regard to the experimental group, all the personalized (optimally filtered) learning materials designed for classroom training were sent to each student through QQ (Version 8.6; QQ Team, 2016) file transmission and were available to download and use immediately. At the same time, all learning materials were uploaded in the QQ group, and therefore the students in the control group were also given free access to their training materials. This means that students from both groups could practice the learning materials which had been used in class activities for perceptual training at any time and anywhere.

Pilot study

In preparing the main experiment, a pilot study was carried out with a small number of first-year non-English major undergraduates. Specifically, procedures were tried out on 12 students, and everything went smoothly. At the end of the pilot study, they were interviewed, and there were no difficulties reported in the use of the filtered materials itself. They did not complain, and they were willing to practice their English pronunciation in such a way using digital filtering.

3.6 Data collection procedures

As mentioned at the beginning of this chapter, a mixed methods research design was adopted, which indicated that both quantitative and qualitative data were collected in the present study. Specifically, quantitative data were collected from the diagnosis of the corrective optimals, questionnaires and tests (i.e. the perception pre- and posttests and the production pre- and posttests). Qualitative data were gathered from students' questionnaires, self-reports and semi-structured interviews. The data collection methods and procedures were as follows:

Prior to the experiment, the students were asked to complete a pronunciation learning questionnaire to provide demographic information as well as personal information on pronunciation learning (see Appendix A).

Next, the students were randomly assigned to the experimental and control groups using different training approaches (i.e. the VTPL approach and the traditional approach).

Then, the students were given the perception pretest (see Appendix B) and the production pretest (see Appendix C) in order to determine whether the differences between the experimental and control groups were statistically significant before perceptual training in terms of the perception and production of the target English vowel sounds.

Before the first training session, the experimental group students' personal corrective optimals were diagnosed based on the native speaker optimals (Asp, 1972) set by Petar Guberina as the starting point.

During training sessions, the students in the experimental group listened to the materials enhanced using digital filtering. The students in the control group were exposed to the same stimuli without any filtering. The whole training process included two sets of activities: in-class activities and out-of-class activities. At the same time, the students were told to record and report their in-class as well as out-of-class pronunciation practice activities (see Appendix L).

After pronunciation training, the students were post-tested using the same perception test (see Appendix B) and production test (see Appendix C) to determine if there were statistically significant differences in the posttest between the two groups concerning the perception and production of the target English vowels.

Students' pronunciation performances (i.e. word-reading, sentence-reading and storytelling) in the production pre- and posttests were stored in .WAV files and each was assigned a 7-digit random number using a double-blind rating design.

What followed was a learning satisfaction questionnaire concerning pronunciation learning. At this stage, the students were asked to rate 9 statements using a 5-point Likert scale, ranging from “strongly agree” to “strongly disagree” (see Appendix D).

Subsequently, in order to get in-depth information about the student’ attitudes and opinions toward learning English pronunciation via the VTPL approach, 16% of students from the experimental group were randomly selected to take part in the semi-structured interviews (see Appendix E).

In the end, in order to corroborate data as well as increase reliability, the results collected from the two methods were compared and combined.

3.7 Data analysis

In this study, the data obtained through quantitative and qualitative methods were analyzed and interpreted in both quantitative and qualitative ways. The quantitative data were imported into JASP (Version 0.12.2; JASP Team, 2020) statistical software for analysis. The qualitative data were analyzed using a thematic analysis procedure (Braun & Clarke, 2006).

3.7.1 Quantitative data analysis

In this study, the quantitative data were gathered from the diagnosis of the corrective optimals, questionnaires, the perception pre- and posttests and the production pre- and posttests.

3.7.1.1 Descriptive statistics

The descriptive statistics were used to describe the basic features of the data in terms of the frequency distributions of the students' personal corrective optimals concerning the target English vowel sounds.

3.7.1.2 Paired samples t-test

The paired samples t-test was used to examine if there were significant differences between the pretest and posttest scores in the experimental and control groups. By using the paired sample t-test, the researcher can statistically conclude whether or not a particular training approach (i.e. the VTPL approach or the traditional approach) is effective for assisting the acceptable production of the target English vowels.

3.7.1.3 Independent samples t-test

The independent samples t-test was performed to determine whether there were significant differences between the means of the experimental and control groups in the pretest and posttest. By using the independent samples t-test, the researcher is able to statistically compare which approach (i.e. the VTPL approach or the traditional approach) is more effective for assisting the acceptable production of the target English vowels.

3.7.2 Qualitative data analysis

In this study, thematic analysis was used to analyze the qualitative data collected from the students' semi-structured interviews. However, this analysis was not used for the self-reports because the numbers of comments received were small, making it difficult to pull out themes. Prior to qualitative data analysis, recordings of semi-structured interviews were transcribed verbatim. Two experienced English teachers were invited to check the data and do the coding in order to make the qualitative part more rigorous and credible.

The data were analyzed by following the basic steps for thematic analysis (Braun & Clarke, 2006). First, for a better understanding of the data, all transcripts were read and reread repeatedly. Second, the data were read word by word to derive initial codes. Next, codes were sorted into potential themes. Then, the candidate themes were reviewed and refined. After that, the themes were defined and named. Finally, the report was produced.

3.8 Summary

This chapter discussed the research methodology of the present study. Research participants, research design and variables were described in sequence. Besides, different kinds of research instruments were presented. In addition, the pedagogic procedures were explained, including the in-class and out-of-class activities. Finally, the methods and procedures for data collection and analysis were given. In the next chapter, the results will be reported based on the research questions.

CHAPTER 4

RESULTS

With reference to the research methodology outlined in the previous chapter, this chapter presents the results of data analysis pertaining to all five research questions raised in Chapter 1. The chapter consists of seven sections. The first section is devoted to the corrective optimals and their comparison with the classical native speaker optimals. The second and third sections are concerned with the quantitative data analysis based on the statistical summary of data derived from both the perception and production pretests and posttests. The fourth section focuses on the results of comprehensibility, fluency and pronunciation testing in sentence-reading and storytelling. The fifth section reports the findings of the satisfaction questionnaire, self-reports and semi-structured interviews. In this section, data analysis covers both quantitative and qualitative aspects. The sixth section provides answers to the research questions for this study. The last section summarizes the chapter.

4.1 Corrective optimals

For the purpose of investigating the corrective optimals for the six target vowel sounds (i.e. /ɪ/, /iː/, /e/, /æ/, /ʊ/, /uː/), students in the experimental group were exposed to six monosyllabic words (i.e. ship, sheep, bed, bad, soot, suit). For each word, the target vowel sound was filtered through six broad and modified corrective optimals, including single bandpass filters and discontinuous multiband filters (i.e. 1 octave, $\frac{1}{2}$ octave, $\frac{1}{3}$ octave, 0-320 Hz + 1 octave, 0-320 Hz + $\frac{1}{2}$ octave and 0-320 Hz + $\frac{1}{3}$ octave).

Table 4.1 below shows the results of the corrective optimals of the six target vowels. The detailed information concerning bandpass filter, octave type, center frequency and percentage of best production are displayed in the table below to help better understand the corrective optimals identified.

Table 4.1 Corrective optimals of the target vowels as tested in this study

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		0-320 Hz + 2419-3048 Hz	0-320 Hz + $\frac{1}{3}$ octave	2715 Hz	33 (89.2%)
		0-320 Hz + 1815-2286 Hz	0-320 Hz + $\frac{1}{3}$ octave	2037 Hz	4 (10.8%)
		1280-2560 Hz	1 octave	1811 Hz	0 (0.0%)
		1440-2880 Hz	1 octave	2037 Hz	0 (0.0%)
/ɪ/	37	1600-3200 Hz	1 octave	2263 Hz	0 (0.0%)
		1760-3520 Hz	1 octave	2489 Hz	0 (0.0%)
		1920-3840 Hz	1 octave	2715 Hz	0 (0.0%)
		1713-2422 Hz	$\frac{1}{2}$ octave	2037 Hz	0 (0.0%)
		2283-3229 Hz	$\frac{1}{2}$ octave	2715 Hz	0 (0.0%)

Table 4.1 Corrective optimals of the target vowels as tested in this study (Cont.)

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		1815-2286 Hz	$\frac{1}{3}$ octave	2037 Hz	0 (0.0%)
/ɪ/	37	2419-3048 Hz	$\frac{1}{3}$ octave	2715 Hz	0 (0.0%)
		0-320 Hz + 1713-2422 Hz	0-320 Hz + $\frac{1}{2}$ octave	2037 Hz	0 (0.0%)
		0-320 Hz + 2283-3229 Hz	0-320 Hz + $\frac{1}{2}$ octave	2715 Hz	0 (0.0%)
		0-320 Hz + 4838-6096 Hz	0-320 Hz + $\frac{1}{3}$ octave	5431 Hz	32 (86.5%)
		0-320 Hz + 4567-6459 Hz	0-320 Hz + $\frac{1}{2}$ octave	5431 Hz	4 (10.8%)
		0-320 Hz + 4435-5588 Hz	0-320 Hz + $\frac{1}{3}$ octave	4978 Hz	1 (2.7%)
		2559-5118 Hz	1 octave	3619 Hz	0 (0.0%)
		2879-5758 Hz	1 octave	4072 Hz	0 (0.0%)
/i:/	37	3200-6400 Hz	1 octave	4525 Hz	0 (0.0%)
		3520-7040 Hz	1 octave	4978 Hz	0 (0.0%)
		3840-7680 Hz	1 octave	5431 Hz	0 (0.0%)
		4186-5920 Hz	$\frac{1}{2}$ octave	4978 Hz	0 (0.0%)
		4567-6459 Hz	$\frac{1}{2}$ octave	5431 Hz	0 (0.0%)
		4435-5588 Hz	$\frac{1}{3}$ octave	4978 Hz	0 (0.0%)
		4838-6096 Hz	$\frac{1}{3}$ octave	5431 Hz	0 (0.0%)
		0-320 Hz + 4186-5920 Hz	0-320 Hz + $\frac{1}{2}$ octave	4978 Hz	0 (0.0%)

Table 4.1 Corrective optimals of the target vowels as tested in this study (Cont.)

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		0-320 Hz + 2419-3048 Hz	0-320 Hz + $\frac{1}{3}$ octave	2715 Hz	34 (91.9%)
		0-320 Hz + 2016-2540 Hz	0-320 Hz + $\frac{1}{3}$ octave	2263 Hz	2 (5.4%)
		0-320 Hz + 1613-2033 Hz	0-320 Hz + $\frac{1}{3}$ octave	1811 Hz	1 (2.7%)
		1280-2560 Hz	1 octave	1811 Hz	0 (0.0%)
		1440-2880 Hz	1 octave	2037 Hz	0 (0.0%)
		1600-3200 Hz	1 octave	2263 Hz	0 (0.0%)
		1760-3520 Hz	1 octave	2489 Hz	0 (0.0%)
		1920-3840 Hz	1 octave	2715 Hz	0 (0.0%)
/e/	37	1523-2154 Hz	$\frac{1}{2}$ octave	1811 Hz	0 (0.0%)
		1903-2691 Hz	$\frac{1}{2}$ octave	2263 Hz	0 (0.0%)
		2283-3229 Hz	$\frac{1}{2}$ octave	2715 Hz	0 (0.0%)
		1613-2033 Hz	$\frac{1}{3}$ octave	1811 Hz	0 (0.0%)
		2016-2540 Hz	$\frac{1}{3}$ octave	2263 Hz	0 (0.0%)
		2419-3048 Hz	$\frac{1}{3}$ octave	2715 Hz	0 (0.0%)
		0-320 Hz + 1523-2154 Hz	0-320 Hz + $\frac{1}{2}$ octave	1811 Hz	0 (0.0%)
		0-320 Hz + 1903-2691 Hz	0-320 Hz + $\frac{1}{2}$ octave	2263 Hz	0 (0.0%)
		0-320 Hz + 2283-3229 Hz	0-320 Hz + $\frac{1}{2}$ octave	2715 Hz	0 (0.0%)

Table 4.1 Corrective optimals of the target vowels as tested in this study (Cont.)

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		0-320 Hz + 1512-1905 Hz	0-320 Hz + $\frac{1}{3}$ octave	1697 Hz	30 (81.1%)
		0-320 Hz + 1209-1523 Hz	0-320 Hz + $\frac{1}{3}$ octave	1357 Hz	4 (10.8%)
		0-320 Hz + 1815-2286 Hz	0-320 Hz + $\frac{1}{3}$ octave	2037 Hz	3 (8.1%)
		960-1920 Hz	1 octave	1357 Hz	0 (0.0%)
		1080-2160 Hz	1 octave	1527 Hz	0 (0.0%)
		1200-2400 Hz	1 octave	1697 Hz	0 (0.0%)
		1320-2640 Hz	1 octave	1867 Hz	0 (0.0%)
		1440-2880 Hz	1 octave	2037 Hz	0 (0.0%)
/æ/	37	1141-1614 Hz	$\frac{1}{2}$ octave	1357 Hz	0 (0.0%)
		1427-2018 Hz	$\frac{1}{2}$ octave	1697 Hz	0 (0.0%)
		1713-2422 Hz	$\frac{1}{2}$ octave	2037 Hz	0 (0.0%)
		1209-1523 Hz	$\frac{1}{3}$ octave	1357 Hz	0 (0.0%)
		1512-1905 Hz	$\frac{1}{3}$ octave	1697 Hz	0 (0.0%)
		1815-2286 Hz	$\frac{1}{3}$ octave	2037 Hz	0 (0.0%)
		0-320 Hz + 1141-1614 Hz	0-320 Hz + $\frac{1}{2}$ octave	1357 Hz	0 (0.0%)
		0-320 Hz + 1427-2018 Hz	0-320 Hz + $\frac{1}{2}$ octave	1697 Hz	0 (0.0%)
		0-320 Hz + 1713-2422 Hz	0-320 Hz + $\frac{1}{2}$ octave	2037 Hz	0 (0.0%)

Table 4.1 Corrective optimals of the target vowels as tested in this study (Cont.)

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		0-320 Hz + 214-269 Hz	0-320 Hz + $\frac{1}{3}$ octave	240 Hz	31 (83.8%)
		0-320 Hz + 415-523 Hz	0-320 Hz + $\frac{1}{3}$ octave	466 Hz	3 (8.1%)
		0-320 Hz + 303-382 Hz	0-320 Hz + $\frac{1}{3}$ octave	340 Hz	2 (5.4%)
		0-320 Hz + 542-682 Hz	0-320 Hz + $\frac{1}{3}$ octave	608 Hz	1 (2.7%)
		170-340 Hz	1 octave	240 Hz	0 (0.0%)
		199-398 Hz	1 octave	282 Hz	0 (0.0%)
		241-482 Hz	1 octave	340 Hz	0 (0.0%)
		270-540 Hz	1 octave	382 Hz	0 (0.0%)
		300-600 Hz	1 octave	424 Hz	0 (0.0%)
/ʊ/	37	330-660 Hz	1 octave	466 Hz	0 (0.0%)
		359-718 Hz	1 octave	508 Hz	0 (0.0%)
		400-800 Hz	1 octave	566 Hz	0 (0.0%)
		430-860 Hz	1 octave	608 Hz	0 (0.0%)
		202-285 Hz	$\frac{1}{2}$ octave	240 Hz	0 (0.0%)
		286-404 Hz	$\frac{1}{2}$ octave	340 Hz	0 (0.0%)
		392-554 Hz	$\frac{1}{2}$ octave	466 Hz	0 (0.0%)
		511-723 Hz	$\frac{1}{2}$ octave	608 Hz	0 (0.0%)
		214-269 Hz	$\frac{1}{3}$ octave	240 Hz	0 (0.0%)
		303-382 Hz	$\frac{1}{3}$ octave	340 Hz	0 (0.0%)

Table 4.1 Corrective optimals of the target vowels as tested in this study (Cont.)

Vowel	N	Corrective optimals	Octave type	Center frequency	Best production
		415-523 Hz	$\frac{1}{3}$ octave	466 Hz	0 (0.0%)
		542-682 Hz	$\frac{1}{3}$ octave	608 Hz	0 (0.0%)
/ʊ/	37	0-320 Hz + 202-285 Hz	0-320 Hz + $\frac{1}{2}$ octave	240 Hz	0 (0.0%)
		0-320 Hz + 286-404 Hz	0-320 Hz + $\frac{1}{2}$ octave	340 Hz	0 (0.0%)
		0-320 Hz + 392-554 Hz	0-320 Hz + $\frac{1}{2}$ octave	466 Hz	0 (0.0%)
		0-320 Hz + 511-723 Hz	0-320 Hz + $\frac{1}{2}$ octave	608 Hz	0 (0.0%)
		0-320 Hz + 302-381 Hz	0-320 Hz + $\frac{1}{3}$ octave	339 Hz	32 (86.5%)
		0-320 Hz + 277-349 Hz	0-320 Hz + $\frac{1}{3}$ octave	311 Hz	5 (13.5%)
		160-320 Hz	1 octave	227 Hz	0 (0.0%)
		180-360 Hz	1 octave	255 Hz	0 (0.0%)
		200-400 Hz	1 octave	283 Hz	0 (0.0%)
		220-440 Hz	1 octave	311 Hz	0 (0.0%)
/u:/	37	240-480 Hz	1 octave	339 Hz	0 (0.0%)
		262-370 Hz	$\frac{1}{2}$ octave	311 Hz	0 (0.0%)
		285-403 Hz	$\frac{1}{2}$ octave	339 Hz	0 (0.0%)
		277-349 Hz	$\frac{1}{3}$ octave	311 Hz	0 (0.0%)
		302-381 Hz	$\frac{1}{3}$ octave	339 Hz	0 (0.0%)
		0-320 Hz + 262-370 Hz	0-320 Hz + $\frac{1}{2}$ octave	311 Hz	0 (0.0%)
		0-320 Hz + 285-403 Hz	0-320 Hz + $\frac{1}{2}$ octave	339 Hz	0 (0.0%)

As can be seen from the table above, the vowels were best produced when they were perceived through discontinuous multiband filters. With regard to the vast majority (81.1%-91.9%) of the students, the corrective optimal determined for /ɪ/ was 0-320 Hz + 2419-3048 Hz (fcenter = 2715 Hz); for /i:/ was 0-320 Hz + 4838-6096 Hz (fcenter = 5431 Hz); for /e/ was 0-320 Hz + 2419-3048 Hz (fcenter = 2715 Hz); for /æ/ was 0-320 Hz + 1512-1905 Hz (fcenter = 1697 Hz); for /ʊ/ was 0-320 Hz + 214-269 Hz (fcenter = 240 Hz); and for /u:/ was 0-320 Hz + 302-381 Hz (fcenter = 339 Hz).

As indicated, a variety of corrective optimals were found to be effective for a small percentage of participants. Apart from the above-mentioned, 0-320 Hz + 1815-2286 Hz (fcenter = 2037 Hz) was determined for /ɪ/. Filters 0-320 Hz + 4567-6459 Hz (fcenter = 5431 Hz) and 0-320 Hz+4435-5588 Hz (fcenter = 4978 Hz) were identified for /i:/. Filters 0-320 Hz + 2016-2540 Hz (fcenter = 2263 Hz) and 0-320 Hz + 1613-2033 Hz (fcenter = 1811 Hz) were determined for /e/. Filters 0-320 Hz + 1209-1523 Hz (fcenter = 1357 Hz) and 0-320 Hz + 1815-2286 Hz (fcenter = 2037 Hz) were identified for /æ/. Filters 0-320 Hz + 415-523 Hz (fcenter = 466 Hz), 0-320 Hz + 303-382 Hz (fcenter = 340 Hz) and 0-320 Hz + 542-682 Hz (fcenter = 608 Hz) were identified for /ʊ/. The filter 0-320 Hz + 277-349 Hz (fcenter = 311 Hz) was determined for /u:/.

The following lists Guberina's optimal octaves: the optimal octave for /ɪ/ was 1600-3200 Hz (fcenter = 2263 Hz); the optimal octave for /i:/ was 3200-6400 Hz (fcenter = 4525 Hz); the optimal octave for /e/ was 1600-3200 Hz (fcenter = 2263 Hz); the optimal octave for /æ/ was 1200-2400 Hz (fcenter = 1697 Hz); the optimal octaves for /ʊ/ were 200-400 Hz (fcenter = 283 Hz), 300-600 Hz (fcenter = 424 Hz) and 400-800 Hz (fcenter = 566 Hz); the optimal octave for /u:/ was 200-400 Hz (fcenter = 283 Hz).

According to the results presented above, together with the notion of optimal octave bands, it can be concluded that the corrective optimals identified so far were actually narrower and finer than the Guberina native speaker optimals. Moreover, the corrective optimals were found to be less uniform and more diverse in comparison with the native speaker optimals determined by Guberina. In other words, this research identified a greater range of individual differences than assumed in the original Guberina study. This section answers the first two research questions.

4.2 Perception performance

To better interpret the results, in the present study, all raw scores collected from the experimental and control groups were converted into percentage scores (i.e. scores out of 100). Effect sizes were also calculated (Cohen's d). According to Cohen (1988), an effect size of about 0.2 is considered small, around 0.5 is considered medium, and 0.8 or above is considered large.

Between-group comparisons

An independent samples t-test was used to examine whether the experimental and control groups performed differently from each other before and after the treatment (see

Table 4.2). As presented below, in the pretest, there were no significant differences between the means of the two groups concerning the perception of the six vowels. The two groups started at the same performance level.

However, in the posttest, statistically significant differences were found between the means of the two groups concerning the perception of the vowels /ɪ/ ($p = 0.004$), /iː/ ($p = 0.002$), /e/ ($p = 0.009$), /ʊ/ ($p = 0.006$) and /uː/ ($p = 0.036$). Effect sizes for these vowels were medium to large for /ɪ/ ($d = 0.68$), /iː/ ($d = 0.73$), /e/ ($d = 0.62$), /ʊ/ ($d = 0.65$) and medium for /uː/ ($d = 0.49$). No significant difference was found between the means of the two groups for the sound /æ/.

Figures 4.1 and 4.2 further illustrate the differences between the means of the two groups in the perception pretest and posttest.

Table 4.2 Perception: independent samples t-test, EG vs. CG, for the target vowels

Test	Vowel	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	/ɪ/	71.08 (8.43)	70.51 (8.57)	0.57	0.772	0.07
	/iː/	70.54 (9.70)	71.03 (9.12)	-0.49	0.823	0.05
	/e/	74.32 (8.67)	74.10 (9.10)	0.22	0.914	0.03
	/æ/	78.38 (9.58)	77.18 (8.87)	1.20	0.573	0.13
	/ʊ/	80.00 (9.13)	79.74 (9.03)	0.26	0.902	0.03
	/uː/	77.57 (8.30)	77.44 (8.50)	0.13	0.946	0.02
Posttest	/ɪ/	78.65 (8.55)	72.56 (9.38)	6.09	0.004	0.68
	/iː/	79.19 (8.62)	72.82 (8.87)	6.37	0.002	0.73
	/e/	80.81 (8.29)	75.64 (8.52)	5.17	0.009	0.62
	/æ/	84.60 (8.69)	81.54 (8.44)	3.06	0.124	0.36
	/ʊ/	85.41 (7.67)	80.26 (8.10)	5.15	0.006	0.65
	/uː/	84.87 (7.31)	80.77 (9.29)	4.10	0.036	0.49

Note. EG = Experimental group; CG = Control group

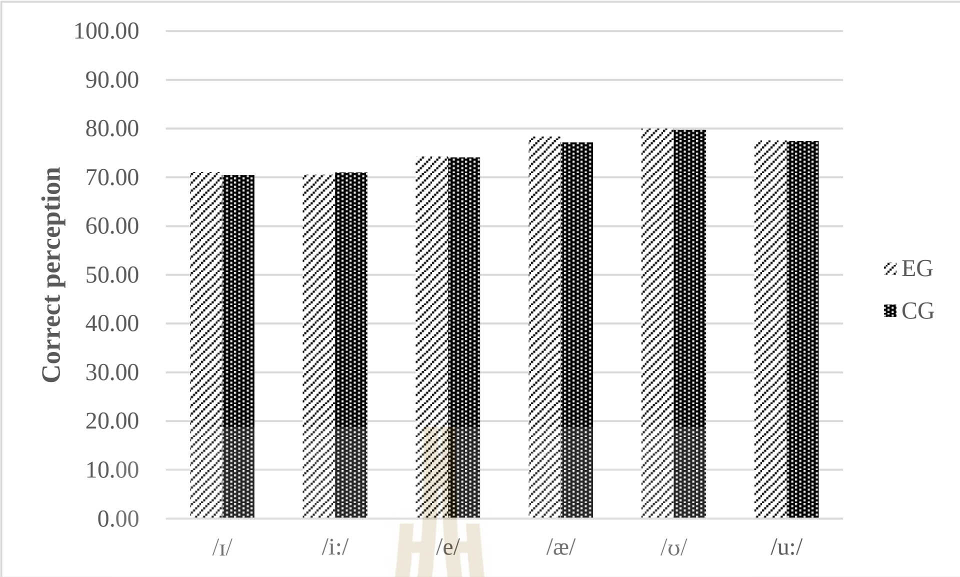

Figure 4.1 Correct perception of the target vowels in the pretest for both groups

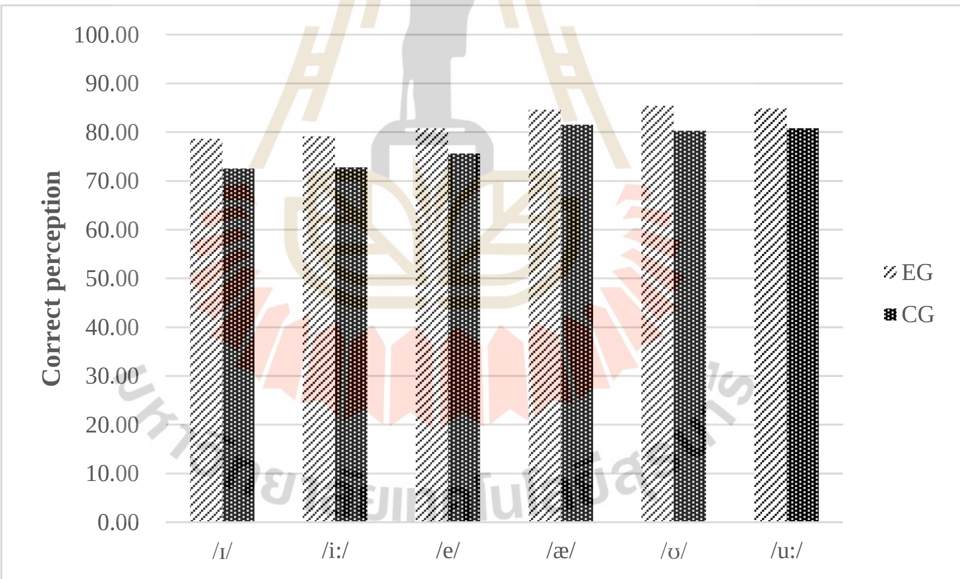

Figure 4.2 Correct perception of the target vowels in the posttest for both groups

Within-group comparisons

A paired samples t-test was conducted to check whether there were significant differences between the perception pretest and posttest scores for the target vowels within each of the two groups (see Table 4.3).

The table below shows that, in the experimental group, there were significant differences in perception scores between pretest and posttest in relation to all the vowels under study. The means in the posttest were statistically significantly higher than in the pretest ($p < 0.001$). Effect sizes were large for /ɪ/ ($d = 0.82$), /i:/ ($d = 1.15$), /ʊ/ ($d = 0.89$) and /u:/ ($d = 0.95$) and medium to large for /e/ ($d = 0.71$) and /æ/ ($d = 0.70$).

In the control group, statistically significant differences were found between pretest and posttest scores for /æ/ ($p = 0.005$) and /u:/ ($p = 0.008$), with medium effect sizes of 0.48 and 0.45. No significant differences were found between pretest and posttest scores for the other four vowels.

Figures 4.3 and 4.4 further illustrate the differences between perception pretest and posttest scores within the two groups.

Table 4.3 Perception: paired samples t-test, pretest vs. posttest, for the target vowels

Group	Vowel	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
EG (N=37)	/ɪ/	71.08 (8.43)	78.65 (8.55)	7.57	< 0.001	0.82
	/i:/	70.54 (9.70)	79.19 (8.62)	8.65	< 0.001	1.15
	/e/	74.32 (8.67)	80.81 (8.29)	6.49	< 0.001	0.71
	/æ/	78.38 (9.58)	84.60 (8.69)	6.22	< 0.001	0.70
	/ʊ/	80.00 (9.13)	85.41 (7.67)	5.41	< 0.001	0.89
	/u:/	77.57 (8.30)	84.87 (7.31)	7.30	< 0.001	0.95
CG (N=39)	/ɪ/	70.51 (8.57)	72.56 (9.38)	2.05	0.146	0.24
	/i:/	71.03 (9.12)	72.82 (8.87)	1.79	0.181	0.22
	/e/	74.10 (9.10)	75.64 (8.52)	1.54	0.205	0.21
	/æ/	77.18 (8.87)	81.54 (8.44)	4.36	0.005	0.48
	/ʊ/	79.74 (9.03)	80.26 (8.10)	0.52	0.623	0.08
	/u:/	77.44 (8.50)	80.77 (9.29)	3.33	0.008	0.45

Note. EG = Experimental group; CG = Control group

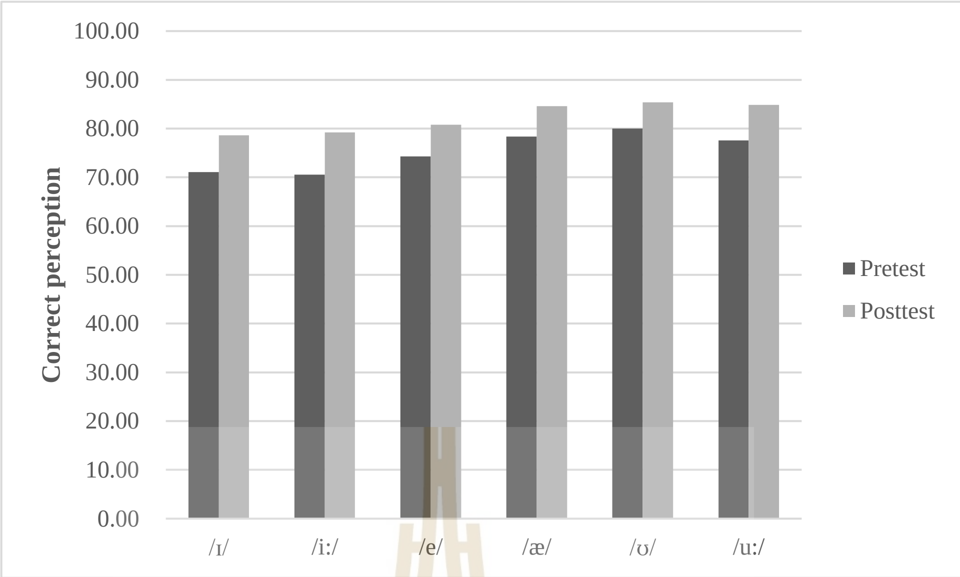

Figure 4.3 Correct perception of the target vowels in the pretest and posttest for the experimental group

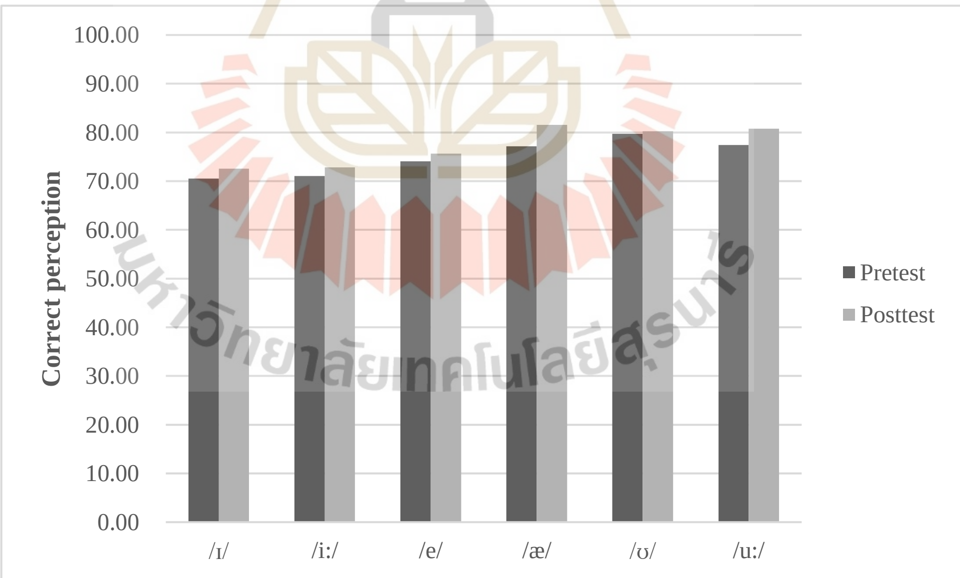

Figure 4.4 Correct perception of the target vowels in the pretest and posttest for the control group

4.3 Production performance

4.3.1 Word-reading

Between-group comparisons

An independent samples t-test was conducted to examine whether the experimental and control groups performed differently from each other before and after training in word-reading (see Table 4.4).

Analysis of the independent samples t-test demonstrated that there were no significant differences between the means of the two groups in the pretest concerning the production of the target vowels at the word level. That is, both the experimental and control groups began from the same starting point.

However, at the end of the treatment, statistically significant differences were found between the means of the two groups for the target vowels /ɪ/ ($p = 0.003$), /iː/ ($p < 0.001$), /e/ ($p < 0.001$), /ʊ/ ($p = 0.010$) and /uː/ ($p = 0.027$). The effect sizes were large for /iː/ ($d = 0.88$) and /e/ ($d = 0.82$), medium to large for /ɪ/ ($d = 0.69$) and /ʊ/ ($d = 0.61$), and medium for /uː/ ($d = 0.52$). At the same time, no significant difference was found between the means of the two groups for sound /æ/.

Table 4.4 Word-reading: independent samples t-test, EG vs. CG, for the target vowels

Test	Vowel	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	/ɪ/	83.51 (7.93)	84.11 (8.53)	-0.60	0.754	0.07
	/i:/	84.59 (8.69)	82.74 (8.68)	1.85	0.354	0.21
	/e/	86.94 (8.44)	86.92 (8.14)	0.02	0.992	0.00
	/æ/	89.55 (7.94)	90.60 (8.02)	-1.05	0.569	0.13
	/ʊ/	88.47 (7.64)	89.15 (7.56)	-0.68	0.700	0.09
	/u:/	86.49 (8.13)	84.79 (8.16)	1.70	0.365	0.21
Posttest	/ɪ/	91.27 (7.71)	85.64 (8.49)	5.63	0.003	0.69
	/i:/	91.08 (7.50)	84.36 (7.73)	6.72	< 0.001	0.88
	/e/	93.42 (6.87)	87.26 (8.13)	6.16	< 0.001	0.82
	/æ/	95.23 (6.51)	93.59 (7.55)	1.64	0.316	0.23
	/ʊ/	94.24 (6.65)	90.18 (6.75)	4.06	0.010	0.61
	/u:/	92.70 (7.02)	88.72 (8.23)	3.98	0.027	0.52

Note. EG = Experimental group; CG = Control group

Figures 4.5 and 4.6 further illustrate the differences between the means of the two groups in the pretest and posttest.

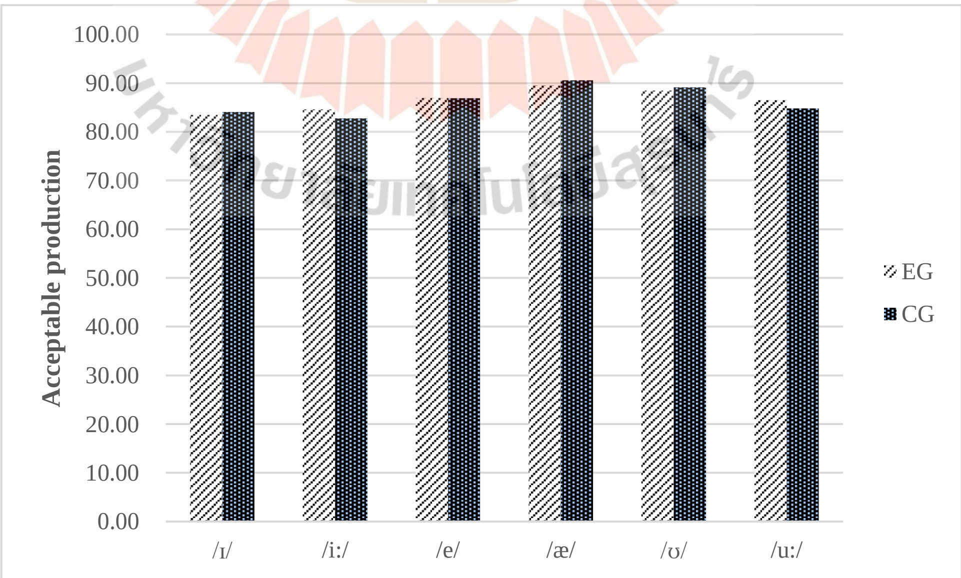

Figure 4.5 Acceptable production of the target vowels in the pretest for both groups in word-reading

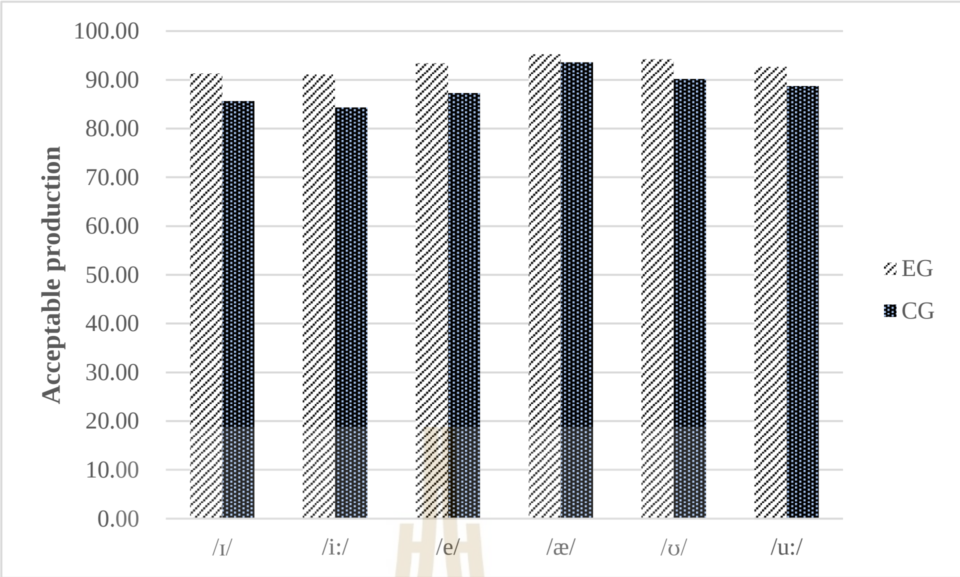

Figure 4.6 Acceptable production of the target vowels in the posttest for both groups in word-reading

Within-group comparisons

A paired samples t-test was used to determine whether statistically significant differences existed between the production pretest and posttest scores for the target vowels within each of the two groups concerning word-reading (see Table 4.5).

The table below shows that, in the experimental group, there were significant differences between pretest and posttest scores in word-reading. Inspection of the means in the two tests indicated that posttest scores were statistically significantly higher than pretest scores ($p < 0.001$). The effect sizes were large for all six vowels, i.e. /ɪ/ ($d = 1.03$), /i:/ ($d = 1.02$), /e/ ($d = 0.94$), /æ/ ($d = 1.05$), /ʊ/ ($d = 0.91$) and /u:/ ($d = 0.94$).

In the control group, statistically significant differences were found between pretest and posttest scores for /æ/ ($p = 0.016$) and /u:/ ($p < 0.001$), and the effect sizes were 0.40 (small to medium) and 0.58 (medium). No significant differences were found between the pretest and posttest scores for the rest of the target vowels. Figures 4.7 and 4.8 further illustrate the differences between pretest and posttest scores within the two groups.

Table 4.5 Word-reading: paired samples t-test, pretest vs. posttest, for the target vowels

Group	Vowel	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
EG (N=37)	/ɪ/	83.51 (7.93)	91.27 (7.71)	7.76	< 0.001	1.03
	/iː/	84.59 (8.69)	91.08 (7.50)	6.49	< 0.001	1.02
	/e/	86.94 (8.44)	93.42 (6.87)	6.48	< 0.001	0.94
	/æ/	89.55 (7.94)	95.23 (6.51)	5.68	< 0.001	1.05
	/ʊ/	88.47 (7.64)	94.24 (6.65)	5.77	< 0.001	0.91
	/uː/	86.49 (8.13)	92.70 (7.02)	6.21	< 0.001	0.94
CG (N=39)	/ɪ/	84.11 (8.53)	85.64 (8.49)	1.53	0.230	0.19
	/iː/	82.74 (8.68)	84.36 (7.73)	1.62	0.215	0.20
	/e/	86.92 (8.14)	87.26 (8.13)	0.34	0.794	0.04
	/æ/	90.60 (8.02)	93.59 (7.55)	2.99	0.016	0.40
	/ʊ/	89.15 (7.56)	90.18 (6.75)	1.03	0.363	0.15
	/uː/	84.79 (8.16)	88.72 (8.23)	3.93	< 0.001	0.58

Note. EG = Experimental group; CG = Control group

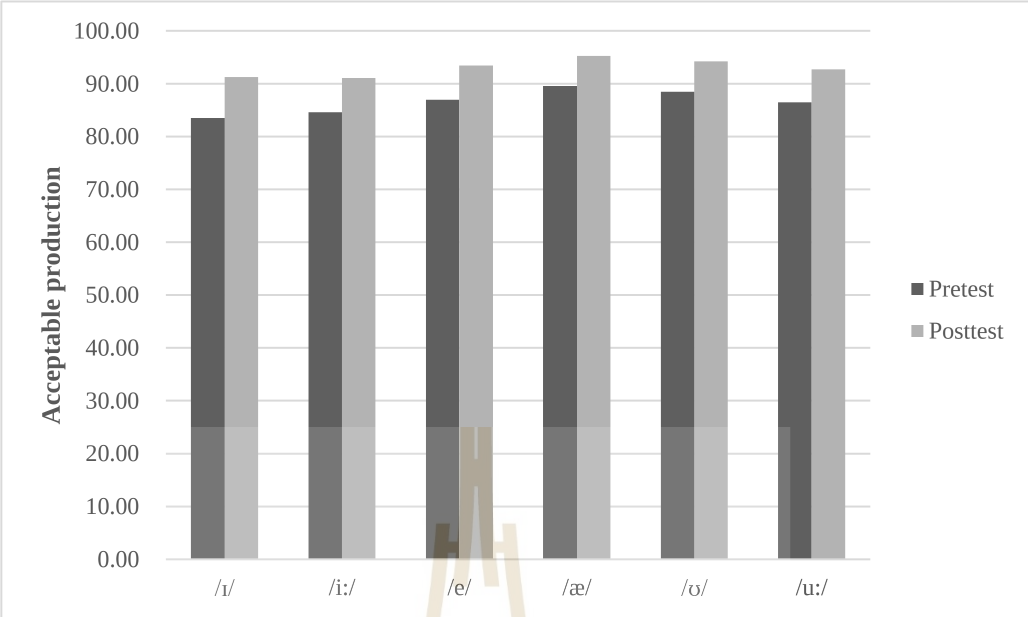

Figure 4.7 Acceptable production of the target vowels in the pretest and posttest for the experimental group in word-reading

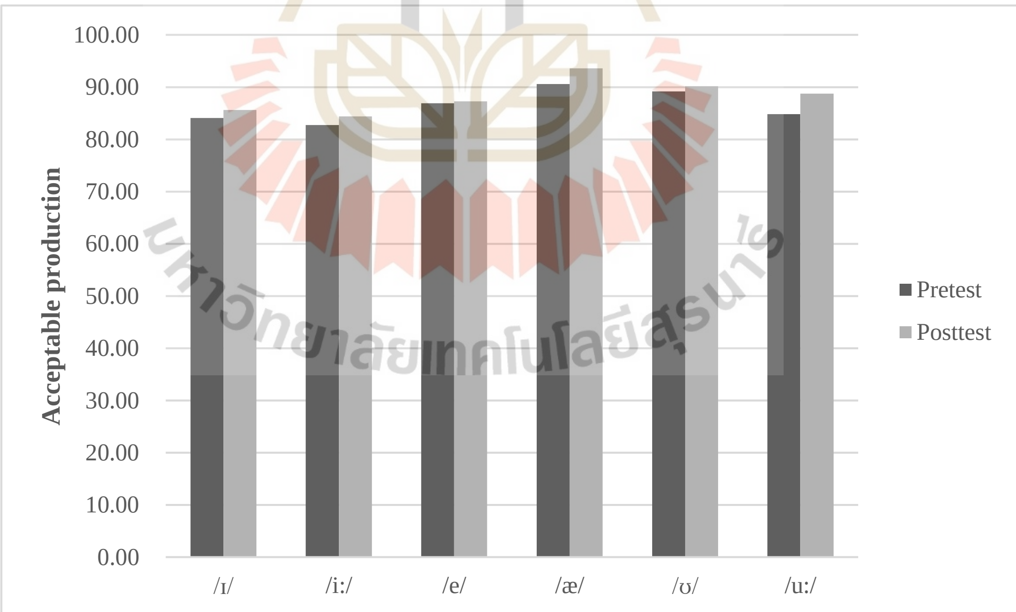

Figure 4.8 Acceptable production of the target vowels in the pretest and posttest for the control group in word-reading

4.3.2 Sentence-reading

Between-group comparisons

An independent samples t-test was performed to compare the differences between the means of the experimental and control groups in the pretest and posttest in terms of vowel production in sentence-reading (see Table 4.6).

Table 4.6 Sentence-reading: independent samples t-test, EG vs. CG, for the target vowels

Test	Vowel	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	/ɪ/	83.15 (6.62)	83.68 (7.09)	-0.53	0.738	0.08
	/i:/	81.53 (7.81)	82.74 (7.25)	-1.21	0.486	0.16
	/e/	81.80 (7.06)	81.88 (7.57)	-0.08	0.961	0.01
	/æ/	82.35 (7.29)	81.28 (6.74)	1.07	0.511	0.15
	/ʊ/	82.97 (7.15)	83.85 (7.12)	-0.88	0.594	0.12
	/u:/	81.17 (7.55)	80.77 (7.36)	0.40	0.815	0.05
Posttest	/ɪ/	88.56 (7.18)	83.42 (7.67)	5.14	0.004	0.69
	/i:/	88.92 (7.12)	83.93 (7.14)	4.99	0.003	0.70
	/e/	87.21 (6.69)	82.65 (8.25)	4.56	0.010	0.61
	/æ/	86.58 (7.05)	84.45 (8.18)	2.13	0.230	0.28
	/ʊ/	88.29 (6.51)	84.19 (6.57)	4.10	0.008	0.63
	/u:/	86.76 (6.83)	83.42 (7.36)	3.34	0.044	0.47

Note. EG = Experimental group; CG = Control group

As can be seen from the table above, there were no significant differences between the means of the experimental and control groups in the pretest concerning the production of the target vowels at the sentence level. That is, they began at the same performance level.

However, in the posttest, statistically significant differences were found between the means of the two groups for the target vowels /ɪ/ (p = 0.004), /iː/ (p = 0.003), /e/ (p = 0.010), /ʊ/ (p = 0.008) and /uː/ (p = 0.044). Effect sizes for these vowels were medium to large for /ɪ/ (d = 0.69), /iː/ (d = 0.70), /e/ (d = 0.61) and /ʊ/ (d = 0.63) and medium for /uː/ (d = 0.47). No significant difference was found between the means of the two groups for the sound /æ/. Figures 4.9 and 4.10 further illustrate the differences between the means of the two groups in the pretest and posttest.

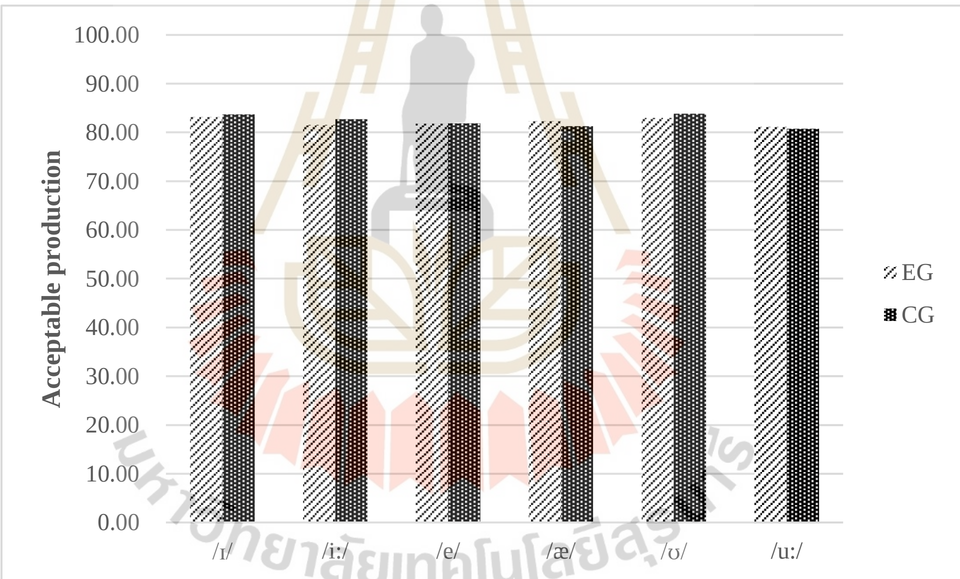

Figure 4.9 Acceptable production of the target vowels in the pretest for both groups in sentence-reading

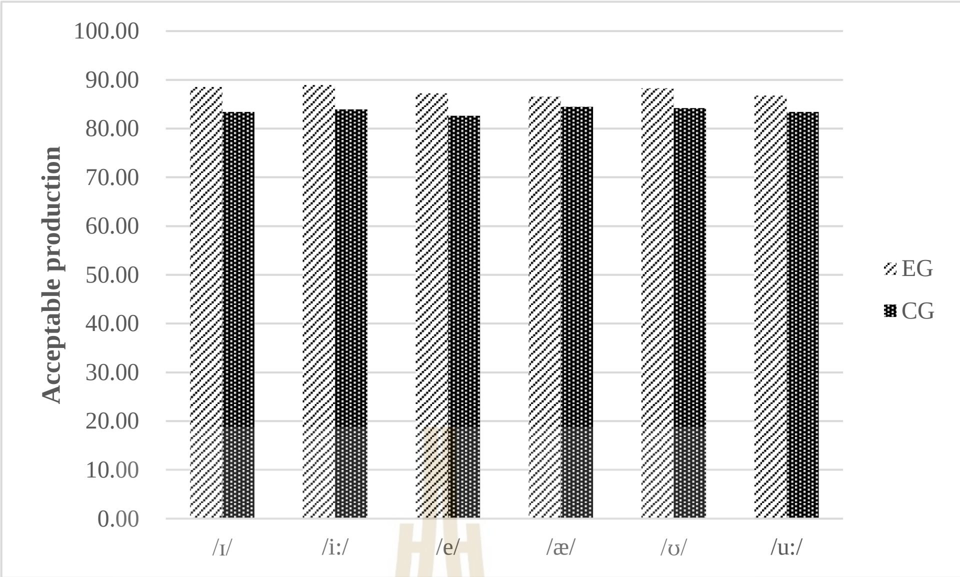

Figure 4.10 Acceptable production of the target vowels in the posttest for both groups in sentence-reading

Within-group comparisons

A paired samples t-test was conducted to check whether there were significant differences between the production pretest and posttest scores for the target vowels within each of the two groups concerning sentence-reading (see Table 4.7).

Table 4.7 Sentence-reading: paired samples t-test, pretest vs. posttest, for the target vowels

Group	Vowel	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
	/ɪ/	83.15 (6.62)	88.56 (7.18)	5.41	< 0.001	0.90
	/i:/	81.53 (7.81)	88.92 (7.12)	7.39	< 0.001	1.26
EG	/e/	81.80 (7.06)	87.21 (6.69)	5.41	< 0.001	0.86
(N=37)	/æ/	82.35 (7.29)	86.58 (7.05)	4.23	0.002	0.54
	/ʊ/	82.97 (7.15)	88.29 (6.51)	5.32	< 0.001	0.68
	/u:/	81.17 (7.55)	86.76 (6.83)	5.59	< 0.001	1.05
	/ɪ/	83.68 (7.09)	83.42 (7.67)	-0.26	0.833	0.03
	/i:/	82.74 (7.25)	83.93 (7.14)	1.19	0.221	0.20
CG	/e/	81.88 (7.57)	82.65 (8.25)	0.77	0.507	0.11
(N=39)	/æ/	81.28 (6.74)	84.45 (8.18)	3.17	0.029	0.36
	/ʊ/	83.85 (7.12)	84.19 (6.57)	0.34	0.719	0.06
	/u:/	80.77 (7.36)	83.42 (7.36)	2.65	0.004	0.48

Note. EG = Experimental group; CG = Control group

As can be seen from the table above, in the experimental group, significant differences were found between pretest and posttest scores in sentence-reading. The means in the posttest were statistically significantly higher than in the pretest for /ɪ/ ($p < 0.001$), /i:/ ($p < 0.001$), /e/ ($p < 0.001$), /æ/ ($p = 0.002$), /ʊ/ ($p < 0.001$) and /u:/ ($p < 0.001$). The effect sizes were large for /ɪ/ ($d = 0.90$), /i:/ ($d = 1.26$), /e/ ($d =0.86$) and /u:/ ($d = 1.05$), medium to large for /ʊ/ ($d =0.68$) and medium for /æ/ ($d = 0.54$).

In the control group, statistically significant differences were found between pretest and posttest scores for /æ/ ($p = 0.029$) and /u:/ ($p = 0.004$) and the effect sizes were 0.36 (small to medium) and 0.48 (medium). No significant differences between pretest and posttest scores were found for other target vowels.

Figures 4.11 and 4.12 further illustrate the differences between pretest and posttest scores within the two groups.

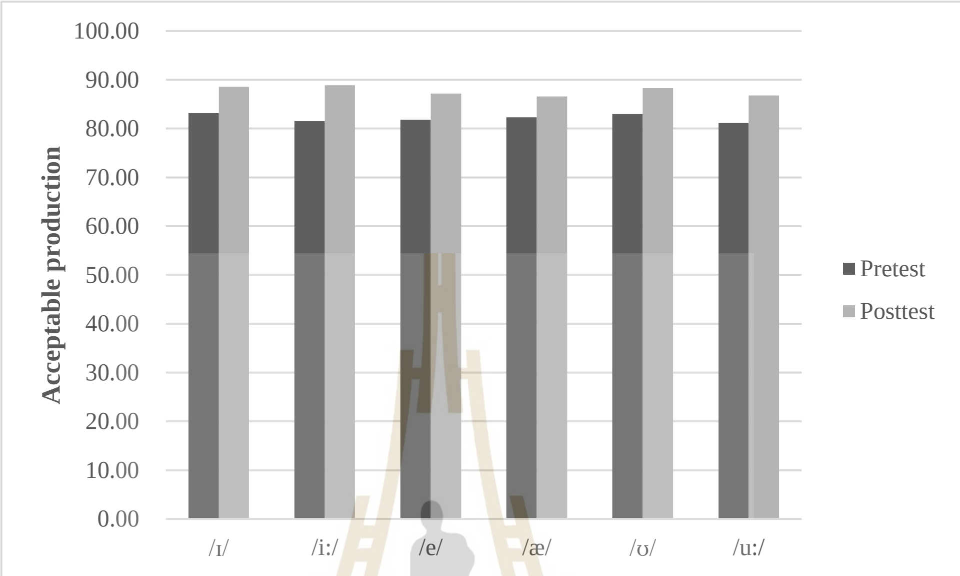

Figure 4.11 Acceptable production of the target vowels in the pretest and posttest for the experimental group in sentence-reading

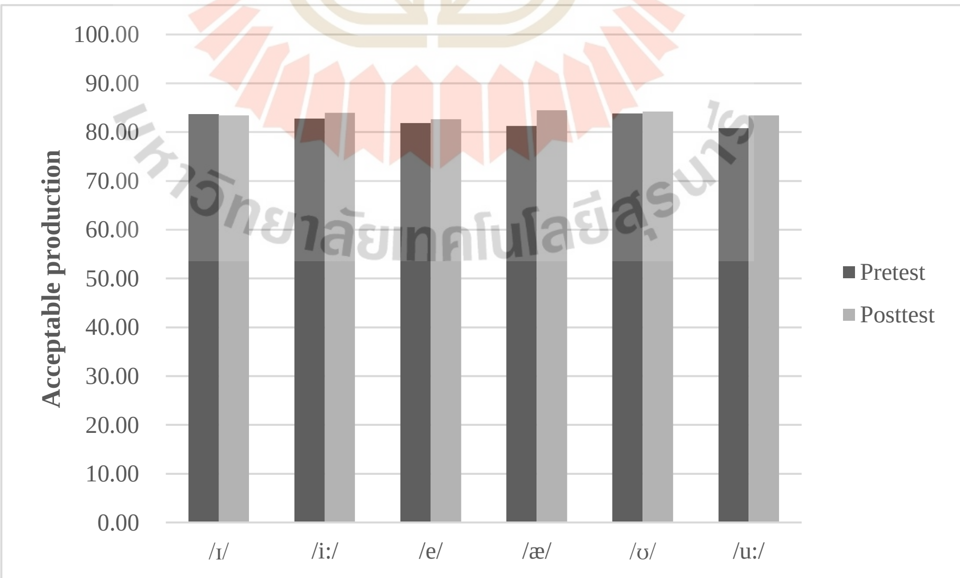

Figure 4.12 Acceptable production of the target vowels in the pretest and posttest for the control group in sentence-reading

4.3.3 Storytelling

Between-group comparisons

An independent samples t-test was conducted to verify whether the experimental and control groups produced the six target vowels differently in the pretest and posttest concerning storytelling (see Table 4.8).

Table 4.8 Storytelling: independent samples t-test, EG vs. CG, for the target vowels

Test	Vowel	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	/ɪ/	69.19 (6.60)	70.77 (7.32)	-1.58	0.328	0.23
	/iː/	69.46 (7.10)	71.28 (7.93)	-1.82	0.297	0.24
	/e/	72.44 (7.06)	73.33 (7.26)	-0.89	0.588	0.13
	/æ/	73.78 (7.34)	72.56 (7.47)	1.22	0.476	0.16
	/ʊ/	73.25 (6.96)	73.33 (7.54)	-0.08	0.961	0.01
	/uː/	72.07 (7.17)	71.62 (7.41)	0.45	0.790	0.06
Posttest	/ɪ/	76.58 (6.65)	71.80 (7.13)	4.78	0.004	0.69
	/iː/	76.76 (6.96)	71.54 (7.61)	5.22	0.003	0.72
	/e/	77.57 (6.83)	73.16 (7.45)	4.41	0.009	0.62
	/æ/	78.92 (7.12)	76.24 (7.47)	2.68	0.113	0.37
	/ʊ/	77.39 (6.34)	73.50 (6.88)	3.89	0.013	0.59
	/uː/	77.93 (6.78)	74.53 (7.16)	3.40	0.037	0.49

Note. EG = Experimental group; CG = Control group

Analysis of the independent samples t-test demonstrated that there were no significant differences between the means of the experimental and control groups in the pretest concerning the production of the target vowels during storytelling. That is, the two groups began from the same starting point.

However, after training, statistically significant differences were found between the means of the two groups for the target vowels /ɪ/ (p = 0.004), /iː/ (p = 0.003), /e/ (p = 0.009), /ʊ/ (p = 0.013) and /uː/ (p = 0.037). The effect sizes for these vowels were medium to large for /ɪ/ (d = 0.69), /iː/ (d = 0.72) and /e/ (d = 0.62) and medium for /ʊ/ (d = 0.59) and /uː/ (d = 0.49). No significant difference was found between the means of the two groups in the posttest for the sound /æ/. Figures 4.13 and 4.14 further illustrate the differences between the means of the two groups in the pretest and posttest.

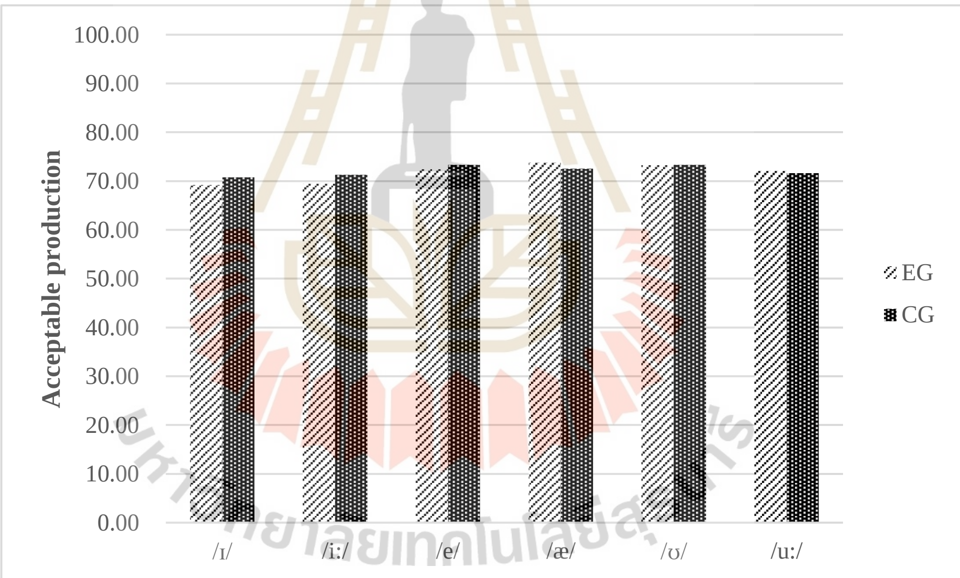

Figure 4.13 Acceptable production of the target vowels in the pretest for both groups in storytelling

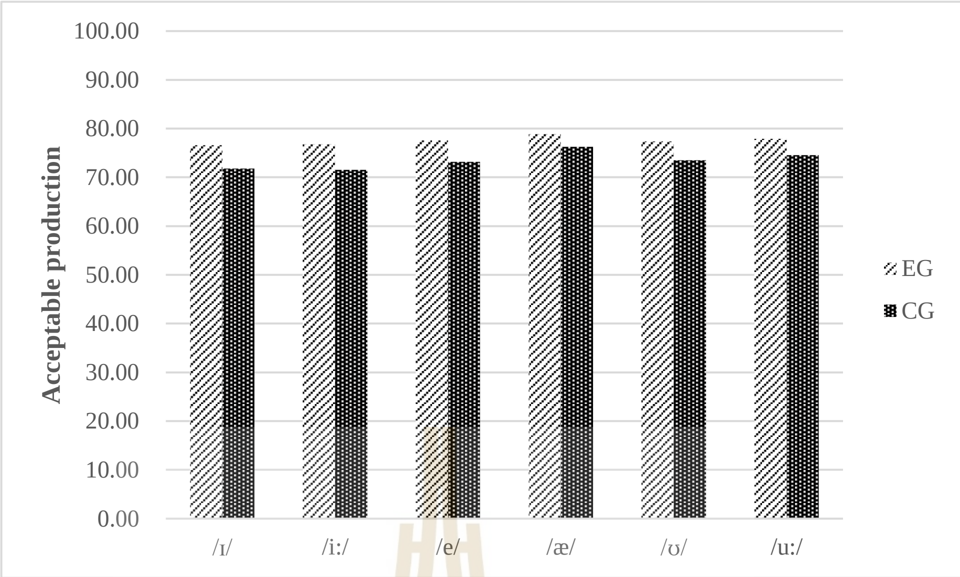

Figure 4.14 Acceptable production of the target vowels in the posttest for both groups in storytelling

Within-group comparisons

A paired samples t-test was used to determine whether there were significant differences between the production pretest and posttest scores for the target vowels within each of the two groups concerning storytelling (see Table 4.9).

Table 4.9 Storytelling: paired samples t-test, pretest vs. posttest, for the target vowels

Group	Vowel	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
EG (N=37)	/ɪ/	69.19 (6.60)	76.58 (6.65)	7.39	< 0.001	1.29
	/i:/	69.46 (7.10)	76.76 (6.96)	7.30	< 0.001	0.98
	/e/	72.44 (7.06)	77.57 (6.83)	5.13	< 0.001	0.66
	/æ/	73.78 (7.34)	78.92 (7.12)	5.14	< 0.001	0.70
	/ʊ/	73.25 (6.96)	77.39 (6.34)	4.14	< 0.001	0.62
	/u:/	72.07 (7.17)	77.93 (6.78)	5.86	< 0.001	0.96
CG (N=39)	/ɪ/	70.77 (7.32)	71.80 (7.13)	1.03	0.460	0.12
	/i:/	71.28 (7.93)	71.54 (7.61)	0.26	0.845	0.03
	/e/	73.33 (7.26)	73.16 (7.45)	-0.17	0.901	0.02
	/æ/	72.56 (7.47)	76.24 (7.47)	3.68	0.006	0.47
	/ʊ/	73.33 (7.54)	73.50 (6.88)	0.17	0.872	0.03
	/u:/	71.62 (7.41)	74.53 (7.16)	2.91	0.015	0.41

Note. EG = Experimental group; CG = Control group

The table above shows that, in the experimental group, there were significant differences between pretest and posttest scores in relation to all the target vowel sounds. Inspection of the means in the two tests showed that posttest scores were statistically significantly higher than pretest scores for all six target vowel sounds ($p < 0.001$). The effect sizes were large for /ɪ/ ($d = 1.29$), /i:/ ($d = 0.98$) and /u:/ ($d = 0.96$) and medium to large for /e/ ($d = 0.66$), /æ/ ($d = 0.70$) and /ʊ/ ($d = 0.62$). In the control group, statistically significant differences were found between pretest and posttest scores for /æ/ ($p = 0.006$) and /u:/ ($p = 0.015$) and the effect sizes were 0.47 (medium) and 0.41 (small to medium). No significant differences between pretest and posttest scores were found for the rest of the target vowels.

Figures 4.15 and 4.16 further illustrate the differences between pretest and posttest scores within the two groups.

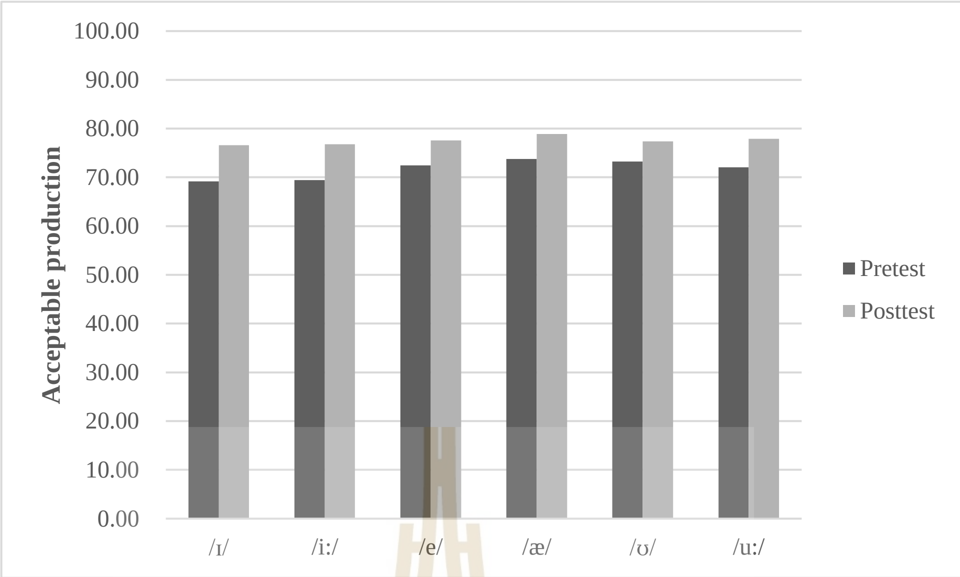

Figure 4.15 Acceptable production of the target vowels in the pretest and posttest for the experimental group in storytelling

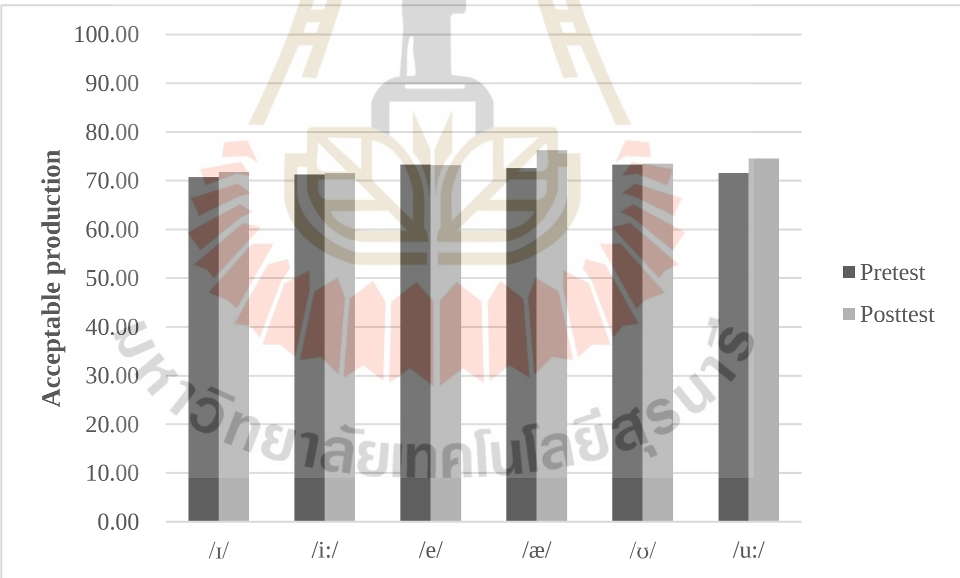

Figure 4.16 Acceptable production of the target vowels in the pretest and posttest for the control group in storytelling

Comparison of the means

As presented in the table below, the highest scores for the target vowels in both groups were found in word-reading, followed by sentence-reading and storytelling (see Table 4.10). This shows that, among the three production tasks, word-reading most likely placed the lowest cognitive load on students. In contrast, storytelling seemed to place the highest cognitive load on students. This could be explained by the fact that the task of storytelling is more demanding, which makes it more difficult for the information required to be easily retrieved from the participants' brains. In other words, it drains learners more to work on storytelling compared to the other two tasks. As indicated, the experimental group outperformed the control group on all three production tasks, which means that the VTPL approach was more favorable to language production and the students in the experimental group using this approach were able to do so much better than the students using the traditional approach in the control group. In other words, students in the experimental group had a better internalization of the target vowels.

Table 4.10 Comparison of the means, EG vs. CG, for the target vowels in word-reading, sentence-reading and storytelling

Vowel	Word-reading		Sentence-reading		Storytelling	
	EG	CG	EG	CG	EG	CG
/ɪ/	91.27	85.64	88.56	83.42	76.58	71.80
/i:/	91.08	84.36	88.92	83.93	76.76	71.54
/e/	93.42	87.26	87.21	82.65	77.57	73.16
/æ/	95.23	93.59	86.58	84.45	78.92	76.24
/ʊ/	94.24	90.18	88.29	84.19	77.39	73.50
/u:/	92.70	88.72	86.76	83.42	77.93	74.53

Note. EG=Experimental group; CG=Control group

Improvement ratio

For a better understanding of the improvements of the experimental and control groups, it is helpful to provide the improvement ratios for the two groups concerning the target vowels (see Table 4.11).

Table 4.11 Improvement ratio, EG vs. CG, for the target vowels in perception, word-reading, sentence-reading and storytelling

Task	Vowel	EG Improvement	CG Improvement	Improvement ratio
Perception	/ɪ/	7.57	2.05	3.69x
	/i:/	8.65	1.79	4.83x
	/e/	6.49	1.54	4.21x
	/æ/	6.22	4.36	1.43x
	/ʊ/	5.41	0.52	10.40x
	/u:/	7.30	3.33	2.19x
Word-reading	/ɪ/	7.76	1.53	5.07x
	/i:/	6.49	1.62	4.01x
	/e/	6.48	0.34	19.06x
	/æ/	5.68	2.99	1.90x
	/ʊ/	5.77	1.03	5.60x
	/u:/	6.21	3.93	1.58x
Sentence-reading	/ɪ/	5.41	-0.26	20.81x
	/i:/	7.39	1.19	6.21x
	/e/	5.41	0.77	7.03x
	/æ/	4.23	3.17	1.33x
	/ʊ/	5.32	0.34	15.65x
	/u:/	5.59	2.65	2.11x
Story telling	/ɪ/	7.39	1.03	7.17x
	/i:/	7.30	0.26	28.08x
	/e/	5.13	-0.17	30.18x
	/æ/	5.14	3.68	1.40x
	/ʊ/	4.14	0.17	24.35x
	/u:/	5.86	2.91	2.01x
Average improvement ratio				8.76x

Note. EG=Experimental group; CG=Control group

The results of the improvement ratio indicated that the experimental group performed much better than the control group concerning both the perception and production tasks. The range of improvement ratios for the target vowels in perception performance is from 1.43x to 10.40x. With regard to production performance, the improvement ratios in word-reading, sentence-reading and storytelling range from 1.58x-19.06x, 1.33x-20.81x, and 1.40x-30.18x, respectively. On average, the experimental group outperformed the control group by a factor of 8.76 times for both the perception and production tasks.

As mentioned previously, at the end of the treatment, statistically significant differences were found between the means of the two groups for five out of six target vowels. At the same time, no significant difference was found between the means of the experimental and control groups for the sound /æ/, however, it is worth pointing out that, while not statistically significant, there is suggestive evidence that the experimental group outperformed the control group.

The results indicated that the VTPL approach helped improve the students' perception and production of the target vowels and was more effective than the traditional approach for assisting the acceptable production of the target English vowels. The results in this section and the previous section give answers to the third and fourth research questions.

4.4 Comprehensibility, fluency and pronunciation testing

To further understand the students' ability to speak English before and after phonetic correction, general production performances of the experimental and control groups concerning comprehensibility, fluency and pronunciation were measured both in sentence-reading and storytelling. The maximum score for each measure is 20.

4.4.1 Comprehensibility, fluency and pronunciation testing in sentence-reading

Between-group comparisons

An independent samples t-test was performed to verify whether the experimental and control groups performed differently in the pretest and posttest (see Table 4.12).

Table 4.12 Sentence-reading: independent samples t-test, EG vs. CG, for comprehensibility, fluency and pronunciation

Test	Measure	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	Comprehensibility	15.43 (1.08)	15.05 (1.36)	0.38	0.181	0.31
	Fluency	12.70 (0.93)	12.67 (1.24)	0.03	0.886	0.03
	Pronunciation	14.16 (1.30)	14.03 (1.48)	0.13	0.689	0.09
Posttest	Comprehensibility	16.24 (0.86)	15.46 (1.16)	0.78	0.001	0.76
	Fluency	13.57 (1.17)	12.85 (0.99)	0.72	0.005	0.67
	Pronunciation	15.27 (1.23)	14.26 (1.07)	1.01	< 0.001	0.88

Note. EG = Experimental group; CG = Control group

As reported in the table above, before training, there were no significant differences between the means of the experimental and control groups concerning comprehensibility, fluency and pronunciation in sentence-reading. That is, both groups started at the same level. However, after treatment, statistically significant differences were found between the means of the two groups in sentence-reading concerning comprehensibility ($p = 0.001$), fluency ($p = 0.005$) and pronunciation ($p < 0.001$), and the effect sizes were large for pronunciation ($d = 0.88$) and medium to large for comprehensibility ($d = 0.76$) and fluency ($d = 0.67$). The results showed that, after training, the experimental group demonstrated greater improvements than the control group.

Within-group comparisons

A paired samples t-test was conducted to check whether there were significant differences in scores for production between pretest and posttest for comprehensibility, fluency and pronunciation in the two groups concerning sentence-reading (see Table 4.13).

Table 4.13 Sentence-reading: paired samples t-test, pretest vs. posttest, for comprehensibility, fluency and pronunciation

Group	Measure	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
EG (N=37)	Comprehensibility	15.43 (1.08)	16.24 (0.86)	0.81	< 0.001	0.98
	Fluency	12.70 (0.93)	13.57 (1.17)	0.87	< 0.001	1.15
	Pronunciation	14.16 (1.30)	15.27 (1.23)	1.11	< 0.001	1.25
CG (N=39)	Comprehensibility	15.05 (1.36)	15.46 (1.16)	0.41	0.064	0.31
	Fluency	12.67 (1.24)	12.85 (0.99)	0.18	0.328	0.16
	Pronunciation	14.03 (1.48)	14.26 (1.07)	0.23	0.322	0.16

Note. EG = Experimental group; CG = Control group

It can be seen from the data that the control group did not improve in any of these aspects. In contrast, the experimental group improved their comprehensibility, fluency and pronunciation scores in a statistically significant way from the pretest to the posttest with regard to sentence-reading ($p < 0.001$), and the effect sizes were large for all three aspects, i.e. comprehensibility ($d = 0.98$), fluency ($d = 1.15$) and pronunciation ($d = 1.25$).

4.4.2 Comprehensibility, fluency and pronunciation testing in storytelling

Between-group comparisons

An independent samples t-test was conducted to verify whether the experimental and control groups performed differently in the pretest and posttest (see Table 4.14).

Table 4.14 Storytelling: independent samples t-test, EG vs. CG, for comprehensibility, fluency and pronunciation

Test	Measure	EG (N=37) Mean (SD)	CG (N=39) Mean (SD)	Difference EG-CG	P-value	Cohen's d
Pretest	Comprehensibility	15.42 (0.93)	15.49 (1.05)	-0.07	0.750	0.07
	Fluency	12.73 (0.84)	12.96 (1.07)	-0.23	0.306	0.24
	Pronunciation	14.32 (1.02)	14.21 (1.17)	0.11	0.638	0.11
Posttest	Comprehensibility	16.30 (0.75)	15.67 (1.10)	0.63	0.005	0.67
	Fluency	13.34 (0.66)	12.87 (0.91)	0.47	0.012	0.59
	Pronunciation	15.15 (0.73)	14.54 (1.08)	0.61	0.005	0.66

Note. EG = Experimental group; CG = Control group

As reported above, in the pretest, there were no significant differences between the means of the experimental and control groups for comprehensibility, fluency and pronunciation in storytelling. This indicated that both groups began at the same starting line.

However, after training, statistically significant differences were found between the means of the two groups concerning comprehensibility (p = 0.005), fluency (p = 0.012) and pronunciation (p = 0.005), and the effect sizes were 0.67 (medium to large), 0.59 (medium) and 0.66 (medium to large) respectively. The results indicated that, after training, the experimental group did significantly better than the control group.

Within-group comparisons

A paired samples t-test was conducted to examine whether there were significant differences between the production pretest and posttest scores for comprehensibility, fluency and pronunciation in the two groups concerning storytelling (see Table 4.15).

Table 4.15 Storytelling: paired samples t-test, pretest vs. posttest, for comprehensibility, fluency and pronunciation

Group	Measure	Pretest Mean (SD)	Posttest Mean (SD)	Difference (posttest-pretest)	P-value	Cohen's d
EG (N=37)	Comprehensibility	15.42 (0.93)	16.30 (0.75)	0.88	< 0.001	1.41
	Fluency	12.73 (0.84)	13.34 (0.66)	0.61	< 0.001	0.95
	Pronunciation	14.32 (1.02)	15.15 (0.73)	0.83	< 0.001	1.08
CG (N=39)	Comprehensibility	15.49 (1.05)	15.67 (1.10)	0.18	0.185	0.22
	Fluency	12.96 (1.07)	12.87 (0.91)	-0.09	0.507	0.10
	Pronunciation	14.21 (1.17)	14.54 (1.08)	0.33	0.061	0.31

Note. EG = Experimental group; CG = Control group

The table above shows that, in the control group, no significant differences were found between the pretest and posttest scores for comprehensibility, fluency and pronunciation. In contrast, in the experimental group, statistically significant differences were found between the pretest and posttest scores for comprehensibility, fluency and pronunciation ($p < 0.001$) and the effect sizes were large, i.e. 1.41, 0.95 and 1.08 respectively.

Improvement ratio

For a better understanding of the improvements of the experimental and control groups, it is necessary to discuss the improvement ratio results for the two groups concerning comprehensibility, fluency and pronunciation (see Table 4.16).

Table 4.16 Improvement ratio, EG vs. CG, for comprehensibility, fluency and pronunciation in sentence-reading and storytelling

Task	Measure	EG Improvement	CG Improvement	Improvement ratio
Sentence-reading	Comprehensibility	0.81	0.41	1.98x
	Fluency	0.87	0.18	4.83x
	Pronunciation	1.11	0.23	4.83x
Story-telling	Comprehensibility	0.88	0.18	4.89x
	Fluency	0.61	-0.09	6.78x
	Pronunciation	0.83	0.33	2.52x
Overall improvement ratio				4.31x

Note. EG=Experimental group; CG=Control group

As can be seen from the table above, the experimental group improved considerably in comparison with the control group in terms of comprehensibility, fluency and pronunciation. The improvement ratio for comprehensibility, fluency and pronunciation in sentence-reading and storytelling ranges from 1.98x to 6.78x. The experimental group was, on average, 4.31 times better than the control group concerning all three aspects. This section gives additional information to answer the third and fourth research questions. The next section answers the last research question.

4.5 Students' opinions of the VTPL approach

4.5.1 Satisfaction questionnaire

As for the investigation of students' satisfaction with pronunciation learning, nine questionnaire items were designed where students in the experimental group declared their degree of agreement (see Table 4.17).

Table 4.17 Results of students' satisfaction with pronunciation learning

	Statement	N	SA	A	U	D	SD	M
1.	The current pronunciation lessons are interesting.	37	5 13.5%	23 62.2%	5 13.5%	3 8.1%	1 2.7%	4
2.	The current pronunciation lessons are helpful.	37	7 18.9%	17 45.9%	12 32.4%	0 0.0%	1 2.7%	4
3.	I feel confident about improving my pronunciation through this approach.	37	6 16.2%	16 43.2%	13 35.1%	1 2.7%	1 2.7%	4
4.	I feel comfortable with the current pronunciation lessons.	37	4 10.8%	21 56.8%	10 27.0%	1 2.7%	1 2.7%	4
5.	I prefer this approach to other approaches in pronunciation learning.	37	7 18.9%	15 40.5%	9 24.3%	4 10.8%	2 5.4%	4
6.	I like to learn pronunciation on my own through this approach.	37	10 27.0%	17 45.9%	7 18.9%	2 5.4%	1 2.7%	4
7.	I think I can make better sense of the pronunciation of the vowels by using this approach.	37	13 35.1%	18 48.6%	3 8.1%	2 5.4%	1 2.7%	4
8.	I think I can identify the pronunciation of the vowels more effectively by using this approach.	37	16 43.2%	16 43.2%	3 8.1%	2 5.4%	0 0.0%	4
9.	I think my motivation to learn pronunciation is strengthened by using this approach.	37	15 40.5%	11 29.7%	10 27.0%	0 0.0%	1 2.7%	4

Note. SA = Strongly agree; A = Agree; U = Undecided; D = Disagree; SD = Strongly disagree; M = Median

Specifically, students in the experimental group were asked to rate their answers using a 5-point Likert scale, ranging from “strongly agree” to “strongly disagree”. Each statement was scored as: 5 = Strongly agree; 4 = Agree; 3 = Undecided; 2 = Disagree; 1 = Strongly disagree.

Statement 1: The current pronunciation lessons are interesting.

As indicated, of the 37 participants who completed the questionnaire, 28 of them agreed that the current pronunciation lessons were interesting, with 5 participants (13.5%) strongly agreeing and 23 participants (62.2%) agreeing, for a total of 75.7% agreeing. In contrast, only 4 participants (10.8%) disagreed with the statement, with 3 participants (8.1%) merely disagreeing and 1 participant (2.7%) strongly disagreeing. Another 5 participants (13.5%) were undecided.

Statement 2: The current pronunciation lessons are helpful.

The participants’ responses indicated that 7 participants (18.9%) strongly agreed and 17 of them (45.9%) agreed with the statement, for a total of 64.8% agreeing that the current pronunciation lessons were helpful. A very low percentage of 2.7% strongly disagreed and another 12 participants (32.4%) were undecided.

Statement 3: I feel confident about improving my pronunciation through this approach.

In response to statement 3, a majority of those (59.4%) surveyed indicated that they felt confident about improving their pronunciation, with 6 participants (16.2%)

strongly agreeing and 16 participants (43.2%) agreeing. A very low percentage of 5.4% disagreed, with 1 participant (2.7%) disagreeing and 1 participant (2.7%) strongly disagreeing. Another 13 participants (35.1%) were undecided.

Statement 4: I feel comfortable with the current pronunciation lessons.

A majority of participants (67.6%) who responded to statement 4 felt comfortable with the current pronunciation lessons, with 4 participants (10.8%) strongly agreeing and 21 participants (56.8%) agreeing. In contrast, a very low percentage of 5.4% disagreed, with 1 participant (2.7%) disagreeing and 1 participant (2.7%) strongly disagreeing. Another 10 participants (27%) were undecided.

Statement 5: I prefer this approach to other approaches in pronunciation learning.

A total of 59.4% agreed with the statement and reported they preferred this approach to other approaches in pronunciation learning. Of these, 7 participants (18.9%) strongly agreed and 15 participants (40.5%) agreed. In contrast, a minority of participants (16.2%) disagreed, with 4 participants (10.8%) disagreeing and 2 participants (5.4%) strongly disagreeing. Another 9 participants (24.3%) were undecided.

Statement 6: I like to learn pronunciation on my own through this approach.

In response to statement 6, 27 participants (72.9%) surveyed reported that they liked to learn pronunciation on their own through this approach, with 10 participants (27%) strongly agreeing and 17 participants (45.9%) agreeing. A minority of them (8.1%) disagreed, with 2 participants (5.4%) disagreeing and 1 participant (2.7%) strongly disagreeing. Another 7 participants (18.9%) were undecided.

Statement 7: I think I can make better sense of the pronunciation of the vowels by using this approach.

A great majority of participants (83.7%) who responded to statement 7 reported that they could make better sense of the pronunciation of the vowels. Of these, 13 participants (35.1%) strongly agreed and 18 participants (48.6%) agreed. In contrast, only 3 participants (8.1%) disagreed, with 2 participants (5.4%) disagreeing and 1 participant (2.7%) strongly disagreeing. Another 3 participants (8.1%) were undecided.

Statement 8: I think I can identify the pronunciation of the vowels more effectively by using this approach.

In response to statement 8, a great majority of participants (86.4%) surveyed reported that they could identify the pronunciation of the vowels more effectively, with 16 participants (43.2%) strongly agreeing and 16 participants (43.2%) agreeing. In contrast, a very low percentage of 5.4% disagreed and another 3 participants (8.1%) were undecided.

Statement 9: I think my motivation to learn pronunciation is strengthened by using this approach.

A majority of participants (70.2%) who responded to statement 9 reported that their motivation to learn pronunciation was strengthened, with 15 participants (40.5%) strongly agreeing and 11 participants (29.7%) agreeing. A very low percentage of 2.7% strongly disagreed and another 10 participants (27%) were undecided.

As shown in the table above, the median of each statement is 4, indicating that the number of participants who agreed with the statements exceeded the number of those who disagreed with the statements. This is a very positive outcome, especially when we consider that the number of people who disagreed is very small. As indicated, many participants held positive attitudes toward pronunciation learning using the VTPL approach. In other words, in general, the participants were satisfied with the training. A more detailed discussion is presented in the next chapter.

4.5.2 Self-reports

Participants from both the experimental and control groups were asked to report on their pronunciation learning process over an 8-week period. Specifically, they were told to fill in the form when each in-class or out-of-class learning activity was completed. As mentioned in the previous chapter, the information in self-report included name, date, place, starting time, ending time, materials, problems, progress and activities before and after learning.

4.5.2.1 Self-reports in the experimental group

Self-reports of 37 participants were collected from the experimental group that trained with the VTPL approach. The reports are analyzed as follows.

According to the self-reports, many of them practiced the target English vowels on both weekdays and weekends. Compared with the dates reported at the beginning of training, they were more willing to give up part of their weekends to practice English pronunciation.

In terms of the place for study, most participants were found to practice their pronunciation on campus, including language lab, dormitory, classroom and library. Compared with the places mentioned at the beginning of the treatment, many participants reported more choices about where to study.

When it came to the starting time and ending time of each practice activity, most students reported spending 90 minutes per week (both in class and out of class) practicing the target English vowels. Few students reported spending more than 90 minutes per week on pronunciation exercises.

With regard to pronunciation learning materials, almost all the participants reported that they only listened to and repeated the filtered training materials provided by the researcher to practice the target English vowels.

Referring to the problems with pronunciation learning, participants said they had certain problems, such as difficulty in following the recordings and confusion with short and long vowels. Moreover, at the beginning of the experiment, many participants reported that the filtered recordings were vaguer in comparison to the natural ones and were more difficult to practice. However, at the end of the experiment, fewer pronunciation problems were reported, as they became accustomed to the new approach.

In terms of pronunciation progress, the results revealed that, in general, participants held positive attitudes toward the VTPL approach. For example, the following comments:

> *Pronunciation learning is not as difficult as before. It's true that I'm making progress. Gradually, I can tell the differences between short and long vowels. I can read more fluently (participant no.37).*

> *My pronunciation has somewhat improved as a result of pronunciation training. I'm more confident when speaking English (participant no.40).*

With regard to the before and after activities, participants in the experimental group reported a variety of curricular and extracurricular activities, ranging from coursework to daily chores.

4.5.2.2 Self-reports in the control group

Self-reports of 39 participants were collected from the control group that trained with the traditional approach. The reports are presented as follows.

Many participants noted that they practiced the vowels under study on both weekdays and weekends. As the training went on, it showed that the students were more willing to learn pronunciation on weekends. This was similar to what had been reported by the experimental group.

Data collected revealed that participants from the control group had preferred places for study. The study places most frequently mentioned included language lab, dormitory, classroom and library, indicating that they were similar to the students from the experimental group.

In terms of starting time and ending time for each learning session, the reported length of time for the great majority of students in the control group was 90 minutes per week, including both in-class and out of class pronunciation learning activities. Several students reported that the length of time spent on this was longer than the expected (90 minutes per week). This indicated that both groups spent a similar length of time practicing the target English vowels.

Referring to the materials that participants used for pronunciation learning, the self-reporting data indicated that they only practiced the unfiltered learning materials concerning the target English vowels, which were provided by the researcher.

With regard to pronunciation learning problems, they reported that various problems occurred during the training process, such as the problem with speech rate, the problem of repeating and difficulty in pronouncing certain vowels and vowel contrasts. Some students reported that they had the same problems both at the beginning and end of pronunciation training. That is, students had just as much difficulty as before in English pronunciation. For example:

> *As for me, some vowel sounds are really indistinguishable, such as long and short vowel sounds and perhaps also consonants. For example, I cannot tell the difference between vowel sounds /e/ and /æ/ during this training (participant no.29).*

In terms of progress in pronunciation training, some participants indicated that they had positive attitudes toward the traditional approach. Among the responses, positive attitudes were expressed by such words as "good" and "like". For example:

> *I think this is pretty good, however a bit boring. But I still like this approach, because I can improve my pronunciation by comparing my speech sounds to standard ones (participant no.46).*

In contrast, some of them classified their attitudes as negative. Negative attitudes were indicated by such words as "not effective" and "tired". For example:

> *This approach is not quite effective, because I feel like I'm not making much progress. After completion of the high-intensity learning tasks, I am exhausted and getting tired of learning (participant no.16).*

Referring to the before and after activities, self-reporting data indicated that participants in the control group took part in different in-class and out-of-class activities, which were similar to the activities reported by the experimental group.

To conclude, the self-reporting data revealed that participants from both the experimental and control groups shared many similarities, e.g. date, place, materials etc. The main difference in the self-reports between the two groups was found in the learning approach that they used, i.e. VTPL approach or traditional approach.

4.5.3 Semi-structured interviews

Semi-structured interviews were used at the end of pronunciation training to generate in-depth data from the experimental group. Six participants (16%) were randomly selected and interviewed. The interview consisted of 9 questions and was conducted in Mandarin Chinese (Putonghua). The results of the semi-structured interviews are presented below.

Theme 1: Distinctiveness of the approach

The majority of the participants (5 out of 6) mentioned the distinctive features of the VTPL approach. They used expressions such as “individual”, “more freedom” and “tailored” to show that they preferred this approach to learn pronunciation. Some of the examples are present below:

> *I think this training deals with individual students' pronunciation problems, so is it more targeted (participant no.32).*
>
> *I prefer the approach that we are using. This approach gives us more freedom to correct our weaknesses in pronunciation (participant no.40).*
>
> *I think the good aspect of the training is that not everyone uses the same piece of learning material. Those recordings are tailored to everyone's pronunciation habits (participant no.43).*
>
> *I like the approach because there are different learning stages in this approach. These different stages will allow us to adapt slowly (participant no.71).*

Theme 2: Feelings of improved performance

All participants agreed that the VTPL approach could bring about better performance in perception and production after the phonetic correction of the target vowel sounds. The examples are given below:

This training is effective. Those confusing vowels may have been pronounced almost the same in the past, but now after a few training sessions, I can tell the difference and read more accurately (participant no.9).

My personal feeling is that it helps me because it can help me to hear more clearly the differences between the sounds. Although I couldn't hear any difference at the beginning, after training a few more times and listening to it for a while, I was gradually able to tell the difference (participant no.32).

Because of the repeated training, you get to know that the parts where you produce speech are different and become aware of the characteristics of each sound. I feel that the sounds that I produce are getting closer to the standard ones (participant no.68).

Theme 3: Boosting confidence

All participants showed their confidence in learning English pronunciation through the VTPL approach. Four participants explained that continuous learning helped enhance confidence. The examples are as follows:

Yes. I'm confident. I think if you continue to practice and follow the recording, you will make progress. After all, as the saying goes, practice makes perfect (participant no.40).

I used to be unconfident in my spoken English, but now I have a certain level of confidence. I tried to avoid facing the problem when it happened to me, but now I can look squarely at it. I look it up in the dictionary and write down some notes so that I won't make the same mistake next time (participant no.68).

Besides the confidence in learning English pronunciation through this approach, one participant also mentioned the concern for more practice.

I am confident, and I have improved. I can produce vowel sounds /e/ and /æ/ better than before. However, I think I still need some training on the vowel sounds /u/ and /u:/ (participant no.32).

Theme 4: Learning preference (individual vs. group learning)

Some participants (4 out of 6) reported that they preferred to learn English pronunciation individually. They noted that they could learn at their own pace and learn more efficiently. The examples are as follows:

I prefer individual learning because it is quieter and you could enjoy more freedom. I feel free to decide what to practice and when to practice (participant no.32).

Because people's situations are different, some people still have a heavy accent and will be ashamed of opening their mouths to speak. This problem does not exist in individual learning. We can learn English pronunciation more efficiently (participant no.68).

Another two participants showed a preference for group learning, and the examples are presented below:

I prefer to learn in groups because when we get together as a group, we can discuss together, correct mistakes together and learn together (participant no.9).

You can work in a group and read to each other because people can exchange ideas, which may be more interesting as compared to the individual learning…I participate in an English learning club after class. We read various words together. There is an elder sister who is good at English. She helps us to read. Sometimes we play games, and it's quite interesting (participant no.40).

Theme 5: Difficulties in learning pronunciation

Several participants (4 out of 6) reported that they were unable to follow the recordings at the beginning of pronunciation training, but were more accustomed to them in the late stages. The examples are presented below:

As for me, when it comes to the drills of contrasting vowels in sentences, the words *that sound similar will be confusing if those sentences are too long. I must listen to them a few more times to follow the recordings (participant no.9).*

I couldn't do it at the beginning, especially for sentences, because there were many similar sounds. In the subsequent training, the situation seemed to remain unchanged. However, in the later stages, I was able to listen and repeat (participant no.32).

I found the confusing vowel sounds in the previous perception test difficult to distinguish. For this test, although I still have some difficulties when the words are read too quickly, in general, they are more obvious than the previous time, and the distinctions of certain sounds are particularly obvious (participant no.40).

Theme 6: Expectations of future training

The participants expressed their hopes for future training, including control of speech rate, improvement in diversity and difficulty of training materials, and integration of multiple functions in the VTPL approach. The examples are as follows:

I hope that this approach could be more helpful by integrating multiple functions into it, including listening, repeating, recording and rating (participant no.9).

If sentences are read too fast, it will be difficult for me to follow. I suggest that the speed of speech should be controlled (participant no.40).

I think that the words provided are not large enough, and this aspect needs to be improved. As students, we need to expand our vocabulary (participant no.68).

I need more words for practice. We need to add more leaning materials in terms of difficulty and diversity (participant no.71).

4.6 Answers to research questions

4.6.1 Answer to research question 1: What are the corrective optimals of Chinese non-English major EFL learners for the following English vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/?

The target vowels (i.e. /ɪ/, /iː/, /e/, /æ/, /ʊ/, /uː/) were best produced when filtered through corrective optimals consisting of discontinuous multiband filters containing a low-frequency component. As indicated, for the majority of the students, corrective optimals were identified as follows: 0-320 Hz + 2419-3048 Hz (fcenter = 2715 Hz) for /ɪ/; 0-320 Hz + 4838-6096 Hz (fcenter = 5431 Hz) for /iː/; 0-320 Hz + 2419-3048 Hz (fcenter = 2715 Hz) for /e/; 0-320 Hz + 1512-1905 Hz (fcenter = 1697 Hz) for /æ/; 0-320 Hz + 214-269 Hz (fcenter = 240 Hz) for /ʊ/; 0-320 Hz + 302-381 Hz (fcenter = 339 Hz) for /uː/. At the same time, there were some other corrective optimals, which were also found to be effective for some of the students in the experimental group: 0-320 Hz + 1815-2286 Hz (fcenter = 2037 Hz) for /ɪ/; 0-320 Hz + 4567-6459 Hz (fcenter = 5431 Hz) and 0-320 Hz + 4435-5588 Hz (fcenter = 4978 Hz) for /iː/; 0-320 Hz + 2016-2540 Hz (fcenter = 2263 Hz) and 0-320 Hz + 1613-2033 Hz (fcenter = 1811 Hz) for /e/; 0-320 Hz + 1209-1523 Hz (fcenter = 1357 Hz) and 0-320 Hz + 1815-2286 Hz (fcenter = 2037 Hz) for /æ/; 0-320 Hz + 415-523 Hz (fcenter = 466 Hz), 0-320 Hz + 303-382 Hz (fcenter = 340 Hz) and 0-320 Hz + 542-682 Hz (fcenter = 608 Hz) for /ʊ/; 0-320 Hz + 277-349 Hz (fcenter = 311 Hz) for /uː/.

4.6.2 Answer to research question 2: Are there any differences between the native speaker optimals and the corrective optimals for Chinese non-English major EFL learners? If yes, what are these differences?

According to the results, native speaker optimals and corrective optimals were different from one another in two ways. First, the corrective optimals turned out to be narrower and finer in comparison with the native speaker optimals as stipulated by Guberina and other classical practitioners of the verbotonal theory. Second, the corrective optimals were more diverse and less uniform than the native speaker optimals.

4.6.3 Answer to research question 3: Is the VTPL approach effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners? In particular, is simple exposure to the corrective optimals alone sufficient to bring about acceptable production of the target vowels?

The experimental group experiencing pronunciation training showed statistically significant improvements in terms of both the perception and production of the target vowels from pretests to posttests. The overall speaking proficiency also improved in a significant way after phonetic correction concerning comprehensibility, fluency and pronunciation. The findings indicated that the simple exposure to the corrective optimals alone was effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners.

4.6.4 Answer to research question 4: Which approach, VTPL or traditional, is more effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners?

The results indicated that, before training, no significant differences were found between the experimental and control groups concerning perception and production tasks, including comprehensibility, fluency and pronunciation. The two groups started at the same performance level. However, after training, the experimental group scored significantly higher than the control group for five out of six vowel sounds on both the perception and production tasks (i.e. word-reading, sentence-reading and storytelling). Moreover, the improvements were significantly greater in the experimental group than in the control group for comprehensibility, fluency and pronunciation both in sentence-reading and storytelling. The findings indicated that the VTPL approach was more effective than the traditional approach for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners.

4.6.5 Answer to research question 5: What are the Chinese non-English major EFL learners' opinions of the VTPL approach to phonetic correction?

According to both qualitative and quantitative aspects of data analysis, a great majority of students who participated in this study had positive attitudes toward using the VTPL approach for learning the target English vowels. The results in the current study suggested that, in general, the students were satisfied with the training.

4.7 Summary

This chapter presented the results of both the quantitative and qualitative data analysis in a systematic and detailed way. Specifically, the quantitative data were analyzed using JASP statistical software. The qualitative data were interpreted using a thematic analysis procedure. According to research findings, the VTPL approach appeared to be a beneficial and effective approach to phonetic correction of the target English vowels for Chinese non-English major EFL learners. In the next chapter, a discussion of research findings will be given.

CHAPTER 5

DISCUSSION

This chapter aims to discuss and interpret the research findings presented in the previous chapter. The first section focuses on a discussion of corrective optimals, including a comparison between the corrective optimals and native speaker optimals. The second and third sections include discussions of perception and production performances. The fourth section discusses comprehensibility, fluency and pronunciation. The fifth section is devoted to a discussion of students' attitudes toward the VTPL approach. The sixth section focuses on the key outcomes of the research. In the last section, a summary of the chapter is given.

5.1 Corrective optimals

As reported in the previous chapter, the target English vowel sounds (i.e. /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/) were best produced when the target vowels were filtered through the corrective optimals determined in this study (i.e. discontinuous multiband filters) (see Section 4.1). Specifically, the combined use of a lowpass filter (below 320 Hz) and a single bandpass filter (i.e. partial octave) was determined to be effective for enabling students to produce acceptable vowel sounds.

According to the results of the study on corrective optimals, there are important findings worth pointing out.

First, the results showed that the majority of students shared the same corrective optimals for each of the six target vowel sounds. This indicated that learners from the same cultural and educational backgrounds tended to give the same perceptual value to the same sound.

Second, not every individual learner had the same set of personal corrective optimals, confirming the existence of significant learner differences in perception of the target vowel sounds. The results suggested that although the learners shared the same basic perceptual mechanisms, personal differences were still possible.

Third, corrective optimals constructed with discontinuous multiband filters were clearly more effective than the single octave-wide bandpass filter optimals traditionally used and the partial octaves (i.e. $\frac{1}{2}$ octave and $\frac{1}{3}$ octave) tested in the present study, showing the value of combining lowpass filtering and partial octave bands.

Fourth, most important, as shown in the results of this study, simple exposure to corrective optimals followed by repetition did actually influence participants' speech productions at the diagnostic stage even with no perceptual training, thus confirming the close connection between perception and production as well as confirming the value of the filtering process (simple exposure to unfiltered language did not produce comparable improvements).

Fifth, it is also important to note that, according to the study's findings, lowpass filtering plays a critically important role in the construction of optimals. As part of the tradition of verbotonal theory, optimals consist of octave-wide bands above the fundamental frequency (F_0). Low frequencies are frequently used in speech treatment. In the current study, only the sounds to be learned were filtered, thus maintaining the intelligibility of the language used. Necessarily, the lowpass filtering was applied only when the target vowel sounds in words and sentences were filtered, and the rest of the sounds remained the same, i.e. lowpass filtering was active only for the duration of the vowel. In order to offer an explanation for the effectiveness of the lowpass filtering in improving vowel production, the following two assumptions are made.

First, lowpass filtering (fundamental frequency F_0) makes vowel sounds more salient. As Hillenbrand and Gayvert (1993) proposed, "perceived vowel quality is strongly correlated with the frequencies of the two or three lowest formants" (p. 694). On the basis of acoustic analysis, the fundamental frequency has been regarded as the main correlate of the pitch patterns of the linguistic prosodic systems of tone and intonation (Abberton & Fourcin, 1997). Similarly, Honorof and Whalen (2005) also pointed out that "fundamental frequency (F_0) carries information about many different linguistic and paralinguistic aspects of the speech signal" (p. 2193). In other words, the fundamental frequency embodies characteristics of the vowel that we are trying to produce. After all, each vowel is never produced in a monotone, and we seldom produce identical sounds. In this sense, the fundamental frequency also helps to reflect the other characteristics of the vowel. As a result, when the intonation is changed or corrected inside each vowel, all the other formants (i.e. F_1, F_2, F_3, etc.) adjust themselves to and become harmonious with the fundamental frequency. This indicates that, without the vibration of F_0, we cannot produce vowels. It also indicates that if the intonation is manipulated, then the formants are also going to get manipulated because it is the fundamental frequency of the voice (F_0), the intonation that generates the formants. Without F_0, the intonation, there would be no formants.

Additionally, the low frequency of the sound stimuli can impact directly on the learner's body as a whole given that the body is sensitive to low frequencies, thus offering different and additional support for the sound to be perceived (Asp, 2006; Lian, 1980).

However, one of the problems in dealing with pronunciation learning is that we have been taught in our academic courses to separate phonemes and intonation from one another and to focus initially and primarily on phonemes with intonation then added as a kind of decoration. This seems based on the simplistic assumption that spoken language consists of a base of individual sounds strung together (the linguistic), with intonation (which is only paralinguistic) plastered onto this base structure. Phoneticians have long-known, however, that this is simply not the case and that single sounds rarely occur in isolation but are constantly in transition: everything is moving, everything is dynamic. Thus, we are usually taught to focus first on individual sounds, and then to move on to intonation, but actually, the individual sounds cannot be separated from the intonation: they are the same thing. Intonation is embodied in the phonemes and the phonemes are embodied in intonation. In our case, exposure to lowpass filtering not only helps to emphasize the individual sounds and intonation but also helps to give sounds some of their dynamic characteristics. As a result, if a significant, perhaps enhanced, intonative component is added to auditory stimuli, learners may become more aware of the characteristics of the sounds that they are studying.

Second, the effectiveness of the lowpass filtering could be explained in relation to cerebral lateralization, i.e. left-hemispheric lateralization and right-hemispheric lateralization. Over the past two decades, numerous research studies have been devoted to the discussion of the issue. Some researchers have stated that language processing depends largely on the left hemisphere of the brain (Cogan et al., 2014; Gut, 2009; Indefrey, 2011; Price et al., 1996; Schiffler; 1992; Watkins, Strafella, & Paus, 2003). As Schiffler (1992) pointed out, the left hemisphere is in charge of "precise, formal verbal learning and logical, analytical thinking" while the right hemisphere is responsible for "imagery and associative-intuitive and prosodic, verbal learning, as well as concrete, holistic-synthetic thinking" (p. 22). Similarly, according to Gut (2009), the left brain is responsible for "decoding the meaning of a linguistic message", whereas the right brain is devoted to "the perception of music and the prosodic characteristics (especially pitch) of language" (p. 189). In a word, the left brain is the master of verbal or language-based information, while the right brain is the master of non-verbal or performance-based information.

In contrast to the above view, a number of researchers have found that the right hemisphere also actively participates in language processing: speech perception and speech production involve both hemispheres of the brain (Alexandrou, Saarinen, Mäkelä, Kujala, & Salmelin, 2017; Byrd & Mintz, 2010; Hickok & Poeppel, 2007; Hultén, Karvonen, Laine, & Salmelin, 2014; Simonyan & Fuertinger, 2015; Silbert, Honey, Simony, Poeppel, & Hasson, 2014). For example, in more recent research, Alexandrou et al. (2017) explored the neural substrates of natural speech perception and speech production with magnetoencephalography by modulating three speech-related central features, i.e. the amount of linguistic content, speaking rate and social relevance. They reported that "the right hemisphere also occupies an important role in processing meaningful speech" and "speech production and speech perception show a notable bilateral overlap" (p. 636). At the same time, the results also revealed that "such an overlap also occurs for natural speech and that this joint activation pattern is particularly salient in the right hemisphere" (p. 636). That is, the notable involvement of the right hemisphere in speech processing is highlighted. In this study, the inclusion of a lowpass frequency component in the corrective optimals capitalizes on the right hemisphere involvement.

Moreover, previous research studies also suggested that the left-right brain activity is culture-dependent (McGilchrist, 2009; Rozin, Moscovitch, & Imada, 2016; Sheikh & Sheikh, 1996; Springer & Deutsch; 1997). For example, Rozin et al. (2016) conducted research to direct attention to a possible cultural-brain parallel. In their research, four lateralized tasks (i.e. ambiguous face-vegetable array, nose identification, form vs function/semantics, and hierarchical processing of letters in letter identification: holistic vs analytic processing) were used to distinguish left versus right hemisphere functioning. The findings showed that there existed “a significant link between Asian origins and rearing, and enhanced, or default right hemisphere processing in three of four cases” (p. 7), indicating that people from eastern (i.e. East Asian and South Asian) cultures put more emphasis on the right hemisphere of the brain as compared to western (i.e. Euro-American) cultures. This would be the case of Chinese EFL learners.

Arguably, successful results could be obtained if we take a brain that is already trained in performing a lot of right brain processing and feed it right brain type signals such as lowpass filtering, because this may influence it more than the brain which has no practice in processing signals through the right brain. The point of having the lowpass filtering is that individuals can actually maximize the normal language processing mechanisms because lowpass filtering acts directly on the right brain. In light of the above-mentioned discussions, it is reasonable to expect that, after perceptual training, the brain can be more activated in speech processing, helping the learners to raise their phonological awareness and produce the target vowel sounds acceptably.

It should also be pointed out, however, that the opposite outcome may be possible: stimulating a right brain that is already linguistically-formed may prevent it from reacting to novel signals and simply rejecting them as the hemisphere is already linguistically formed. Thus, the signals might not be recognized. This realization opens up a new set of valuable and original research projects.

Corrective optimals vs. native speaker optimals

As mentioned in Chapter 3, the determination of the corrective optimals was based on the traditional native speaker optimals. From the research results, the corrective optimals and the native speaker optimals differ from one another in two aspects.

First, the corrective optimals for Chinese non-English major EFL learners were shown to have narrower and finer filtering in comparison to the native speaker optimals. In terms of the native speaker optimals, the six vowels have traditionally been determined as optimal when produced through the filtering of the full octave bands. As for the corrective optimals, all six target vowel sounds were best perceived and produced when filtered through discontinuous multiband filters containing a low-frequency component. Specifically, the majority of the students produced the target vowel sounds best when the sounds were filtered through a discontinuous multiband filter 0-320 Hz + $\frac{1}{3}$ octave or 0-320 Hz + $\frac{1}{2}$ octave, indicating the effectiveness of the combined use of a lowpass filter (below 320 Hz) and a single bandpass filter ($\frac{1}{3}$ octave or $\frac{1}{2}$ octave).

Second, the corrective optimals for Chinese non-English major EFL learners turned out to be more varied and less uniform when compared to the native speaker optimals. Take /e/ sound, for example, the native speaker optimal for this sound was found to be 1600-3200 Hz (fcenter = 2263 Hz). In contrast, more corrective optimals were diagnosed for /e/ sound, including 0-320 Hz + 2419-3048 Hz (fcenter = 2715 Hz), 0-320 Hz + 2016-2540 Hz (fcenter = 2263 Hz) and 0-320 Hz + 1613-2033 Hz (fcenter = 1811 Hz). This indicates the existence of perceptual differences between the native speakers of English and Chinese non-native speakers of English. Therefore, in order to help learners in non-English speaking countries produce acceptable English sounds, the classical verbotonal theory needs to be adjusted and changed to cater to new social demands. Henceforth, the focus will be on acceptable performance rather than native speaker performance.

5.2 Perception performance

Between-group comparisons indicated that, in the pretest, no significant differences in perception were found between the means of the experimental and control groups, indicating that the two groups started at the same level. In the posttest, statistically significant differences were found between the means of the two groups for the vowel sounds /ɪ/ ($p = 0.004$), /iː/ ($p = 0.002$), /e/ ($p = 0.009$), /ʊ/ ($p = 0.006$), and /uː/ ($p = 0.036$), with the effect size ranging from 0.49 (medium) to 0.73 (medium to large). Yet, no significant difference was found between the means of the two groups for the sound /æ/. In 5 out of 6 instances, the experimental group outperformed the control group after training.

Within-group comparisons indicated that, in the experimental group, statistically significant differences were found between perception pretest and posttest scores for all six vowels ($p < 0.001$), with the effect size ranging from 0.70 (medium to large) to 1.15 (large). In the control group, statically significant differences between pretest and posttest scores were found only for /æ/ ($p = 0.005$) and /u:/ ($p = 0.008$) and the effect sizes were medium.

The results showed that repeated exposure to the filtered target vowels could, in and of itself, without the intervention of teachers or anyone else, result in an improvement in pronunciation. The findings of the present study on improved perception are in agreement with and confirm the findings of several other studies. For instance, Miner and Danhauer (1977) found that exposure to optimal octave bands contributed to correct perception and identification of the vowel sounds. Similarly, Asp (2006) also pointed out that the optimal octaves were effective for establishing correct perception.

One possible explanation for the positive results relates to the use of lowpass filtering. Guberina and Asp (1981), who used the verbotonal approach for rehabilitating people with communication problems, pointed out that the lowpass filtering accounted for the enhanced rhythm and intonation of the language. Moreover, according to Laroy (1995), the quality of speech sounds (phonemes) was greatly influenced by the suprasegmental features of rhythm and intonation. As Jakobson and Waugh (2002) pointed out, "the prosodic features are a property of phonemes when functioning as syllabics and thus are primarily a property of vowels" (p. 146). Lowpass filtering helps to improve vowel perception by highlighting the prosodic features built into every set of vowel production.

5.3 Production performance

In this section, the discussion is devoted to the three aspects of speech production tested, i.e. word-reading, sentence-reading and storytelling. They are detailed as follows.

5.3.1 Word-reading

Between-group comparisons indicated that, there were no significant differences in word-reading between the means of the two groups before pronunciation training. That is, the two groups started at the same level. After training, there were statistically significant differences between the means of the two groups for the target vowels /ɪ/ ($p = 0.003$), /iː/ ($p < 0.001$), /e/ ($p < 0.001$), /ʊ/ ($p = 0.010$), and /uː/ ($p = 0.027$), with the effect sizes ranging from 0.52 (medium) to 0.88 (large). However, in the posttest, no significant difference was found between the means of the two groups for the sound /æ/, the experimental group outperformed the control group in the 5 other vowels (i.e. 5 out of 6 vowels).

Within-group comparisons indicated that, in the experimental group, statistically significant differences in word-reading were found between pretest and posttest scores for these target vowels ($p < 0.001$) and the effect sizes were all large. In the control group, statistically significant differences between pretest and posttest scores were found only for /æ/ ($p = 0.016$) and /u:/ ($p < 0.001$) and the effect sizes were 0.40 (small to medium) and 0.58 (medium).

The better performance of the students in the experimental group demonstrates the link between perception and production. According to Guberina and Asp (1981), each speaker is a producer, and at the same time, a receiver of speech. The way that we produce auditory information is a reflection of the way that we perceive speech. Thus, it is reasonable to conclude that if the student's perception changes, his or her speech will also change. If we have corrected the student's speech, it is because we have corrected his or her perception. In the current study, improvements in perception did result in improvements in production, confirming that perception and production are closely related.

When speaking in English, it is important to stress the correct syllable in each word as well as the sentence in order to achieve clear communication. This is known as word stress. In English, every word has a word stress, which is only on a vowel. Word stress helps us to make sure which version of a particular word we are hearing. For example, PREsent is the noun that refers to a thing that you give to somebody as a gift and preSENT is the verb that means offering something to someone. Although the two words share the same root, the word stress falls on different syllables. According to Carmen (2010), English native speakers use word stress so naturally that they do not even notice they use it. If non-native speakers who speak English to native or other competent speakers of English do not use word stress properly, they will encounter two problems. First, they will find it difficult to understand competent speakers of English, especially those who speak fast. Second, the competent speakers of English find it difficult to understand them. In the current study, interview data showed that, at the beginning, students reported having more difficulty figuring out the difference between contrasting vowel sounds, however, after training, the students said they had fewer problems and were able to tell the difference. Taking into account the fact that word stress is always on a vowel, together with the students' improvements in the production of the target vowels, it is reasonable to conclude that after training the students' use of word stress had also improved comparatively.

5.3.2 Sentence-reading

Between-group comparisons indicated that, in the pretest, no significant differences in sentence-reading were found between the means of the two groups. They started at the same level. In the posttest, significant differences were found between the means of the experimental and control groups for /ɪ/ ($p = 0.004$), /iː/ ($p = 0.003$), /e/ (p

= 0.010), /ʊ/ ($p = 0.008$), and /u:/ ($p = 0.044$), with the effect sizes ranging from 0.47 (medium) to 0.70 (medium to large). However, no significant difference was found between the means of the two groups for the sound /æ/. This means that the experimental group performed much better than the control group in 5 out of 6 instances.

Within-group comparisons indicated that the experimental group improved significantly in sentence-reading from the pretest to the posttest concerning the production of the target vowels /ɪ/ ($p < 0.001$), /i:/ ($p < 0.001$), /e/ ($p < 0.001$), /æ/ ($p = 0.002$), /ʊ/ ($p < 0.001$), and /u:/ ($p < 0.001$), with the effect sizes ranging from 0.54 (medium) to 1.26 (large). In the control group, statistically significant differences between pretest and posttest scores were found only for /æ/ ($p = 0.029$) and /u:/ ($p = 0.004$) and the effect sizes were 0.36 (small to medium) and 0.48 (medium).

In this study, three kinds of filtered sentences (i.e. optimal sentences, contrast-embedded sentences and non-optimal sentences) were designed to help raise the students' phonological awareness for both the experimental and control groups. During the self-managed training sessions, experimental group students went through the digitally-filtered sentences of varying difficulty, ranging from load-lightening activities to load-increasing activities. The good results confirmed the effectiveness of the corrective optimal filters and evidenced that students' awareness of English pronunciation has been enhanced, and they were able to produce the target vowels acceptably concerning sentence-reading.

According to interview data, at the beginning of training, students in the experimental group reported their difficulty in following the sentences because they were read too fast. In contrast, at the end of the treatment, the students noted they were more accustomed to the rapid spoken English. Surprisingly, they were able to quickly retrieve and process not only the target vowel sounds but many other language sounds (i.e. vowels and consonants) too. This shows that changes in perception lead to changes in the brain, providing additional support for the impact of verbotonal theory on neuroplasticity.

5.3.3 Storytelling

Between-group comparisons indicated that, in the pretest, there were no significant differences in storytelling between the means of the experimental and control groups, showing that the two groups started at the same level. In the posttest, however, there were statistically significant differences between the means of the two groups for 5 out of 6 target vowels /ɪ/ ($p = 0.004$), /iː/ ($p = 0.003$), /e/ ($p = 0.009$), /ʊ/ ($p = 0.013$), and /uː/ ($p = 0.037$), with the effect sizes ranging from 0.49 (medium) to 0.72 (medium to large). No significant difference was found between the means of the two groups for the sound /æ/. Clearly, the experimental group outperformed the control group.

Within-group comparisons indicated that, in the experimental group, statistically significant differences in storytelling were found between pretest and posttest scores for the six target vowels ($p < 0.001$), with the effect sizes ranging from

0.62 (medium to large) to 1.29 (large). In the control group, statistically significant differences were found only for /æ/ ($p = 0.006$) and /u:/ ($p = 0.015$) and the effect sizes were 0.47 (medium) and 0.41 (small to medium).

As storytelling is more akin to real-life communication, the VTPL approach is also likely to be especially valuable for real-life communication, which is more demanding than word-reading and sentence-reading. It works better most likely because the knowledge is better internalized and more retrievable as a result of the way in which it was learned. There are a number of studies that lend support to the findings of the present study in terms of longer connected speech. For instance, a study by Zhang (2006) reported that the students assisted by the verbotonal-based approach did much better in conversations (i.e. short dialogues) than the control group concerning the aspects of intelligibility, intonation and pitch. In He's (2014) research, she found that the students working with the verbotonal-based approach did perform significantly better in oral interviews as compared to the students using the traditional approach. Similarly, a more recent study by Yang (2016) suggested that the verbotonal-based approach was a better choice for improving the students' overall performances in oral interviews and helping them produce a flow of meaningful speech.

The effectiveness of the VTPL approach in facilitating carryover to longer connected speech could be explained by the fact that this approach puts emphasis on both vowel sounds and intonation. As aforementioned, individual sounds are not isolated from the intonative context of their productions. In reality, intonation is an integral part of what we call a sound, and a sound is an integral part of intonation, and all vowels are a product of fundamental frequency (i.e. intonation). Furthermore, the melody influences and overrides the formants. That is, the fundamental frequency necessarily embodies characteristics of the vowel that we are trying to produce. Traditionally, we talk about sounds and intonation as being separate from each other, and then we connect them. This approach is based entirely on our categorization and conceptualization of the relationship between vowel sounds and prosody. We look at F_1, F_2, F_3, etc. and we decide that these are somehow different from F_0, yet, without F_0, there would be no F_1, F_2, F_3, etc. In other words, the distinction between individual sounds and intonation is a false distinction.

The integration of both features explains why the students in the experimental group could produce longer connected speech (i.e. storytelling) that can be understood more easily. It actually says that, as an outcome of training, the face-to-face longer natural utterances are more successfully produced and also the differences between the experimental and control groups are greater (see Table 4.11). This means that the sounds of utterances can be retrieved and produced faster in the experimental group. This is an indication that the phonemes the students are trying to learn are somehow more deeply embedded in the experimental group than in the control group. That is to say, they are better managed in the brains of the experimental group. Exactly how this happens cannot be addressed by the current study. Importantly, it is in face-to-face mode where these skills are really needed in real life so as to competently maintain spoken interaction. Moreover, the fact that the sounds can be retrieved and produced faster is also an indication of improved phonological working memory in the experimental group.

In addition, when the lowpass filter is added to the target vowel sounds, the prosodic features of the sounds are enhanced, leading to better awareness-raising and meaning-making and better internalization by the learner of the characteristics of the sounds being learned and greater capacity to recall successful production techniques.

The results indicated that the two groups started at the same perceptual and production level. However, after training, as is shown by the improvement ratio, the experimental group outperformed the control group by a factor of 8.76 times on average concerning all perception and production tasks taken together, indicating that the VTPL approach was more powerful when compared to the traditional approach in helping the students correctly perceive and acceptably produce the target vowel sounds.

5.4 Comprehensibility, fluency and pronunciation

This section is concerned with discussions of comprehensibility, fluency and pronunciation in sentence-reading and storytelling. The first subsection focuses on a discussion of comprehensibility. The second subsection is devoted to a discussion of fluency. The third subsection includes a discussion of pronunciation.

5.4.1 Comprehensibility

As an essential aspect of L2 proficiency, most recent studies define comprehensibility in Munro and Derwing's (1995a) sense. From their point of view, comprehensibility is defined as listener's judgements concerning the effort required to understand an L2 speech production, rather than the listener's understanding of the content (Derwing et al., 1998; Derwing et al., 2008; Munro & Derwing, 1995a). Over the years, a distinction has grown between comprehensibility and intelligibility. According to Levis (2006), the difference between comprehensibility and intelligibility depends on whether the concept of intelligibility is understood in its narrow or broad sense. Intelligibility is assessed through the inspection of listeners' accuracy of L2 learners' utterances (narrowly speaking). More broadly, the term is used interchangeably with comprehensibility and refers to the ease of understanding (Munro & Derwing, 1995a). In this study, comprehensibility coveys the same meaning as intelligibility and refers to the ease of general understanding.

Between-group comparisons indicated that, in the pretest, no significant differences in sentence-reading and storytelling were found between the means of the two groups concerning comprehensibility, showing that the two groups were at the same comprehensibility level before training. However, after treatment, there existed significant differences in sentence-reading ($p = 0.001$) and storytelling ($p = 0.005$)

between the means of the two groups and the effect sizes were medium to large. That is, the experimental group significantly outperformed the control group.

Within-group comparisons indicated that the experimental group improved their comprehensibility from the pretest to the posttest in a statistically significant way in terms of sentence-reading and storytelling ($p < 0.001$) and the effect sizes were large. However, no such improvement occurred in the control group.

According to previous research, a series of studies have been conducted to investigate the impact of segmentals and suprasegmentals on the measure of comprehensibility. Crowther, Trofimovich, Saito and Isaacs (2015) argued that comprehensibility was associated with both segmentals and suprasegmentals. That is to say, not only segmental but also suprasegmental features can affect comprehensibility. Importantly, previous research found that suprasegmentals affected the way segmentals were pronounced (Sewell, 2016). Derwing (2008) also noted that suprasegmental features appeared to have more impact on comprehensibility than segmental features. Taking into consideration the improvements obtained by the experimental group, it seems reasonable to conclude that the VTPL approach is unexpectedly effective in improving comprehensibility, even though no training was conducted at the prosodic level. We can surmise from previous research studies that the lowpass filtering may be significantly responsible for the improved management of comprehensibility as per He (2014) and Yang (2016). That is, the use of even small amounts of the low-frequency band (below 320Hz) to filter the target vowel sounds may have helped to provide enough prosodic assistance to improve speech comprehensibility significantly.

5.4.2 Fluency

In recent years, the fluency dimension of speech has attracted much attention from scholars in the field of L2 speech production (Derwing & Munro, 2009; Derwing et al., 1998; Kang, 2010; Saito et al.,2015; Trofimovich & Isaacs, 2012). De Jong (2016) argued that fluency was one of the most critical dimensions of speech that made an L2 production performance successful. In terms of fluency evaluation, some scholars suggested that the assessment of fluency depended on several concrete measures of speech. According to Trofimovich and Isaacs (2012), they include filled (non-lexical) pauses, unfilled (silent) pauses, pause errors, repetitions/self-corrections, articulation rate and mean length of run (MLR). In the present study, to make the judgement more manageable, the notion of fluency is understood as general fluency when speaking English.

Between-group comparisons indicated that, in the pretest, there were no significant differences in fluency between the means of the experimental and control groups in terms of sentence-reading and storytelling. In other words, they started at the same fluency level. However, after training, there existed significant differences between the means of the two groups with regard to sentence-reading ($p = 0.005$, medium to large effect) and storytelling ($p = 0.012$, medium effect): the experimental group outperformed the control group.

Within-group comparisons indicated that, the experimental group improved their fluency from the pretest to the posttest in a statistically significant way concerning both sentence-reading and storytelling ($p < 0.001$) and the effect sizes were large. In contrast, no significant improvement was found within the control group.

In other words, the experimental group showed more fluent speech production as compared to the control group, indicating that the VTPL approach was more effective than the traditional approach in enhancing the students' fluency. This is a welcome but unexpected result. One possible explanation for this surprising result is that while both groups focused on the production of individual sounds in identical connected speech contexts, the procedures adopted by the experimental group used low-frequency input that seems to have had a collateral impact.

As we know, many learners have problems with their English and struggle to become fluent speakers. However, they find it difficult to produce sentences in a fluent way when they speak, even though they study hard and learn lots of grammar and vocabulary. One possible explanation for this refers to the learners' lack of flow in their speech production that emerges from their inability to quickly retrieve from memory and produce a rapid stream of correct grammatical, lexical and phonological structures. When it comes to L2 oral ability, for better fluency, the concept of flow cannot be overlooked. It is widely accepted that flow is basically comprised of four constructs, including contractions, reductions, linking and flap sound. In this sense, better flow enables better fluency. In terms of the good experimental results for fluency, it is logical to conclude that the students' flow in the experimental group has also improved comparatively after training.

5.4.3 Pronunciation

As a medium of communication used by people speaking different languages, the English language has now moved to a new status as a global language and worldwide lingua franca (Galloway, 2017; House, 2014; Jenkins, 2015; Matsuda, 2012; Northrup, 2013; Pan, 2015; Rose & Galloway, 2019; Seidlhofer, 2011; Walker, 2010). In response to this trend, English pronunciation learning and teaching should no longer focus on helping learners to speak English like a native. Instead, it should be more focused on helping learners produce sounds which fall within the range of acceptable realizations of sounds for most speakers of English. The current study, in line with this idea, focuses holistically on acceptable production of the sounds of English.

Between-group comparisons indicated that, in the pretest, there were no statistically significant differences in pronunciation between the means of the two groups concerning sentence-reading and storytelling. In other words, they started at the same level. However, in the posttest, significant differences were found between the means of the two groups with regard to sentence-reading ($p < 0.001$, large effect) and storytelling ($p = 0.005$, medium to large effect). That is, the experimental group did significantly better than the control group.

Within-group comparisons indicated that, in the experimental group, significant differences in pronunciation were found between pretest and posttest scores concerning sentence-reading and storytelling ($p < 0.001$, large effect), while no such difference was observed in the control group.

In other words, the students in the experimental group significantly improved their pronunciation after being exposed to the VTPL approach. This is an indication that after training the experimental group could produce more acceptable English sounds (i.e. vowels and consonants) than the control group. As Lian (1980) argued, the successful learning of pronunciation happens when the learners are able to defeat their "deafness" and develop a "feel" of body and phonation. After training, students in the experimental group were able not only to feel better the target vowel sounds but also felt better the surrounding sounds (i.e. other vowels and consonants). This is probably because after being exposed to the filtered target vowel sounds, students' perceptions of these sounds would most likely be corrected, and thus leading to changes in their brains which would further act on their perceptions and productions of the English vowels and consonants.

The results of the present study indicated that the two groups started at the same performance level in sentence-reading and storytelling in terms of comprehensibility, fluency and pronunciation. In contrast, after treatment, as is shown by the improvement ratio, the experimental group was on average 4.31 times better than the control group concerning all three aspects. Although we did not set out to explicitly improve comprehensibility, fluency and pronunciation, a welcome side effect of pronunciation training for the six vowel sounds of the study resulted in all three measures improving. In contrast, the control group did not improve in any of these aspects. This is an important outcome and additional benefit to the VTPL approach.

To summarize, the VTPL approach turns out to be a valuable approach for enabling Chinese non-English major EFL learners to generate acceptable productions of the target English vowel sounds.

5.5 Students' attitudes toward the VTPL approach

The qualitative aspect of results obtained from the current study indicated that a great majority of students in the experimental group showed positive attitudes toward the VTPL approach which implied that most of the students were satisfied with this approach. A detailed discussion is presented below.

First, according to the results, a majority of students (75.7%) thought that the current pronunciation lessons were interesting. They reported that pronunciation learning assisted by the VTPL approach was good, effective and impressive. Arguably, this is likely to reduce their affective filter (Krashen, 1982) and enable them to learn the target vowels better. As is shown in the results, several students (10.8%) thought the lessons were not interesting and reported they were afraid of the filters and not interested in the filtered sounds.

Second, a high percentage of students (64.8%) reported that the current pronunciation lessons were helpful in assisting the production of the target vowels. Two possible explanations would account for the helpfulness. One explanation was that, in the present study, each student was optimally supported using the corrective optimals, which is connected to precision language education (Lian & Sangarun, 2017). Another explanation was that the pronunciation training was well organized to include awareness-raising activities of varying difficulty, ranging from optimal to non-optimal as per standard verbotonal theory. At the same time, some of the students (32.4%) were uncertain about the helpfulness of this approach. As noted, they found themselves confused by the unusual features of the procedures used in this study, although it clearly worked for them.

Third, a majority of students (59.4%) indicated that they felt confident about improving their pronunciation thought the VTPL approach. As reported, they became more confident and less worried and could face their problems better. However, as indicated in the qualitative data, some students (35.1%) were uncertain about their confidence in using this approach to improve their pronunciation. The concern they expressed was that although they thought it worked for those vowels, they were not sure if this approach would work for other English vowels and consonants.

Fourth, a high percentage of students (67.6%) felt comfortable with the current pronunciation lessons. A possible explanation for this was that the learning protocol created in this study allowed students to work entirely on their own. That is, they could take charge of their learning in a self-managed learning environment. However, as indicated, a minority of students (27%) were undecided. The reason they mentioned was that they were not accustomed to the filtered sounds.

Fifth, a majority of students (59.4%) reported they preferred the VTPL approach to other approaches in pronunciation learning. As indicated, they expressed their good feelings toward the VTPL approach. They also noted that this approach worked much better for them in comparison to the traditional approach concerning pronunciation learning because the learning materials provided were more tailored to meet individual learner's pronunciation learning needs. In contrast, a minority of participants (16.2%) did not show their preference for the VTPL approach. They reported that they had not been exposed to other approaches, and therefore, they were unable to say which approach was most suitable for them to use.

Sixth, a high percentage of students (72.9%) reported that they liked to learn pronunciation on their own through the VTPL approach. They reported that this approach was effective in addressing individual differences as well as assisting individual learning. One explanation was that, in the current study, each student was diagnosed and supported at a personal level. Another explanation was that each student had free access to the personalized learning materials designed for classroom training. Therefore, they could enjoy more freedom and decide on their own about what to practice and when to practice.

Seventh, a great majority of participants (83.7%) reported that they could make better sense of the pronunciation of the vowels by using the VTPL approach. One explanation was that the sounds have been internalized and stayed in students' phonological working memory. Another explanation was that students' "deafness" has been defeated and therefore were able to develop a better "feel" of body and phonation.

Eighth, a very high percentage of students (86.4%) stated that they could identify the pronunciation of the vowels more effectively. They reported that personal corrective optimals were helpful, and they were able to listen and speak better. They also noted that they were more able to tell the difference between the contrasting vowel sounds.

Ninth, a majority of students (70.2%) reported that their motivation to learn pronunciation was strengthened. A possible explanation for this would be that the students had the desire to learn after an initial improvement in the pronunciation of the target vowel sounds. A minority of students (27%) were undecided about whether or not this approach would strengthen their motivation to learn. They explained that they encountered some problems during the learning process, such as difficulty in following the recordings, confusion with short and long vowels and the fast speech rate and therefore, they rated their answers as undecided.

This is a very positive outcome, taking into account the number of people disagreed is relatively small. As is shown in the results, a majority of students expressed good feelings toward the VTPL approach. This is an indication that in general students were satisfied with the treatment. The number of undecided participants is interesting as it does not indicate satisfaction but confusion or puzzlement at some of the more unusual features of the procedures used.

5.6 Key outcomes of the research

As the results of the above discussions, a number of key outcomes have become apparent. They are summarized below.

First, the classical notion of the single, octave-wide, bandpass filter optimals as originally defined by Guberina was redefined to include narrower and finer digital filtering, thus enriching one of the verbotonal theory's long-standing beliefs and practices. As indicated in this study, an optimal can and should contain a low-frequency or fundamental frequency component as it clearly assists in generating a better-quality optimal (i.e. better-quality perception and production).

Second, although we did not set out to improve the students' performance on prosody, a welcome but unexpected side effect of training for the six English target vowels resulted in a significant improvement in their prosody as evidenced by the significant improvements in comprehensibility and fluency, both of which depend on prosody.

Third, it is important to point out that, in this study, a simple self-managed teacherless training system was developed. The whole system was student-centered rather than teacher-centered. No teaching was performed in any form during the training. In other words, the students were involved in self-managed learning, and the whole process was teacherless. The students had total freedom to learn at their own speed. However, fundamentally, this system depended primarily on the automatic nature of the pronunciation training. Basically, if we feed the brain the right signal, then the brain will do the work for us, without any further intervention.

Fourth, a helpful diagnostic protocol was created to provide guidance on the steps to be taken when conducting a diagnosis of the corrective optimals for the sounds of English for Chinese non-native speakers of English. This protocol could also hopefully be used or adapted to other language sounds and other contexts.

Fifth, the present study indicates that while an improvement in a learner's prosody helps to improve his or her pronunciation (He, 2014; Yang, 2016; Zhang, 2006) an improvement in a learner's pronunciation (as per the VTPL approach) helps to improve his or her prosody, confirming the reciprocity between prosody and pronunciation training in the verbotonal approach.

Sixth, the distinction between sounds and intonation was redefined. Or, to put it another way, the study made explicit the fact that intonation is an integral part of language sounds thus demonstrating that the common distinction between intonation and language sounds is a false distinction created by the current discourse in TESOL (Teaching English to Speakers of Other Languages) that separates language sounds (the linguistic) from intonation/prosody (the paralinguistic). This separatist discourse clearly prioritizes the "linguistic" over the "paralinguistic" as the terms themselves indicate.

Seventh, optimal learning sequences were created and implemented to maximize the pronunciation learning experience under the experimental conditions. Specifically, this refers to the gradual shift from optimal to non-optimal built into this study as per standard verbotonal practice. In the present study, to assist the acceptable production of the target English vowels, several phonetic learning environments were constructed and used throughout the pronunciation learning process, progressing from optimal to non-optimal environments. Specifically, the students practiced the target vowels in an optimal learning environment first, and then they moved to a neutral learning environment and after that, they practiced the vowel sounds in a non-optimal learning environment. After repeated exposure to both the optimal and non-optimal phonetic environments, the students' perceptions of these vowel sounds would be more effectively reorganized.

Eighth, typical use of low frequencies in the verbotonal approach focuses only on the development of acceptable prosodic speech productions. In this perspective, entire sentences are filtered, thus rendering them unintelligible. These unintelligible sentences are then fed to the learners according to protocols established for that purpose (He, 2014; Lian, 1980, 2004, 2014; Yang, 2016). This procedure works well for sensitizing learners to the prosodic patterns of language. However, it works less well when focusing on individual sounds as the phonological information is obliterated. In this study, to make the target English vowels more salient and act more effectively on students' perceptual systems, the vowel sounds and only the vowel sounds in the sentence stream were filtered.

5.7 Summary

This chapter opened with the discussion of corrective optimals and native speaker optimals. Next, it focused on the interpretation of perception performance and production performance. Then, the discussion was devoted to comprehensibility, fluency and pronunciation, followed by the discussion of the students' attitudes toward the VTPL approach. After that, the key outcomes of the research were presented. In the next chapter, the summary of the study, implications, limitations, as well as recommendations for future research, will be presented.

CHAPTER 6

CONCLUSION

The chapter consists of four sections. The first section summarizes the research findings. The second section discusses the implications. The third section is concerned with the limitations of the present study. In the last section, some recommendations for future research are presented.

6.1 Summary of the study

Based on the verbotonal theory of phonetic correction, the current study investigated the corrective optimals of the target English vowels for Chinese non-English major EFL learners and also investigated whether the use of the corrective optimals alone would be sufficient to bring about acceptable production of the target vowels in the selected sample using a self-managed approach.

Participants in this study were 76 first-year non-English major undergraduates form a local university in China. They were randomly assigned to the experimental and control groups. Both groups were pretested for perception and production performances. No significant differences were found between them. The two groups went through exactly the same training procedures and materials with only two differences: (a) the experimental group went through a diagnostic phase to determine each person's corrective optimals and (b) members of the experimental group listened to optimally filtered materials while the members of the control group did not. Otherwise, their activities were identical. Variables were strictly controlled. Hopefully, extraneous/unexpected variables were eliminated. It is important to note that there was no teaching involved in any form other than a short training period to introduce them to the VTPL approach and its activities (no content teaching). All participants were involved in a self-managed learning environment.

The present study was conducted to answer the following five research questions:

(1) What are the corrective optimals of Chinese non-English major EFL learners for the following English vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/?

(2) Are there any differences between the native speaker optimals and the corrective optimals for Chinese non-English major EFL learners? If yes, what are these differences?

(3) Is the VTPL approach effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners? In particular, is simple exposure to the corrective optimals alone sufficient to bring about acceptable production of the target vowels?

(4) Which approach, VTPL or traditional, is more effective for assisting the acceptable production of the target English vowels in Chinese non-English major EFL learners?

(5) What are the Chinese non-English major EFL learners' opinions of the VTPL approach to phonetic correction?

To answer these questions, a mixed methods research design was adopted. The study was quasi-experimental in nature. The quantitative data were collected through the diagnosis of the corrective optimals, questionnaires and tests (i.e. the perception pre- and posttests and the production pre- and posttests). The qualitative data were collected from students' questionnaires, self-reports and semi-structured interviews. The quantitative data collected were analyzed using JASP (Version 0.12.2; JASP Team, 2020) statistical software. Specifically, descriptive statistics were used to explore the frequency distributions of the students' personal corrective optimals concerning the target English vowel sounds. Both paired samples t-tests and independent samples t-tests were conducted to examine the intragroup differences as well as the between-group differences in and between the pretest and posttest. With regard to the qualitative data, a thematic analysis focusing on themes was processed (Braun & Clarke, 2006). A brief summary of the major findings is given below.

First, according to the results of corrective optimals identified in this study, it can be concluded that the students produced the target vowel sounds best when the sounds were filtered through personal corrective optimals consisting of the single bandpass filters enriched by the addition of a lowpass filter (below 320 Hz), i.e. discontinuous multiband filters.

Second, the corrective optimals were found to be narrower and finer in comparison with the native speaker optimals traditionally used in typical verbotonal practice (Guberina, 1972). Besides, the corrective optimals were more diverse and less uniform as compared to the traditional optimal octaves.

Third, the results of the students' perception and production performances indicated that, from pretests to posttests, the experimental group that trained with the VTPL approach showed statistically significant improvements concerning both perception and production of the target vowels. After training, the experimental group's overall speaking proficiency also significantly improved in terms of comprehensibility, fluency and pronunciation. That is, simple exposure to the corrective optimals alone was effective for assisting the acceptable production of the target English vowels in the selected sample.

Fourth, based on the results of the students' perception and production performances, it was clear that, the experimental group outperformed the control group for five out of six vowel sounds on all perception and production tasks. The experimental group also outperformed the control group in terms of comprehensibility, fluency and pronunciation. That is, the VTPL approach was more effective in comparison with the traditional approach not only strictly in the perception and production of the target vowels but more generally too (a welcome surprise and a reflection of the power of the system).

Fifth, as is shown in the results of the present study, a high percentage of participants were positive about learning the target English vowels through the VTPL approach. These results indicated that, in general, the students were satisfied with the treatment. Although some students felt uncertain about the procedures due to their unfamiliarity, the vast majority of students, even the doubters, improved.

Here are some additional important findings emerging from this research.

First, undoubtedly, the most significant finding to emerge from this study is the redefinition or at least refinement of the notion of optimals. Ultimately, this redefinition will impact on native speaker optimals as well. This redefinition changes both some fundamental tenets of verbotonal theory and opens up avenues for new research into corrective optimals which may, in turn, give rise to a complete or at least very large set of research-based corrective optimals for learners of English and other foreign languages in due course.

Second, another significant finding to emerge from the study is the unexpected effectiveness of corrective optimals in helping to improve intonation. This will shed light on intonation training as well as pronunciation training.

Third, it seems that these optimals will be able to be applied in the context of teacherless pronunciation learning, thus enabling students to learn in a fully self-managed environment.

Fourth, the diagnostic protocol created in this study can help identify the corrective optimals for other English vowels and consonants as well, thus assisting the learners' acceptable production of all English sounds.

Fifth, the present study helps to strengthen the idea of the reciprocity between intonation and pronunciation training, thus giving additional evidence for future intonation training and pronunciation training.

Sixth, an important finding to emerge from the study is the redefinition of the distinction between sounds and intonation which will help remove the false distinction created by current discourse in pronunciation learning that separates sounds (the segmentals) and intonation/prosody (the suprasegmentals).

Seventh, optimal sequences designed in this study can help sensitize and reinforce the learners' perceptions of the target English vowel sounds. This will shed light on the design of load-lightening and load-increasing activities concerning awareness-raising and meaning-making in pronunciation learning and teaching.

Eighth, to make the target vowels sounds more salient, the filtering is limited to individual sounds rather than on the whole word or sentence and it works quite well, indicating that filtering can work successfully at the sound level.

6.2 Implications

According to the findings of this study, the following five implications can be extrapolated. They are presented below.

6.2.1 Verbotonal theory

This study enriches verbotonal theory by refining the notion of optimal away from a monolithic one size fits all set of frequency bands controlled, perhaps, by the aspiration to discover or establish universal features specific to each language (i.e. a linguistic model) toward a more nuanced concept where the learner, not the language, is central. If there are commonalities in optimal frequency bands, they are due not so much to the language as to the cultural environments of learners that have resulted in the construction of similar understandings and perceptual systems in individuals. This clearly demonstrates that perception is a property of the listener and not a property of the language. Verbotonalism has always recognized the centrality of the learner during the act of perception but, somehow, the reliance on single, monolithic, optimal frequency bands contradicts this. The present study eliminates that contradiction by redefining and refining the notion of optimal to bring into line with a major principle of verbotonalism that an act of perception is an act of individual, personal, meaning-making.

6.2.2 Precision language education

These findings encourage us to try and extend our understanding of the concept of optimality in phonetics to language education in general. Clearly, the idea of optimization works quite well in terms of sounds, but we still do not know whether it works for other aspects of language learning, such as listening, reading and writing etc. It may be possible to create optimized input conditions there too. To get into this frame of mind, we might ask ourselves questions like: Does it work for something else? What might we optimize, and how do we optimize conditions? These are important questions to ask because we need to identify as precisely as possible what we will do in order to be able to have a maximally effective intervention process, in the spirit of precision language education (Lian & Sangarun, 2017). The precision language education project is necessarily heavily dependent on optimization, even though it may not be expressed in those terms.

With time, and in a precision language education perspective, it will be valuable to develop a streamlined self-diagnostic process based on our work but, further refined to enable fine-tuning of the corrective optimals required by each learner. Such a refinement might be based on developments in artificial intelligence (AI) and may even free the diagnostic process from human intervention, thus reducing our reliance on people, freeing them to do something else and encouraging autonomy in language learners. Some prototype work is already being researched by Andrew Lian (personal communication, 2020) using TensorFlow to produce simple but useful expert systems for pronunciation verification.

6.2.3 Theory of learning

Results have important implications for the theory of learning in terms of meaning-making and awareness-raising. Learning has been traditionally understood as simple knowledge transmission: remembering things. In this sense, classroom teaching was centered on habit formation and on the teacher talking about information. However, in the present study, the promising results demonstrated that we could create optimal conditions requiring no teacher intervention and supervision as opposed to traditional learning and teaching systems where teachers constantly intervene. This means that it is possible for students to learn something correctly without actually being taught by simply being provided with an appropriate learning approach, such as the VTPL approach which acts directly on the learners' perceptual system and induces neuroplastic change.

Importantly, taking into account the effectiveness of the teacherless system, the present study helps open up a new possibility for learners in non-English speaking countries. This is really important because nowadays we have a large and increasing number of people learning English throughout the world (Godwin-Jones, 2018) and it is important to keep learning standards at the highest possible levels and to distribute these across all environments. By involving students in the teacherless system, we are able to lighten the face-to-face loads of teachers by removing some materials and activities from the classroom and putting them directly into the hands of learners.

6.2.4 Teacher education

Some of the implications emerging from the findings relate specifically to teacher education programs, such as the Master of Arts in TESOL. The present study effectively defines a different mindset for dealing with pronunciation difficulties encountered by learners, thus triggering specific interventions designed to target and respond to students' specific pronunciation learning problems. Importantly, the results confirm that people are different because people are in fact affected differently by the various signals to which they are exposed (in this case digital filtering). It also shows that, first and foremost, they are individuals who nevertheless have certain characteristics that are similar to those of others by virtue of their culture, education and other sociolinguistic milieu. These findings, when put together, argue for learners being thought of not as a monolithic, somehow homogenous, block but as individuals with significant differences between them. In order to deal in the most effective way with students' learning problems and difficulties, both in-service teachers and future teachers should be able to understand this phenomenon so as to create a structure where a diagnosis conducted at the individual level is possible as a prelude to further intervention. That is to say, in order to help people achieve successful learning outcomes, teachers should rethink what they teach and how they teach as a function of the individuals that they are teaching and as a result of the new knowledge created in this research. In effect, this means rethinking and redesigning the notion and the role of the syllabus in both classroom and non-classroom settings and perhaps attaching greater value and importance to work done outside the classroom.

In addition, as we enter the 21st century and go further into the information age, with the rapid development of computer science and technology, now, more and more educational institutions and schools are seeking new possibilities to use technology to improve their language teaching quality and students' learning. As Liu (2009) states, learning and teaching foreign languages through technology seems to be a new trend in foreign language education. Similarly, according to Quaglia and Lande (2016), the teacher who has no idea of how to use technology will be replaced by the teacher who does know technology. Under these circumstances, teachers will need to have a good knowledge of technology as well as language pedagogy, and they will need to be able to handle both equally well. This means that potentially a new kind of teacher will become necessary for teaching English and other languages in the new era.

More importantly, though, the theoretical and practical discoveries made together with the clear learning success achieved should create an intellectual obligation to include these findings in teacher education programs, thus challenging some of the current preconceptions of language learning and teaching.

6.2.5 Policymakers and administrators

Last but not least, the study has significant implications for policymakers and administrators for English curriculum reforms in China and around the world. Policymakers and administrators are responsible for making various policies and policy decisions and thus imposing ways of thinking on the teacher population. Policymakers and administrators have a choice. They can play a positive role in pedagogic innovation or they can impede progress. Specifically, they should be clear about how the policies they enact could positively affect the classroom, and what kind of implementation problems these policies might encounter. That is, people who set the syllabus must understand these things and therefore lead the change in syllabus construction. In general, though, policymakers and administrators tend to opt for conservative solutions. The findings of this study have provided empirical support for the idea that pronunciation learning can be enhanced with optimized practice. Further, with regard to curriculum development, both policymakers and administrators should pay attention to pronunciation learning in English curriculum and syllabus design, providing perceptual training for ESL/EFL learners, and using their data to continuously evaluate and improve the learning process. In other words, they should keep abreast of the research world and change their policies according to documented research outcomes.

6.3 Limitations of the study

Although this study yielded some promising results and valuable insights into the verbotonal theory of phonetic correction of the target English vowels for Chinese EFL learners and, by extension, the way the brain and perception work, it is necessary to note some limitations of the present study.

First, due to some practical constraints, it was impossible to conduct the present study using large groups of students. In other words, the sample size of 76 students might not be large enough for fully generalizable results. Furthermore, taking into consideration that the population of this study was representative of only one specific group of students, caution is recommended when generalizing the results to other groups, i.e. students from other universities or the general public etc.

Second, another limitation of this study is that no delayed posttest was given to the participants. It was impossible to examine the pronunciation improvement for these participants over time. Therefore, the present pretest-posttest design may miss some aspects of the retention of the learning effect some time after the experiment.

6.4 Recommendations for future research

In relation to the aforementioned limitations and research findings, the following recommendations may be taken into consideration for future research.

First, as the sample size used in this study was quite small, it is recommended that a replication study with a larger sample size be conducted by collecting more data at various universities and elsewhere to make the estimate more exact and convincing and, more important, to help more learners. Theoretically, this will enrich the verbotonal theory in general and the notion of optimals specifically and will yield a larger and more refined set of corrective optimals of value to the general population of learners.

Second, the present study focused on only six English vowels. Admittedly, this is far from adequate for learning English pronunciation as a whole, and therefore, it is suggested that future researchers should use the notion of optimals to diagnose the optimal filters of sounds (both vowels and consonants) that have been uninvestigated so far in this and previous research.

Third, with regard to future research, a pretest-posttest-delayed posttest design should be conducted on the participants to investigate the comparative retention by learners. This may help identify the change of learners' pronunciation proficiency over time. More important, students' perceptual and production performances at different periods of time may provide more valuable evidence on the training of English pronunciation.

Fourth, it is worth pointing out that the present study is not about sounding British or imposing some kind of colonial model of languages, e.g. native speaker (NS) models. It is about producing sounds which fall within the range of acceptable realizations of sounds for most competent speakers of English. In order to communicate, we cannot just produce any sounds at all. They must conform to some standards of pronunciation in order for them to function as phonemes. Therefore, for future research, it will be necessary to explore the corrective optimals of English sounds in other non-English speaking countries so as to help more non-native speakers of English produce the English vowels and consonants acceptably.

Fifth, in terms of pronunciation learning, it will be helpful and beneficial to include more parameters into the mix of variables used to define optimals. These might include intonation, movements, duration, rhythm, tension, loudness, etc. It will be interesting for future researchers to modify several variables at the same time. For example, they can change the loudness levels between lowpass and bandpass filters (i.e. stronger vs. weaker, weaker vs. stronger).

Last but not least, ideally, it would be interesting and valuable to design some kind of software that works for filtering the sounds in real time. In this context, sounds could be dealt with automatically to cater to learners' different learning needs without any human intervention, thus improving the self-management capabilities of our language-learning systems.

This thesis has attempted to review an original approach to pronunciation learning. Its results and implications provide some potentially valuable outcomes and offer some reason for optimism in fields (both language learning and verbotonalism) whose principles have been relatively unchallenged for some time.

REFERENCES

REFERENCES

Abberton, E., & Fourcin, A. (1997). Electrolaryngography. In M. J. Ball & C. Code (Eds.), *Instrumental Clinical Phonetics* (pp. 119-148). London: Whurr.

Ahn, S. Y. (2016). Exploring language awareness through students' engagement in language play. *Language Awareness, 25*(1-2), 40-54.

Alexandrou, A. M., Saarinen, T., Mäkelä, S., Kujala, J., & Salmelin, R. (2017). The right hemisphere is highlighted in connected natural speech production and perception. *NeuroImage, 152*, 628-638.

Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. A. (2014). *Introduction to research in education* (9th ed.). Belmont: Wadsworth Cengage Learning.

Asher, J. J. (1993). *Learning another language through actions: The complete teacher's guidebook* (4th ed.). Los Gatos: Sky Oaks Productions.

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, *53*(5), 334-341.

Asp, C. W. (1972). *The effectiveness of low-frequency amplification and filtered-speech testing for preschool deaf children.* Knoxville: The University of Tennessee.

Asp, C. W. (2006). *Verbotonal speech treatment*. San Diego: Plural Publishing.

Asp, C. W., & Kline, M. (2012). Timeless verbotonal corrective body movements. In A. Dulčić (Ed.), *Man and speech: scientific and professional monograph of* *the 7th International Symposium of the Verbotonal System* (pp. 102-113). Zagreb: Poliklinika SUVAG.

Asp, C. W., Kline, M., & Koike, K. J. (2012). Verbotonal worldwide. In R. Goldfarb (Ed.), *Translational speech-language pathology and audiology: Essays in honor of Dr. Sadanand Singh* (pp. 319-326). San Diego: Plural Publishing.

Audacity Team. (2016). Audacity (Version 2.1.2) [Computer software]. Retrieved from https://www.audacityteam.org/

Baker, A. (2006). *Ship or sheep? Student's book: An intermediate pronunciation course* (3rd ed.). Cambridge: Cambridge University Press.

Baker, W. (2015). *Culture and identity through English as a lingua franca: Rethinking concepts and goals in intercultural communication*. Berlin: De Gruyter Mouton.

Ball, M. J., & Rahilly, J. (1999). *Phonetics: The science of speech*. New York: Oxford University Press.

Barkhuizen, G., Benson, P., & Chik, A. (2014). *Narrative inquiry in language teaching and learning research*. New York: Routledge.

Baumgarten, N., & House, J. (2010). I think and I don't know in English as lingua franca and native English discourse. *Journal of Pragmatics*, *42*(5), 1184-1200.

Blom, L. A., & Chaplin L. T. (1988). *The moment of movement: Dance improvisation*. Pittsburgh: University of Pittsburgh Press.

Bohn, O.-S., & Munro, M. J. (Eds.). (2007). *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins Publishing Company.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature, 391*,756.

Bourdieu, P. (1990). *The logic of practice*. Palo Alto: Stanford University Press.

Boyer, S. (2002). *Understanding English pronunciation: An integrated practice course*. Glenbrook: Boyer Educational Resources.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brown, A. (1992). Twenty questions. In A. Brown (Ed.), *Approaches to pronunciation teaching* (pp. 1-17). London: Macmillan.

Bunge, M., & Ardila, R. (1987). *Philosophy of psychology*. New York: Springer-Verlag.

Burnham, D. K., & Mattock, K. (2010). Auditory development. In J. G. Bremmer & T. D. Wachs (Eds.), *The Wiley-Blackwell handbook of infant development, Volume 1: Basic research* (2nd ed.) (pp. 81-119). Malden: John Wiley & Sons.

Byrd, D., & Mintz, T. H. (2010). *Discovering speech, words, and mind.* Chichester: Wiley-Blackwell.

Cai, C. (2018). Effects of explicit/implicit instruction and frequency on pronunciation accuracy of Chinese EFL learners. *Modern Foreign Languages, 41*(4), 517-528.

Cai, D., Zhu, L., & Chen, J. (2015). Research on the influence of Zhaotong dialect on English majors' pronunciation. *Journal of Zhaotong University, 3*, 100-103.

Cai, J. (2011). A study of the objective of college English teaching. *Foreign Language World, 1*, 21-29.

Cameron, S. (2018). *Perfecting your English pronunciation* (2nd ed.). New York: McGraw-Hill Education.

Carlet, A., & Cebrian, J. (2015). Identification vs. discrimination training: Learning effects for trained and untrained sounds. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: The University of Glasgow.

Carmen, R. (Ed.). (2010). *Spoken English: Flourish your language*. Chandigarh: Abhishek Publications.

Carter, R. (2003). Key concepts in ELT: Language awareness. *ELT Journal*, *57*(1), 64-65.

Caspers, J. (2010). The influence of erroneous stress position and segmental errors on intelligibility, comprehensibility and foreign accent in Dutch as a second language. In J. van Kampen & R. Nouwen (Eds.), *Linguistics in the Netherlands 2010* (pp. 17-29). Amsterdam: John Benjamins Publishing Company.

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.

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

Chan, A. Y. W. (2011). The perception of English speech sounds by Cantonese ESL learners in Hong Kong. *TESOL Quarterly, 45*(4),718-748.

Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum, 13*(1), 67-85.

Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed.) (pp. 310-324). Cambridge: Cambridge University Press.

Chang, S. (2019). A study on the phonetic learning strategies of the first-year English majors. *Journal of Jinzhou Medical University (Social Science Edition), 17*(2), 108-112.

Chen, C. P. (2011). Career meaning making: An essential in life. In M. McMahon & M. Watson (Eds.), *Career counselling and constructivism: Elaboration of constructs* (pp. 39-55). New York: Nova Science.

Chen, F. (2015). The influence of dialect on acquisition of English pronunciation in northwest Hubei. *Journal of Qiqihar University (Philosophy & Social Science Edition), 5*, 110-113.

Cheng, B., Zhang, Y., & Zhang, X. (2017). Brain mechanisms in phonetic learning and their practical applications in foreign accent reduction. *Foreign Language Education, 38*(4), 62-66.

Chun, C. W. (2015). *Power and meaning making in an EAP classroom: Engaging with the everyday*. Bristol: Multilingual Matters.

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014). Sensory-motor transformations for speech occur bilaterally. *Nature, 507*, 94-98.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge: Cambridge University Press.

Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students* (3rd ed.). London: Routledge.

Cong, L. (2013). The problems in college English pronunciation teaching and their countermeasures. *Journal of Changchun Normal University (Humanities and Social Sciences), 32*(5), 166-167.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks: Sage.

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2015). Second language comprehensibility revisited: Investigating the effects of learner background. *TESOL Quarterly, 49*(4), 814-837.

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal, 30*(1), 13-45.

De Jong, N. H. (2016). Fluency in second language assessment. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 203-218). Berlin: De Gruyter Mouton.

Derwing, T. M. (2003). What do ESL students say about their accents? *The Canadian Modern Language Review, 59*, 547-566.

Derwing, T. M. (2008). Curriculum issues in teaching pronunciation to second language learners. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 347-369). Amsterdam: John Benjamins Publishing Company.

Derwing, T. M. (2017). L2 fluency development. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 246-259). New York: Routledge.

Derwing, T. M., & Munro, M. J. (2001). What speaking rates do non-native listeners prefer? *Applied Linguistics*, *22*(3), 324-337.

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

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

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

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). 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., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*(4), 665-679.

Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, *34*(2),183-193.

Domjan, M. (1998). *The principles of learning and behavior* (4th ed.). Pacific Grove: Brooks/Cole Publishing Company.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah: Lawrence Erlbaum.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York: Routledge.

Edwards, J. G. H., Zampini, M. L., & Cunningham, C. (2019). Listener proficiency and shared background effects on the accentedness, comprehensibility and intelligibility of four varieties of English. *Journal of Monolingual and Bilingual Speech*, *1*(2), 333-356.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Escudero, P. (2007). Second language phonology: The role of perception. In M. C. Pennington (Ed.), *Phonology in context* (pp. 109-134). New York: Palgrave Macmillan.

Eskey, D. (2005). Reading in a second language. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 563-580). Mahwah: Erlbaum.

Félix-Brasdefer, J. C., & Koike, D. A. (Eds.). (2012). *Pragmatic variation in first and second language contexts: Methodological issues.* Amsterdam: John Benjamins Publishing Company.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning, 20*(5), 409-442.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47-65.

Flege, J. E. (1995a). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium: York Press.

Flege, J. E. (1995b). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics, 16*, 425-442.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106*(5), 2973-2987.

Fouz-González, J. (2019). Podcast-based pronunciation training: Enhancing FL learners' perception and production of fossilised segmental features. *ReCALL, 31*(2), 150-169.

Fu, Q. (2018). A research on negative transfer of Xiang dialect in learning English vowels from segmental level. *Journal of Xiangnan University*, *39*(4), 70-74.

Galloway, N. (2017). *Global Englishes and change in English language teaching: Attitudes and impact.* London: Routledge.

Gassin, J. (1990). The notion of synchrony in second language learning. In M. A. K. Halliday, J. Gibbons & H. Nicholas (Eds.), *Learning, keeping and using language: Selected papers from the eighth world congress applied linguistics* (pp. 433-443). Amsterdam: John Benjamins Publishing Company.

Gattegno, C. (1987). *The science of education part 1: theoretical considerations.* New York: Educational Solutions.

Gilakjani, A. P. (2016). English pronunciation instruction: A literature review. *International Journal of Research in English Education, 1*(1), 1-6.

Gilakjani, A. P. (2017). English pronunciation instruction: Views and recommendations. *Journal of Language Teaching and Research, 8*(6), 1249-1255.

Gilakjani, A. P., & Ahmadi, M. R. (2011). Why is pronunciation so difficult to learn? *English Language Teaching, 4*(3), 74-83.

Godwin-Jones, R. (2018). Reflections on global English instruction: New roles and approaches. In W. Lin, I. J. Weng, & R. Godwin-Jones (Eds.), *Internationalizing English language education in globalized Taiwan* (pp. 253-272). Taipei: Tung Hua.

Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. New York: Cambridge University Press.

Goldstein, I. (2014). *Language and ESOL methodology-a unique perspective*. Singapore: Partridge Publishing Singapore.

Golombek, P., & Jordan, S. R. (2005). Becoming "black lambs" not "parrots": A poststructuralist orientation to intelligibility and identity. *TESOL Quarterly, 39*(3), 513-533.

González-Bueno, M. (1997). Voice-onset-time in the perception of foreign accent by native listeners of Spanish. *International Review of Applied Linguistics in Language Teaching, 35*(4), 251-267.

Gowans, S. (2012). *EFL communication strategies in second life: An exploratory study*. Norderstedt: Books on Demand.

Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a foreign language'*. London: British Council.

Guberina, P. (1972). *Case studies in the use of restricted bands of frequencies in auditory rehabilitation of the deaf*. Zagreb: Institute of Phonetics, Faculty of Arts, University of Zagreb.

Guberina, P. (1989). *Verbotonal method in our time*. Retrieved from http://verbotonal.utk.edu/Documents/Guberina's%20Speech.pdf

Guberina, P., & Asp, C. W. (1981). *The Verbo-tonal method for rehabilitating people with communication problems*. New York: World Rehabilitation Fund, International Exchange of Information in Rehabilitation.

Gut, U. (2009). *Introduction to English phonetics and phonology*. Frankfurt am Main: Peter Lang.

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

Hai, M. (2018). Mongolian college students' English pronunciation learning difficulties and countermeasures. *Journal of Chifeng University (Natural Science Edition), 34*(1), 167-168.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Hang, J. (2012). A study on the verbotonal method in Japanese tones teaching. *Journal of Changchun Education Institute, 28*(11), 124-125.

Harris, S. B. (2018). (Ed.). *Effective teaching: Educators perspective of meaning making in higher education*. Charlotte: Information Age Publishing.

Hawkins, E. W. (1999). Foreign language study and language awareness. *Language Awareness, 8*(3-4), 124-142.

He, B. (2014). *Improving the English pronunciation of Chinese EFL learners through the integration of CALL and verbotonalism*. Retrieved from http://sutir.sut.ac.th:8080/sutir/handle/123456789/5370

He, J. (2016). Study of gender differences in pronunciation learning beliefs and strategies of English majors. *Research in Teaching, 39*(1), 18-22.

He, M. (2019). A survey on pronunciation learning strategies of non-English majors and the corresponding countermeasures. *Journal of Jinzhou Medical University (Social Science Edition)*, *17*(4), 77-80.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393-402.

Hillenbrand, J., & Gayvert, R. T. (1993). Vowel classification based on fundamental frequency and formant frequencies. *Journal of Speech and Hearing Research. 36*(4), 694-700.

Hong, T. (2017). Investigation and reflection of college English pronunciation teaching in independent colleges. *Journal of Yanbian Institute of Education, 31*(6), 56-58.

Honorof, D. N., & Whalen, D. H. (2005). Perception of pitch location within a speaker's F0 range. *The Journal of the Acoustical Society of America, 117*(4), 2193-2200.

House, J. (2009). Subjectivity in English as lingua franca discourse: The case of you know. *Intercultural Pragmatics, 6*(2), 171-194.

House, J. (2010). The pragmatics of English as a lingua franca. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 363-390). Berlin: De Gruyter Mouton.

House, J. (2014). Managing academic discourse in English as a lingua franca. *Functions of Language, 21*(1), 50-66.

Hu, X. (2016). On English pronunciation teaching methods for non-English majors. *Journal of Ningbo Institute of Education*, *18*(4), 78-80.

Hultén, A., Karvonen, L., Laine, M., & Salmelin, R. (2014). Producing speech with a newly learned morphosyntax and vocabulary: An magnetoencephalography study. *Journal of Cognitive Neuroscience, 26*, 1721-1735.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 1-16.

Iverson, P., & Evans, B. G. (2007). Auditory training of English vowels for first-language speakers of Spanish and German. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1625-1628). Saarbrücken: Saarland University.

Jakobson, R., & Waugh, L. R. (2002). *The sound shape of language* (3rd ed.). Berlin: De Gruyter Mouton.

Jarosz, A. (2019). *English pronunciation in L2 instruction: The case of secondary school learners*. Cham: Springer Nature.

JASP Team. (2020). JASP (Version 0.12.2) [Computer software]. Retrieved from https://jasp-stats.org/

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*(1),83-103.

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

Jenkins, J. (2006). The spread of EIL: A testing time for testers. *ELT Journal, 60*(1), 42-50.

Jenkins, J. (2009). English as a lingua franca: Interpretations and attitudes. *World Englishes, 28*(2), 200-207.

Jenkins, J. (2012). English as a lingua franca from the classroom to the classroom. *ELT Journal, 66*(4), 486-494.

Jenkins, J. (2015). *Global Englishes: A resource book for students* (3rd ed.). London: Routledge.

Jensen, E. (2005). *Teaching with the brain in mind* (2nd ed.). Alexandria: Association for Supervision and Curriculum Development.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America, 119*(2), 1118-1130.

Johnson, R. (2001). *Elemental movement: A somatic approach to movement education*. Boca Raton: Universal-Publishers.

Jones, D. (2011). *Cambridge English pronouncing dictionary*. Cambridge: Cambridge University Press.

Jordan, A., Carlile, O., & Stack, A. (2008). *Approaches to learning: A guide for teachers*. Maidenhead: Open University Press.

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

Kapurani, E. (2016). Results of implementing communicative language teaching method in English language learning in 9-year elementary schools (Albanian context). *European Journal of Language and Literature Studies*, *4*(1), 56-64.

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

Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness, 19*(3), 171-185.

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

Koike, K. J. (2012). *Optimal filter perception of speech sounds: Implications to hearing aid fitting through verbotonal rehabilitation* [PowerPoint slides]. Retrieved from https://fdocuments.in/document/optimal-filter-perception-of-speech-sounds-phonemic-audiogram-adjustment.html

Kovač, M. M., & Vickov, G. (2019). The influence of pre-task planning on speed and breakdown fluency. *Journal for Foreign Languages, 11*(1), 87-102.

Kramsch, C. (1998). The privilege of the intercultural speaker. In M. Byram & M. Fleming (Eds.), *Language learning in intercultural perspective: Approaches through drama and ethnography* (pp. 16-31). Cambridge: Cambridge University Press.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. London: Routledge.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 11850-11857.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831-843.

Lai, Y. H. (2010). English vowel discrimination and assimilation by Chinese-speaking learners of English. *Concentric: Studies in Linguistics, 36*(2), 157-182.

Lapteva, O. (2011). *Speaker perception and recognition. An integrative framework for computational speech processing*. Kassel: Kassel University Press.

Laroy, C. (1995). *Pronunciation*. Oxford: Oxford University Press.

Lee, A. H., & Lyster, R. (2017). Can corrective feedback on second language speech perception errors affect production accuracy? *Applied Psycholinguistics, 38*(2), 371-393.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology, 46*(6), 1093-1096.

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, TESOL and applied linguistics: Challenges for theory and practice* (pp. 245-270). New York: Palgrave Macmillan.

Levis, J. M. (2016). The interaction of research and pedagogy. *Journal of Second Language Pronunciation, 2*(1), 1-7.

Levis, J. M., & Sonsaat, S. (2019). Quality teacher education for pronunciation teaching in L2 classrooms. In J. D. D. M. Agudo (Ed.), *Quality in TESOL and teacher education: From a results culture towards a quality culture* (pp. 213-222). Oxford: Routledge.

Li, X. (2015). The influence of dialect on pronunciation acquisition of English learners. *Journal of Lanzhou Jiaotong University*, *2*, 82-84.

Li, Z. (2014). Visualization of pronunciation testing and evaluation. *Journal of Sichuan International Studies University*, *30*(6), 98-103.

Lian, A.-P. (1980). *Intonation patterns of French (Teacher's book)*. Melbourne: River Seine Publications.

Lian, A.-P. (1987). Awareness, autonomy and achievement in foreign language learning. *Revue de Phonétique Appliquée*, *82-84*, 167-184.

Lian, A.-P. (1993). Awareness, autonomy and achievement. In A-P. Lian, D. L. Hoven, and T. J. Hudson (Eds.), *Audio-video computer enhanced language learning and the development of listening comprehension skills* (pp. 25-41). Brisbane:

An Australian Second Language Learning Project, Commonwealth of Australia.

Lian, A.-P. (2000). From first principles: Constructing language learning and teaching environments. In M.-S. Lin (Ed.), *Selected papers from the ninth international symposium on English teaching* (pp. 49-62). Taipei: Crane Publishing.

Lian, A.-P. (2004). Technology-enhanced language learning environments: A rhizomatic approach. In J.-B. Son (Ed.), *Computer-assisted language learning: Concepts, contexts and practices* (pp. 1-20). New York: iUniverse.

Lian, A.-P. (2011). Reflections on language learning in the 21st century: The rhizome at work. *Rangsit Journal of Arts and Sciences, 1*(1), 3-15.

Lian, A.-P. (2013). *Plenary session: Awareness-raising in a technology-supported language-learning environment: Concepts, challenges and solutions.* Keynote speech presented at The Globalization and Localization in Computer-Assisted Language Learning (GLoCALL) 2013 Conference, The University of Danang, Da Nang, Vietnam.

Lian, A.-P. (2014). On-demand generation of individualised language learning lessons. *Journal of Science, Ho Chi Minh City Open University, 1*(9), 25-38.

Lian, A.-P., & Lian, A. B. (1997). The secret of the Shao-Lin monk: Contribution to an intellectual framework for language-learning. *On-CALL, 11*(2), 2-18.

Lian, A.-P., & Pineda, M. V. (2014). Rhizomatic learning: “As… when… and if…” A strategy for the ASEAN community in the 21st century. *Beyond Words, 2*(1), 1-28.

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.

Lian, A.-P., & Sussex, R. D. (2018). Toward a critical epistemology for learning languages and cultures in 21st century Asia. In A. Curtis & R. D. Sussex (Eds.), *Intercultural communication in Asia: Education, language and values* (pp. 37-54). Cham: Springer International Publishing.

Lin, X. (2017). Comparative analysis of English and Chinese phonetics and its teaching implications. *Overseas English*, *12*, 216-217.

Lin, Y. (2019). Research on English pronunciation teaching reform under flipped classroom model – a case study of Chongqing University of Arts and Sciences. *College English Teaching and Research*, *1*, 68-73.

Lippi-Green, R. (1997). *English with an accent: Language, ideology and discrimination in the United States*. London: Routledge.

Little, D. (1997). Language awareness and the autonomous language learner. *Language Awareness*, *6*(2-3), 93-104.

Liu, J. (2009). A survey of EFL learners' attitudes toward information and communication technologies. *English Language Teaching*. *2*(4), 101-106.

Liu, Q., & Hu, Y. P. (2010). Needs analysis of oral English teaching for Chinese science majors. *Journal of Xi'an International Studies University, 18*(1), 86-89.

Liu, S. (2016). A practical study on the English pronunciation and intonation contest in college English pronunciation teaching. *Foreign Language Teaching*, *3*, 49-54.

Liu, S., & Niu, Z. (2018). An empirical research on the improved multidimensional pronunciation teaching assessment model. *Foreign Language Teaching*, *4*, 62-68.

Logan, J. S., & Pruitt, J. S. (1995). Methodological issues in training listeners to perceive non-native phonemes. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross language research* (351-377). Timonium: York Press.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*(3), 251-285.

Long, M. H. (2007). *Problems in SLA*. Mahwah: Lawrence Erlbaum Associates.

Low, E. L. (2015). *Pronunciation for English as an international language: From research to practice*. London: Routledge.

Lucas, M. W., & Yiakoumetti, A. (2019). Cross-linguistic awareness-raising practices can enhance written performance in EFL classes in Japanese universities. *International Journal of Bilingual Education and Bilingualism*, *22*(8), 958-973.

Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer Assisted Language Learning*, *29*(3), 451-476.

Ma, L. (2019). On the negative transfer of Chinese northern dialect to English phonetics and the teaching strategies. *Foreign Language Education & Research*, *7*(2), 11-18.

Mason, J. (1998). Enabling teachers to be real teachers: Necessary levels of awareness and structure of attention. *Journal of Mathematics Teacher Education*, *1*(3), 243-267.

Matsuda, A. (Ed.). (2012). *Principles and practices of teaching English as an international language*. Bristol: Multilingual Matters.

McArthur, T. (2002). *Oxford guide to world English*. Oxford: Oxford University Press.

McGilchrist, I. (2009). *The master and his emissary: The divided brain and the making of the western world*. New Haven: Yale University Press.

Meador, D., Flege, J. E., & MacKay, I. R. A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition, 3*(1), 55-67.

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

Mildner, V., & Bakran, J. (2001). Acoustic correlates of phonetic correction. *Clinical Linguistics & Phonetics, 15*(1-2), 151-155.

Mildner, V., & Tomić, D. (2007). Effects of phonetic speech training on the pronunciation of vowels in a foreign language. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1665-1668). Saarbrücken: Saarland University.

Miner, R., & Danhauer, J. L. (1977). Relationship between formant frequencies and optimal octaves in vowel perception. *Journal of the American Audiology Society, 2*(5), 163-168.

MOE (Ministry of Education). (2000). *Teaching syllabus for English majors*. Beijing: Foreign Language Teaching and Research Press.

MOE (Ministry of Education). (2003). *English curriculum standards for senior high schools*. Beijing: People's Education Press.

MOE (Ministry of Education). (2007). *College English curriculum requirements*. Beijing: Foreign Language Teaching and Research Press.

MOE (Ministry of Education). (2011). *English curriculum standards for compulsory education*. Beijing: Beijing Normal University Publishing Group.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, *25*(3), 481-520.

Mortimer, E., & Scott, P. (2003). *Meaning making in secondary science classrooms*. Maidenhead: Open University Press.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation and instruction. *Studies in Second Language Acquisition*, *21*(1), 81-108.

Muijs, D. (2004). *Doing quantitative research in education with SPSS*. London: Sage.

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

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289-306.

Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 377-396). Chichester: Wiley Blackwell.

Murphy, J. M. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 25*(1),51-75.

No, G. (2003). Intelligibility and English pronunciation in elementary schools. *English Teaching, 58*(1), 201-218.

Northrup, D. (2013). *How English became the global language*. New York: Palgrave Macmillan.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Ohata, K. (2004). Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL Journal, 6*(4), 1-19.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*(3), 261-285.

Ozawa, S. (2019). Effects of Japanese university students' characteristics on the use of an online English course and TOEIC scores. *CALICO Journal, 36*(3), 225-239.

Pan, L. (2015). *English as a global language in China: Deconstructing the ideological discourses of English in language education*. Cham: Springer International Publishing.

Pawlak, M., Waniek-Klimczak, E., & Majer, J. (Eds.). (2011). *Speaking and instructed foreign language acquisition*. Bristol: Multilingual Matters.

Peng, N. (2014). A study on the differences in the use of pronunciation learning strategies between English and non-English majors. *Foreign Language Research, 2*, 105-110.

Peng, N., & Wang, W. (2014). A study of non-English majors' pronunciation learning strategies. *Foreign Language and Literature, 30*(4), 152-156.

Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives.* London: Palgrave Macmillan.

Phillips, P. P., & Stawarski, C. A. (2008). *Data collection: Planning for and collecting all types of data*. San Francisco: Pfeiffer.

Pleuger, J. (2001). *How to teach modern languages and survive!* Clevedon: Multilingual Matters.

Price, C. J., Wise, R. J. S., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., Frackowiak, R. S. J., & Friston, K. J. (1996). Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain, 119*(3), 919-931.

Qian, H. (2015). A comparative study of English and Chinese phonetics and of English phonetic teaching from the perspective of language transfer theory. *Journal of Taizhou University, 37*(1), 20-23.

Qiang, N., & Wolff, M. (Eds.). (2011). *The lowdown on China's higher education*. Newcastle upon Tyne: Cambridge Scholars Publishing.

QQ Team. (2016). QQ (Version 8.6) [Computer software]. Retrieved from https://im.qq.com/

Quaglia, R. J., & Lande, L. L. (2016). *Teacher voice: Amplifying success.* Thousand Oaks: Corwin.

Rao, S. N. (2002). *Educational psychology*. New Delhi: New Age International (P) Limited.

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

Renard, R. (1975). *Introduction to the verbo-tonal method of phonetic correction* (B. Morris, Trans.). Paris: Didier.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Timonium: York Press.

Rokita-Jaskow, J. (2008). Training foreign language teachers of young learners in Poland: Principles and problems. In I. R. Gay, A. J. M. Guijarro & J. I. A. Hernandez (Eds.), *New trends in English teacher education* (pp. 213-222). Castilla-La Mancha: Universidad de Castilla-La Mancha.

Rose, H., & Galloway, N. (2019). *Global Englishes for language teaching*. Cambridge: Cambridge University Press.

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review, 65*(3), 395-412.

Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research, 2*, 49-60.

Rozin, P., Moscovitch, M., & Imada, S. (2016). Right: Left: East: West. Evidence that individuals from East Asian and South Asian cultures emphasize right hemisphere functions in comparison to Euro-American cultures. *Neuropsychologia*, *90*, 3-11.

Rvachew, S., & Brosseau-Lapré, F. (2012). *Developmental phonological disorders: Foundations of clinical practice*. San Diego: Plural Publishing.

Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *36*, 1-25.

Sandhaas, B. (1989). Models, methods and basic elements of intercultural learning-an educational approach. In P. Funke (Ed.), *Understanding the USA: A cross-cultural perspective* (81-109). Tübingen: Gunter Narr Verlag.

Sangarun, P. (2014). Developing reading comprehension skills in another language: A computer-based 21st century prototype. *Beyond Words*, *2*(2), 121-153.

Schiffler, L. (1992). *Suggestopedic methods and applications*. Philadelphia: Gordon and Breach Science Publishers.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*(4), 357-385.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206-226.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and* *awareness in foreign language learning* (pp. 1-63). Honolulu: Second Language Teaching & Curriculum Center, University of Hawaii at Manoa.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of the Fourth International Conference CLaSIC 2010* (pp. 721-737). Singapore: Centre for Language Studies, Faculty of Arts and Social Sciences, National University of Singapore.

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In W. M. Chan, K. N. Chin, S. Bhatt, & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27-50). Boston: De Gruyter Mouton.

Schwienhorst, K. (2008). *Learner autonomy and CALL environments*. New York: Routledge.

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

Seidlhofer, B. (2001). Pronunciation. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 56-65). Cambridge: Cambridge University Press.

Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford: Oxford University Press.

Sewell, A. (2016). *English pronunciation models in a globalized world: Accent, acceptability and Hong Kong English*. New York: Routledge.

Shang, C. (2016). Using Praat for pronunciation instruction. *Journal of Mudanjiang University*, *25*(4), 148-151.

Sheets-Johnstone, M. (2011). *The primacy of movement* (2nd ed.). Amsterdam: John Benjamins Publishing Company.

Sheikh, A. A., & Sheikh, K. S. (Eds.). (1996). *Healing east and west: Ancient wisdom and modern psychology*. New York: Wiley.

Shi, M. (2014). A needs analysis survey of the college English pronunciation teaching. *Journal of Chongqing University of Science and Technology (Social Sciences Edition)*, *9*, 144-146.

Shinohara, Y., & Iverson, P. (2018). High variability identification and discrimination training for Japanese speakers learning English /r/-/l/. *Journal of Phonetics*, *66*, 242-251.

Shintani, N., Saito, K., & Koizumi, R. (2019). The relationship between multilingual raters' language background and their perceptions of accentedness and comprehensibility of second language speech. *International Journal of Bilingual Education and Bilingualism*, *22*(7), 849-869.

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), E4687-E4696.

Silverstein, A., Silverstein, V. B., & Nunn, L. S. (2001). *Hearing*. Brookfield: Twenty-First Century Books.

Simonyan, K., & Fuertinger, S. (2015). Speech networks at rest and in action: Interactions between functional brain networks controlling speech production. *Journal of Neurophysiology, 113*(7), 2967-2978.

Song, X. (2015). Review and reflections on the methods of teaching English sounds. *Journal of Weinan Normal University, 16*, 58-62.

Springer, S. P., & Deutsch. G. (1997). *Left brain, right brain: Perspectives from cognitive neuroscience* (5th ed.). New York: Freeman.

Stern, H. H. (1992). *Issues and options in language teaching*. Oxford: Oxford University Press.

Sung, C. C. M. (2016). Does accent matter? Investigating the relationship between accent and identity in English as a lingua franca communication. *System*, *60*, 55-65.

Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, *42*(1), 143-167.

Swan, M., & Smith, B. (Eds.). (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge: Cambridge University Press.

Swann, P. (1992). Computer assisted language learning for English as a foreign language. *Computers & Education, 19*(3), 251-266.

Szpyra-Kozłowska, J. (2015). *Pronunciation in EFL instruction: A research-based approach*. Bristol: Multilingual Matters.

Tang, J. (2020). Design of flipped classroom hybrid teaching mode of English pronunciation under MOOC. *Journal of Bengbu University*, *9*(2), 109-112.

Tang, Y., & Ge, R. (2019). A study on the perception and production of English front vowels by Changchun English learners. *Journal of Educational Institute of Jilin Province*, *35*(11), 150-155.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks: Sage.

Tavakoli, P. (2019). Automaticity, fluency and second language task performance. In Z. Wen & M. J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 39-52). Amsterdam: John Benjamins Publishing Company.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: Sage.

Thir, V. (2016). Rethinking pronunciation teaching in teacher education from an ELF perspective. *Vienna English Working Papers*, *25*, 1-29.

Thomson, C. K. (1996). Self-assessment in self-directed learning: Issues of learner diversity. In R. Pemberton, E. S. L. Li, W. W. F. Or & H. D. Pierson (Eds.), *Taking control: Autonomy in language learning* (pp. 77-92). Hong Kong: Hong Kong University Press.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal*, *28*(3), 744-765.

Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ELT Journal*, *56*(3), 240-249.

Tomić, D., Kiš, A., & Mildner, V. (2011). Chinese L1-Croatian L2 – difficulties and success. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 2006-2009). Hong Kong: Department of Chinese, Translation & Linguistics, City University of Hong Kong.

Trazo, S. P., & Abocejo, F. T. (2019). International phonetic alphabet (IPA) front vowel sound recognition of beginner foreign learners. *European Journal of Education Studies*, *5*(12), 183-196.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*(1).1-30.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition, 15*(4), 905-916.

Trubetzkoy, N. S. (1969). *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley: University of California Press. (Original work published 1939).

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 80-91.

Valeo, A. (2013). Language awareness in a content-based language programme. *Language Awareness, 22*(2), 126-145.

Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. London: Longman.

Wagener, D. (2006). Promoting independent learning skills using video on digital language laboratories. *Computer Assisted Language Learning, 19*(4-5), 279-286.

Wahiduzzaman, S. M. (2017). Effect of good English pronunciation in art of speaking during oral communication: a study among tertiary level students in Bangladesh. *MIST Journal of Science and Technology, 5*(1), 51-57.

Walker, R. (2010). *Teaching the pronunciation of English as a lingua franca*. Oxford: Oxford University Press.

Wang, H., & Heuven, V. J. van. (2004). Cross-linguistic confusion of vowels produced and perceived by Chinese, Dutch and American speakers of English. In L. Cornips & J. Doetjes (Eds.), *Linguistics in the Netherlands 2004* (pp. 205-216), Amsterdam: John Benjamins Publishing Company.

Wang, X. (2018). Design of error speech automatic recognition system for English pronunciation. *Modern Electronics Technique, 41*(10), 179-182

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*(4), 539-552.

Wang, Y., & Shi, X. (2015). A phonetic teaching mode of college English and its practice. *Journal of Tangshan Teachers College*, *37*(6), 158-160.

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.

Wen, B., & Chen, Z. (2019). A study of English & Chinese phonetic contrastive pedagogical model and its feasibility – based on contrast of vowels. *Journal of Shantou University (Humanities & Social Sciences Edition), 35*(1), 46-55.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535.

White, J. L., & Horst, M. (2012). Cognate awareness-raising in late childhood: Teachable and useful. *Language Awareness, 21*(1-2), 181-196.

Wiltsher, C. (2005). Fundamentals of adult learning. In J. P. Wilson (Ed.), *Human resource development: Learning & training for individuals & organizations* (2nd ed.) (pp. 201-219). London: Kogan Page.

Wong, J. W. S. (2015). The impact of L2 proficiency on vowel training. In J. A. Mompean & J. Fouz-González (Eds.), *Investigating English pronunciation: Trends and directions* (pp. 219-239). Basingstoke: Palgrave Macmillan.

Wu, Y. (2013). Application research of the verbotonal method in Hunan Japanese learners – based on the perspective of second language acquisition. *Journal of Heihe University, 4*(5), 69-72.

Wu, Z. (2020). Negative transfer of mother tongue on college students' study of English phonetics and corresponding teaching strategies. *Education Teaching Forum, 6*, 224-225.

Xia, N. (2019). A phonological investigation of English pronunciation teaching. *Journal of Kaifeng Institute of Education, 39*(1), 65-67.

Xiang, Q. (2019). Investigation and analysis of English phonetics teaching in junior middle schools. *English Teachers, 19*(15), 38-42

Xu, Y., & Zeng, Y. (2017). A study of pronunciation features and score predictors for the read-aloud task. *Journal of Xi'an International Studies University, 25*(2), 69-74.

Yang, Y. (2016). *Improving the phonological working memory and English speaking skills of Chinese primary EFL learners with a verbotonal-based approach.* Retrieved from http://sutir.sut.ac.th:8080/sutir/handle/123456789/6753

Zacharias, N. T., & Manara, C. (Eds.). (2013). *Contextualizing the pedagogy of English as an international language: Issues and tensions*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Zenotz, V. (2012). Awareness development for online reading. *Language Awareness, 21*(1-2), 85-100.

Zhang, C. (2013). A study on the phonetic problem in college English teaching for non-English majors. *Journal of Jimei University (Education Science Edition), 14*(1), 122-124.

Zhang, D. (2012). Chinese primary school English curriculum reform. In J. Ruan & C. B. Leung (Eds.), *Perspectives on teaching and learning English literacy in China* (pp. 67-83). Dordrecht: Springer Science & Business Media.

Zhang, F. (2006). *The teaching of Mandarin prosody: A somatically-enhanced approach for second language learners* (Unpublished doctoral dissertation). University of Canberra, Canberra, Australia.

Zhang, F., & Yin, P. (2009). A study of pronunciation problems of English learners in China. *Asian Social Science*, *5*(6), 141-146.

Zhang, J. (2013). Perception of RP English monophthongs by non-English majors. *Journal of Qiqihar University (Philosophy & Social Science Edition)*, *1*, 138-141.

Zhang, L., & Ma, J. (2015). Research on innovative method for English pronunciation teaching based on network platform. *Journal of Luoyang Institute of Science and Technology (Social Science Edition)*, *30*(4), 88-91.

Zhang, W., & Wang, D. (2017). The effect of pronunciation correction on English majors' production of English vowels. *Journal Beijing International Studies University*, *4*, 56-69.

Zhang, Y. (2019). On the reform of English phonetics teaching in colleges and universities from the perspective of constructivism. *Education Teaching Forum*, *2*, 139-140.

Zhao, Y. (2019). Several theoretical issues of second language phonology and phonological acquisition. *Foreign Language Research*, *6*, 61-70.

Zheng, T., & Liu, Q. (2018). Influence of negative transfer of mother tongue on Chinese English learners' pronunciation. *Theory and Practice in Language Studies*, *8*(11), 1478-1484.

Zhi, N., & Li, A. (2020). Phonetic training based on visualized articulatory model. *Journal of Foreign Languages*, *43*(1), 59-74.

Zhou, W., & Song, H. (2015). A thumbnail sketch of Chinese college students' productive phonetic competence in English: An acoustic perspective. *Foreign Languages and Their Teaching*, *3*, 1-7.

APPENDICES

APPENDIX A

Pronunciation Learning Questionnaire

(English and Chinese)

Dear Student,

This questionnaire is designed to gather information about how you, as a Chinese non-English major EFL learner, go about the learning of English pronunciation. Data collected in this survey will help provide valuable information for improving pronunciation instruction for Chinese EFL learners. Please take a few minutes to complete the survey below. Your time and help will be greatly appreciated. Your information will be kept confidential. Thank you very much for your cooperation.

I. General Information

Name________________

Gender ☐ male ☐ female

Major________________

Age____________

Place of birth________________

Do you belong to any minority group? If yes, please specify _________________

Do you speak any Chinese dialects other than Mandarin Chinese?

If yes, please specify ______________

College entrance exam score for English ______________

II. Pronunciation Learning Information

1. How long have you studied English as a foreign language? _______ years

2. Have you been taught English pronunciation explicitly by English teachers in school?

   ☐ Yes ☐ No

   If no, please specify how you have learned to pronounce English. ______________

   If yes, from whom? Check all that apply.

   ☐ primary school English teacher

   ☐ junior middle school English teacher

   ☐ senior middle school English teacher

   ☐ college English teacher

   What? Check all that apply.

   ☐ consonants ☐ vowels ☐ stress ☐ intonation ☐ rhythm

   others (please specify) _____________

   How? Check all that apply.

   ☐ articulatory explanation

   ☐ listen and repeat after the model

   ☐ minimal pair drills

   others (please specify) _____________

3. Are you confident in English pronunciation?

   ☐ very confident ☐ confident ☐ somewhat confident ☐ not confident

☐ not confident at all

4. Are you confident in speaking English?

☐ very confident ☐ confident ☐ somewhat confident ☐ not confident

☐ not confident at all

5. Are you motivated to learn English pronunciation?

☐ strongly motivated ☐ motivated ☐ somewhat motivated

☐ not motivated ☐ not motivated at all

6. Have you ever been misunderstood because of your English pronunciation?

☐ Yes ☐ No

If yes, please specify. ____________

7. Do you want to improve your English pronunciation?

☐ Yes ☐ No

If yes, what? Check all that apply.

☐ consonants ☐ vowels ☐ stress ☐ intonation ☐ rhythm

others (please specify) ____________

语音学习调查问卷

亲爱的同学，你好！

本调查问卷旨在了解非英语专业大学生的英语语音学习情况。你所提供的信息对于提高大学英语语音教学质量非常宝贵。请根据自己真实的情况和意见填写，回答结果保密。谢谢你的支持！

I. 个人信息

姓名______________

性别 ☐ 男 ☐ 女

专业______________

年龄______________

出生地______________

你是否是少数民族？如是，请填写 ______________________

除普通话外，你是否使用方言？如是，请填写 ______________________

高考英语成绩_____________

II. 语音学习信息

1. 你学习英语多长时间了？ _______ 年

2. 学校老师是否教授过英语语音知识？

☐ 是 ☐ 否

如否，请说明你是如何学习英语语音的。______________________

如是，谁教的？请在合适的选项前打勾。

☐ 小学英语老师

☐ 初中英语老师

☐ 高中英语老师

☐ 大学英语老师

教了什么？请在合适的选项前打勾。

☐ 辅音 ☐ 元音 ☐ 重音 ☐ 语调 ☐ 节奏

其它（请予以说明）______________

怎么教的？请在合适的选项前打勾。

☐ 发音部位讲解 ☐ 听音模仿 ☐ 最小配对练习

其它（请予以说明）______________

3. 你对自己的英语语音是否感到自信？

☐ 非常自信 ☐ 自信 ☐ 有点自信 ☐ 不太自信 ☐ 完全不自信

4. 你对自己的英语口语是否感到自信？

☐ 非常自信 ☐ 自信 ☐ 有点自信 ☐ 不太自信 ☐ 完全不自信

5. 你是否有学习英语语音的动机？

☐ 比较强烈 ☐ 一直 ☐ 有时 ☐ 很少 ☐ 完全没有

6. 你的英语语音是否造成过他人误解？

☐ 是 ☐ 否

如是，请予以说明。__________________

7. 你是否想提高自己的英语语音能力?

☐ 是　　☐ 否

如是，在哪方面?请在合适的选项前打勾。

☐ 辅音　☐ 元音　☐ 重音　☐ 语调　☐ 节奏

其它（请予以说明）____________________

APPENDIX B

Perception Pre/Posttest

Instructions: You will hear 90 items. In each item you will hear two words. Sometimes the two words are the same. Sometimes they have one sound that is different. You will listen once only to each item. Please circle the correct answer on the answer sheet provided.

1	a. bin-bean c. bean-bean	b. bean-bin d. bin-bin	2	a. beat-bit c. bit-beat	b. bit-bit d. beat-beat
3	a. din-din c. dean-dean	b. din-dean d. dean-din	4	a. bed-bad c. bad-bad	b. bed-bed d. bad-bed
5	a. marry-marry c. merry-marry	b. marry-merry d. merry-merry	6	a. met-met c. mat-mat	b. met-mat d. mat-met
7	a. cooed-cooed c. cooed-could	b. could-could d. could-cooed	8	a. wooed-wooed c. wood-wooed	b. wooed-wood d. wood- wood
9	a. pooled-pooled c. pulled-pulled	b. pooled-pulled d. pulled-pooled	10	a. sheep-sheep c. sheep-ship	b. ship-sheep d. ship-ship
11	a. seat-seat c. seat-sit	b. sit-sit d. sit-seat	12	a. pit-pit c. pit-peat	b. peat-peat d. peat-pit
13	a. man-men c. man-man	b. men-man d. men-men	14	a. gas-gas c. guess-guess	b. gas-guess d. guess-gas
15	a. lend-land c. lend-lend	b. land-land d. land-lend	16	a. fool-full c. full-full	b. fool-fool d. full-fool
17	a. pulled-pulled c. pooled-pooled	b. pulled-pooled d. pooled-pulled	18	a. soot-soot c. suit- suit	b. soot-suit d. suit-soot
19	a. seat-sit c. seat-seat	b. sit-sit d. sit-seat	20	a. ship-sheep c. sheep-sheep	b. ship-ship d. sheep-ship
21	a. lick-lick c. leak-leak	b. leak-lick d. lick-leak	22	a. and-and c. end-and	b. and-end d. end-end
23	a. met-met c. mat-met	b. mat-mat d. met-mat	24	a. guess-guess c. guess-gas	b. gas-gas d. gas-guess
25	a. kook-kook c. cook-cook	b. kook-cook d. cook-kook	26	a. soot-soot c. soot-suit	b. suit-suit d. suit-soot
27	a. wood-wood c. wooed-wooed	b. wooed-wood d. wood-wooed	28	a. chip-cheap c. cheap-cheap	b. chip-chip d. cheap-chip
29	a. lip-lip c. leap-leap	b. lip-leap d. leap-lip	30	a. sin-sin c. sin-seen	b. seen-seen d. seen-sin
31	a. beg-bag c. bag-bag	b. beg-beg d. bag-beg	32	a. land-land c. lend-lend	b. land-lend d. lend-land
33	a. marry-marry c. marry-merry	b. merry-merry d. merry-marry	34	a. kooky-kooky c. kooky-cookie	b. cookie-kooky d. cookie-cookie
35	a. cook-cook c. kook-kook	b. cook-kook d. kook-cook	36	a. hood-hood c. hood-who'd	b. who'd-who'd d. who'd-hood
37	a. green-grin c. grin-grin	b. green-green d. grin-green	38	a. sick-sick c. sick-seek	b. seek-seek d. seek-sick

39	a. seat-sit c. seat-seat	b. sit-seat d. sit-sit	40	a. had-had c. had-head	b. head-head d. head-had
41	a. shall-shall c. shell-shall	b. shall-shell d. shell-shell	42	a. vat-vat c. vet-vet	b. vet-vat d. vat-vet
43	a. pull-pull c. pool-pool	b. pull-pool d. pool-pull	44	a. hood-hood c. who'd-who'd	b. hood-who'd d. who'd-hood
45	a. pulling-pooling c. pulling-pulling	b. pooling-pooling d. pooling-pulling	46	a. heat-hit c. hit-heat	b. hit-hit d. heat-heat
47	a. peak-peak c. pick-pick	b. peak-pick d. pick-peak	48	a. knit-knit c. neat-neat	b. knit-neat d. neat-knit
49	a. pen-pan c. pan-pan	b. pen-pen d. pan-pen	50	a. vat-vat c. vet-vet	b. vat-vet d. vet-vat
51	a. axe-axe c. X-axe	b. X-X d. axe-X	52	a. soot-soot c. soot-suit	b. suit-suit d. suit-soot
53	a. cooed-cooed c. could-could	b. cooed-could d. could-cooed	54	a. fool-fool c. full-fool	b. full-full d. fool-full
55	a. did-did c. did-deed	b. deed-deed d. deed-did	56	a. teen-teen c. tin-tin	b. teen-tin d. tin-teen
57	a. sick-sick c. seek-seek	b. sick-seek d. seek-sick	58	a. pet-pat c. pet-pet	b. pat-pat d. pat-pet
59	a. jam-jam c. gem-gem	b. jam-gem d. gem-jam	60	a. shall-shall c. shall-shell	b. shell-shell d. shell-shall
61	a. pulling-pulling c. pooling-pooling	b. pulling-pooling d. pooling-pulling	62	a. kooky-kooky c. kooky-cookie	b. cookie-cookie d. cookie-kooky
63	a. kook-kook c. kook-cook	b. cook-cook d. cook-kook	64	a. feel-feel c. fill-fill	b. feel-fill d. fill-feel
65	a. cheap-cheap c. chip-chip	b. cheap-chip d. chip-cheap	66	a. sheep-sheep c. sheep-ship	b. ship-ship d. ship-sheep
67	a. said-said c. said-sad	b. sad-sad d. sad-said	68	a. axe-axe c. X-X	b. axe-X d. X-axe
69	a. gem-gem c. jam-jam	b. gem-jam d. jam-gem	70	a. could-could c. could-cooed	b. cooed-cooed d. cooed-could
71	a. pull-pull c. pool-pull	b. pool-pool d. pull-pool	72	a. cooed-cooed c. cooed-could	b. could-could d. could-cooed
73	a. eat-eat c. it-it	b. it-eat d. eat-it	74	a. feel-feel c. feel-fill	b. fill-fill d. fill-feel
75	a. pick-pick c. pick-peak	b. peak-pick d. peak-peak	76	a. set-set c. set-sat	b. sat-sat d. sat-set
77	a. hem-hem c. ham-hem	b. ham-ham d. hem-ham	78	a. bed-bed c. bad-bad	b. bed-bad d. bad-bed
79	a. wood-wooed c. wooed-wooed	b. wood-wood d. wooed-wood	80	a. pooling-pooling c. pooling-pulling	b. pulling-pulling d. pulling-pooling
81	a. kooky-kooky c. kooky-cookie	b. cookie-cookie d. cookie-kooky	82	a. fit-fit c. feet-feet	b. fit-feet d. feet-fit
83	a. neat-neat c. knit-neat	b. knit-knit d. neat-knit	84	a. feet-feet c. feet-fit	b. fit-fit d. fit-feet
85	a. bet-bet c. bet-bat	b. bat-bat d. bat-bet	86	a. bad-bad c. bed-bed	b. bad-bed d. bed-bad
87	a. hem-hem c. ham-hem	b. hem-ham d. ham-ham	88	a. hood-hood c. hood-who'd	b. who'd-who'd d. who'd-hood
89	a. full-full c. full-fool	b. fool-fool d. fool-full	90	a. pull-pull c. pull-pool	b. pool-pool d. pool-pull

APPENDIX C

Production Pre/Posttest

I. Word-reading Instructions: Please read aloud, and in a natural way, the following 60 words as clearly as possible. Please stop for 1 or 2 seconds after each word.

1 sit	21 pooling	41 wooed
2 bag	22 suit	42 pill
3 teen	23 wood	43 soot
4 pool	24 pull	44 kook
5 eat	25 peel	45 pan
6 bend	26 shooed	46 cooed
7 feet	27 mat	47 should
8 and	28 who'd	48 fit
9 tin	29 seat	49 beat
10 guess	30 sat	50 look
11 live	31 it	51 band
12 pen	32 end	52 beg
13 cheap	33 feel	53 cook
14 x	34 gas	54 hit
15 man	35 bit	55 men
16 fool	36 axe	56 fill
17 set	37 chip	57 pulling
18 heat	38 hood	58 full
19 Luke	39 could	59 leave
20 vet	40 vat	60 met

II. Sentence-reading Instructions: Please read aloud the following 9 sentences as clearly and naturally as possible.

1. We'd like three cups of tea and coffee, three cheese pizzas and eleven meat pizzas.

2. Could you tell me where is the great city that everyone wants to visit, please?

3. He believes that being happy and healthy is the only important thing.

4. Luke gives me a how-to-cook book which is the ideal book for everyone looking to cook better.

5. I could use either cookies or pudding instead of using sugar on Tuesday noon.

6. Julie understood her full-time position couldn't be applied for during the month of June.

7. Fred was happy when he met Ted and knew that his fat cat was not missing.

8. That man is reading the sentence on the wall and it says "east or west, home is best".

9. I wish I could tell that story a little better on this February morning.

III. Storytelling Instructions: Please tell me three stories. For each story, you should use as many words containing the target vowel sounds as possible. Specifically, these three stories should focus on /ɪ/-/iː/, /e/-/æ/, and /ʊ/-/uː/ respectively. For your convenience, you can consult the following lists. However, you can also use your own words that contain the target vowel sounds. Please state your ideas as clearly as possible.

List A		List B		List C	
/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/
sit	seat	bed	bad	look	Luke
ship	sheep	end	and	cook	kook
it	eat	pet	pat	pull	pool
fill	feel	said	sad	full	fool
hit	heat	met	mat	could	cooed
kip	keep	shell	shall	wood	wooed
lick	leak	guess	gas		
sick	seek	bend	band		
bin	bean	pen	pan		
chick	cheek	beg	bag		

APPENDIX D

Pronunciation Learning Satisfaction Questionnaire

(English and Chinese)

Dear Student,

This questionnaire is designed to gather information about your opinions after learning English pronunciation through this approach. The questionnaire consists of nine statements. Please read each statement carefully. After you have done so, please put a check (√) in one of the five columns that best indicates your level of opinion. If you have any additional comments, please add them at the end. Please use the following scale to select your answer: 5=Strongly agree; 4=Agree; 3= Undecided; 2=Disagree; 1= Strongly disagree.

Statement	**Level of Opinion**				
	5	**4**	**3**	**2**	**1**
1. The current pronunciation lessons are interesting.					
2. The current pronunciation lessons are helpful.					
3. I feel confident about improving my pronunciation through this approach.					
4. I feel comfortable with the current pronunciation lessons.					
5. I prefer this approach to other approaches in pronunciation learning.					
6. I like to learn pronunciation on my own through this approach.					
7. I think I can make better sense of the pronunciation of the vowels by using this approach.					
8. I think I can identify the pronunciation of the vowels more effectively by using this approach.					
9. I think my motivation to learn pronunciation is strengthened by using this approach.					
Others (please specify)					

语音学习满意度调查问卷

亲爱的同学，你好！

本调查问卷旨在了解你使用当前方法学习英语语音的满意程度。本问卷共包括九个陈述。请仔细阅读每一个陈述。请在最能够代表你观点的选项处打勾（√）。参照以下等级量表，选择你对下面陈述的赞成程度：5=完全同意；4=同意；3=不确定；2=不同意；1=完全不同意。如有其他陈述，请附加在本问卷最后一栏。

陈述	赞成程度				
	5	4	3	2	1
1. 这个方法学习语音很有趣。					
2. 这个方法对于语音学习很有帮助。					
3. 我有信心通过这个语音学习方法提高语音能力。					
4. 我能够适应这个语音学习方法。					
5. 相对于其它语音学习方法，我更喜欢这个语音学习方法。					
6. 我愿意利用这个语音学习方法单独学习语音。					
7. 通过这个语音学习方法，我能够更好地感知元音。					
8. 通过这个语音学习方法，我能够更好地识别元音。					
9. 通过这个语音学习方法，我的语音学习动机得到增强。					
其它（请予以说明）					

APPENDIX E

Semi-structured Interview

(English and Chinese)

1. What do you think of this approach?
2. Do you prefer this approach or some other approach to learn the target English vowels? Why?
3. What were your feelings about using this approach for learning the target English vowels?
4. Does this approach help you to identify better the target English vowels? Why?
5. Does this approach help you to learn the target English vowels? Why?
6. In this study, you learned English pronunciation on your own. Do you prefer to learn English pronunciation individually, in pairs or in groups? Why?
7. Are you confident of improving your English pronunciation through this approach? Why?
8. Did you have any trouble learning English pronunciation with this approach?
9. What comments or suggestions do you have about learning English pronunciation with this approach?

半结构化访谈

1. 你对于使用这个方法学习英语元音是怎么看的？
2. 你喜欢这个还是其它语音学习方法纠正这些英语元音？为什么？
3. 你对于使用这个语音学习方法学习这些英语元音的感受如何？
4. 这个语音学习方法是否能够帮助你更好地识别这些英语元音？为什么？
5. 这个语音学习方法是否能够帮助你更好地学习这些英语元音？为什么？
6. 当前大家自主学习语音，你更希望以个人，两人小组还是多人小组的形式学习英语语音？为什么？
7. 你是否有信心通过这个语音学习方法提高自己的英语语音能力？为什么？
8. 你在使用这个语音学习方法学习英语语音的时候是否碰到什么困难？
9. 你对于使用这个语音学习方法学习英语语音有什么评价或建议？

APPENDIX F

The Index of Item-Objective Congruence (IOC) Analysis for Pronunciation Learning Questionnaire

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	0	1	√
5	1	1	1	1	√
6	1	1	1	1	√
7	1	1	1	1	√
Total	7	7	6	7	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 7

R = 7 + 7 + 6 + 7 = 27 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 27/4 = 6.75

Percentage: 6.75/7 x 100% = 96.43%

APPENDIX G

The Index of Item-Objective Congruence (IOC) Analysis for Perception Pre/Posttest

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	1	1	√
5	1	1	1	1	√
6	1	1	1	1	√
7	1	1	1	1	√
8	1	1	1	1	√
9	1	1	1	1	√
10	1	1	1	1	√
11	1	1	1	1	√
12	1	1	1	1	√
13	1	1	1	1	√
14	1	1	1	1	√
15	1	1	1	1	√
16	1	1	1	1	√
17	1	1	1	1	√
18	1	1	1	1	√
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	1	1	0	1	√
26	1	1	1	1	√
27	1	1	1	1	√
28	1	1	1	1	√
29	1	1	1	1	√
30	1	1	1	1	√
31	1	1	1	1	√
32	1	1	1	1	√
33	1	1	1	1	√

34	1	1	1	1	√
35	1	1	1	1	√
36	1	1	1	1	√
37	1	1	1	1	√
38	1	1	1	1	√
39	1	1	1	1	√
40	1	1	1	1	√
41	1	1	1	1	√
42	1	1	1	1	√
43	1	1	1	1	√
44	1	1	0	1	√
45	1	1	1	1	√
46	1	1	1	1	√
47	1	1	1	1	√
48	1	1	1	1	√
49	1	1	1	1	√
50	1	1	1	1	√
51	-1	1	1	1	√
52	1	1	1	1	√
53	1	1	1	1	√
54	1	1	1	1	√
55	1	1	1	1	√
56	1	1	1	1	√
57	1	1	1	1	√
58	1	1	1	1	√
59	1	1	1	1	√
60	1	1	1	1	√
61	1	1	1	1	√
62	1	0	1	1	√
63	1	1	1	1	√
64	1	1	1	1	√
65	1	1	1	1	√
66	1	1	1	1	√
67	1	1	1	1	√
68	1	1	1	1	√
69	1	1	1	1	√
70	1	1	1	1	√
71	1	1	1	1	√
72	1	1	1	0	√
73	1	1	1	1	√
74	1	1	1	1	√

75	1	1	1	1	√
76	1	1	1	1	√
77	1	1	1	1	√
78	1	1	1	1	√
79	1	1	1	0	√
80	1	1	1	1	√
81	1	1	1	1	√
82	1	1	1	1	√
83	1	1	1	1	√
84	1	1	1	1	√
85	1	1	1	1	√
86	1	1	1	1	√
87	1	1	1	1	√
88	1	1	1	1	√
89	1	1	1	1	√
90	1	1	1	1	√
Total	88	89	88	88	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 90

R = 88 + 89 + 88 + 88 = 353 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 353/4 = 88.25

Percentage: 88.25/90 x 100% = 98.06%

Appendix H

The Index of Item-Objective Congruence (IOC) Analysis for Production Pre/Posttest

I. Word-reading

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	1	1	√
5	1	1	1	1	√
6	1	1	1	1	√
7	1	1	1	1	√
8	1	1	1	1	√
9	1	1	1	1	√
10	1	1	1	1	√
11	1	1	1	1	√
12	1	1	1	1	√
13	1	1	1	1	√
14	-1	1	1	1	√
15	1	1	1	1	√
16	1	1	1	1	√
17	1	1	1	1	√
18	1	1	1	1	√
19	1	1	0	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	1	1	1	1	√
26	1	1	1	0	√
27	1	1	1	1	√
28	1	1	0	1	√
29	1	1	1	1	√
30	1	1	1	1	√
31	1	1	1	1	√
32	1	1	1	1	√
33	1	1	1	1	√

34	1	1	1	1	√
35	1	1	1	1	√
36	1	1	1	1	√
37	1	1	1	1	√
38	1	1	1	1	√
39	1	1	1	1	√
40	1	1	1	1	√
41	1	1	1	1	√
42	1	1	1	1	√
43	1	1	1	1	√
44	1	1	1	1	√
45	1	1	1	1	√
46	1	1	1	1	√
47	1	1	1	1	√
48	1	1	1	1	√
49	1	1	1	1	√
50	1	1	1	1	√
51	1	1	1	1	√
52	1	1	1	1	√
53	1	1	1	1	√
54	1	1	1	1	√
55	1	1	1	1	√
56	1	1	1	1	√
57	1	1	1	1	√
58	1	1	1	1	√
59	1	1	1	1	√
60	1	1	1	1	√
Total	58	60	58	59	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 60

R = 58 + 60 + 58 + 59 = 235 (Scores from experts)

N = 4 (Numbers of experts)   IOC = 235/4 = 58.75

Percentage: 58.75/60 x 100% = 97.92%

II. Sentence-reading

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	1	0	√
5	1	1	1	1	√
6	1	1	1	1	√
7	1	1	1	1	√
8	1	1	1	1	√
9	1	1	0	1	√
Total	9	9	8	8	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 9

R = 9 + 9 + 8 + 8 = 34 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 34/4 = 8.5

Percentage: 8.5/9 x 100% = 94.44%

III. Storytelling

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
Total	1	1	1	1	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 1

R = 1 + 1 + 1 + 1 = 4 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 4/4 = 1

Percentage: 1/1 x 100% = 100%

Average percentage: (97.92% + 94.44% + 100%)/3 = 97.45%

Appendix I

The Index of Item-Objective Congruence (IOC) Analysis for Pronunciation Learning Satisfaction Questionnaire

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	1	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	1	1	√
5	1	1	0	1	√
6	1	1	1	0	√
7	1	1	1	1	√
8	1	1	1	1	√
9	1	1	1	1	√
Total	9	9	8	8	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/N)

Item number: 9

R = 9 + 9 + 8 + 8 = 34 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 34/4 = 8.5

Percentage: 8.5/9 x 100% = 94.44%

Appendix J

The Index of Item-Objective Congruence (IOC) Analysis for Semi-structured Interview

Item	Expert 1	Expert 2	Expert 3	Expert 4	Result
1	1	1	1	0	√
2	1	1	1	1	√
3	1	1	1	1	√
4	1	1	1	1	√
5	1	1	1	1	√
6	1	1	1	1	√
7	1	1	1	1	√
8	1	1	1	1	√
9	1	1	1	1	√
Total	9	9	9	8	

Notes:

1. 1 = certain that the item is congruent with the objective

2. 0 = uncertain that the item is congruent with the objective

3. -1 = certain that the item is not congruent with the objective

Result of IOC:

(IOC = ∑R/ N)

Item number: 9

R = 9 + 9 + 9 + 8 = 35 (Scores from experts)

N = 4 (Numbers of experts)

IOC = 35/4 = 8.75

Percentage: 8.75/9 x 100% = 97.22%

Appendix K

Rubrics for Production Pre/Posttest

Dear Teacher,

Thank you for your participation in helping the present study. I would like you to evaluate whether or not the students' productions of the target English vowels (/ɪ/, /iː/, /e/, /æ/, /ʊ/ and /uː/) in word-reading, sentence-reading and storytelling are acceptable and also assess their overall speaking proficiency in sentence-reading and storytelling concerning comprehensibility, fluency and pronunciation.

The scoring rubrics for comprehensibility, fluency and pronunciation are listed below. Please rate the students' recordings holistically, taking into consideration the following five pronunciation variables, namely, sounds (vowels and consonants), rhythm, word stress, intonation, and speech rate. The maximum score for each measure is 20. Please spend some time listening carefully to each recording before making your final decision.

I. Rubrics for comprehensibility

Description	**Score (sentence-reading & storytelling)**
Not understandable	0-3
Slightly understandable	4-7
Significantly understandable	8-11
Mostly understandable	12-15
Fully understandable	16-20

II. Rubrics for fluency

Description	Score (sentence-reading & storytelling)
Intrusive	0-3
Intermediate	4-7
Good	8-11
Advanced	12-15
Native-like	16-20

III. Rubrics for pronunciation

Description	Score (sentence-reading & storytelling)
Poor	0-3
Fair	4-7
Good	8-11
Very good	12-15
Excellent	16-20

APPENDIX L

Self-report

(English and Chinese)

Dear Student,

This self-reporting form is designed to collect general information about your pronunciation learning process. Please complete the form below. Your time and help will be greatly appreciated. Your information will be kept confidential. Thank you very much for your cooperation.

Name__________________

Date __________________

Place __________________

Starting time__________________ Ending time__________________

Please name the materials that you used. _______________________

What materials did you use for pronunciation learning? Check all that apply.

☐ filtered ☐ natural

Did you have any problems with pronunciation learning?

☐ yes ☐ no

If yes, please specify.

____

____

Do you feel that you are making progress? Please explain your answer.

What were you doing before and what will you do after completing this exercise (not necessarily learning activities)?

学习记录

亲爱的同学，你好！

本记录旨在收集培训期间语音学习情况。请在每次语音学习完成后及时填写本记录（包括课内及课外语音学习）。请根据自己的真实情况填写，回答结果保密。谢谢你的支持！

姓名__________________

日期__________________

地点__________________

开始时间__________________　结束时间__________________

你所使用的语音学习材料____________________________________

你使用哪种音频材料训练？（可多选）

☐ 过滤音　　☐ 自然音

你在语音学习过程中是否遇到困难？

☐ 是　　☐ 否

如是，请予以说明

____

____

____

____

你是否感觉语音有进步？请予以说明

参加这个培训之前你在干什么？培训之后你准备干什么？请具体说明

CURRICULUM VITAE

Fengwei Wen was born in November 1981, in Jiangxi Province, China. He obtained his Bachelor of Arts degree in English Education from Jiangxi Normal University in 2004. He graduated from Guizhou University with a Master of Arts degree in Foreign Linguistics and Applied Linguistics in 2007.

He is currently a lecturer in English at the School of Foreign Languages, Gannan Normal University, China. His research interests include applied linguistics and computer-assisted language learning and teaching.