การปรับปรุงความจำเชิงปฏิบัติการด้านระบบเสียงและทักษะการพูดภาษาอังกฤษของนักเรียนระดับประถมศึกษาชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศโดยวิธีเวอร์เบอร์โทนอล

นางหยาน แยง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2559

IMPROVING THE PHONOLOGICAL WORKING MEMORY AND ENGLISH SPEAKING SKILLS OF CHINESE PRIMARY EFL LEARNERS WITH A VERBOTONAL-BASED APPROACH

Yan Yang

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 2016

IMPROVING THE PHONOLOGICAL WORKING MEMORY AND ENGLISH SPEAKING SKILLS OF CHINESE PRIMARY EFL LEARNERS WITH A VERBOTONAL-BASED APPROACH

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

Thesis Examining Committee

(Assoc. Prof. Dr. Pannathon Sangarun)
Chairperson

(Assoc. Prof. Dr. Anchalee Wannaruk)
Member (Thesis Advisor)

(Prof. Dr. Andrew Lian)
Member

(Asst. Prof. Dr. Harald Kraus)
Member

(Asst. Prof. Dr. Arjuna Chaiyasena)
Member

(Dr. Adcharawan Buripakdi)
Member

(Prof. Dr. Sukit Limpijumnong)
Vice Rector for Academic Affairs and Innovation

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

หยาน แยง : การปรับปรุงความจำเชิงปฏิบัติการด้านระบบเสียงและทักษะการพูดภาษาอังกฤษของนักเรียนระดับประถมศึกษาชาวจีนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศโดยวิธีเวอร์เบอร์โทนอล (IMPROVING THE PHONOLOGICAL WORKING MEMORY AND ENGLISH SPEAKING SKILLS OF CHINESE PRIMARY EFL LEARNERS WITH A VERBOTONAL-BASED APPROACH)
อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.อัญชลี วรรณรักษ์, 335 หน้า

ในบรรดาทักษะทางภาษาสี่ทักษะ ทักษะการพูดได้รับการระบุว่าเป็นทักษะที่แย่ที่สุดสำหรับผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ ระดับประถมศึกษาของจีน วิธีการออกเสียงแบบเวอร์โบโทนอล(verbotonal) ซึ่งได้รับการพัฒนาในขั้นต้น เพื่อเป็นการบำบัดสำหรับผู้ที่มีปัญหาการได้ยิน ได้รับการพิสูจน์ว่ามีประสิทธิภาพในการเรียนรู้ภาษาต่างประเทศ อย่างไรก็ตาม ในบริบทของจีน ก็ไม่เคยมีการนำมาใช้ในการสอนทักษะการพูดให้กับนักเรียนระดับประถมศึกษาและมักใช้ในการสอนการออกเสียงแทนการพูด ดังนั้น การศึกษาในปัจจุบันจึงได้พัฒนาวิธีการใช้ verbotonal-based approach (VTA) ซึ่งส่วนใหญ่เป็นการพัฒนาทักษะการพูดของผู้เรียนภาษาจีนระดับประถมศึกษาและเสริมสร้างความเข้มแข็งในการสร้างหน่วยความจำเชิงสัทศาสตร์ (PWM) (Guberina, 1972; Lian, 1980; ASP, 2006) ขั้นตอนการสอนของ VTA ซึ่งมีการคิดค้นพัฒนาขึ้นอย่างมากจากการศึกษาก่อนหน้านี้โดย Andrew Lian (1980) ประกอบด้วยการรับรู้ด้านความรู้สึกในชั้นเรียนและการเสริมแรงนอกชั้นเรียน การทดลองกึ่งทดลองซึ่งดำเนินการเป็นเวลา 18 สัปดาห์ได้ดำเนินการเพื่อตรวจสอบประสิทธิภาพของ VTA ในรูปแบบการเรียนรู้ทักษะการพูดและการเพิ่ม PWM ในขณะเดียวกันยังได้มีการตรวจสอบความสัมพันธ์ระหว่าง PWM กับทักษะการพูด

กลุ่มตัวอย่างได้มาจากโรงเรียนประถมศึกษา Yiliang Anjiaqiao (YAPS) นักเรียนชั้นประถมศึกษาปีที่ 3 ที่มีอายุเฉลี่ย 9 ปี จากการแจกแจงแบบสุ่ม กลุ่มหนึ่งเป็นกลุ่มทดลองและได้รับ VTA ในขณะที่อีกกลุ่มหนึ่งเป็นกลุ่มควบคุมและเข้ารับการฝึกทักษะการพูดแบบดั้งเดิม

หลังการทดลอง กลุ่มทดลองมีพัฒนาการพูดและการอ่านอย่างมีนัยสำคัญ ได้แก่ การอ่านคำ อ่านประโยค ร้องเพลง การสัมภาษณ์ปากเปล่า รวมทั้งทักษะย่อย 5 คำ ได้แก่ คำศัพท์ไวยากรณ์ การออกเสียง ความคล่องและความเข้าใจ กลุ่มทดลองมีประสิทธิภาพดีกว่ากลุ่มควบคุมในทุกด้าน

กลุ่มทดลองยังมีการปรับปรุงอย่างมีนัยสำคัญในแง่ของสมรรถนะด้าน PWM โดยรวม นอกจากนี้ผู้เข้าร่วมในกลุ่มทดลองยังมีความคืบหน้าในการทำซ้ำคำที่ไม่เป็นคำยาวและคำที่ดูไม่เป็นคำที่มีความคล้ายคลึงกันต่ำ ในทางตรงกันข้ามไม่มีการปรับปรุงด้าน PWM และการทดสอบย่อยอย่างมีนัยสำคัญในกลุ่มควบคุม

ผลการวิจัยแสดงให้เห็นว่า VTA มีประสิทธิภาพไม่เพียงแต่ในการพัฒนาทักษะการพูดของผู้เรียนเท่านั้น แต่ทักษะภาษาด้านอื่น ๆ ก็ดีขึ้นและมีประสิทธิภาพมากขึ้นด้วย

สาขาวิชาภาษาต่างประเทศ
ปีการศึกษา 2559

ลายมือชื่อนักศึกษา_____
ลายมือชื่ออาจารย์ที่ปรึกษา_____
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

YAN YANG : IMPROVING THE PHONOLOGICAL WORKING MEMORY AND ENGLISH SPEAKING SKILLS OF CHINESE PRIMARY EFL LEARNERS WITH A VERBOTONAL-BASED APPROACH.
THESIS ADVISOR : ASSOC. PROF. ANCHALEE WANNARUK, Ph.D., 335 PP.

SPEAKING SKILLS/ PHONOLOGICAL WORKING MEMORY/ CHINESE/ PRIMARY/ EFL LEARNERS/ VERBOTONAL

Among the four language skills, speaking was identified as the poorest skill for Chinese primary EFL learners. The verbotonal approach, developed initially as a therapy for people with hearing difficulties, has also been shown to be effective in the learning of foreign languages. However, in the Chinese context, it has never been used to teach speaking skills to primary school students and it is commonly used to teach pronunciation rather than speaking. Thus, the present study developed a verbotonal-based approach (VTA) mainly for improving the speaking skills of Chinese primary EFL learners, and simultaneously, ameliorating their phonological working memory (PWM) (Guberina, 1972; Lian, 1980; Asp, 2006). The pedogogical procedures of VTA which were devised heavily on the basis of a previous study carried out by Andrew Lian (1980) consisted of in-class sensitization and out-of-class reinforcement sessions. A quasi-experiment which spanned 18 weeks was performed to investigate the effectiveness of VTA as a manner of learning speaking skills and enhancing PWM. In the meantime, the relationship between PWM and speaking skills was also examined.

The sample was drawn from the Yiliang Anjiaqiao Primary School (YAPS). Eighty Grade 3 students with an average age of nine years. Randomly, one class formed the experimental group, and received VTA, while the other constituted the control underwent the traditional way of learning speaking skills.

After the intervention, the experimental group advanced significantly in both overall speaking proficiency and individual tests: word-reading, sentence-reading, singing, oral interview as well as the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility. Addtionally, the experimental group outperformed the control group in all aspects tested.

The experimental group also improved substantially in terms of overall PWM capacity. Besides, participants in the experimental group showed considerable progress in repeating lengthy nonwords and nonwords with low wordlikeness. On the contrary, no significant improvement in terms of overall PWM capacity and subtests was detected in the control group.

The research evidenced that VTA was demonstrably effective in improving not only the speaking skills of young learners but their other language skills as well and that it did so more effectively than the other approaches used.

School of Foreign Languages

Academic Year 2016

Student's Signature_____

Advisor's Signature_____

Co-advisor's Signature_____

ACKNOWLEDGEMENTS

First and foremost, my special thanks are extended to my supervisor, Professor Dr. Andrew Lian, for the great support he has given me in completing this study. It has been through his patience, inspiration, continual encouragement and professional experience that I was able to step on the amazing journey of research and discover my field of interest. He has devoted his time and energy to reading, revising and polishing my drafts. And we have spent long hours together discussing the project and were excited about the surprising and promising findings. These precious moments will be cherished in the deepest part of my heart and never fade from memory.

I am also greatly indebted to my other supervisor, Associate Professor Dr. Anchalee Wannaruk. Without her constructive comments, the completion of this study would be impossbile. Her encouraging words and immense help have pulled me through the arduous study life. She also assisted me in translating the title of my dissertation into Thai. All that she did for me has filled me with warm memories.

Heartfelt thanks are also expressed to my thesis examination committee: Associate Professor Dr. Pannathon Sangarun (who served as chair), Assistant Professor Dr. Harald Kraus, Assistant Professor Dr. Arjuna Chaiyasena and Assistant Professor Dr. Adcharawan Buripakdi for their valuable suggestions.

I would particularly like to acknowledge Dr. Sirinthorn Seepho, Dr. Issra Pramoolsook, and many other faculty and staff members of the School of Foreign Languages, Suranaree University of Technology (SUT). They have always been helpful and enabled me to experience a memorable and rewarding life in Thailand.

Deepest thanks go to the students and teachers in the Yiliang Anjiaqiao Primary

School. It is their active participation and cooperation that has turned the implemention of the pedogogical experiment into a success.

I am also sincerely grateful to all the Chinese experts who have suggested and validated the data colletion instruments and data analysis techniques of the study and performed the blind rating task. Without their painstaking efforts, the study could not have been accomplished smoothly.

Many thanks are given to my home university, Kunming University of Science and Technology, for allowing me this opportunity to engage in professional development and providing me with financial support.

I greatly appreciate all my international and Chinese friends who kept me company in SUT. Their friendship has made my life here not so stressful, tedious and boring but relaxing, joyful, and enriching.

I owe particular thanks to my husband, Associate Professor Dr. Han Chunhua. He shouldered all the reponsibilty to take take of the whole family during my study abroad. His love, understanding and tolerance have soothed me and tided me over the dreary and oppressive times.

Words of thanks are also uttered to my daughter, Han Xinyang, who is a fourth-grader and as lovely as an angel. Her merry laugter, childlike innocence and imaginative ingenuity have always been and will continue to be the inspiration and sunshine of my life.

Most of all, I wish to extend my countless thanks to my parents, Mr. Yang Chunwu and Mrs. Pang Zhimin. Over the years, their selfless love and wholehearted assistance have always been there and propeled me forward. I will do my utmost to be the pride in their eyes.

Enormous thanks are also articulated to my parents-in-law, Mr. Han Jiafu and Mrs. Meng Xueying. They have always rendered a helping hand to me whenever I am in need.

Yan Yang

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

Page

LIST OF TABLES (Continued)

Page

LIST OF TABLES (Continued)

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ACTFL	American Council on the Teaching of Foreign Language Guidelines
	ALM the Audio-lingual Method
ANCOVA	an Analysis of Covariance
CEFR	Common European Framework of Reference
CALL-VT	Computer-Assisted Language Learning-Verbotonalism
ETS	Educational Testing Service
ESOL	the Cambridge Exams in English for Speakers of Other Languages
ELT	English Language Teaching
EFL	English as a Foreign Language
IELTS	the International English Language Testing System
ITEP	the International Test of English Proficiency
LTM	Long-term Memory
LTS	Long-term Store
L1	the First Language
L2	the Second Language
MOE	the Chinese Ministry of Education
NWR	Nonword Repetition
OET	Occupational English Test
PWM	Phonological Working Memory
PTE	Pearson Tests of English

LIST OF ABBREVIATIONS (Continued)

RIT	the Rasch Item Response
RCM	the Relative Contribution Model
STM	Short-term Memory
STS	Short-term Store
SEA	Somatically Enhanced Approach
SPSS	Statistical Package for the Social Sciences
TOEFL	the Test of English as a Foreign Language
TCLE	Trinity College London Exams
TPR	the Total Physical Response Method
WM	Working Memory
VT	Verbotonalism
VTA	the Verbotonal-based Approach
YAPS	the Yiliang Anjiaqiao Primary School

CHAPTER 1

INTRODUCTION

The present study aims at improving the English speaking skills and phonological working memory (PWM) of Chinese primary EFL learners using a Verbotonal-based approach (VTA). This introductory chapter first presents the background of the study. Then, the statement of the problem, the significance, objectives, research questions, and delimitations of the study are outlined. Lastly, a number of terms that are important to the study are defined, and a summary of the whole chapter is laid out.

1.1 Background of the study

In order to situate the present study in its context, it is valuable to have a glimpse of the history of English language education in Chinese primary schools. Since China's opening up to the world in 1978, there has been an escalating demand for citizens with a good mastery of English. The importance of English as a tool for engaging in international political, economic, and cultural activities has been increasingly recognized by the Chinese leadership. At the national level, English proficiency has been regarded as playing a significant role in promoting the development of science and technology, industrialization, modernization and economic growth while, at the

individual level, it has been thought of as bringing about a host of economic, professional, career and educational opportunities conducive to improving one's well-being (Adamson, 2001; Hu, 2002a; Cortazzi & Jin, 2003; Yi, 2009; Zhang, 2014). As a consequence, English language education in China has been accorded more and more attention. As early as 1962, the Chinese Ministry of Education (hereinafter referred to as MOE) issued a document approving the teaching of English in Grades 4 and 5 in the 5-year primary schools selected to pilot new school systems, and in Grades 5 and 6 in the 6-year primary schools in urban areas. What should be pointed out here is that, at that time, there was no unified English curriculum or syllabus issued by the government to direct the teaching of English in primary schools. However, in 1966, because of social and political factors, English language education in China suffered from stagnation and was mainly offered to secondary and college students. It was not until 1978 that the first unified national English syllabus entitled *English Syllabus for Ten-Year Full-Time Primary and Secondary Schools* was promulgated (MOE, 1978; Hu, 2003). The syllabus mandated English as a compulsory subject in the curriculum of primary schools in China and it also designated the starting point of English learning for students: Primary 3. However, the nation-wide provision of English language education was frustrated by many a problem such as the severe shortage of qualified teachers and appropriate teaching materials (Hu, 2002a). In recognition of the constraints of limited resources for English language education as well as disparities in socioeconomic and regional development, the Chinese government suspended the compulsory requirement

of expanding English language teaching (hereafter referred to as ELT) in primary schools during the 1990s (Hu, 2002b). However, during this period, primary schools in some economically advanced regions and cities such as Zhejiang and Shanghai still provided English instruction to students from Grade 3 or even lower grades.

Stepping into the new era, China continues to enjoy phenomenal economic growth, and simultaneously, faces the challenges of the ever-increasing internationalization and globalization. To meet the demands resulting from social development, in January 2001, the MOE issued a directive to embark on the universal provision of English in primary schools. It was required that primary schools in cities and counties offer English classes at Primary 3 level, beginning in autumn of 2001 and the primary schools located in rural areas were required to do so in the following year (MOE, 2001a). Meanwhile, with the efforts of a panel of nine experts, *the English Curriculum Standards for Full-time Compulsory Education and Senior High Schools (Trial Version)* was promulgated to guide the teaching of English in both primary and middle schools in China (MOE, 2001b; Zhang, 2012). Unlike its predecessor, this version of the English curriculum standards was devised on the same educational philosophy as that for other subjects in primary and secondary schools. According to the curriculum standards, English language education is regarded as one facet of quality education. Collaborative and experiential learning are highly valued but, at the same time, students' learning idiosyncrasies are given recognition and value. The curriculum specifies that the ultimate aim of basic English language education is to promote

students' comprehensive language competence which, according to the curriculum standards, is made up of five aspects: (1) linguistic knowledge, (2) language skills, (3) learning strategies, (4) cultural awareness, and (5) emotional attitudes. To put it in a more detailed way, **linguistic knowledge** involves knowledge of communicative functions, topics, grammar, vocabulary and pronunciation. In the past, there was a belief that linguistic knowledge was at the core of English learning. For a very long time, English teachers from primary schools to universities in China had paid too much attention to the teaching of linguistic knowledge. **Language skills** incorporate listening, speaking, reading and writing skills. At that stage, policy-makers, educators, researchers and English teachers began to realize the significance of teaching English as a means of communication and interaction. It is the first time in the history of the development of the English curriculum in China that listening and speaking skills have received as much attention as the other two literacy skills. **Learning strategies** range from cognitive, meta-cognitive, communicative strategies to resource strategies. As specified by the curriculum standards, cognitive strategies are a mental process or procedure for completing a particular cognitive task. Metacognitive strategies refer to students' conscious management, monitoring and evaluation of their learning processes and outcomes. Communicative strategies mean strategies students use to resolve the problems they come across in communication. Recourse strategies refer to students' capabilities for searching for and utilizing a wide range of resources such as multi-media resources, to facilitate their study. **Cultural awareness** comprises both

knowledge and understandings of different cultures, especially the cultures of their motherland and English speaking countries, and intercultural communication awareness and competence. **Emotional attitudes** denote students' motivation, interest, attitudes and self-efficacy in learning English. The standards specify that the principal aims of primary English teaching are to cultivate students' interest, stimulate their motivation, elevate their confidence and help them establish a positive attitude towards English learning. Guided by the directive and curriculum standards, the nation-wide expansion of English into primary schools in China has gained momentum since 2001. In 2011, the MOE slightly revised the 2001 standards. The *2011 English Curriculum Standards for Compulsory Education* specifies nine levels of English competence which the primary and secondary school students are supposed to reach after a certain number of years of learning. Level 1 is the standard set for students from Grade 3 and Grade 4, Level 2 for students from Grade 5 and Grade 6 (see Table 1.1). Moreover, the development of students' autonomous learning has been included as one of the objectives in English learning in the 2011 standards.

Fourteen years have elapsed since the implementation of the universal provision of English in primary schools. However, the outcomes of primary English language education appear to be unsatisfactory and discouraging in spite of the huge investment in teaching and learning resources and infrastructures. On average, a student learns English for 10 years or even more in China. But many students are incapable of communicating in English. This phenomenon is notoriously referred to as "dumb" or

"mute English" (Hu 2004; Zhang, 2011; Zhang, 2012). The low efficiency of English language education prompted policymakers, researchers and educators to examine and reflect on the status-quo of English teaching, identify the problems and seek out solutions.

Table 1.1 2011 English curriculum standards-levels of English competence

Level	Standards
Level 1 (Grade 3 – Grade 4)	Be interested in and enjoy listening to English; be able to play games and perform tasks (such as coloring and matching) according to teachers' directions in English, engage in simple role-plays in English, sing simple English songs, read and understand short stories with the aid of pictures, deliver short self-introduction and express emotions and feelings in simple English, and spell letters and the keywords required by the syllabus; learn and understand foreign cultures and customs
Level 2 (Grade 5-Grade 6)	Develop a continuing interest in the learning of English; be able to express greetings, provide personal information in simple English, act out learned dialogues and songs, read, understand and retell learned stories with the help of pictures, and write

	simple sentences based on picture cues; participate actively in and have a positive attitude toward the learning of English; learn and understand foreign cultures and customs

With regard to ELT in the primary schools, there are many problems emerging. According to the 2011 curriculum standards, English learning should be communication-oriented. But the disappointing fact is that among the four macro-skills (listening, speaking, reading and writing), speaking is identified as the most difficult for Chinese primary school students, with listening being the second (Zhang, 2011, Li, 2016). The major challenges which hinder English speaking instruction in primary schools in China are a great shortage of qualified teachers and a severe lack of suitable learning materials (Hu, 2004; Wang, 2007; Teng, 2010). In addition, the overly large class size, inadequate facilities for instruction and insufficient coordination between primary and secondary English courses also pose a serious challenge to the teaching and learning of speaking in primary schools in China (Hu, 2004; Li, 2001; Zhang, 2012, Li, 2016). Under these circumstances, more and more parents choose to send their children to after-school training centers to learn English. English training has become the backbone of the Chinese educational training industry. It was estimated that, by the end of 2015, there had been more than 50,000 English training centers in China, with the aggregated market value adding up to RMB 30 billion (Zhang, 2011). The

prosperity of the English training industry also serves as powerful evidence for the disappointing outcomes of in-school English teaching.

In order to address the intractable problems, researchers, educators and teachers in English language education in China are beginning to invest more effort in exploring progressive English teaching approaches and methods (Zhang, 2012). Nevertheless, the fact is that the problems appearing in the teaching and learning of English speaking in primary schools have not been resolved successfully despite the adoption of the "latest" approaches from the mainstream. Clearly, there is a desperate need for more sufficiently developed teaching approaches.

To develop a suitable language teaching approach, it is valuable to engage in a fundamental rethink of the nature of language learning. The language learning process, like that of learning in general, is essentially a process of personal/individual meaning-making (Lian, 2004). As Lian argues, in the learning process, individuals' perception of incoming information into meaning hinges on their personal operational history, which includes their world knowledge, strategies and much more that cannot be potentially defined or described. These operational histories "*necessarily hold our entire functional personal knowledge and understandings and the things that we have learned: the operational sum total of our experience"* (Lian, 2014, p.13). Obviously, the operational histories of each of us are different. Most of all, the differences in personal operational histories necessarily make learning individual-specific. Meanwhile, the learning experience should be complexified rather than simplified

(Lian, 2004). Hence, language teaching should be able to identify the idiosyncratic characteristics of individuals' learning and try to address them differently. Subjecting all learners to a centralized, standardized and systematized teaching approach is unlikely to be a good choice. More importantly, any language learning or teaching system should be able to raise learners' awareness and change their perceptions of the foreign symbol system.

Taken together, the present study aims to develop a verbotonal-based approach (VTA) to improve the English speaking skills of Chinese primary EFL learners. The focus of the study is to raise students' awareness and change their perceptions of the characteristics of the phonological system of English. Meanwhile, PWM, being a temporary verbal-acoustic storage system, has been found to be essential for both native and second language development (Gathercole & Adams, 1994; Gathercole & Baddeley, 1989, 1990; Gathercole et al., 1991, 1992; Michas & Henry, 1994; Baddeley, 2012; 2013; Baddeley et al., 2014). There is compelling evidence indicating that PWM is an effective predictor of foreign language skills (Service, 1992; Gathercole & Baddeley, 2014). Although working memory (WM) has been considered as a stable property and constant trait of human beings, studies suggest that it can be improved by explicit and implicit training (Klingberg, 2010). The verbotonal-based approach developed in the present study maintains that learning is a process of meaning-making and regards neuroplasticity of the human brain as a very important principle in the teaching of English speaking. The primary goal of the approach is to provide optimal stimulation

to the brain of learners and thus induce neuroplasticity of their brain, which has been recognized as a necessary condition for ameliorating WM capacity (Guberina & asp, 1981; Hempel et al., 2004; Klingberg, 2010). In addition to that, the verbotonal-based approach proposed in the current research places emphasis on suprasegmental features and incorporates body movements into the learning of speaking. These procedures are also considered to be of essential significance to the improvement of PWM. Thus, to enhance students' PWM using VTA becomes another objective to pursue in the current study.

1.2 Statement of the problem

Being a rural primary school, the Yiliang Anjiaqiao Primary School (YAPS) is located in Anjiaqiao Village, Gucheng Town, Yiliang County, Kunming City, Yunnan Province, China. As of September 1st, 2015, there were altogether 10 classes from Grade 1 to Grade 6, the full-time students amounting to 400. On average, there were approximately 40 students in a class. Abiding by the national requirements, YAPS provides English classes to students of Grade 3 and above. And among the 15 full-time teachers, four teachers are responsible for English language teaching. The number of teaching hours for English for students of all grades is three academic periods (45 minutes/period) per week. There are widespread complaints about the poor speaking performances of students from parents, teachers and students themselves.

In order to have a clear picture of the problems facing both teachers and students in the teaching and learning of speaking skills, the researcher conducted a semi-structured interview with 16 students and all 4 English teachers in YAPS. These students, between 8-12 years of age, were sampled randomly from Grade 3 to Grade 6, 4 coming from each grade and 8 being male and 8 female. The semi-structured interview questionnaires were constructed by the researcher based on previous research in this field. To achieve acceptable validity and reliability of the instruments, a small-scale pilot study was first carried out, in which 4 students (1 from each grade, 2 males and 2 females) and 2 teachers (1 teaching Grades 3 and the other Grade 6) were interviewed. Then, the researcher adapted the questionnaire based on the response and suggestions from the interviewees. After that, the adapted questionnaire was sent to 5 experts specializing in English language education, 1 coming from YAPS, and the other 4 from Kunming University of Science and Technology, for content validity evaluation. Item-objective congruence (IOC) measure developed by Rovinelli and Hambleton (1977), a procedure in which all experts conduct an independent rating of the individual items of the questionnaire on the degree to which they measure what they are supposed to measure, was applied to validate the construct of the interview questionnaires. To be more detailed, all experts evaluated each item independently by giving it a rating of 1 (if the item is congruent with the objective), -1 (if the item is not congruent with the objective), or 0 (the congruence of the item is unclear). Corresponding revisions and modifications were made on the basis of the feedback from the experts. The result of

the calculation of the IOC for the interview questionnaire for students was 87% (see Appendix C) and that for the teachers was 86% (see Appendix F), indicating a high content validity with reference to the minimum acceptable validity index-50% established by Rovinelli and Hambleton. The researcher then conducted the finalized semi-structured interviews to both the teachers and the selected students (see Appendix A and Appendix D). Through verbatim transcription of the interviews and content analysis of the data, both the students' and teachers' perceptions of the problems existing in the learning and teaching of speaking skills were summarized as follows:

1) Student-related problems. Student-related problems which hindered the learning of speaking in YAPS included students' poor foundation of linguistic knowledge and relative indifference toward speaking. Most of the students considered that their speaking proficiency was at a low level, whereas a few thought of their speaking ability as very good. This was also confirmed by the 4 teachers who commented that there was a wide disparity in the speaking ability of the students. The two groups of participants reported that students' incorrect pronunciation, limited size of vocabulary, and low-level of grammatical accuracy were the major obstacles to their learning of English-speaking skills. And according to all the teachers, students' poor pronunciation was the biggest problem obstructing the learning of speaking.

2) Teacher-related problems. Teacher-related problems identified were mainly made up of two parts: teachers' insufficient expertise in teaching speaking and inappropriate teaching approaches. All four teachers thought that their speaking ability was not good enough for them to teach speaking skills. All of them taught English in a very traditional way. That is, during the 45-minute academic period, they would spend some 15 minutes teaching the pronunciation of words chiefly by the teacher reading the words aloud and asking the students to repeat them, another 20 minutes reading the sentences, sometimes acting out the dialogues or singing songs or watching videos, and the last 10 minutes was spent spelling the keywords specified by the syllabus. Most of the time, they were the models for students to learn the pronunciation of words and to speak English. Further, as there was a large disparity in the listening and speaking skills of students, in order to make themselves understood by most of the students, the teachers tended to use more Chinese than English as the language medium of teaching. All the teachers reported that they had a heavy workload, and seldom had the chance to receive professional training. They hardly ever had the opportunity to participate in research projects.

3) School-related problems. School-related problems identified as influencing learning to speak involved the overly large class size, and the

lack of favorable environments for practicing the language and the assessment system for speaking proficiency. As stated before, generally speaking, the class capacity was around 40 students. Facing such a large class size and limited teaching time, it was difficult for the teacher to identify the problems of each student and to adopt flexible teaching methods based on students' individuality. In addition, the school rarely organized activities for the students to practice speaking. Meanwhile, there was no English speaking test conducted with any grade level in YAPS. So the weight of learning to speak English had been underestimated by some of the students.

4) The after-school problem. The after-school problem which impeded the efficient learning of speaking skills was infrequent speaking practice out of class. Typically, the students were required to practice English for 30 minutes after class from Monday to Friday. The practice mainly centered on reading and reciting learned dialogues, and memorizing words and phrases. Students hardly had a chance to put what they had learned in class into use in real-life situations. Therefore, they were apt to forget the learned material and lose motivation.

As stated previously, in order to address the problems influencing the learning and teaching of English speaking skills in YAPS, a verbotonal approach (VTA) was proposed in the present study. VTA is a holistic system which lays stress on the

multi-sensory, multi-modal, and multi-channel way of learning to speak English. The starting point of the approach is to raise students' awareness and change their perceptions of the characteristics of the phonological system of English. And once students' perceptions of the sounds of English has been changed, they will become more sensitive to the phonetic specificities of the language. Situational dialogue which requires collaboration will be used in the approach to highlight the pragmatics of the spoken language and enhance students' social and conversational skills. Most importantly, the approach attaches importance to experiential and autonomous learning. During the whole learning process, the teacher only acts as facilitator and catalyst. Students are empowered to self-adjust their learning and be free from any intervention and instruction from the teacher.

1.3 Significance of the study

From a theoretical standpoint, the present study is expected to enrich the theories in education, psycholinguistics, neuroscience and linguistics. From a practical perspective, it may have implications for pedagogy and provide empirical evidence for the understanding and training of the working memory system of human beings. Specifically, the significant points of the current research are outlined as follows.

First, the verbotonal approach (VTA) developed in the current study proposes an integrative and holistic way of learning to speak in English (or any foreign language). The speech stimulation procedures advocated by the approach follow the

developmental patterns which have been observed in children's acquisition of their mother tongue. In this sense, the implementation of VTA in the teaching of speaking skills will deepen our understanding and thinking about the nature of language as well as language learning and teaching.

Second, the verbotonal system bases its theoretical foundation on the belief that language learning is a process of meaning-making. With a holistic way of learning to speak being the focus, this approach is designed to raise students' awareness of the foreign language phenomenon, change their way of perceiving and producing speech and optimize their learning of speaking. In the whole learning process, the students' vestibular, tactile and proprioceptive senses are stimulated and orchestrated simultaneously. In so doing, students' acquisition of the speaking skills are enhanced.

Third, VTA stresses the importance of students' sensitivity to the suprasegmental features such as stress, rhythm, loudness, pitch, and intonation, in other words, the prosodic patterns of the target language. Different languages have different prosodic patterns. These prosodic patterns organize the vowel and consonant sounds of a particular language in a way which is unique to that language (Lian, 1980). Studies show that prosody assists learners in parsing continuous speech and recognizing words, and provides cues to syntactic structure, grammatical boundaries and sentence type (Nooteboom, 1997). Besides, Mcnerney and Mendelsohn (1992) argue that suprasegmentals can aid intelligibility and make learners' speech more easily comprehensible. Therefore, this approach is likely to be more effective than other

approaches in the learning of speaking as it puts great emphasis on suprasegmental features.

Fourth, the inclusion of whole-body movements in the learning of speaking is another significant point of the present study. Evidence shows that movement promotes students' thinking as it helps their brains grow and develop (Diamond, 2000). Simply standing up and moving increases blood flow and oxygen to the brain by 15% (Jensen, 1998). More importantly, evidence substantiates that there is a dynamic interchange between cognition and action (Condon & Ogston, 1966; Ballard et al., 1997; Barsalou, 1999). According to Condon and Ogston (1966), synchronization between body movements and speech occurs in both intra-individual and interactional behavior. This implies that, in conversation, the body motion of the speaker is synchronous with his speech rhythms. Meanwhile, the body movements of the listener also spontaneously move in rhythm with those of the speaker (Condon & Ogston, 1966; 1971; Shockley et al., 2003; Richardson et al., 2008). In light of these ideas, it is paramount to engage body movements in the teaching and learning of speaking skills.

Fifth, as stated previously, phonological working memory (PWM) capacity predicts performances in a wide range of cognitive and linguistic tasks. Although it has long been an article of faith that working memory skills are the result of exceptional ability and are difficult to change and improve, a few studies suggest that they can be acquired through training (Klingberg, 2010). In the current research, the verbotonal-based approach is implemented to enhance PWM capacity. The findings of the present

study may shed light on the investigation of the functional brain networks underlying working memory system and the intertwined relationship between speaking skills and PWM.

1.4 Objectives of the study

This research was intended to improve the speaking skills of EFL learners in the Yiliang Anjiaqiao Primary School (YAPS), and at the same time, to enhance their phonological working memory (PWM). To be more specific, the following five objectives were targeted:

1) To develop an approach based on Verbotonalism (VTA) for learning to speak English;
2) To enhance PWM through VTA;
3) To determine the effectiveness of VTA in the learning of speaking skills;
4) To examine the effectiveness of VTA in PWM enhancement;
5) To investigate students' and the teacher's opinions of VTA as a manner of learning to speak English and training PWM.

1.5 Research questions

The above-mentioned objectives are accordingly translated into five research questions as follows:

1) Is VTA effective for learning to speak in English? If yes, in what ways?

2) Is VTA effective for enhancing PWM? If yes, in what ways?

3) Is there a relationship between students' PWM performance and speaking skills? If yes, in what ways?

4) What are the students' opinions of VTA as a way of learning to speak English and enhancing PWM?

5) What are the teacher's opinions of the implementation and outcomes of VTA as a way of learning to speak English and enhancing PWM?

1.6 Delimitations of the study

Delimitations are defined as the boundaries and scope of a study (Simon, 2011). Even though delimitations are in the control of the researcher, they might affect the external validity, which refers to the extent to which the findings of a research project can be generalized and applied to a larger population in different settings of a study (Gall et al., 1996). The delimiting factors are explicated as follows to carve out the own domain of the present study.

1) The subjects for the research comprised 80 students from Grade 3 in YAPS. The sample was drawn on the basis of accessibility, convenience and availability. The control and experimental groups were selected on the basis of class enrollments. As it is known, in any form of research, it would be ideal to test the entire population. In practice, however, due to budget, time and feasibility, it is impossible to include every individual of the

population. As Piaget (1973) suggests, because of neuroplasticity, children at a younger age, especially before puberty, are able to learn to speak a foreign language more easily because their neural maturational process is still undergoing evolution, and is thus optimal to receive varied neural stimulation. So the present study focused on Grade 3 students who had just learned English for half a year and those from other grades were excluded.

2) Baddeley and Hitch (1974) proposed that working memory could be divided into three subsystems: the phonological loop, the visuospatial sketchpad and the central executive. Over the years, Baddeley and Hitch's three-component model has been shown to be valuable. However, according to Baddeley (2000), there are still a number of phenomena that could not be taken into account by the original model. This prompted him to propose a fourth component of WM, the episodic buffer, which was assumed to provide temporary retention of multidimensional episodes or chunks. As the phonological loop is concerned with storage and retrieval of verbal and acoustic information and contributes significantly to the learning of a foreign language, so the present study only concentrated on the training of phonological working memory (PWM) instead of the whole system of working memory (WM).

1.7 Definitions of key terms

1) EFL learners: EFL learners in the current study refer to Grade 3 students in the Yiliang Anjiaqiao Primary School (YAPS), Kunming City, Yunnan Province, China, who had just learned English as a compulsory subject as required by the government for one semester.

2) Comprehensive language competence: Competence in a second language is one of the key fundamental concepts in the field of Applied Linguistics. However, its definition has always been a topic of intense debate (Iwashita, 2007). There are two models that are well-known in the literature for defining language competence. The first was advanced by Canale and Swain (1980), which regards language competence as consisting of grammatical, sociolinguistic and strategic knowledge and skills. The second model, proposed by Bachman and Palmer (1996), claims that language competence is composed of linguistic (vocabulary, grammar, and phonology, etc.) and non-linguistic knowledge (pragmatic, sociolinguistic and functional knowledge, etc.) and strategic competence (cognitive and metacognitive strategies). In the spirit of the two models, the national English curriculum standards stipulate that the principal aim of primary English teaching is to improve students' comprehensive language competence which involves language skills (listening, speaking, reading and writing), linguistic knowledge (vocabulary, grammar,

phonology, spelling, etc.), learning strategies (cognitive, metacognitive, resources and communicative strategies), emotional attitudes (interest, motivation and attitudes) and cultural awareness (cultural knowledge and intercultural communication competence).

3) Filtered sentences: Filtered sentences in this study refer to sentences filtered through a low-pass filter. The cutoff point is set at about the upper end of the fundamental frequency of a sound (F_0), which in speech is determined by the rate of vibration of the vocal cords in the larynx. Such a procedure of filtering deliberately removes vowel and consonant sounds and makes the intonation and rhythm patterns salient. In the meantime, auditory training with low-pass filters enables the listeners to restructure their rhythm and intonation skills and this lays the foundations for the perception of segmental elements, grammatical structure, and emotions of a speaker (Kent & Read, 1992; Asp, 2002).

4) Phonological working memory (PWM): Working memory (WM) is concerned with the transient and temporary storage and processing of information which is presumed to be of essential importance for various kinds of complex cognitive activities. As early as 1974, Baddeley and Hitch proposed a three-component model of the human working memory system. The first component is the phonological loop concerning the retention of verbal and acoustic information. The second is the

visuospatial sketchpad taking charge of visual signals, and the third is an attention-control system with limited capacity, the central executive. As mentioned earlier, a fourth component, the episodic buffer was later proposed by Baddeley in 2000, which was assumed to provide temporary storage of multidimensional episodes or chunks. PWM, as an integral part of working memory, as touched upon earlier, is the system which is in charge of temporary storage and retrieval of verbal-acoustic information. In other words, it is the capacity of the verbal component of working memory.

5) Verbotonalism: The verbotonal system was proposed by Petar Guberina from the Institute of Phonetics of the University of Zagreb. It originally applied auditory-based strategies to treat hearing-impaired children and adults by maximizing their listening skills, and at the same time, developing intelligible speech through binaural listening (Asp, 2006). It has also been adopted in the teaching of foreign languages. It holds the view that condition such as hearing impairment (deafness) can be thought of not so much as a physical deficit but as a different way of organizing audition. So its primary aim is to nurture a profound, optimal and lasting conditioning of audition and restructure the perception of learners' brain by changing the way they perceive and produce language.

6) Speaking skills, speaking ability, speaking proficiency, and speaking

performance are used interchangeably in the present study to denote learners' accurate and effective use of linguistic and sociolinguistic resources to achieve a range of communicative functions.

7) Self-directed or self-adjusted learning: Self-directed or self-adjusted learning means that during the learning process, students take the initiative to manage, regulate and assess their own learning activities.

1.8 Summary

In this chapter, an introduction to the background of the research was presented, including the description of the history and status-quo of primary English instruction in China and the problems facing the students and teachers in the learning and teaching of English speaking. After that, the significance, objectives, research questions and limitations of the study were outlined. And the last part of this section explained the definitions of the important terms appearing in the thesis. In the forthcoming chapter, a critical review of relevant theories and previous studies in this field will be provided and the theoretical framework of this study will be expounded.

CHAPTER 2

LITERATURE REVIEW

This chapter critically reviews the literature pertinent to the study. It is composed of six sections. In the first section, the notion of learning and a postmodern view of learning are discussed. In the subsequent section, previous studies on the componential parts of speaking proficiency as well as speaking instruction and assessment are presented. The third section profiles the relationship between phonological working memory (PWM) and language learning. The fourth section provides an introduction to Verbotonalism, and reviews previous research on the verbotonal approach. In the fifth section, the theoretical framework that informs the present study is explicated. And the last section constitutes the summary of the whole chapter.

2.1 Theories of learning

2.1.1 What is learning?

To develop a well-founded approach for language learning and teaching, a question emerges as the starting-point for investigation. The question is: what is learning? As learning is very complex in nature, to arrive at a generally accepted definition of the concept becomes completely impossible (Illeris, 2003, 2009). Hence,

understandings of learning from different perspectives and epistemological beliefs have been generated. For example, Jarvis (2009) suggested that:

> Human learning is the combination of processes throughout a lifetime whereby the whole person-body (genetic, physical and biological) and mind (knowledge, skills, attitudes, values, emotions, beliefs and senses)-experiences social situations, the perceived content of which is then transformed cognitively, emotively or practically (or through any combination) and integrated into the individual person's biography resulting in a continually changing (or more experienced) person (p. 25).

Illeris (2007) referred to learning as *"any process that in living organisms leads to permanent capacity change and which is not solely due to biological maturation or ageing"* (p. 3). Kolb (1984) claimed that *"learning is the process whereby knowledge is created through the transformation of experience"* (p. 38). Obviously, this is not an exhaustive list. It deserves mention that although these modern understandings of learning are much more thorough and profound than the traditional conceptualization of learning as the acquisition of knowledge and skills, they are still deficient in taking the nature of learning into full account. In the researcher's view, learning involves first and foremost understanding. In other words, if we intend to learn something, we have to understand it. To put it another way, every act of learning implies an act of understanding, otherwise, what is to be learned remains outside the individual's sphere of (either conscious or unconscious) awareness. As Lian (2011) argues, to learn, first, we have to make sense of what is happening. One generates meanings and understandings of the outside world internally on the basis of his/her personal history

which consists of his/her entire functional personal knowledge and understandings and the things that he/she has learned (Lian, 2004)). Because of the different biological and cultural backgrounds as well as idiosyncratic experiences, each individual's personal history is doomed to be individualized, and helps them make sense of the world, but at the same time, limits their understandings and perceptions (Lian & Pineda, 2014). As a consequence, we never perceive the world as it is; we perceive it as we are (Lian, 2004).

Taking the above ideas into account, there are two basic principles regarding learning. **The first** principle is that learning should be able to raise learners' awareness of the things to be learned and to change learners' perception mechanisms and bring into their personal history that which had previously been unconsciously excluded and ignored by them. As maintained by Andrew Lian (2000), to some degree, learning *"is an act of violence"* in that it violates *"the regularities inferred by the individual on the basis of the past"* (p.3). In this regard, to learn is to change one's personal history (Lian, 2001). As mentioned before, because of the differences in personal history, each learner's learning needs, learning approaches and learning interests are individual-specific. Their ways of perceiving and processing information are inevitably different. Any teaching or learning system should respect learners' individuality and idiosyncrasy and make provision for this. **The second** principle is that learning is an active process of knowledge construction. It requires active participation, initiative, creativity and self-organization on the part of the learner. Learners' awareness of their active role in learning should be raised. Learning needs to be organized around the learner's active

effort to resolve problems emerging from the learning process. Teachers' role is to support and facilitate that construction and present problems that will stimulate learners' curiosity, arouse their interest and activate their motivation. The capability of the teacher/tutor to make use of teaching skills to improve students' learning determines the quality and success of any instructional method which is aimed to develop and challenge students' thinking or reasoning skills and help them to self-adjust their own learning (Barrows, 1992, Cunningham & Duffy, 1996).

In the case of foreign/second language learning, the learners' ultimate aim is to function competently in the target language and culture. Like learning in general, language learning is also a matter of awareness-raising, understanding and meaning management (Lian, 2000). In learning a foreign language, to perceive, parse, understand and produce these new language signals effectively, learners, most of all, have to generate meanings from them. These meanings are not the ones which can be found in dictionaries or textbooks, but are constructed by meaning-making mechanisms generated in the heads of each individual learner. Therefore, any effective language teaching or learning system must be able to foster learners' sensitivity to and "feel" towards the target language by raising their awareness of the language phenomena of interest, including rhythmic and melodic aspects, to help learners use their personal *"meaning-making mechanisms (the 3Ms)"* (Lian, 2001, p.2) to modify their understandings of the new symbolic system by "*enabling them to confront, contrast and contest (the 3Cs) their personal understandings, beliefs and personal logics*

against what they are observing" (ibid), and, as a result, create an environment in which learners take on responsibility for their own learning.

So far, we have engaged in a fundamental thinking of learning in general and language learning in particular. These thoughts are consistent with the postmodern view of learning, which will be discussed in the part to come.

2.1.2 A postmodern view on learning

It is difficult to pin down the precise definition of postmodernism, although an extensive literature attempts to do so (Lyotard, 1984; Foster, 1985; Connor, 1989; Boyne & Rattansi, 1990; Lash, 1990; Wakefield, 1990; Best & Kellner, 1991; Featherstone, 1991; Harvey, 1991; Bauman, 1992; Crook et al., 1992; Rosenau, 1992; Smart, 1992; Seidman & Wagner, 1992). Marshall (1992) thinks that postmodernism is probably something that is unnamable. All we can say is that postmodernity is so complicated and multiform that it refuses lessened and simplistic explanation and explication (Usher & Edwards, 1994). As Smart (1994) states:

> The postmodern problematic has been invoked to distinguish an historical period, an aesthetic style, and a change in the condition of knowledge; to conceptualize difference-a distinctive form beyond the modern-as well as similarity-a variant of the modern or its limit form; and to describe affirmative or reactionary and critical or progressive discourses and movements. (p.164)

So in the light of Smart's opinions, postmodernism can be seen as not only a system of ideas and thoughts that confront contemporary social and cultural phenomena, but also a way of practicing - there is wide application of postmodern approaches to architecture,

art, and literature (Usher & Edwards, 1994). When it comes to learning, over a long period of time, learning is treated as incompatible with postmodern thoughts and concepts because, historically, educational theories and practices are created in the context of modernity. And the very purpose of modern education is to achieve modernity's *"*grand narratives*"*, and the enlightenment and emancipation of human beings' minds (ibid, p.30). Although education in the modern sense purports to cultivate rational individuals who are able to be self-directed and self-motivated in educational process, the actual reality is that there is power, observation, surveillance and oppression penetrating every corner of teaching and learning (Foucault, 1980). As a consequence, we are disillusioned to discover that the theoretical rhetoric of modern education runs counter to its practical reality. Hence, postmodern learning theories and practices have begun to gain prominence in the field of education (Usher & Edwards, 1994). The following space will be reserved to profile the major characteristics of learning from a postmodern perspective and its implications and guidance to the development of the teaching model in the present study.

Postmodern learning is first characterized by uncertainty. Postmodern scholars object strongly to placing human actions, including learning, on any foundational certainty (Usher & Edwards, 1994; Briton, 1996; Usher et al., 1997; Bagnall, 1999). This loss of certainty or certainty of uncertainty in the age of postmodernism serves as a point of departure in exploring human beings' learning (Kang, 2007). As learning is always closely associated with the production and dissemination of knowledge, we will

concentrate on postmodern ideas and thoughts about knowledge. Traditionally, knowledge is seen as something that is certain, true, valid and stable. So it is necessary and legitimate for a competent teacher who possesses it not only to impart it to others who do not possess it but to prescribe *"what to do, how to behave, what ends to pursue and by what means"* (Bauman, 1992, p.9). So according to Bauman (1987), human beings are subject to this kind of knowledge and knowledge-imparters, and have *"the infinite capacity to be acted upon, shaped, perfected"* by them (p.68). Similarly, Foucault (1979), one of the prominent postmodern scholars, argued that power was pervasive in every aspect of human life and the construction and impartment of knowledge in the modern sense was inescapably the product of power. For him, individuals' acquisition of the *"power knowledge"* is a process of being oppressed, coerced, and controlled (Foucault, 1979, p.89). As they are deprived of the rights to realize their inherent rationality, express themselves freely and develop themselves fully, they are more likely to be *"docile bodies"* and *"obedient souls"* (ibid), which contradicts the very objective of education. On the contrary, postmodern thoughts and insights deny the existence of precisely prescribed, powerful and stable knowledge and embrace the notion that knowledge is tentative and contextual (Kilgore, 2001). Human beings are oftentimes required to deal with the uncertain, the undiscovered, the unanticipated, the uncategorized, and the unpredicted and unpredictable real life problems. To solve these problems, they have to take new forms of learning to suffice their specific needs as and when they arise. This uncertainty and unpredictability lead

to a new form of learning which occurs in a *"just in time, just enough* and *just for me"* fashion (Lian & Pineda, 2014).

The second characteristic of postmodern learning is multiplicity. The postmodernist view of learning posits that learning is a process of individual meaning-making (Lian, 2004). Meaning is never *"out there"* to be discovered but created and interpreted internally by individuals on the basis of their personal representational and logical systems which are constituted by their past experience (Lian & Pineda, 2014, p.13). As each individual's past experience is different, their act of meaning-making and interpreting is always individualized and shaped and changed by the social-historical context. Meaning in this sense is constructed through a process of convergence from multi-channel experience. If we are to understand something just from one perspective, we will be unable to understand it at all (Lian, 2004). That is why there is no "real" meaning, and "real" truth and "real" reality. As Derrida (1982) puts it, truth is multileveled, and there is no possibility of arriving at *"the one"* truth because it is dependent on comprehension and interpretation (p.13). Usher (1989) also argued that there was no reality but the experience of it. Reality does not stand outside language discourses through which we interpret reality and our experience. Therefore, any effective and efficient learning system should take the multiplicity of learning into full consideration, and respect the differences, plurality and heterogeneity of learning needs, learning styles and learning approaches.

The third characteristic of postmodern learning is its emphasis on creativity. As the majority of modern educational systems are dominated by a culture of standardized teaching and testing, they are flattening and stifling some of the basic skills and processes that creative achievement depends on, and turning highly imaginative children into conformist adults (Land & Jarman, 1992; Vint, 2005; Azzam, 2009; Robinson, 2011; Lian & Pineda, 2014). As stated above, postmodern scholars regard learning as a dynamic process of meaning-making. Human beings are no longer passive recipients of knowledge but are active creators of it. This idea finds its origin in constructive postmodernism's emphasis on creativity. Constructive postmodernism stresses the importance of human beings' internal creativity, initiatives, active achievement and active devotion (Griffin, 1997). An educational system should enable each individual to have the opportunity to develop their creativity and to benefit from that of others.

The fourth characteristic of postmodern learning is its encouragement of autonomy and empowerment. Autonomy and empowerment are the *"defining traits"* of postmodernism, and learner-centeredness, dialogic teaching, and learners' self-direction, self-governance, self-reflection and self-evaluation are strongly emphasized in postmodern pedagogy (Bauman, 1992, P.202). Postmodernism advocates that individuals have the innate propensities and capacities to develop themselves fully. So they should be empowered to obligate and control themselves from an insidious source or a source that is natural to them (Usher & Edwards, 1994). In other words, individuals

should be encouraged to become autonomous, and any external imposition and intervention should be eliminated. In education practices, to encourage autonomy and empowerment, learners need to be seen as active knowing subjects and be given independence and freedom to take charge of their own learning. The key question in education is to create a learning environment which can optimize self-determined learning (Hiemstra, 1994).

Postmodernism offers concepts and insights for thinking anew the efficacy of education at the personal and structural level (Usher & Edwards, 1994). However, postmodern approaches to language learning are somewhat underdeveloped (Lian, 2004). Thus, a postmodern view of learning was adopted as the basic intellectual framework to guide the design of the new learning approach (VTA) of this study. In VTA, the teacher was no longer viewed as the one who possessed "real" knowledge and should impart it to or impose it on the students who did not have this knowledge. The pedagogic procedures of VTA valued students' creativity and respected their individual meaning-making mechanisms and multiple needs in learning. Students who experienced VTA were encouraged to generate their own understandings of English, and empowered to negotiate meaning and resolve the problems emerging in the whole learning process. In this sense, the whole study was enveloped in a post-modern nebula. To locate the current research in the context of the previous studies, the forthcoming sections will be allotted to review related literature on the instruction of speaking skills and enhancement of phonological working memory (PWM).

2.2 Speaking skills instruction

2.1.1 The construct of speaking proficiency

As Harley et al. (1990) suggest, the conceptualization of the nature of language proficiency is of paramount importance to the development of language teaching methods. Although speaking proficiency is acknowledged as the chief goal of foreign language teaching (Batool, et al., 2015), its precise definition seems to be elusive. The term has been used quite differently from researcher to researcher, and there has been no consensus reached about what speaking proficiency entails (Galloway 1987; McNamara, 1996). Clark (1975) describes it as learners' capability for using the language to engage in real-life communicative activities whilst Van Lier (1989) proposes that speaking proficiency involves those aspects of communicative competence which are measured in oral interviews. It should be mentioned that one of the most popular and influential notions of speaking proficiency was given by the American Council in the *Teaching of Foreign Language Guidelines* (1985, 1999) (ACTFL), which considers speaking proficiency as communicative growth, and holds that it is made up of four factors including function, content, context, and accuracy. Quite a few studies attempting to investigate the relative contribution that individual factors of speaking performance made to the overall proficiency with reference to ACTFL Scale have emerged (Iwashita, et al., 2007).

For example, Adams (1980) conducted research to unravel the construct of speaking proficiency. In his study, 834 participants were required to take a Foreign

Service Institute (FSI) Oral Interview Test. Through the testers' ratings, the five subskills including vocabulary, accent, grammar, comprehension, and fluency, which were thought of as influencing speaking performance were identified. Adams also investigated the relationship between the overall global proficiency and the five subskills. It was found that vocabulary and grammar discriminated most and accent and fluency least between different levels of speaking proficiency.

Following Adams's work, Higgs and Clifford (1982) investigated raters' perceptions of the different contributions that each of the five component subskill features including vocabulary, grammar, pronunciation, fluency and sociolinguistic factors made to the overall notion of speaking proficiency. This is known as the Relative Contribution Model (RCM). More specifically, in their study, fifty foreign language teachers, instead of rating the speaking performances of learners, rated the relative contribution of the above-mentioned five subskills to the six different proficiency levels of the FSI rating scale. Their study revealed that vocabulary and grammar were the most important factors affecting all levels of proficiency. And as the proficiency level advanced, the role of other features, namely, pronunciation, fluency and sociolinguistic factors, also became increasingly important. They then presented the RCM to a group of experienced teachers to elicit their opinions and insights of the relative weight of the subskills in overall proficiency. According to the teachers, vocabulary and pronunciation were the most important skills at the lower levels, fluency and grammar being the least important. All subskills became equally important at the higher levels of proficiency.

McNamara (1990) administered a test named Occupational English Test (OET) to health professionals. A speaking test in the form of a situational role-play as an integral part of the whole test was given to the candidates in two sessions. Trained raters rated the two-session speaking performance in terms of the global score (overall communicative effectiveness) and five analytic scales (intelligibility, fluency, comprehension, appropriateness, and resources of grammar and expression). Meanwhile, the Rasch Item Response (RIT) model was applied to identify the construct of overall communicative effectiveness. They found that resources of grammar and expression contributed greatly to the variance of the overall communicative effectiveness (68% and 70% of variance explained in the two sessions). They reached the conclusion that resources of grammar and expression were the most significant determinants of the score of overall communicative effectiveness.

In light of Higgs and Clifford's (1982) RCM model, De Jong and Van Ginkel (1992) tested the speaking proficiency of 25 Dutch speakers of French as a foreign language. In their study, the 25 participants were required to perform separate tasks for pronunciation, sentence completion, picture description, strategic skills, and structured conversation. The rating for global proficiency was given after all other components had been rated and raters based the global rating on participants' performances on all the sections. Their findings echoed those of Higgs and Clifford's study, which showed that the relative contribution of the components varied between levels of global proficiency. Specifically, at the lower level, pronunciation contributed most to overall

proficiency. As the level went up, all the subskills played an equal part in determining the global proficiency.

It is worthy of mentioning that some researchers made an effort to explore the nature of speaking proficiency through the comparison of the scores of the transcripts of the oral interviews and the ratings of the speaking performance of learners. For example, Magnan (1988) analyzed the relationship between grammatical errors appearing in the transcripts of oral interviews carried out with 40 college students and the overall oral proficiency interview (OPI) ratings. And their study showed that there was a significant negative correlation between the percentages of grammatical errors and speaking proficiency scores. Similarly, Douglas (1994) administered a speaking test to six Slovak speakers of English. Two raters rated the performance from the aspects of comprehensibility, grammar and pronunciation. Transcripts of the speaking performances were scored in terms of grammar, vocabulary, fluency, content and rhetorical organization. Then, the actual speaking performances of the participants who received similar scores on the test were compared. It was found that there was very little relationship between the scores on the test and the speaking proficiency of subjects. The researcher concluded that although scores on speaking tests were more likely to be used as evidence for both learners' language proficiency and the second language acquisition process, similar ratings on a speaking test are shown to represent qualitatively different learner performances.

More recently, Iwashita et al. (2007) examined the nature of English speaking proficiency in the context of a large project in order to develop a rating scale for TOEFL iBT. The data of their study were initially collected as part of the piloting of prototype tasks for TOEFL iBT. The speaking performance representing five different tasks and five different proficiency levels (200 performances in all) were double-rated by trained Educational Testing Service (ETS) test development staff, using a range of measures of grammatical accuracy and complexity, vocabulary, pronunciation, and fluency. The results demonstrated that all the features contributed to differentiate overall speaking performances, with vocabulary and fluency being the strongest determinants across all levels.

De Jong et al. (2012) conducted an in-depth study of the componential structure of second language (L2) speaking proficiency. One hundred and eighty-one nonnative and fifty-four native speakers of Dutch participated in their study. All the participants performed eight speaking tasks. Performances in the speaking tasks, the dependent variable, were rated by a panel of judges on functional adequacy. The independent variables including linguistic knowledge (vocabulary and grammar), linguistic processing skills (four reaction time measures obtained in three tasks: picture naming, delayed picture naming and sentence building) and pronunciation skills (speech sounds, word stress, and intonation) were also assessed separately. The results showed that all the linguistic skills were significantly related to functional adequacy of speaking (76% of the variance explained). And the researcher drew the conclusion that L2 speaking

proficiency was composed of linguistic knowledge components and language-processing and pronunciation skills.

Clearly, the above is not an exhaustive list of literature analyzing factors which influence speaking proficiency. As mentioned before, although it is difficult to reach a precise definition of speaking proficiency, a general agreement can be obtained by reviewing the literature, that is, speaking proficiency includes learners' accurate and effective use of linguistic and sociolinguistic features to achieve a range of communicative functions. Furthermore, the above-mentioned studies reveal that factors such as vocabulary, grammar, pronunciation, fluency and comprehensibility have been identified as the principal determinants of the overall speaking proficiency of an L2 speaker. And these factors contribute differently to overall speaking performances at different levels. Factors such as vocabulary and pronunciation are found to be the most important at the lower level, and all the factors are observed to make an equal contribution to global proficiency at the higher level.

2.2.2 A historical overview of teaching speaking skills

For centuries, formal education has been recognized as the principal means to learn a second/foreign language. Thus, a variety of teaching methods and approaches have been devised (Ariza, 2002). There is a general consensus that speaking is the most important and rewarding aspect of language learning (Nunan, 1991; Hall & Austin, 2004). However, speaking has been a neglected skill in foreign language education (Ur, 1996) and the teaching of speaking tends not to be an objective in its own right (Alonso,

2014). As Bygate (2009) says, *"speaking is often seen as how it can contribute to language acquisition in general"* (p.4). As a consequence, the following part will be allocated to review the teaching of speaking implemented in the influential methods and approaches practiced in foreign language teaching in general.

To begin with, the grammar-translation method (also named as the classical method) which was in vogue in the 16th century, holds the belief that different sections of human's brain are responsible for the manipulation and processing of different kinds of knowledge and studying varied subjects is an effective way to exercise the brain (Ariza, 2002). Thus, it is believed that learning a second/foreign language can activate and develop the functional part of the brain which is assumed to be earmarked for languages. According to this method, the primary goal for learning a language is not for speaking and/or communicating but for the exercise of the mind and being able to read in that language (Fotos, 2005). In practice, this method places emphasis on the learning of grammatical rules through rote memorization and recitation, translation being used as a means to facilitate teaching and comprehension (Chang, 2011). Teaching is conducted primarily in the students' first language. And little attention is paid to the correction of pronunciation of the target language (Liu & Shi, 2007). Even though students spent several years learning a foreign language, they were found not to be able to speak that language (Ariza, 2002).

Totally different from the grammar-translation method, the direct method, also known as the "natural" or "conversational" method, was underpinned in 1884 by the

German scholar and psychologist F. Frankel who theoretically justified the method by clarifying the direct association between forms and meaning in the target language (Batool et al., 2015). According to this method, the principle aim for students to learn a language is to be able to speak. So the instruction and classroom activities are conducted using the target language, the use of the native language being avoided in this method. Speaking is taught chiefly through conversational activities revolving around specific topics. In addition, the importance of teaching pronunciation by mimicking that of their teachers, the primacy of oral practice, the significance of the use of objects, visuals, and realia and the necessity of learner-centered instruction are all emphasized in the teaching of speaking (Stern, 1983). Another point deserving mention is that both the theoreticians and practitioners of the method advocated the *"avoidance of written or printed words until the pupil's pronunciation was so sound that it would not be influenced by seeing how the words were spelt"* (Richardson, 1983, p.38). Obviously, the direct method offers an exciting and interesting way for the students to learn speaking through activities, and at the same time, it turns out to be effective in freeing students from the inhibitions that are often associated with speaking a foreign language, especially at the early stages (Batool et al., 2015). However, the method is also subject to criticism because of its strong reliance on the teachers' expertise, for example, a native-like accent and fluency in the target language (Richards & Rodgers, 2014).

Another popular method adopted for teaching speaking is entitled the audio-lingual method (ALM). The origin of this method dates back to the Second World War. Due to America's involvement in the war, the American government realized that there was an urgent need for people who could communicate in foreign languages. In response to this, the audio-lingual method focusing on aural/oral skills, was developed on the basis of behavioristic psychology (Nagaraj, 2003; Anggraeni, 2007). In the belief that language is primarily speech, not writing, the method makes an endeavor to improve students' oral proficiency and thus enable them to communicate effectively in the foreign language learned. The teaching of speaking is implemented by letting the students sit in language laboratories listening to and repeating recordings of dialogues. Besides, teachers organize lots of pattern drills to enhance students' mastery of sentence structures which are embedded in the dialogues. For instance, in a substitution drill, the original structure might be: I wanted to go to the park. Students are then asked to think out other words such as "supermarket", "department store" or "hospital" to take the place of "park". As in the direct method, the use of the native language in teaching is also not allowed in ALM. In this method, it is assumed that, through repetition, mimicry, memorization, recitation and drilling, students would form a habit and be capable of speaking the language automatically and effortlessly in genuine communicative situations. Nevertheless, the outcomes of this method fell short of expectations and students tended not to be able to put what they had learned in class into actual use outside the classroom. As analyzed by Ariza (2002), *"the audio-lingual was too*

prescriptive" and "*students were taught a script"* (p.65). So there was a lack of true communication in the foreign language classroom.

In recognition that the essential purpose of language learning is communication, a more holistic teaching approach, the communicative approach, was developed and became pervasive in the field of foreign language teaching (Campbell & Wales, 1970; Habermas, 1970; Clark, 1972). According to this approach, cultivating learners' communicative competence should be at the forefront of teaching a foreign language (Hymes, 1972). The teaching of speaking under the communicative principle follows a functional approach and stresses the pragmatic functions of speech (Alonso, 2014). The speaking instruction using this approach is characterized by the employment of a wide range of interactional activities and real-world tasks that are meaningful and focus on oral communication instead of grammatical patterns (Efrizal, 2012). Meanwhile, in the communicative approach, the interactive and cooperative learning between learners is strongly emphasized, and there are also opportunities for them to go outside of the classroom and take part in real-life communicative activities (Littlewood, 1981). The approach holds that plentiful exposure to language in actual use and plenty of opportunities to use it for real communication are crucial for the development of speaking skills (Harmer, 1998). As can be seen from the above discussion, the communicative approach boasts a vast number of strong points over its counterparts. However, there is still controversy about its little emphasis on grammatical patterns (Efrizal, 2012).

It is undeniable that the development and implementation of teaching methods should take students' learning styles, psychology, motivation and affection into account. Suggestopedia, introduced by Bulgarian psychiatrist-educator Georgi- Lozanov, is one of the methods which takes the above-mentioned factors into relatively full consideration (Felix, 1993). As Brownlee (1980) comments, Suggestopedia is very suitable for the teaching of speaking to young learners because there are audio-visual displays and physical activities used in the learning process. Meanwhile, a series of activities under the principle of desuggestive-suggestive communicative psychotherapy are employed by the teacher to motivate and encourage the students in learning. The key features of the methods in teaching speaking are as follows. 1) Comfortable classroom environment: The classroom is usually darkened. Then, soothing and soft music such as Baroque music is played to create an appealing and relaxing learning atmosphere. It is believed that Baroque music, with its 60 beats per minute and its special rhythm can create a state of relaxed concentration and thus facilitate the reception and retention of new-coming information (Lian, 1980; Bancroft, 2005). Besides, the students are imbued with the idea that they can be successful in learning because this method is based on the belief that if the students think they can learn, they will. 2) The use of songs, drama and games to practice speaking: It offers a more enjoyable and interesting way for the students to learn speaking compared with other methods mentioned above. Especially, the employment of creative dramatics in teaching not only stimulates students' interests in learning but also provides them with

more opportunities to practice speaking both in and out of class. 3) Tolerance of mistakes: To foster students' self-efficacy in learning, the teacher tends not to deliberately correct students' mistakes (Racle, 1976). However, this method is questioned because it is difficult for the teachers to apply it in large classrooms (ibid).

Besides suggestopedia, the total physical response (TPR) method proposed by psychologist James Asher (1969) has also proven to be effective for the teaching of foreign languages to young learners. Asher believes that foreign language learning should follow the developmental pattern that children learn their native language. TPR takes the position that one can learn best when being involved both physically and mentally (Larsen-Freeman, 1986). In teaching speaking skills using TPR, a teacher first gives a series of commands and students are required to respond non-verbally to them. In order to facilitate comprehension, the teacher models the performances. Then, after practicing many times, the teacher simply issues the commands, and students respond by themselves. Last, it is the students who present the commands orally, and the teacher responds nonverbally (Widodo, 2005). As in some other methods, in TPR, teachers usually utilize teaching aids such as images, pictures, and objects to make learning fun and easier. The biggest strength of this method is that it can encourage students with different levels of oral proficiency to participate actively in the classroom activities because at the beginning stages, the students are only required to respond to the commands nonverbally. But there are still some drawbacks in the method. As introduced before, teachers serve as both the verbal and non-verbal model for the

students. So this leads to a high demand for teachers' expertise, especially language proficiency. In other words, if there is a lack of qualified teachers, the adoption of TPR should be treated with caution.

2.2.3 Speaking skills assessment

Speaking skills assessment, as an implicit goal of a foreign language course, is critical in determining the learning outcomes. As Higgs (1987) claims, if there are no reasonably designed evaluation procedures, all the teaching goals, theories and methods, educational philosophies and pedagogical orientation will be futile and impractical. However, assessing the speaking performances of language learners in a practical, plausible and reliable manner is a very challenging task (Ferrara, 2008; Wolf et al., 2008; Cubillos, 2010). Underhill (1987) also states that a speaking test is different in nature from the tests of other language skills because it includes techniques and criteria which are exclusive to it.

With regards to techniques, which are referred to as means used to *"elicit a sample of language that can be scored"* (Fulcher, 2003, P. 50), Underhill (1987) in his book *Testing Spoken Language: A Handbook of Oral Testing Techniques* made a long list of the techniques that could be employed in the measurement of oral proficiency. They range from discussion, conversation, oral report, role-play, interview, description and re-creation, form-filling, question and answer, reading blank dialogue, use of a picture or picture story, giving instructions, explanation, retelling a story or text from aural or written stimulus, reading aloud, translating or interpreting, sentence completion

from oral or written stimulus, sentence correction and transformation, to sentence repetition. He thinks that a single technique marked on a global proficiency criterion will just lead to a quick but approximate assessment whereas a sequence of techniques will produce *"a more accurate and consistent result"* (Underhill, 1987, p. 231). He also points out that the choice of specific techniques in designing an oral test depends on the purposes and resources of the program.

Varela and Martínez (2013) examined the oral test modules of some of the best known English proficiency tests worldwide such as the Cambridge Exams in English for Speakers of Other Languages (ESOL), the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL), Trinity College London Exams (TCLE), Pearson Tests of English (PTE) and the International Test of English Proficiency (ITEP). The researcher found that the majority of the speaking examinations involved activities which intended to reproduce real-life situations, and although there were a large number of speaking testing techniques available, almost all the speaking tests were found to make use of very similar models and types of tasks. To be more detailed, these speaking tests typically start from an introductory interview, which aims to warm the candidates up through asking questions about their family, life experience and educational background. Subsequently, the tests turn to a more complex and independent discussion of a topic. At this stage, the candidates are given some time to prepare and then express their opinions. These tests usually end with an interactive task in which the candidates are required to conduct a real conversation with the examiner.

In addition to techniques, the criteria adopted to evaluate speaking skills are also of essential importance to a speaking test because they will have a significant impact on the feasibility, validity and reliability of the test (Fulcher & Davidson, 2007; Xi, 2008; Varela & Martínez, 2013). Despite the fact that there are apparently a large variety of criteria available for assessing speaking performances, the most frequently used ones by the majority of speaking tests are range, accuracy, speaking delivery, promptness of response and the use of different language functions (Varela & Martínez, 2013). Taking the internationally recognized language proficiency tests as an example, Cambridge ESOL exams establish the criteria in terms of grammatical resource, lexical resource, discourse management, pronunciation and interactive communication. TOEFL focuses on speaking delivery, use of language and topic development, and IELTS selects fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation as the criteria (Galaczi, 2005; Varela & Martínez, 2013).

It deserves mentioning that the criteria outlined by the Council of Europe in *the Common European Framework of Reference* (CEFR) are commonly referred to and *"used as a basis in speaking assessment"* (Luoma, 2004, P. 71). As stated above, there exists a variety of criteria being used to assess speaking proficiency. However, according to CEFR guidelines, it is reasonable for any practical assessment system to reduce the number of possible criteria to a feasible number because more than 5 criteria will lead to cognitive overload. Therefore, it formulates five analytic criteria, namely, range, accuracy, fluency, interaction and coherence (Council of Europe, 2001) for

speaking assessment. **Range** is concerned with the differing linguistic forms including words, phrases, sentence patterns and formulaic expressions used by speakers in speaking, in other words, the linguistic repertoire of a speaker; **accuracy** includes the appropriate use of grammatical and syntactic rules as well as accurate pronunciation; **fluency** refers to the capability of expressing oneself spontaneously and fluently by producing stretches of language with a natural colloquial flow; **interaction** is involved with the ability to initiate and control discourse including getting and keeping the floor smoothly, taking turns appropriately, keeping the conversation going and ending it when it is time to do so, and taking advantage of non-verbal features, and interact with other speakers skillfully by relating one's own contributions naturally to that of others; **coherence** denotes the ability to create coherent and cohesive discourse by the accurate use of a variety of organizational patterns, connectors and other cohesive devices. The appropriate selection of analytic rating criteria can ensure relatively reliable and effective assessment, and more importantly, present specific information which is helpful for learning development towards a higher level (Barton & Collins, 1997).

In the field of assessing the speaking ability of primary school students, Efthymiou (2012) applied portfolios to assess the speaking skills of Greek students. He constructed the analytic scales in terms of grammatical accuracy, fluency, scope of vocabulary, pronunciation and content. Westerveld and Moran (2011), in their study, set verbal productivity, syntactic complexity, grammatical accuracy and verbal fluency as the criteria. Srikaew et al. (2015) designed an analytic rating scale to assess English

speaking performances of Grade 6 Thai EFL learners. The assessment criteria in his study included vocabulary, syntax, cohesion, pronunciation, ideational function and fluency. The scale was examined and showed a high reliability. And according to the results of the data analysis, fluency proved to be the greatest variance component of the composite score of the analytic rating scale, followed by ideational function, cohesion, vocabulary and syntax, and pronunciation respectively.

In terms of research into speaking skills assessment in primary schools in China, it should be pointed out that there is no speaking test conducted in most of the schools due to the complexity of measuring speaking proficiency (Teng, 2010; Zhang, 2011). So, to date, there has been a severe shortage of empirical studies on speaking skills assessment. It can be found in the literature that only Teng (2010) made a comparison of the speaking evaluation criteria and rubrics adopted in the primary schools between China and foreign countries such as America and Chile. She discussed the applicability and adaptability of these foreign speaking evaluation criteria and rubrics in the Chinese context and argued that the teaching of speaking should be aimed at improving students' linguistic and socio-linguistic knowledge and skills.

According to Richards and Renandya (2002), the development of a language teaching pedagogy should involve issues of assessment because an effective speaking assessment system can serve as evidence for the outcomes of learning and teaching, and hence, feedback on the effectiveness of the teaching process. Surveying the techniques and criteria used in speaking assessment gives implications to the design of

speaking tests as well as the development of analytic scales to be used in the present study.

2.2.4 Studies on the teaching of English speaking skills in the primary schools of China

As preceding discussion indicates, ushering in the new era, English language education has been universally established in primary schools in China since 2001. However, the discouraging fact is that the English-speaking skills of primary school students have been generally considered as poor (Dong 2003; Hu 2007; Chen 2008; Hu, 2013; Zhang, 2014, Li, 2016). Although there are a large number of studies on primary English instruction, only a few of them focus deliberately and exclusively on the teaching of speaking skills.

Going over the existing literature, it can be concluded that research on the teaching and learning of speaking revolve mainly around the identification of the challenges and problems faced by teachers and learners, and exploration of teaching methods which attempt to improve students' speaking skills. Studies show that the speaking proficiency of students is generally low. And there are mainly three major problems or challenges existing in the teaching and learning of speaking (Hu, 2007; Teng, 2010; Zhang, 2012, Zhang, 2016). According to Zhang (2012), the severe scarcity of qualified teachers is the major challenge in teaching speaking in the primary EFL classrooms in China. The nationwide popularization of the English class casts a demand for a substantial number of primary school English teachers (Wang 2002; Hu 2005,

2007). And at a rough estimate, more than 300,000 teachers should be added to the existing pool of 200,000 (Wang, 2002). To add to the complexity, the quality of primary school teachers is in no way optimistic, even in the socioeconomically advantaged regions (Dong 2003; Shu & Zhang, 2003; Lin 2005; Wang 2007; Hu 2009; Wang 2009; Du 2010). Teachers are provided with very few opportunities to receive training in speaking pedagogy and professional development by the schools (Wang 2009). Also, teachers themselves do not have enough time to improve their expertise for speaking instruction because of their heavy workload (20 or more teaching hours per week) (Fang et al. 2004; Fu 2006; Zhang, 2011).

The second problem lies in the employment of inappropriate teaching methods by the teachers. Some of the teachers, especially those in the rural primary schools are inclined to accord speaking a secondary emphasis in teaching and still adopt very traditional pedagogy which attaches importance to the instruction of grammar and translation (Liu, 2006). In this teacher-centered instruction model, the question-and-answer exercise plays a dominant role in classroom activities. The monotonous teaching is found to demotivate students, although most of the students show very strong enthusiasm in learning to speak English at the very beginning. Moreover, in teaching pronunciation, teachers are usually the models for the students to learn to pronounce words. This causes the poor pronunciation on the part of the students (Zhang, 2012). The reason is that, the pronunciation of some teachers is inadequate. Even for those teachers, who are qualified in teaching pronunciation, they are unable to identify

the problems each student meets in learning pronunciation and perform immediate correction for them because of the limited teaching hours and large class size, a challenge that will be discussed in the forthcoming part. According to Nation and Newton (2008), if students can acquire a clear pronunciation, it will be easier for them to become familiar with the sentence patterns and grammatical rules of the target language. In view of what has already been discussed, it's unsurprising that most Chinese primary school students are not confident in speaking English, reporting that their problems are chiefly concerned with poor pronunciation and grammar, and limited vocabulary (Zhang, 2012).

The third problem lies in the large class size and lack of learning environment for the students to practice speaking. The huge number of the primary school students leads to the large class size. Li (2007) points out that in some of the primary schools in rural areas and small cities in China, there tends to be around 100 students in a class. So under such circumstances, it is not easy for the teachers to organize activities that enable each student to have an equal chance to speak in class. To make the situation worse, out of class, there are barely any chances for students to use English in real-life communicative situations (Chen, 2008; Zhang, 2012).

In respect to studies on the development of appropriate methods for teaching speaking, since the introduction of the communication-oriented teaching approach from the west, a growing number of linguists, educators and language teachers have accorded their attention to the search for teaching methods which are suitable for Chinese primary

school students. Some researchers suggest that teachers should abandon the traditional methods which focus only on pure delivery of linguistic knowledge, and implement situational, communicative and task-based pedagogies in which a variety of activities and tasks such as role-playing, debating, games, competition, substitution, singing songs, chanting, simulation, storytelling, English corner, and arts and crafts in English, etc. are organized for the students to practice speaking both in and out of class (Liu, 2001; Tang, 2003; Yi, 2009; Tian, 2014; Zhang, 2014). Other researchers realize the significance of resorting to technology to teach speaking. For example, Liu (2013) advocates that teachers should take advantage of multimedia resources to make their teaching more attractive, interesting and exciting. Bai et al. (2013) proposed the application of electronic schoolbag in teaching. Liang and He (2010) conducted an experiment to examine the effectiveness of the employment of multimedia resources in improving students' listening and speaking skills. In their research, students in the experimental group received the audiovisual presentation of the sentences, whilst those in the control group were only exposed to audio materials. The results showed that the experimental group outperformed the control group in both phonetic perception and simulation.

Reviewing what has been done so far in the research on speaking instruction in Chinese primary schools, most studies in the literature concern themselves mainly with exposing challenges and problems in the teaching and learning of speaking. There is a severe scarcity of empirical studies aiming to improve students' speaking skills.

More importantly, it can be observed that, the verbotonal approach has never been implemented in the teaching of EFL speaking in the primary schools in China. Thus, the present study sets out to develop an approach based on Verbotonalism, to attempt to solve, at least partly, problems in the teaching/learning of speaking skills for young children, and at the same time, seek to improve their' PWM, which will be discussed in the following section.

2.3 Phonological working memory (PWM) and language

2.3.1 Working memory (WM)

The concept of WM evolved from that of short-term memory (STM), the temporary retention of information over a short period of time (Baddeley, 1974, 2003). As early as 1949, in his book *The Organization of Behavior*, Hebb (1949, revised in 2005) first made a distinction between long-term memory (LTM), which involved stable and durable changes in the structure of neurons in the brain, a process termed as long-term potentiation, and short-term memory (STM), which he attributed to transient electrical activity (Baddeley, 2007). After that, the empirical studies carried out by Brown (1958) and the Petersons (1959) substantiated the fact that the effects of verbal stimulation would decay over a few seconds if the rehearsal of the stimulus was prevented. And forgetting of the verbal stimulation progressed at differential rates according to the amount of controlled rehearsal of the stimulus. So their studies presented evidence for the existence of a temporary STM system which they contrasted

with LTM. They put forward that the STM storage system was an important but neglected aspect of language acquisition. Inspired by their research, more and more scholars set their foot in exploring the two separate types of memory. It arose as a heated topic during the 1960s.

The term "working memory" was coined by Miller et al. (1960, cited in Baddeley, 2000), in their influential book *Plans and the Structure of Behaviour*, and has been used in a wide range of subjects such as cognitive psychology, neuroscience, cognitive science, education, psychiatry and paleoanthropology, etc. It was used by Atkinson and Shiffrin (1968) in their paper *Human Memory: A Proposed System and Its Control Processes.* In their paper, they differentiated the short-term store from the long-term store (STS & LTS). They posited that information came in from the environment into the temporary STS, which was the subject's WM, and was then transferred to a fairly durable repository-LTS. They pointed out that information entering into the STS decayed completely and was lost within a period of about 30 seconds, but by means of rehearsal, a factor of crucial importance in experiments on human memory, a limited amount of information could be maintained in STS if the subject intended to do so. They assumed WM as a single unitary store which was indispensable not only for long-term storage of information, but also for many other high-level cognitive activities, for example, reasoning and comprehension.

Although Atkinson and Shiffrin's work has been recognized as the most influential two-component model (Baddeley, 2000, 2003), it has been criticized mainly

for three reasons (Baddeley, 2012). According to Baddeley (2012), the first is its assumption that the mere retention of information in WM will guarantee long-term learning. This was identified as incorrect by Craik and Lockhart's (1972) study, which demonstrated that processing played a vital role in learning, and the more elaborate the processing was, the better the learning outcome would be. The second lies in its belief that WM is essential for access to LTM. However, neuropsychological evidence demonstrated that patients who had deficiencies in WM still had an unimpaired LTM (Shallice & Warrington, 1970). The third is its claim that WM was crucial in performing cognitive activities. According to Atkinson and Shiffrin's two-component model, patients with impaired WM would lead to cognitive deficits. However, related research indicated that patients with impaired WM were observed to still be able to engage in normal long-term learning and cognitive activities. These merged into a paradox.

In order to tackle the paradox, Baddeley and Hitch (1974) first distinguished WM from STM and then proposed a three-component model of WM. According to them, STM refers to the merely temporary retention of information, whereas WM implies both the storage and processing of information (Baddeley, 2000). As stated by Baddley (2010), by choosing the term "working memory", they intended to emphasize its role in both maintenance and manipulation of in-coming information, *"allowing it to play an important role in cognition more generally, hopefully providing a framework and a set of techniques that could be applied practically to the wide range of activities for which working memory might be important"* (p.138). In their model, the unitary WM model

was separated into three components, which were considered as working together as part of a unified working memory system that served the function of facilitating the performance of a range of complex tasks. The three components are 1) the phonological loop, responsible for verbal and acoustic information; 2) the visuospatial sketchpad, for visual materials; and 3) the central executive, for controlling attention.

Over the years, Baddeley and Hitch's three-component model has been shown to be useful, resilient and successful. However, according to Baddeley (2000), there are still a number of phenomena that cannot be taken into account by the original model. This drove him to propose a fourth component of WM, the episodic buffer, which is assumed to provide temporary retention of multidimensional episodes or chunks. Baddeley (2010) maintained that the episodic buffer served as a platform in which all the components of WM could *"interact through participation in a multidimensional code, and can interface with information from perception and LTM"* (p.138). Conscious awareness is believed to be the primary mode of retrieval from the buffer. The revised model (see Figure 2.1) is different from the original one mainly in that its emphasis is on the processes of integrating information, instead of on the isolation of the components. In this perspective, the revised model is better in dealing with *"the complex aspects of executive control in working memory"* (Baddeley, 2000, p. 1) and can be applied across a wide range of cognitive activities.

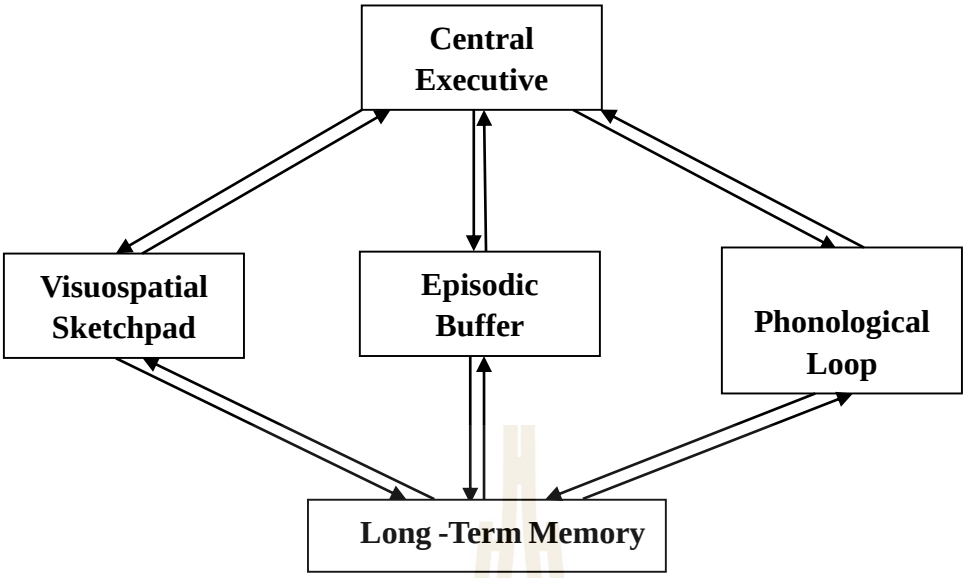


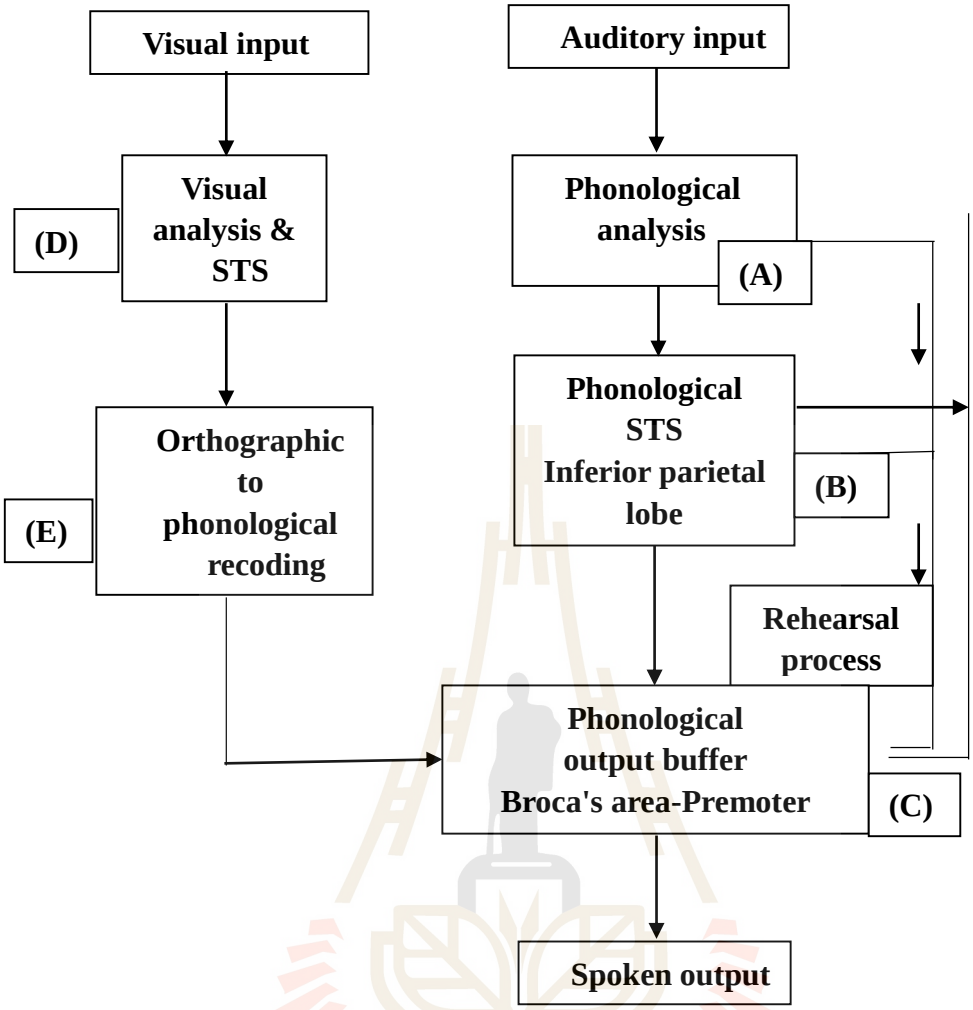
Figure 2.1 The working memory model

2.3.2 Phonological working memory (PWM) and the measures of PWM

The most widely studied and best-developed aspect of WM is the phonological loop (Baddeley et al., 1998; Baddeley, 2003, 2007, 2010) because it is the most tractable system to study (Baddeley, 2012). The phonological loop, being an important component of WM as discussed in the section above, is specialized for the temporary storage of verbal-acoustic information over brief intervals of time, which is referred to as PWM (Baddeley & Hitch, 1974; Baddeley & Larsen, 2007). Baddeley and Hitch (1974) proposed that the loop consisted of two subsystems (see Figure 2.2), a limited-capacity store which maintained verbal information over short periods of time, and a subvocal rehearsal system which not only held verbal-acoustic information in the store, but was also responsible for registering visual information within the store, provided the items can be named. Hence, if a subject is presented visually a sequence of digits

for immediate recall, the subjects tend to subvocalize them first. So in this case, the storage of these digits will depend principally on their phonological characteristics (Baddeley, 2003).

The model proved to be capable of taking account of a great deal of data from normal adult subjects, children, and neuropsychological patients (Gathercole & Baddeley, 1993; Baddeley, 1997; Baddeley et al., 1998). Studies which used the popular tools such as the phonological similarity effect, the word length effect, articulatory suppression, irrelevant sound effects and retaining serial order to investigate the loop presented evidence for the two-component model (Baddeley, 2012). Besides, neuroimaging studies further manifested that the phonological store was associated with Brodmann area 44, the cortical area, whereas the subvocal rehearsal system appeared to be associated with Broca's area (Brodmann areas 6 and 40) (Baddeley, 2003). Hence, PWM was believed to be the activation mainly of the left hemisphere of the brain. However, a few scholars suggested that there were synchronous activities occurring in the right hemisphere under certain conditions (ibid).



Cited from Baddeley (2003, p.193)

Figure 2.2 A proposed structure for the phonological loop

One of the most widely used measures of PWM is the **auditory span task** (Daneman & Carpenter, 1980; Engle & Oransky, 1999; Melby-Lervåg & Hulme, 2013) which aims to measure the maximum length of sequence of verbal items that each subject can correctly recall (Baddeley, 2012; Gathercole & Baddeley, 2014). In the auditory span task, the researcher usually speaks aloud a series of items ((e.g., letters, words, digits) which the subjects are required to repeat immediately in correct serial

order. At least, two lists of verbal items are provided at each list length, usually beginning with length two. If the two sequences at each length are correctly repeated by the subjects, the length of the next list is increased by one, and a further two lists are presented. If subjects repeat both of the two lists at one length incorrectly, no further lists will be offered. If they are able to recall only one of the first two lists, a third list at that same length will be given. If subjects can repeat the third list successfully, they can progress to lists at the next length. If they fail to repeat the third list, then the testing will end. Hence, their PWM is scored as the maximum length at which they correctly recall at least two lists (Gathercole, 1995; Gathercole et al., 1999). As the auditory span task is effective in measuring the capacity of subjects' phonological loop, it has been adopted as a subtest in a wide range of standardized ability test batteries (Baddeley et al., 1998), for example, the Wechsler Intelligence Scale for Children (Wechsler, 1974), Wechsler Adult Intelligence Scale (Wechsler, 1981), and the British Abilities Scales (Elliott et al., 1983), etc.

In addition to the auditory span task, there is another very popular measure of PWM, **nonword repetition (NWR)** (Gathercole & Baddeley, 1994; Gathercole, 1995). Evidence shows that PWM holds the key for the learning of new vocabulary (Gathercole & Baddeley, 1990a, 1990b; Baddeley et al. 1992; Baddeley & Wilson, 1993; Gathercole et al., 1997; Baddeley, 2013; Thorn, 2013; Baddeley et al., 2014). There are also studies demonstrating that NWR is the best predictor of the acquisition of vocabulary both in foreign and native languages (Gathercole et al., 1994; Gathercole,

1995, 2006; Rispens & Baker, 2012). Baddeley et al. (2014) found that the ability to repeat an unfamiliar phonological form – in a sense, the first step in learning a new word – is largely dependent upon the capacity of PWM. Their findings led to the hypothesis that NWR performance was more sensitive to phonological loop function than the digit span measure (Baddeley et al., 1998; Gathercole, 2006; Baddeley et al., 2014). As explicated by Gathercole (1995), *"the phonological form of a nonword is necessarily unfamiliar, it has been argued that the listener has to rely heavily upon the capacity of their phonological memory system to encode and maintain the novel phonological sequence for sufficiently long"* (p.83).

A NWR test should seek to avoid the constraints of long-term lexical and phonotactic knowledge (Gathercole, 1995; Gathercole et al., 1999). That is to say, nonwords need to be devised with syllables that do not appear as real words in the ambient language and the CV (consonant and vowel) combinations should conform to the phonotactic rules and prosodic patterns of the language in question but cannot be highly predictable (Dollaghan & Campbell, 1998) because nonwords with high wordlikeness will result in the reliance on the long-term lexical knowledge to facilitate performance on a NWR test (Gathercole, 1995). The most influential NWR test is The Children's Test of Nonword Repetition established by Gathercole et al. (1994). The test comprises 40 nonword stimuli such as *diller*, *frescovent*, *stopograttic*, and *versatrationist*, 10 each containing two, three, four and five syllables. All the 40 nonwords are rated in terms of their wordlikeness, 20 being high-wordlike such as

hampent, prindle and *commerine*, 20 being low-wordlike such as *tafflest, skiticult* and *frescovent*. In the test, the experimenter usually speaks aloud each nonword with his/her mouth obscured. Subjects repeat the nonword immediately after hearing it. The experimenter then scores each repetition dichotomously (correct or incorrect). The number of nonwords correctly repeated by the subjects is calculated as their NWW performance (Gathercole, 1995).

Given that nonword repetition has been recognized as a relatively pure measure of the capacity of the phonological loop (Baddeley et al., 1998), the present study adopted NWR as the measure of PWM. And all the nonwords in the NWR test were selected from The Children's Test of Nonword Repetition invented by Gathercole et al. (1994).

2.3.3 Review of research on PWM and language learning

Related studies have revealed a close association between PWM and the acquisition of both native and foreign languages (Gathercole & Baddeley, 1989; Service, 1992; Gathercole et al., 1992; Michas & Henry, 1994; Kohonen, 1995; Avon et al., 1998; Gathercole et al., 1999; Dufva & Voeten, 1999; Bowey, 2001; Baddeley, 2012; Baddeley et al., 2014). There is extensive literature examining the relationship between PWM and L1 vocabulary development. For example, Gathercole and Baddeley's (1989) study showed that subjects with good PWM skills have been consistently found to have greater vocabulary knowledge and better new word learning ability in their native language than those with poorer memory performance. Their findings were confirmed

by a series of studies conducted later (Michas & Henry, 1994; Gathercole, et al., 1997; Avons et al., 1998; Baddeley et al., 1998; Gathercole et al., 1999; Baddeley, 2012). It should be noted that Gathercole and his colleagues (1992) revealed the reciprocal relationship between PWM skills and vocabulary knowledge. There are also several studies investigating the link between PWM abilities and L1 reading skills. For instance, Gathercole (1995) found that PWM performances played a vital role in the development of reading skills in the early stages. Similar results were also obtained by other researchers such as Muter (1998), Seigneuric et al. (2000), Swanson and Howell (2001), Kintsch and Rawson (2005), and MacFarlane et al. (2012). In addition, the correlation between L1 language production and PWM capacity interests a few scholars. As early as 1989, Levelt discussed the significance of the short-term storage of information in language production. Adams and Gathercole (2000) studied the relationship between verbal and visual short-term memory and language development, more particularly, spoken language. Their findings showed that children with good verbal and visual short-term memory abilities were capable of using a larger number of different words in their speech, and they could produce longer sentences and more varied sentence structures.

In terms of the correlation between PWM and foreign language learning, a large amount of evidence has indicated that PWM is crucial in L2 vocabulary learning. Baddeley et al. (1988) conducted a study with a neuropsychological patient, P.V., who had an impaired phonological loop. She had no difficulty in learning native language

vocabulary but had problems in learning foreign language words. The important role of PWM in foreign vocabulary learning was further underpinned by the research carried out by Papagno et al. (1991), which revealed that normal subjects' ability to learn foreign language words was inhibited when their PWM was interfered by using the treatment of articulatory suppression. Masoura and Gathercole (1999) also examined the relationship between PWM skills and children's performance in learning both native and foreign language words. They found that children's ability to learn foreign rather than native vocabulary highly depended upon their PWM. Besides, several studies demonstrated that PWM appeared to be more closely linked to early stages of vocabulary learning, when learners had smaller long-term linguistic knowledge (Masoura & Gathercole, 2005). Taking a look at the more recent literature, Farnia and Geva (2011) studied the concurrent and longitudinal relationships between phonological awareness and PWM and vocabulary learning. In their study, mutual facilitation was found to exist between phonological awareness, PWM, and vocabulary learning. Apart from vocabulary learning, PWM skills have also been found to predict other aspects of foreign language learning such as reading comprehension (Swanson & Howell, 2001; Swanson et al., 2004), writing skills (Kormos & Sáfár 2008; Mackie et al., 2013), listening comprehension (Service & Kohonen 1995; Dufva & Voeten 1999; Kormos & Sáfár 2008), and grammatical structures (Williams & Lovatt, 2003; O'brien et al., 2006; French & O'Brien, 2008).

Compared with the above-discussed aspects of foreign language learning, studies on the correlation between PWM and speech production remain to be insufficient (Martin & Ellis, 2012). One typical study in this field was conducted by O'Brien et al. (2006). Their study examined the role of PWM in L2 speech production in terms of lexical, narrative, and grammatical abilities by English-speaking adults learning Spanish. Their findings suggested that PWM was an effective predictor of the development of narrative skills at earlier stages of L2 learning and grammar learning at later stages. O'Brien et al. (2007) took a further step to investigate the link between PWM and L2 oral fluency gains in native English speakers learning Spanish at their home university or abroad in an immersion context. They found that PWM contributed significantly to the development of L2 oral fluency.

As discussed above, PWM makes a significant contribution to various aspects of both native and foreign language learning. Most researchers attribute the close association between PWM and language learning to the function of the phonological loop for constructing stable, permanent memory representations of novel phonological material (Gathercole & Baddeley, 1993). These memory representations are considered to be crucial to language learning (Gathercole, 1990; Martin & Ellis, 2012).

2.3.4 Training and plasticity of working memory (WM)

WM capacity predicts performances in general cognition (Shipstead et al., 2012). However, the limited capacity of WM has long been regarded as a fundamental and constant trait of human memory system (Ericcson et al,. 1980; Klingberg, 2010).

This limit constrains human ability to process information and engage in a wide range of cognitive activities (Ericcson et al,. 1980). Thus, the training of WM has fascinated many scholars (Sternberg, 2008; Klingberg, 2010). It is believed that, if an individual's WM can be improved through training, a series of related abilities will benefit (Shipstead et al., 2012).

The earliest attempt to train WM was made by Earl Butterfield and colleagues (1972). They sought to improve WM in learning-disabled subjects by teaching them to use subvocal rehearsal strategies. However, the training approach proposed by them was only observed to lead to improvement in participants' performance in trained tasks, and no transfer of the training effects to non-trained tasks or to everyday performance. Similarly, a later study by Ericsson et al. (1980) indicated that, during the course of 230 hours of training in the laboratory, S.F., a college student, with average memory skills and intelligence, gained the ability of increasing his digit span from 7 to 80 digits by the use of strategies such as *"retrieval structure"* (grouping numbers) or *"mnemonic associations"* in the long term memory (p. 1181), (to associate numbers with running times, ages or dates). It was found that, after only a few hundred hours of practice, S.F.'s memory for random digits was comparable to that of the life-time memory experts. However, as argued by the researchers themselves, this training-related improvement tended to be material-specific because when the participant was presented a sequence of letters after 3-month practice, the maximum number of letters he could correctly recall was only 6.

These early training studies seem to support the view that the average capacity of WM is about three to four units, and that it is impossible to improve the WM capacity through training (Ericsson et al., 1980). However, later studies demonstrated that WM training could improve performances on complex cognitive tasks by inducing brain plasticity, which can ultimately lead to changes in the function and morphology of human brain (Buschkuehl et al., 2012). Klingberg (2010) in a review of WM training literature, divided the training approaches into explicit training (explicit teaching of strategies, such as rehearsal, chunking and meta-cognitive strategies), and implicit training (repetition of WM tasks engaging general WM mechanisms, feedback and often gradual adjustment of the difficulty of the task). He found that the implementation of the implicit training method created by him and his colleagues brought about improvements in performance in non-trained tasks which depend on WM and control of attention, and he attributed this transfer of training effects to *"training induced plasticity in an intraparietal-prefrontal network that was common for WM and control of attention"* (p. 322). Thus, he drew the conclusion that implicit training could strengthen general WM capacity. His findings tie in nicely with those of prior research which has identified transfer effects of WM training (Dahlin et al., 2008; Li et al., 2008; Holmes et al., 2009; Karbach & Kray, 2009; Mackey et al., 2011).

Although there is large pool of evidence substantiating that working memory (WM) training produces far-reaching transfer effects on non-trained tasks, the underlying mechanisms that induce this transfer remain unclear (Buschkuehl et al.,

2012). Many researchers resort to brain imaging to gain insight into such mechanisms (ibid). Langer et al. (2013) conducted the first longitudinal study to investigate the functional brain networks underlying WM, using graph-theoretical network analyses. Sixty-six young adults participating in the study were divided into the experimental and control groups. The experimental group practiced three WM tasks and the active control group practiced three tasks with low WM demand. Electroencephalography (EEG) recording showed that WM performance was significantly associated with power in the theta frequency, and WM training was observed to increase theta power. In addition, the brain network of subjects with better WM performance tended to show more small-world topology. Their findings serve as evidence for the WM training-induced plasticity of the functional brain network.

Takeuchi et al. (2013) also carried out a study to investigate whether WM training altered resting-state brain mechanisms, which are believed to reflect intrinsic brain activity and connectivity. In their study, subjects received 4 weeks of WM training and brain image was scanned before and after the training to identify changes of functional network and regional cerebral blood flow during rest. Results demonstrated that WM training engendered plasticity in neural mechanisms and could alter intrinsic brain activity and connectivity.

Although the outcomes of WM training seem to be promising, studies on the training of working memory using training tasks that do not resemble memory tasks, in other words, far transfer tasks, have been found to be insufficient (Shipstead et al.,

2012). According to Shipstead and colleagues (2012), *"increased WM capacity is the mechanism through which far transfer is hypothesized to occur"* (p. 645), because the use of far transfer tasks can *"rule out the possibility that results can only be attributed to task-specific learning"* (p. 628). Thus, the present study made an endeavor to enhance phonological working memory (PWM), as stated previously, one important aspect of WM, which is mainly in charge of verbal-acoustic material, of the primary school students using the verbotonal-based approach. As will be discussed in the following section, Verbotonalism regards neuroplasticity of the human brain as a first and foremost principle both in rehabilitating the hearing-impaired and in the teaching of foreign languages, and its primary goal is to provide optimal stimulation to the brain of subjects and thus induce neuroplasticity of the brain, which has been recognized as a necessary condition for improving WM capacity (Guberina & Asp, 1981; Hempel et al., 2004; Klingberg, 2010). In addition to that, its emphasis on suprasegmental features and body movements is also assumed to be conducive to the improvement of PWM, as there are studies showing that interventions such as mindfulness or meditation training (Fabbro et al., 1999; Van Vugt & Jha, 2011; Zeidan et al., 2010), physical exercise (Lachman et al., 2006), as well as long-term training on musical instruments (Jones, 2006; George & Coch, 2011) may affect WM capacity.

2.4 Verbotonalism

2.4.1 General introduction to Verbotonalism

In 1939, Petar Guberina, a linguist who specialized in *"*the linguistics of speech*"*, discussed in his PhD dissertation at the Sorbonne University, Paris, the impact of intonation on the meaning of a phrase, and that of spoken language on the learning of written language. On the basis of these theoretical assumptions, he conceived a teaching approach, the verbotonal approach, which aimed to teach people of normal hearing to speak before reading a language (Renard & Vlasselaer, 1976; Asp, 2006, revised in 2012). Shortly after its establishment, a few institutions all over the world began to implement the theoretical principles and practices of the verbotonal approach to the rehabilitation and training of subjects with hearing impairments, or with speech or communication disorders as well as to the teaching of foreign languages (Guberina & Asp, 1981).

Verbotonalism is fundamentally a theory of meaning-making and it maintains that *"speech is a social event"* (Guberina & Asp, 1981, p.1). The verbotonal approach is based on the premise that spoken language serves as the basis for the development of other linguistic and cognitive skills, for example, writing, reading, spelling, and analytic skills (ibid). Verbotonalists think that speaking is a sociological encounter, and one's speech reflects the way he/she makes sense of the world. Speaking is the chief mode of communication used by individuals to express their opinions, argue for stances, make explanations, and exchange information with others. And through the exchange of

meanings in the process of communication, social reality is created, maintained and modified. Each individual speaker is a producer, and at the same time, a perceiver of speech. The auditory and visual information in one's production mirrors how he/she perceives speech. In this sense, there are two points which are worthy of attention here. The first is that individuals' perceptions of the same sound are different because of their different personal organisation systems (Lian & Lian, 1997). The second is that individuals' perceptions of the sounds of the target language are influenced by their "phonological sieve" resulting from the phonological system of their native languages: a mechanism for keeping sounds which are recognized and rejecting those which are not (Lian & Lian, 1997; Zhang, 2012). The process of learning a foreign language is almost analogous to hearing-impaired children's or adults' learning of their native language. Obviously, learners' perceptual and productive systems have been trained and attuned to perceive and produce their native language or other languages that they know well. So, when they begin to learn a foreign language, they are liable to be insensitive and "deaf" to the sounds of the new language. In order to defeat this "deafness", the verbotonal system uses octave-band filters to determine the optimal frequency region for the perception of a particular phoneme (Guberina & Asp, 1981). This is to raise learners' awareness of new language, and change their perceptions of the foreign sounds by creating the optimal field of hearing for them. Once one's perception changes, his/her speech will also change. If we have corrected one's speech, we have corrected his/her perception (Guberina & Asp, 1981). Therefore, the central

preoccupation of the verbotonal approach in the teaching of a foreign language is to raise learners' awareness and restructure, i.e., reorganize, the way they perceive language signals.

The verbotonal system considers neuroplasticity of the human brain a principle of paramount importance in the application of the approach to the rehabilitation of hard of hearing subjects and the teaching of foreign languages, and its primary aim is to provide optimal stimulation to the brain of each individual (Guberina & Asp, 1981) in order to restructure the brain. As stated before, the capability of the human brain to change by exposing to optimal stimulation is called neuroplasticity (Asp et al., 2012c). Hearing impairment is diagnosed as a peripheral disorder. The verbotonal approach is aimed to rehabilitate the auditory perception within the brain by means of rewiring and restructuring the brain's neural networks (Asp et al., 2012b). The method supports that the perception and production of speech is a multi-sensory and whole-body experience. Asp et al. (2012b) proposes that *"As an infant, the vibrotactile channel (vestibular) is a more dominant sense than hearing (cochlear), whereas hearing becomes more dominant in the older child. With the adding of proprioception, the human brain goes through developmental neuroplasticity as the developing infant matures to childhood and then adolescence"* (p. 321). As it is known, the human brain receives neural information mainly from the senses. The simultaneous activation of learners' visual, auditory, vibrotactile, vestibular, proprioceptive senses in the learning process will

facilitate the restructuring of the neural connectivity of their brain, and thus engender the plastic changes of their brain (Asp et al., 2012a).

2.4.2 Previous studies on Verbotonalism

Since its advent, the verbotonal system has been studied and applied mainly in the following areas: rehabilitation of children and adults with hearing impairment, be it peripheral or central, diagnostic therapy, treatment of speech language disorders and communication problems such as articulation, stuttering, aphasia and autism, and the teaching of foreign languages largely through the Structuro-Global-Audio-Visual (SGAV) methodology (an adaptation of the principles of Verbotonalism to language teaching). It should be noticed that Verbotonalism has been relatively neglected as an area of research for the past 30 years. Few research studies have been performed, especially in the area of foreign/second language education. This accounts for the lack of recent references in this dissertation. However, the references quoted, even though apparently dated, contain the fundamental principles which underpin the work done here and have served to develop approaches, structures and techniques used in the present study.

In the field of foreign language teaching, the verbotonal has been viewed as an effective approach to phonetic correction. For instance, as early as 1980, Andrew Lian, in his book *Intonation Patterns of French* (*Teacher's Book*), introduced Verbotonalism and performed it in the teaching of French pronunciation. He devised sensitization and reinforcement sessions to create optimal situations to reinforce students' perceptual and

articulatory abilities. The sensitization session is primarily composed of steps such as *"relaxation phase"*, *"audition of filtered sentences", "the importance of movement and gesture", "humming along", "interpretation of the intonation patterns", "mouthing the words", "repetition on a background of filtered patterns", "humming the patterns"* , *and "adding the words to them",* etc. The reinforcement session consists of *"self-testing"* and *"sensitization and intensive practice"* (Lian, 1980, p.3). The implementation of these pedagogical procedures designed on the basis of Verbotonalism placed the students in conditions conducive to the acquisition of the prosodic features of French by removing the interference of vowels and consonants and enhancing the correct auditory and articulatory habits.

Still in the teaching of French pronunciation, Alazard, Astésano and Billières (2011), established, thirty-one years later, a multiphonia corpus consisting of audio-video classroom recordings and compared two approaches of phonetic correction: the traditional articulatory approach, and the verbotonal approach. In their pedagogical experiment, the teacher first helped the learners familiarize themselves with the prosodic structure of the target language through the repetition of prosodic patterns using logatoms and facilitating gestures (for example rising hand movement for salient syllables). In a second phase, the prosodic structure was used to facilitate phoneme perception and re-production. The experiment demonstrates that prosodic cues help access the lexicon and segment of the speech flow in a second language. Pedagogical gestures have a facilitating impact on the reproduction and memorization of prosodic

cues. So, their study serves as a very good example of the benefit of verbotonalism in the teaching of second language pronunciation over the articulatory approach.

Hu and Uno (2005) developed a teaching method based on Verbotonalism to teach Japanese beginners learning the voice tones of Chinese. Thirty-five university students in a basic Chinese class were selected as the subjects. After 7 weeks of training, 90 minutes a week, the students were given a test, including the task of dictating and pronouncing the voice tones of monosyllables and two-syllable words. Results showed that this approach was useful for distinguishing and learning to pronounce 4 different kinds of voice tones, and students' performance on almost all tasks was significantly improved. Besides, Hang (2012) implemented the verbotonal system in the teaching of Japanese to Chinese learners. Echoing previous studies, her findings again supported the effectiveness and usefulness of the verbotonal approach. Zhang (2006) also applied a verbotonal-based approach, which she called *"the Somatically Enhanced Approach (SEA)"* to the teaching of Mandarin prosody to Australian students studying Mandarin (Modern Standard Chinese) (p.2). The 22 students in her experiment were divided into experimental and control groups. The control group was taught using a nonmulti-sensory but communicative approach, whereas the experimental group received the pedagogic treatment based on Verbotonalism-the multi-sensory communicative approach (SEA). The pedagogical procedures in her study were composed of two parts: *"the Sensitization Session"* and *"Repetition Exercises"* (p. 150). The sensitization session included the steps of *"relaxation", "humming", "clapping to the rhythm of the*

sentences", "incorporation of movement and gesture", "mouthing the words", and *"adding words to the intonation patterns"* (p. 155). And the repetition exercises were primarily concerned with *"checking for meaning"* (p. 165). Results indicated that the order of difficulty of the four Mandarin tones was shown to be similar for the two groups of participants. At the beginning stages of learning, learners' first language interference was observed to be the primary cause of errors, instead of *"the physical difficulty of articulating particular phonemes (or any features of Universal Grammar)"* (p. 3). She commented that the verbotonal-based approach (SEA) was effective in the teaching of Mandarin tones by lessening the influence of learners' first language.

In the Chinese EFL context, He Bi (2015) studied the effectiveness of the verbotonal system in the teaching of pronunciation to Chinese university EFL learners. She combined a CALL-based (Computer-Assisted Language Learning) autonomous structure with the verbotonal system of corrective phonetics and named it as *"CALL-VT"* (He, 2015, P.3). Ninety-five Chinese first-year English major students enrolled in a pronunciation course participated in the study. The students were from two intact classes, one group was randomly designated as the control group, and the other the experimental group. The pedagogical intervention was made up of *"classroom and out-of-classroom activities"* (ibid). The classroom activities were aimed at making students become sensitive to the sounds of English, whereas the out-of-class activities attempted to develop students' autonomous learning of pronunciation. After the pedagogical intervention, the experimental group outperformed the control group in every part of

the areas tested: phonemes, word-reading, passage-reading, and oral interview. Meanwhile, feedback from native speakers of English raters showed that the experimental group did a better job in pronunciation, comprehensibility and fluency. Furthermore, two discoveries from the result of the experiment were really surprising, interesting and thought-provoking: 1) The experimental group's pronunciation of individual phonemes was significantly better than that of the control group without any training in phoneme-production, whereas the control group did have training in phoneme-production, and 2) The experimental group outperformed the control group in terms of comprehensibility and fluency *"in a natural setting"* (He, 2015, P.5).

As Alazard et al. (2012) put it, although the verbotonal approach proves to be of great value both in didactics and speech therapy, it has only achieved quite limited visibility, largely because it is poorly understood and inadequately implemented. It nevertheless has a long and respected history within the limits of the research programs undertaken. However, to date, this approach still remains confidential to a small group of international experts, so its validity remains to be demonstrated to a larger audience in order to, one day, be included in comprehensive L2 teaching methods. Given all this, the current research developed a verbotonal-based approach for improving the English-speaking skills and strengthening PWM of Chinese primary EFL learners.

2.5 Theoretical framework of the present study

The primary concern of the present study is to improve the English speaking skills of Chinese primary EFL learners, and at the same time, to enhance their phonological working memory. The whole study is based on the principle that language learning is essentially an act of meaning-making (Lian, 2000). When language signals come into us, they are not necessarily received *"as they are"* (ibid, p.1). We only accept those signals which are meaningful to us. If individuals have problems in recognizing and discriminating the sounds of a foreign language, it is not that they are unable to do so but that they cannot make sense of them. Given all of this, Verbotonalism serves as a point of departure for the current study, which assumes that each individual speaker is both a producer and a perceiver of speech, more importantly, the verbal and visual information in his/her production mirrors how he/she perceives speech. If we intend to improve his/her speech, we need to change his/her perception (Guberina & Asp, 1981).

The verbotonal-based approach proposed in the present study is a holistic system which emphasizes the importance of the multi-sensory, multi-modal, and multi-channel way of learning to speak. The starting point of the approach is to raise students' awareness of the characteristics of the phonological system of English. Once students' perceptions of the sounds of English have been changed, they will become more sensitive to the phonetic specificities of the language. Situational dialogues which require collaboration will be used in the approach to highlight the pragmatics of spoken language and enhance students' social and conversational skills. Most importantly, the

approach stresses the importance of experiential and autonomous learning. During the whole learning process, the teacher only acts as facilitator and catalyst. Students are empowered to be self-directed in learning and free from any intervention and instruction from the teacher. These considerations are coherent with the aforementioned postmodern view of learning. Besides, the basic theories and practices of the verbotonal approach have been found to be the fundamental and necessary conditions for improving one's PWM capacity. Taken together, verbotonalism and Postmodernism, however seemingly different from each other, prove to be intertwined and mutually reinforcing in epistemological thinking. They constitute the over-arching theoretical framework of the current research (see Figure 2.3).

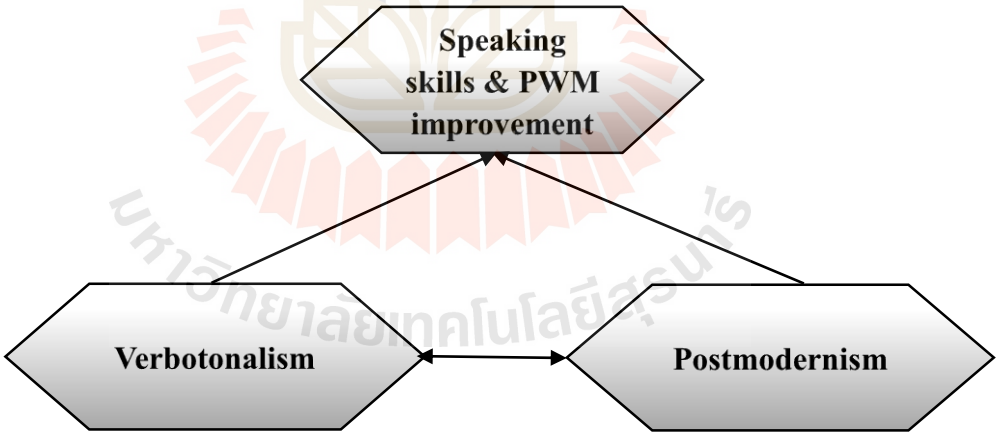


Figure 2.3 Theoretical framework of the present study

2.6 Summary

In this chapter, the basic principles of learning, and learning from a postmodern perspective were discussed. Subsequently, previous studies on the teaching of speaking skills, PWM and language learning, the training of PWM, and Verbotonalism were reviewed. Finally, the theoretical framework that informs the present study was expounded.

CHAPTER 3

METHODS AND MATERIALS

This chapter describes the methods and materials of the current study. It begins with research design including participants, variables, learning materials, pedagogical procedures. Then, instruments used, which include semi-structured interviews, questionnaires, and pretests and posttests for speaking and PWM (phonological working memory) are explicated. Besides, the procedures and techniques for collecting and analyzing data are described. Finally follows the results of the pilot study.

3.1 Research design

The present study was conducted using a quasi-experimental design. According to Seliger and Shohamy (1989), a quasi-experimental design has two prominent strengths. The first is, as a quasi-experimental study is *"constructed from situations that already exist in the real world, and are probably more representative of the conditions found in educational contexts"* (P. 148), it is more likely to have external validity. And the second is, quasi-experimental designs are less intrusive and disruptive, so they are ideal for the researcher to gain access to the target population. However, as pointed out by Punch (2013), a quasi-experimental research is carried out under conditions in which it is difficult for the researcher to manipulate many of the extraneous variables and

randomize the assignment of participants, so there is a clear threat to internal validity. In order to cope with the problem, the researcher resorted to statistical techniques to remove the influence of possible extraneous factors, in other words, control variables, by identifying them, measuring them, and drawing out their effects statistically.

3.1.1 Participants

A population is referred to as the *"aggregate of all cases that conform to some designated set of specifications"* (Isidor, 1976, p.419). Methodologically speaking, it is impossible and impractical for a researcher to collect data from the entire population. A researcher can draw precise inferences on a population based on a relatively small number of cases (a sample) when the sample is capable of representing the whole population. For the present study, the participants were sampled from Grade 3 students in YAPS (the Yiliang Anjiaqiao Primary School). As required by the directive issued by the MOE in 2001, English classes are provided to students from Grade 3 and above in this primary school. It should be emphasized here again that Verbotonalism holds neuroplasticity of the human brain as the principle of paramount importance in the teaching of foreign languages. As stated in Chapter 1, due to the stronger neuroplasticity of the brain, children with younger age, especially before puberty, are capable of learning to speak a foreign language more easily because their neural maturational process is still undergoing evolution, and is thus optimal to receive varied neural stimulation (Piaget, 1973). For this reason, the sample of the present study was drawn from Grade 3 students who had learned English for only one semester. Eighty third-

grade students, aged on average 8 to 9 years, and enrolled in two intact classes, participated in the study. Based on random assignment, all 40 students from Class 201301 formed the experimental group, who received the intervention (see Section 3.1.4), and all 40 students from Class 201302 constituted the control group, that followed the traditional way of learning speaking as previously taught. Therefore, both groups were the same size (40). In the traditional approach, there was specific and explicit teaching for segmental pronunciation as well as spelling and translation skills. Language input was from the teacher and native-speaker recordings. Students spent most of the time in class reading and spelling words, and memorizing and reciting sentences. The two groups were allowed 3 academic periods (45 minutes/period) a week for learning English in class. Classes for both groups were mostly scheduled in the morning. The English proficiency level of the students from the two groups was similar. In addition, both groups were assigned to do 150 minutes of out-of-class speaking practice per week. The control group, Class 201302, was free to practice speaking skills with the methods they normally used, but had no access to VTA materials and training. As for the experimental group, the out-of-class practice focused on the repetition of the steps in the pedagogical treatment (see Section 3.1.4). Time on task was the same for the groups.

3.1.2 Variables

The variable that the researcher wishes to explain is the dependent variable. The variable that the researcher expects to explain change in the dependent variable is

referred to as the independent variable (Punch, 2013). In this case, the independent variable is also called the explanatory variable. It is the presumed cause of changes in the values of the dependent variable. The dependent variable is supposed to be caused or influenced by the independent variable (Seliger & Shohamy, 1989). According to the research objectives and questions, the independent variable of the present study is: the treatment applied to the learning of speaking skills and the dependent variables are (a) students' speaking skills, and (b) students' phonological working memory capacity.

3.1.3 Learning materials

There are more than 30 types of English textbooks offered for use in Chinese primary schools (Zhang, 2011). Each school is given freedom in the selection of textbooks by the government. The textbooks used for all grades in YAPS were published by *People's Education Press,* the most influential textbook publishing house in China. The third-graders used *English (Grade 3) (Vol.1)* in their first semester, and *English (Grade 3) (Vol.2)* in their second semester. Video resources were provided with the textbooks. These videos were cartoons based on the text of each unit. The textbooks and accompanying video resources constituted the primary learning materials for students in the present study. Besides that, students could choose any materials they wanted for self-learning outside class. These formed another source of learning materials.

3.1.4 Pedagogic procedures

The pedagogic procedures of the present study were heavily based on a modified version of previous research carried out by Andrew Lian (1980). They comprised two parts: in-class sensitization and out-of-class reinforcement sessions.

In-class sensitization session

The primary aim of the session was to create the best optimal field of hearing of English for students and to cultivate their "feel" for and sensitivity to the language. In order to achieve these goals, first, students' awareness of the phonetic features of the foreign sounds needed to be raised. Then, their perceptual systems attuned to perceive their native language needed to be changed. Thirdly, their "deafness" to the foreign sounds needed to be defeated. Given all of this, the in-class sensitization session consisted of 8 steps.

Step 1: Relaxation phase. It was suggested to students that they engage in 5-10 minutes of mind-calming activities to nurture a sensation of well-being conducive to learning. An explanation about these unusual and interesting procedures was made by the teacher. The classroom was darkened, music of any kind which the students thought would help them get rid of anxiety and nervousness was played as the background music, such as Yoga, Buddhism, Zen and Baroque music, etc. With music in the background, students were free to do any activity which could relax them physically and/or spiritually. Mind-calming activities adopted by the students on the basis of their own choices ranged from Yoga moves, Yoga breathing, meditation, resting on the desk,

leaning against the chair with eyes closed, deep breathing, free dance or even walk slowly in the classroom.

Step 2: The students listened to the filtered sentences. These sentences were filtered through a low pass digital filter. The filtered materials were played 15 times consecutively. Such a procedure made the prosodic structures, salient by removing the interference of vowel and consonant sounds. In so doing, students' hearing load of novel sounds was lightened and they would be more sensitive to the phonetic specificities of English. Prosody has been recognized as a right-hemispheric faculty (Sammler et al., 2015). Some experimentation suggests that the prosody of a particular language is difficult to acquire after grammar and lexicon. Therefore, language learning should start with prosody, not finish with it (Martin, 2016). The filtered materials in the present study were played using good quality amplifiers and loudspeakers with a good bass response and the volume was as loud as possible in order to enhance students' perceptions of the language input through the body's natural sensitivity to low frequencies (Lian, 1980).

Step 3: Students continued to listen to the filtered sentences for another 10 times. In this phase, students repeated the filtered patterns by "humming" them in chorus. At the same time, they were required to use body movements and gestures to express their personal perceptions of the rhythm of the filtered sentences. In order for the students to understand the requirements well, the teacher usually presented her version of body movements and gestures to them in a non-prescriptive way. That is, the teacher did not

impose the model on the students and did not expect them to follow it. That was made clear to everybody. As for the students, some drew the shape of the intonation curve in the air, some beat out or clapped out the rhythm of the sentences, some came down on every syllable or at the end of every rhythm group with feet, some stretched arms at each rise in pitch, and others danced to the rhythm. The synchrony of humming and body movements can create proprioceptive stimulation to both hemispheres of the listener's brain (Guberina & Asp, 1981). The ability of the human brain to change with optimal stimulation is called neuroplasticity. Body movements are shown to be able to accelerate the neuroplasticity of the brain (ibid).

Step 4: Following the humming and movement exercises, students were encouraged to guess the meaning of the prosodic patterns and to discriminate between them on the basis of what they had previously studied. In addition, the key words specified by the *Curriculum Standards* for Grade 3 students in the second semester were written down on the blackboard by the teacher. The students were then guided to guess the stress patterns of the keywords. The guessing of meaning and discrimination of intonation patterns can, on the one hand, actively engage students in learning and, on the other hand, enable them to have a better understanding of the specificity of English intonation.

Step 5. Mouthing the words. At this stage, the students were exposed to the unfiltered sentences for the first time. The sentences were played 10 times continuously, and the students were required to only mouth the words without producing any sound.

Mouthing the words serves to reinforce students' reception and production of sounds. Besides, in this phase, the teacher led the students to discover whether their previous "guesses" were correct or not and to explore the function of the intonation patterns in normal language context and discuss the specificity of English intonation and rhythm.

Step 6. Articulation of words to the intonation patterns. The students were first required to hum the intonation patterns as they did in Step 3, and then began to add words to the "tune" produced by themselves. It is believed that when we hear things repeatedly, we can hear things that we have never heard before since in the process of hearing, we generate internal representations of the things heard according to our personal representational and logical system (Lian, 2004). By this stage, after hearing both the filtered and unfiltered sentences for many times, the intonation patterns had been somehow internalized and students' perceptual systems tailored for the perception of their L1 had been somewhat changed. It was therefore the right time for the students to articulate words. Unfiltered sentences continued to be played 10 times and students repeated the sentences in chorus.

Step 7. The video in which all the original sentences were acted out in a cartoon was played to the students 3 times. As the video presented sentences functioning in a communicative context, it was expected that this step would not only reinforce students' perceptions of speech, but also familiarize them with the use of these sentences in realistic communicative situations. This phase enables students to have more effective understandings of the language in use.

Step 8. Situational dialogue. In this step, students were divided into groups of 3 or 4. First, each group member attempted the speech dialogue that they had watched in the video by role-playing one or more of the characters in the situation. Then, all groups were encouraged to recreate and make fun of the original dialogue, in other words, to create a parody. In this context, creating a parody meant that students pretended to act out the original dialogue in the textbook by imitating it but turning it into a joke. It was valuable because a successful parody implied that the students had internalized the essential structures and spirit of the dialogue that they were imitating. The teacher came to each group to offer help when need arose. This procedure is of paramount importance mainly in two ways. First, the use of situational dialogues highlights the pragmatics of spoken language. (Guberina & Asp, 1981). This early development of pragmatics can increase students' conversational and social skills that are critical for successful communication. Second, the performing of the parody makes learning enjoyable, and thus stimulates students' motivation towards learning. There are many lines of evidence suggesting that children learn most efficiently, effectively and lastingly when they are intrinsically motivated to do so (Berns et al., 2013).

Out-of-class reinforcement session

The students from the experimental group were assigned to do reinforcement exercises out of class. As stated previously, both the control and the experimental groups were asked to do 150 minutes of out-of-class speaking practice per week with the assistance of their parents, if appropriate. The control group practised speaking with

the approaches normally used: reading, memorizing and reciting words and sentences. As for the experimental group, the reinforcement session was mainly the repetition of the steps in the sensitization part. But the materials could go beyond the textbook and be anything that students had access to. All the parents of the students in the experimental group were villagers in Anjiaqiao Village. They were trained by the researcher to filter the materials chosen by the students. For those whose parents were unable to perform the filtering task, the teacher offered help. In addition, as required by YAPS, parents were expected to do their utmost to assist their children to learn especially when in the lower grades. In this study, parents were expected to cooperate with the students to role-play the situational dialogue by acting out some characters. For some parents whose English proficiency level was not good enough, they just watched their children's performances thus playing a supportive, if not active, role. Their children played all the roles themselves.

The role of the teacher

Obviously, teachers' background, personality, teaching methods and teaching styles can influence teaching outcomes. So, as a way of minimizing differences, the same female teacher, who had been teaching English in YAPS for 4 years, took care of both the control and experimental groups. As mentioned before, the control group received the traditional way of instruction, whereas the experimental group learned speaking skills via VTA. What is worth mentioning is that in the pedagogical procedures, the teacher acted only as facilitator, supporter and organizer. That is, in the

experiment, the students were seen as whole, capable and resourceful. So the teacher's role was not to teach but to oversee the whole learning environment, and monitor students' performances and provide help when needed. The present study was based on the principle that learning is an active process of meaning-making and occurs optimally when learners are free to engage and assimilate information by themselves (Grolnick & Ryan, 1987). Thus, students in the experimental group were empowered to self-adjust learning both in and out of class. In order for the teacher to have a good understanding of the theories and practices of the teaching approach, the researcher first explained the design of the experiment, and then modeled a class for her.

3.2 Instruments

The research objectives and questions of this study determined the types of data to be collected. The next step in the process is to select appropriate instruments for collecting high quality data since the quality of data elicited will exert a great impact on the quality of the research (Seliger & Shohamy, 1989). Therefore, in this section, the instruments used by the researcher to collect data are described, which include semi-structured interviews, questionnaires, observations, the pretest and posttest for speaking skills, and the pretest and posttest for PWM.

3.2.1 Questionnaires

In second language acquisition research, questionnaires are used most frequently to gather information on phenomena which cannot be easily observed, such

as opinions, attitudes, beliefs and motivation (Oppenheim, 2000). Questionnaires are also very effective procedures for the researcher to elicit the background information of subjects, for example, age, language learning background, and number of years for learning a certain language, etc. (Seliger & Shohamy, 1989). A questionnaire (see Appendix G) was constructed and administered to the students in the experimental group to obtain their background information and investigate their attitudes and opinions of the pedagogical treatment in the present study. The development of the questionnaire was based on the findings of a number of previous studies related (Teng, 2010; He, 2014; Zhang, 2014) and the suggestions of the five experts who were invited to validate the questionnaire. As mentioned in Chapter 1, the five experts were specializing in English language education, 1 coming from YAPS, and other 4 from Kunming University of Science and Technology.

The questionnaire was presented in two parts: 1) background information, and 2) students' perceptions and attitudes towards the learning of speaking and training of PWM using VTA. The first part included age, gender, and the amount of time spent on learning to speak English each week after class. Such information is necessary as it provides insight into each participant's background in learning English.

The second part included 12 statements regarding students' beliefs and opinions of learning speaking skills and training PWM by means of VTA. The questionnaire used a 5-point Likert scale which required the students to respond to the 12 items by ticking the choice from "strongly agree", "agree", "undecided", "disagree" and "strongly disagree".

"Strongly agree" was assigned a score of 5 points, while "strongly disagree" received a weight of 1 point. So higher agreement on each item was reflected in higher scores.

It is important to stress that the validity and reliability of the instrument for data collection needs to be rigorously tested to ensure that high quality data can be gathered (Oppenheim, 2000). The afore-mentioned 5 experts specializing in English language education checked the relevance and clarity of the 12 statements and validated the content of the questionnaire using IOC (see Appendix I). Subsequent revisions were made based on the suggestions of these experts. The revised questionnaire was tried out in the pilot study and the Cronbach's Alpha Coefficient was high enough (α= 0.95) to prove the reliable internal consistency of all the statements within the questionnaire. Then, the final version of the questionnaire was distributed to the experimental group.

3.2.2 Semi-structured interviews

The interview, as the "*gold standard of qualitative research*" (Silverman, 2000, p. 51), is a data collection method that provides different perspectives of exploring people's experience and viewpoints. The aim of the interview is to acquire information by actually conducting a talk with the interviewees. Thus, it has been described as a *"professional conversation"* (Kvale, 1996, p.5) or *"conversation with a purpose"* (Burgess, 1984, p.102). The obvious advantage of an interview is that it is personalized and flexible, and enables the researcher to obtain in-depth information and data that have not been predicted (Seliger & Shohamy, 1989). Interviews are classified as three types in terms of their degree of explicitness and structure: structured, open and semi-

structured (Kvale & Brinkmann, 2009). A typical semi-structured interview is made up of specific and defined questions which are determined by the researcher in advance, but at the same time, it is allowed to proceed in unexpected directions where new areas might be explored (Seliger & Shohamy, 1989). Because of this flexibility, among the three types of interviews, the semi-structured interview is the most commonly used interview option (Kvale & Brinkmann, 2009).

In the current study, in order to triangulate the source of data, and probe qualitative information of the effectiveness of VTA in the learning of speaking and PWM training, semi-structured interviews both for the students (see Appendix J) and the teacher (see Appendix M) were developed. Based on previous studies in this field and the objectives of the current research, the researcher first identified the key topics that needed to be covered in the semi-structured interviews, and then formulated 10 specific questions for the students' and teacher's interviews respectively. After that, the researcher trialed the interview with one student from the experimental group and the English teacher who took care of both the control and the experimental groups. Revisions were made according to their suggestions and comments. Furthermore, to ensure the quality of the interviews, the relevancy and clarity of questions were validated by the abovementioned 5 experts using IOC (see Appendix L and Appendix O). The finalized semi-structured interviews were conducted with 6 students from the experimental group and the teacher who taught both the experimental and control groups after the pedagogical intervention.

3.2.3 Observations

Observations are used in qualitative research to gather accurate and naturalistic data which reflect the reality of the situation as the participants see it (Lodico et al., 2010). Gold (1958) classified four types of observation based on the degree of participation of the observer: complete participant, participant as observer, observer as participant, and complete observer. **Complete participant** means that the observer completely participates in the group's activities, and his/her role of being a researcher is not known to the group members. In the "**participant as observer**" situation, the observer also actively takes part in the activities but each member of the group is aware of his/her role. **Observer as participant** denotes that the observer doesn't participate in the group's activities but still has a connection to the group. **Complete observer** refers to that the observer only conducts observations and does not participate in activities. And he/she is not a member of the group.

In this study, observations were conducted to obtain a complete understanding of VTA as a way of teaching speaking skills and training of PWM. The researcher's role is observer as participant. In order to ensure accurate and unbiased observations, the researcher first made a systematic and careful examination of the phenomena being observed. Then, an observational protocol specific to the objectives of the present study was developed. Moreover, both descriptive and reflective field notes were recorded by the researcher.

3.2.4 Pretest and posttest for speaking skills

As discussed in Section 2.1.3, given the complexity of assessing speaking ability, most primary schools in China do not conduct English-speaking tests (Teng, 2010; Zhang, 2011). As in most other primary schools, an English-speaking test is not required in YAPS. So, the tests used in the present study were constructed by the researcher. As it is known, an appropriate language test should not only present evidence for learning outcomes, but also identify students' problems in learning. Most importantly, it should be able to clarify teaching objectives and provide information of the effectiveness of the teaching approach (Bachman & Palmer, 1996). So the design of speaking tests which are appropriate for test takers, and at the same time, for research use in the present study, is paramount. Thus, in the following part, the development, validation and piloting of the test papers used in the pretest and posttest phases are explicated.

The design of the test papers was based on a number of sources. Firstly, the findings of previous studies on the construct of speaking proficiency and speaking assessment, which had been presented in the second chapter, provided insight not only into the goal of the teaching of English speaking but also the speaking skills that needed to be tested in the present study. As discussed before, it is difficult to obtain a precise definition of speaking proficiency. However, it is generally agreed that speaking proficiency includes learners' accurate and effective use of linguistic and sociolinguistic features to achieve a range of communicative functions. Furthermore, linguistic

elements such as lexical richness, grammatical accuracy, pronunciation skills, fluency and comprehensibility have been shown to be the major determinants of the overall speaking proficiency of an L2 speaker. So the tests in the present study were designed mainly to test students' speaking skills in terms of vocabulary, grammar, pronunciation, fluency and comprehensibility.

In addition, the standard level (level 1) specified by the *2011 English Curriculum Standards for Compulsory Education* for primary students, the textbooks used by third-graders in YAPS and speaking tests used by many primary schools in some cities and provinces in China such as Beijing, Jiangsu, and Qingdao, were referred to in order to decide the linguistic content, topic areas, and the speaking tasks of the tests. As presented in the first chapter, according to the curriculum standards, after one to two years of English study, the speaking skills of the third and fourth graders are supposed to meet "Level 1". Specifically, they are supposed to be interested in and enjoy listening to English; be able to play games and perform tasks (such as coloring and matching) according to teachers' directions in English, engage in simple role-plays in English, sing simple English songs, read and understand short stories with the aid of pictures, deliver short self-introduction and express emotions and feelings in simple English, and spell letters and the keywords required by the syllabus; and learn and understand foreign cultures and customs. In light of these standards and with reference to the test papers used in some primary schools in China, four speaking tasks were formulated in both the pretest and posttest papers of the present study: Word-reading,

Sentence-reading, Singing, and Role-play. In addition, the textbooks used by Grade 3 students in YAPS, E*nglish (Grade 3) (Vol.1) (Vol.2),* were also referred to in order to ensure that the linguistic content, topic areas and speaking tasks used in the pretest and posttest were appropriate for the students.

The pretest and posttest papers were pre-piloted among 8 students (2 groups) in YAPS. Revisions were made as a result of inappropriate level of difficulty of the role-play task. Both the students and teacher commented that the role-play task was difficult for the students to accomplish and was unable to elicit their speaking performances effectively. Consequently, the role-play task was replaced by the oral interview task. Thus, the final versions of these two test papers (see Appendix P and Appendix R) both consisted of 4 parts: Word-reading (20%), Sentence-reading (20%), Singing (20%), and Oral interview (40%). The **Word-reading** part comprised 40 words. The **Sentence-reading** part included 10 sentences covering various intonation and rhythm patterns. In the **Singing** part, the students were required to sing an English song. These reading aloud and singing tasks were the tasks students normally performed in their class, aiming at checking their' pronunciation, which is a critical aspect of speaking proficiency because intelligible pronunciation is key to successful communication (Levis & Grant, 2003). And in the **Oral interview** part, the students were given a topic and required to use 4-6 sentences to talk on the topic, then, answer a question about the topic using 1-3 sentences. This part aimed to elicit more authentic language in a natural, conversational style. As the experiment was conducted with the third-graders who were

just at the beginning stage of English learning, so the tasks designed in the tests were mainly based on their course content. Also, to ensure that the difficulty level of these test tasks was appropriate for the students, all the words, sentences, the English song and topic in the pre-test papers were selected from the textbook E*nglish (Grade 3) (Vol.1)* used by the students in the first semester, and those in the post-test papers were from E*nglish (Grade 3) (Vol.2)* used by them in the second semester.

Reliability and validity are critically important for a test. According to Bachman and Palmer (1996), they are *"essential measurement qualities"* (p. 19), because they can justify for the use of test scores as a basis for making inferences and interpretations. The content validity of the test papers was checked by 5 experts using IOC. The results of the IOC analysis (see Appendix Q and Appendix S) indicated that these two test papers were valid to be used in the tests. Then, the pretest was tried out among 40 Grade 3 students twice with a time interval of 2 weeks, and the posttest paper was trialed twice among 40 grade 4 students in YAPS, to check the test-retest reliability of the tests. The Pearson correlation coefficient of the test and retest of the pretest (r= 0.84, p=0.00) and posttest (r= 0.81, p=0.00) were sufficiently high, manifesting that the two tests were highly reliable. The students in the experimental and control groups undertook the tests before and after the experiment to check whether there had been any improvement in their speaking skills, and also to compare the performances of the two groups. And the tests were audio recorded for blind rating.

Raters

Previous studies reveal that although many raters' characteristics have influence on their evaluations of language assessments (Reed & Cohen, 2001), being a native speaker or non-native speaker rater does not significantly affect final ratings (Kim, 2009; Brooks, 2013). For this reason, the raters who were invited to rate students' speaking skills in both the pretest and posttest were three Chinese expert raters. One of these had been teaching primary English for more than 9 years in YAPS. The other two were professional English speaking ability raters from the Faculty of Foreign Languages and Cultures of Kunming University of Science and Technology. The researcher first familiarized the raters with the objectives of the assessment, description of the speaking tasks in the tests, the rating procedures, and the assessment criteria. Then, each of them was given the recordings of the same 10 students and asked to assign both an overall proficiency score and subskill scores independently. A post hoc inter-rater reliability was performed to analyze the extent to which these raters agreed on the scores awarded to the 10 students. As can be seen from Table 3.1, all the correlations are statistically significant, serving as solid evidence for the high reliability of the assessment scores used in the study.

Table 3.1 Correlation matrix on inter-rater reliability

	Rater 3	Rater 1
Rater 1	0.78 **	
Rater 2	0.83 **	0.85 **

**p=<0.05

3.2.5 Pretest and posttest for PWM

As discussed in 2.2.2, nonword repetition (NWR) performance is a significant predictor of language learning, more importantly, it is more effective in measuring PWW than auditory digit span (Baddeley et al., 1998; Gathercole, 2006; Baddeley et al., 2014) because one's capacity to repeat an unfamiliar phonological item is heavily dependent on his PWM system to encode and maintain the novel phonological sequence for sufficiently long (Gathercole, 1995; Baddeley et al., 1998). Thus, NWR was employed in the current study to measure students' PWM before and after the intervention.

Both the pretest and posttest (see Appendix T and Appendix U) were adapted from The Children's Test of Nonword Repetition established by Gathercole et al. (1994). There were 20 nonwords, 5 each containing two, three, four and five syllables, in the pretest and posttest respectively. According to Gathercole and his colleagues, the phoneme sequences in all these nonwords conform to the phonotactic rules of English and within each number of syllables, the items were coined to be in accordance with the dominant syllable stress structures in English for words of the same length (ibid).

More specifically, the stress pattern of the two-syllable nonwords is strong-weak, and that of the three-syllable nonwords is strong-weak-weak. There is no single dominant stress structure for four- and five-syllable English words. By constructing the nonwords in this way, the phoneme sequences within each item are *"all phonotactically and prosodically legal"* (p. 106). In addition, studies show that nonwords with high wordlikeness will lead to the use of long-term lexical knowledge to facilitate performance on a NWR test (Gathercole, 1995). Therefore, a balance of the distribution of nonwords with high and low wordlikeness in both pretest and posttest papers in the present study was kept, with 10 being high-wordlike, and 10 low-wordlike.

In the current research, each student was asked to listen to a cassette recording of a male speaker reading aloud the nonword sequence in an American English accent. On the recording, a three-second interval occurred between the end of one nonword and the onset of the next. Each participant made the repetition attempts individually. As these were essentially a memory test, allowance was made for any regional accent or other cases in which a participant was observed to consistently misarticulate one phoneme as another. The tests were audio recorded for scoring. Each individual's repetition attempt was scored dichotomously (phonologically correct or incorrect) by two raters. The total number of nonwords that a student correctly repeated was calculated as his/her PWM test score.

3.3 Data collection

3.3.1 General procedures

This quasi-experiment was carried out in a natural English learning context in the Yiliang Anjiaqiao Primary School (YAPS). As mentioned earlier, all the participants were third-graders. The experiment spanned 4 months, from March to July, 2016, the second semester for Grade 3 students. For the third-graders in YAPS, E*nglish (Grade 3) (Vol.1),* published by *People's Education Press* in China was the textbook used in their first semester, and E*nglish (Grade 3) (Vol.2)* was the textbook used in their second semester. So, in the present study, the textbook used by both the control and the experimental groups was E*nglish (Grade 3) (Vol.2)*. The textbook comprised 6 units which were supposed to be finished within one semester. Each unit covered 8 parts: *Let's Talk, Let's Learn, Let's Chant*, *Let's Spell, Start to Read, Let's Check, Let's Sing*, and *Story Time*. In order to improve students' ability to produce speech by first enhancing their perceptions of the new language phenomena, the materials in each part were filtered using Audacity (Version 2.1.0) (see Figure 3.1), a multi-track audio editor and recorder software, at the optimal frequency, around 320 Hz. This means that all frequencies above 320 Hz were removed from the audio recording and only frequencies of 320 Hz or below were maintained. This resulted in a recording consisting exclusively of the prosodic features of the voice. The filtered sentences sounded very much like someone humming the sentences. To optimize students' perceptions of the rhythm and intonation patterns embedded in a natural stretch of speech, the sentences in all these

parts were filtered in their entirety. And in contrast, the spectrograms of the same sentence (“Hello! How are you?”), one unfiltered (left) and the other filtered (right), are displayed in Figure 3.2.

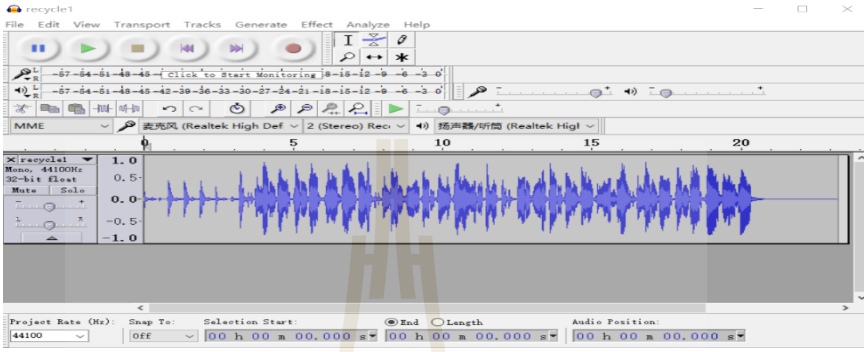


Figure 3.1 Audacity (Version 2.1.0)

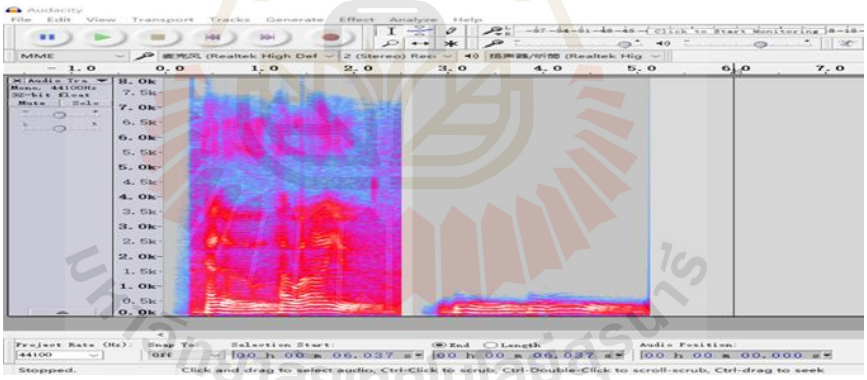


Figure 3.2 Spectrograms of an unfiltered and filtered sentence

As the focus of the study was to investigate the effectiveness of the verbotonal-based approach in both learning to speaking English and in improving PWM, the data collected included quantitative data of students’ performances in speaking and PWM as well as students’ attitudes towards VTA, and qualitative data of students’ and the

teacher's in-depth understandings of this approach in improving speaking and PWM skills. The specific procedures used to collect these data will be exposited in the following part.

3.3.2 Specific procedures

Pretests for speaking skills and PWM. To check the speaking proficiency level and PWM capacity of the two groups before the experiment and to compare the performances between them, both groups received pretests for speaking skills and PWM. Then, the two intact groups were randomly assigned into a control group and an experimental group. After that, VTA was implemented in the experimental group for a period of 18 weeks. The control group was confined to traditional way of teaching and learning of speaking skills and not allowed to have access to any of the pedagogical intervention and filtered materials given to the experimental group.

Posttests for speaking skills and PWM. After 18 weeks of experimentation, posttests for speaking and PWM were administered to both the control and experimental groups to determine if there had been any improvement within each group, and more importantly, to check whether there had been statistically significant differences between the two groups in terms of speaking skills and PWM after the intervention.

Students' and the teacher's opinions on the effectiveness of VTA as a way of improving speaking and PWM skills. During the experimentation, the researcher conducted observations to get a complete understanding of VTA. Then, following the

posttests, questionnaires and follow-up interviews were carried out with the students in the experimental group to obtain their attitudes towards and in-depth understandings of the pedagogical treatment. As the students were third-graders who had problems in understanding some questions in the questionnaires and interviews, sometimes, inevitably, there was a call for language simplicity and clarity. That is to say, the researcher had to explain certain questions in a way which was understandable and acceptable by the students. For example, the researcher asked *"Do you think you know a lot of words?"* to clarify the question *"what do you think of your vocabulary size?"* Also, to triangulate the source of data, a semi-structured interview was administered with the teacher who was in charge of the two groups of students in the experiment to gather her opinions on VTA. In addition, to comply with research ethics, all participants and their parents were informed of the purpose of the study and informed consent from them and their parents was obtained. Additionally, the confidentiality and anonymity of all the participants were protected.

3.4 Data analysis

Data analysis involves *"sifting, organizing, summarizing, and synthesizing the data"* so as to obtain the results and conclusions of the research (Seliger & Shohamy, 1989, p. 201). The data collected in the present study included both quantitative and qualitative data. The data analysis techniques employed to analyze the two types of data will be presented in the parts to come.

3.4.1 Analysis of quantitative data

The analysis of quantitative data is almost always dependent on statistics (ibid). As noted earlier, the quantitative data gathered in this study included students' scores in the pretests and posttests for speaking skills and PWM as well as students' attitudes towards VTA as a way of learning to speak in English and training PWM. After the collection of these data, descriptive statistics were first applied using SPSS (Version 16) to get basic information of the data. Then, inferential statistics including a paired samples t-test, an independent-samples t-test and Pearson correlation coefficient were utilized to conduct subsequent analyses. A statistician from the Faculty of Environmental Engineering of Kunming University of Science and Technology was invited to suggest and validate the data analysis techniques and procedures.

3.4.1.1 A paired samples t-test

A paired samples t-test was applied to examine whether there were statistically significant differences between the pretest and posttest for each group, thus, to determine the effectiveness of VTA in speaking and PWM skills improvement.

3.4.1.2 An independent-samples t-test

An independent-samples t-test was employed to decide whether there were statistically significant differences between the two groups of students before and after the experimentation in terms of speaking and PWM performances, and thus to compare the performances between the groups.

3.4.1.3 Pearson correlation coefficient

Pearson correlation coefficient was employed to reveal the relationship between PWM capacity and overall speaking skills, individual tests including word-reading, sentence-reading, singing and oral interview and subskills comprising vocabulary, grammar, pronunciation, fluency and comprehensibility.

3.4.2 Analysis of qualitative data

Qualitative data are data collected in a non-numerical form (Silverman, 2000). The qualitative data gathered in this study included students' and the teacher's opinions of VTA in improving speaking and PWM performances as well as the researcher's observations. These data were first transcribed verbatim. Then, they were organized, coded and categorized. Subsequently, the researcher explored the relationships across categories and identified themes. Finally, the categories and themes were summarized and interpreted.

3.5 Piloting

A pilot study was conducted to check the feasibility of the experiment. The pilot study spanning two weeks (from March, 1 to 15, 2016) was carried out in Class 201301, the third-graders in their second semester, of YAPS, by the teacher who was responsible for both the control and experimental groups. The researcher explained the purpose of the experiment and detailed the pedagogical procedures to her. Then, the researcher sent the filtered materials (*Unit 1 Welcome back to school*) to the teacher. After two weeks

of teaching, questionnaires were distributed to the students to obtain their opinions of the teaching approach. From Table 3.2, we can obtain that most of the students thought VTA was effective (mean=4.01) in the learning of speaking skills. VTA enabled them to memorize the sound of words more quickly (mean=3.56) and accurately (mean=3.73) than before, and they would rather learn to speak English using VTA than traditional approaches (mean=4.25). In addition, a semi-structured interview was administered to the teacher to obtain her perceptions of and suggestions on VTA. In her opinion, VTA was more helpful and interesting to the teaching of speaking than traditional approaches mainly because of two reasons. First, as the students were encouraged to conduct self-learning, students' initiative in, curiosity about and motivation towards learning were sparked. Also, the workload of the teacher was lightened, and the class atmosphere was more active than ever before. Second, VTA engaged body movements in the learning process, and thus made learning easy, enjoyable and fun. The teacher also posed her suggestions on the out-of-class session activities. Initially, in the out-of-class session, the parents were supposed to cooperate with their children by playing a role in a situation dialogue. But during the pilot study, the teacher found that almost all the parents could not fulfil the task due to their insufficient English proficiency level. So the research design was revised based on the suggestions of the teacher. In the revised version, parents were still required to cooperate with their children to role-play the situational dialogue by acting out some characters. But if their English proficiency level was inadequate for them to do so, they could just watch their child's performance. Their child was suggested to play all the roles themselves.

Table 3. 2 Descriptive statistics for students' opinions on VTA (pilot study)

	N	Mean	S.D.
1. VTA is effective in the learning of speaking.	40	4.01	0.79
2. I believe that my speaking skills are improved by using VTA.	40	3.95	0.82
3. My confidence in speaking English is improved via VTA.	40	3.86	0.93
4. My interest in learning to speak English is improved via VTA.	40	4.34	0.63
5. I am happy with VTA in the learning of speaking.	40	4.37	0.61
6. VTA is interesting in the learning of speaking.	40	4.15	0.73
7. I prefer VTA to traditional approaches in the learning of speaking.	40	4.25	0.68
8. I would like to learn to speak English via VTA in class.	40	4.41	0.59
9. I would like to learn to speak English via VTA on my own.	40	3.43	1.05
10. My learner autonomy is improved via VTA.	40	3.72	0.99
11. I am able to remember the sound of words more quickly than before via VTA.	40	3.56	1.01
12. I am able to remember the sound of words more accurately than before via VTA.	40	3.73	0.98
Valid N	40		

3.6 Summary

This chapter presented the methods and materials of the current study. The research design which included participants, variables, learning materials, pedagogical procedures was described. Then, instruments used to collect data were introduced. Finally came the data collection and analysis procedures. The next chapter will describe the results of this study.

CHAPTER 4

RESULTS

This chapter sets out to reveal the findings of the current research. It starts with the results of the statistical analyses of students' performances on both speaking skills and phonological working memory. Then, students' and the teacher's opinions of the effectiveness of VTA on their speaking proficiency gain and phonological working memory improvement are provided. After that, results from the observations bring the chapter to a conclusion.

4.1 Assessment of speaking skills

As discussed in Section 3.2.4, as a result of the difficulties existing in assessing speaking skills, most of the primary schools in China, especially those in the rural areas, do not set English speaking tests. Like most other primary schools, the Anjiaqiao Primary School (YAPS) has never held any English speaking tests for students. Therefore, the speaking tests in this study were designed specifically for eliciting speaking performances, along with satisfying the research purposes. To guarantee the validity and reliability of the tests, the test papers were validated and piloted. Results showed that the speaking tests employed in the current research are sufficiently valid and reliable (see Section 3.2.4). Students' speaking performances before and after the

experimentation were audio recorded and numbered for blind rating by three Chinese expert raters (see Section 3.2.4). In the following subsections, the pre-test and post-test statistical analyses will be provided

4.1.1 Pretest results

The construct of the pretest will be reiterated here. The pretest paper (see Appendix P) was constituted by four parts respectively: Word-reading (20%), Sentence-reading (20%), Singing (20%), and Oral interview (40%). The word-reading component included 40 words. The sentence-reading part selected 10 sentences covering various intonation and rhythm patterns. In the singing part, the students needed to sing an English song. In the oral interview part, the students were required to talk about a topic using 4-6 sentences, then, answer a question that was pertinent to the topic using 1-3 sentences. As mentioned earlier in the previous chapter, this part was devised in hopes of eliciting more authentic language in a natural and conversational style. The pretest was administered to the students from the experimental and control groups before the pedagogical treatment. Both descriptive and inferential statistics were performed to analyze the data from the pretest. The results are presented as follows:

Total score

Descriptive statistics give a fresh view of the data and serve as the basis for analysis (Dey, 2003). So, a descriptive analysis of the total scores of the pretest was first applied. Table 4.1 shows the descriptive statistics for the total scores of the two groups of students in the pretest for speaking skills. It can be obtained that the test

statistics followed a normal distribution. As can be seen from the following table, the Skewness and Kurtosis of the total scores of the experimental and control groups were -0.02, -0.51, -0.09, and -0.22 respectively. The standard deviations of the two sets of data were 10.72 and 8.14. These statistics all satisfied the requirements of normality assumptions (Morgan et al, 2000), and thus could be analyzed by parametric tests such as a paired samples t-test or an independent samples t-test. Figure 4.1 displays the histograms of the two data sets with normal distribution curves.

Table 4.1 Descriptive statistics for the total scores in the pretest

Item	Experimental group	Control group
Number	40	40
Mean	54.01	53.81
Std. Deviation	10.72	8.14
Skewness	-0.02	-0.09
Kurtosis	-0.51	-0.22
Range	58.50	32.00
Minimum	24.50	37.00
Maximum	83.00	69.00

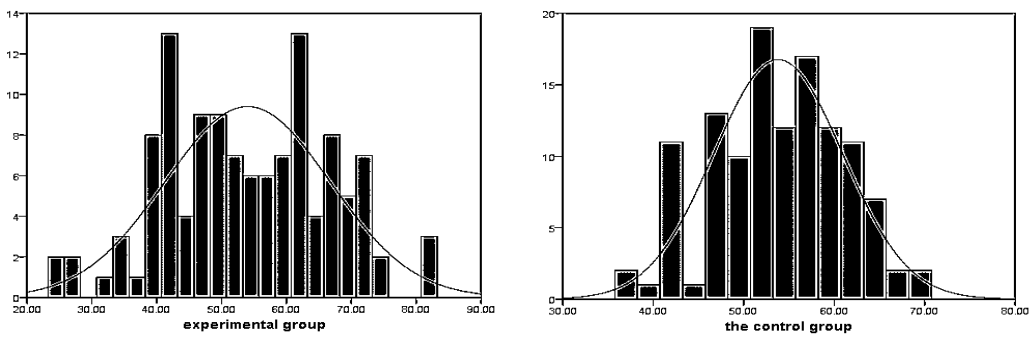


Figure 4.1 Histograms of the distribution of the total scores in the pretest

The significance level set in the present study was 0.05. An independent samples t-test was performed to compare the speaking performances of the two groups of students. Results demonstrated that there was no statistically significant difference between the experimental and control groups ($t=0.15$, $p=0.88>0.05$). That is to say, their speaking proficiency level was more or less the same before the treatment. Another point should be noted here is that, the variability of students' speaking skills within the experimental group (std. deviation=10.72, range=58.50) was greater than that of the control group (std. deviation=8.14, range=32.00).

Statistical analyses were also applied to the four individual parts to examine if there had been any differences between the two groups.

Part 1 Word-reading

The pretest started with the word-reading part. The test-takers were asked to read the 40 words that were all selected from the textbook, E*nglish (Grade 3) (Vol.1),* which was used by Grade 3 students in their first semester in YAPS. The overall maximum score of this part added up to 20.

Table 4.2 presents the results of the descriptive statistics for word-reading. An independent samples t-test was performed to see whether the two groups of students' performances in reading these words were significantly different or not. It can be found that there was a statistically significant difference between the performances of the experimental and control groups ($t=-2.69$, $p=0.01<0.05$). Specifically, the control group outperformed the experimental group in reading the 40 words before the experiment.

Table 4.2 Descriptive statistics for word-reading in the pretest

Item	Experimental group	Control group
Number	40	40
Mean	13.64	15.19
Std. Deviation	2.98	2.07
Range	11.50	9.00
Minimum	8.50	10.00
Maximum	20.00	19.00

Part 2 Sentence-reading

This part included 10 sentences covering various intonation and rhythm patterns. All these sentences were collected from the textbook *English (Grade 3) (Vol.1).* The overall maximum score of this part was 20.

Table 4.3 shows the descriptive statistics for the sentence-reading part. The results of the independent samples t-test indicated that there was no statistically significant difference between the two groups of students in reading the sentences (t=-0.58, p=0.57>0.05). In other words, the two groups' ability to read the 10 sentences were almost at the same level before the pedagogical intervention.

Table 4.3 Descriptive statistics for sentence-reading in the pretest

Item	Experimental group	Control group
Number	40	40
Mean	12.50	12.70
Std. Deviation	3.30	2.40
Range	13.00	8.00
Minimum	6.00	8.00
Maximum	19.00	16.00

Part 3 Singing

The song of the singing part was chosen from *Recycle 1* of the textbook, *English (Grade 3) (Vol.1)*. The overall maximum score of this part totaled 20.

Table 4.4 provides the results of the descriptive statistics for the singing part. The results of the independent samples t-test analysis demonstrated that there was no statistically significant difference between the two groups of students (t=1.15, p=0.25>0.05). To put it another way, the pre-test singing performances of the experimental and control groups were very similar.

Table 4.4 Descriptive statistics for singing in the pretest

Item	Experimental group	Control group
Number	40	40
Mean	11.70	10.95
Std. Deviation	3.78	1.63
Range	14.00	7.00
Minimum	5.00	7.00
Maximum	19.00	14.00

Part 4 Oral interview

The oral interview component was designed to assess students' speaking skills in terms of pronunciation, grammar, vocabulary, fluency and comprehensibility. The three Chinese expert raters were required to give two sets of scores. The first was the overall proficiency score, and the second was the scores for each of the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility. The overall

maximum score for this part amounted to 40, and the score for each subskill was 30. The statistical analyses of the oral interview part will be presented in the following sequence: overall proficiency score, and scores for vocabulary, grammar, pronunciation, fluency and comprehensibility.

Overall proficiency score

Table 4.5 shows the descriptive statistics for the overall proficiency scores of the two groups of students. An independent samples t-test was applied to compare the performances between the experimental and control groups in the oral interview. The results indicated that there was no statistically significant difference between the experimental and control groups in the oral interview part ($t=1.49$, $p=0.14>0.05$) of the pretest.

Table 4.5 Descriptive statistics for overall proficiency scores in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	16.40	15.05
Std. Deviation	4.59	3.43
Range	19.00	14.00
Minimum	5.00	9.00
Maximum	24.00	23.00

The results of the statistical analyses of the five sub-skills will be provided as follows:

Vocabulary

Table 4.6 presents the statistics for vocabulary. Students' ability to use vocabulary in terms of lexical richness and accuracy was rated by three Chinese expert raters independently. The results of the independent samples t-test demonstrated that there was no statistically significant difference between the two groups (t=1.31, p=0.19>0.05). That is, the two groups of students started at the same level for vocabulary use before the treatment.

Table 4.6 Descriptive statistics for vocabulary in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	13.32	12.12
Std. Deviation	4.24	3.92
Range	16.00	17.00
Minimum	3.00	2.00
Maximum	19.00	19.00

Grammar

Table 4.7 displays the statistics for grammar. The results of the independent samples t-test analysis showed that there was no statistically significant difference between the two groups (t=-0.03, p=0.98>0.05). Specifically, the experimental and control groups performed at the same level for grammar before the treatment.

Table 4.7 Descriptive statistics for grammar in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	11.57	11.60
Std. Deviation	4.40	4.10
Range	17.00	16.00
Minimum	2.00	2.00
Maximum	19.00	18.00

Pronunciation

Table 4.8 shows the statistics for pronunciation. The independent samples t-test analysis demonstrated that there was no statistically significant difference between the two groups (t=-0.46, p=0.65>0.05). Specifically, the two groups of students started at the same level in terms of pronunciation before the treatment.

Table 4.8 Descriptive statistics for pronunciation in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	17.70	18.00
Std. Deviation	3.29	2.56
Range	14.00	11.00
Minimum	9.00	12.00
Maximum	23.00	23.00

Fluency

Table 4.9 provides the statistics for fluency. The independent samples t-test analysis indicated that there was no statistically significant difference between the two groups of students (t=-1.13, p=0.26>0.05). Specifically, both groups performed at the same level for fluency before the treatment.

Table 4.9 Descriptive statistics for fluency in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	11.60	12.70
Std. Deviation	4.41	4.28
Range	17.00	17.00
Minimum	2.00	2.00
Maximum	19.00	19.00

Comprehensibility

Table 4.10 provides the statistics for comprehensibility. The independent samples t-test was performed to compare the performances between the two groups. It was obtained that there was no statistically significant difference between the two groups (t=-1.07, p=0.29>0.05). Specifically, the two groups of students were at the same level in terms of comprehensibility before the treatment.

Table 4.10 Descriptive statistics for comprehensibility in the oral interview part of the pretest

Item	Experimental group	Control group
Number	40	40
Mean	18.35	19.12
Std. Deviation	3.64	2.78
Range	15.00	12.00
Minimum	9.00	12.00
Maximum	24.00	24.00

4.1.2 Posttest results

The construct of the posttest was the same as that of the pretest (see Appendix R). It was also constituted by four parts: Word-reading (20%), Sentence-reading (20%), Singing (20%), and Oral interview (40%). All the materials were selected from the textbook *English (Grade 3) (Vol.2)*, which was used in the second semester by Grade 3 students in YAPS. The posttest was administered to the students from the experimental and control groups after the pedagogical treatment. Both descriptive and inferential statistics including a paired samples t-test and an independent samples t-test were utilized to analyze the data from the posttest. The results are presented as follows.

Total score

A descriptive analysis of the total scores of the posttest was first performed to check if these statistics followed a normal distribution and thus were able to be analyzed by parametric tests. Table 4.11 displays the descriptive statistics for the total scores of

the two groups of students in the pretest and posttest. It can be seen that the test statistics followed a normal distribution. The Skewness and Kurtosis of the total scores of the experimental and control groups were -0.53, -0.09, -0.37, and -0.84 respectively. The standard deviations of the two sets of data were 13.27 and 8.46. These statistics all met the requirements of normality assumptions (Morgan et al, 2000). Figure 4.2 shows the histograms of the two data sets with normal distribution curves.

A paired samples t-test was first employed to compare the performances between the pretest and posttest for each group, and thus to check for any improvements in speaking skills within the experimental and control groups respectively. Results showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-13.43$, $p=0.00<0.05$) and no significant difference for the control group ($t=0.18$, $p=0.86>0.05$). Specifically, the experimental group improved significantly after the treatment whereas the control group's performance did not change. Then, an independent samples t-test was applied to check whether there had been any difference between the speaking proficiency of the two groups in the posttest. Results indicated there was a statistically significant difference between the performances of the experimental and control groups ($t=6.49$, $p=0.00<0.05$). That is to say, the experimental group significantly outperformed the control group after the intervention. We can see from Table 4.11 that the variability of students' speaking skills within the experimental group (std. deviation=13.27, range=59.00) was greater than that of the control group (std. deviation=8.46,

range=31.00). This finding echoes the pretest where there was a wider variation in students' speaking skills in the experimental group (std. deviation=10.72, range=58.50) than in the control group (std. deviation=8.14, range=32.00).

Table 4.11 Descriptive statistics for the total scores in the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	54.01	70.23	53.81	54.09
Std. Deviation	10.72	13.27	8.14	8.46
Skewness	-0.02	-0.53	-0.09	-0.09
Kurtosis	-0.51	-0.37	-0.22	-0.84
Range	58.50	59.00	32.00	31.00
Minimum	24.50	37.00	37.00	39.00
Maximum	83.00	96.00	69.00	70.00

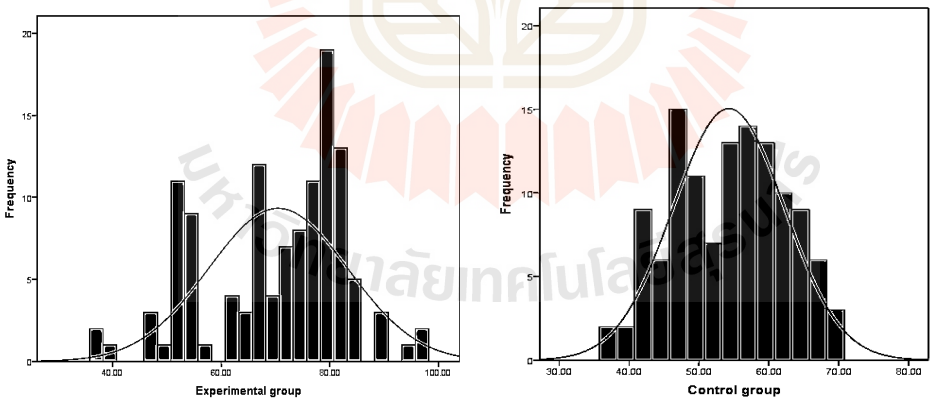


Figure 4.2 Histograms of the distribution of the total scores in the posttest

A descriptive analysis was applied to the four individual components to discern where the differences situated.

Part 1 Word-reading

Table 4.12 Descriptive statistics for word-reading in the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	13.64	17.25	15.19	15.76
Std. Deviation	2.98	2.72	2.07	2.11
Range	11.50	10.00	10.00	9.00
Minimum	8.50	10.00	9.00	10.00
Maximum	20.00	20.00	19.00	19.00

Table 4.12 presents the descriptive statistics for word-reading in the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for both the experimental (t=-8.40, p=0.00<0.05) and control groups (t=-5.31, p=0.00<0, 05). Specifically, both of the two groups progressed significantly in word-reading after the pedagogical intervention. An independent samples t-test was applied to see whether the two groups' performances in reading these words were different or not. It was found that there was a statistically significant difference between the performances of the experimental and control groups (t=2.74, p=0.01<0.05). That is, the experimental group caught up the difference and went ahead of the control group in reading the same 40 words after the experiment.

Part 2 Sentence-reading

Table 4.13 Descriptive statistics for sentence-reading in the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	12.50	16.25	12.70	12.85
Std. Deviation	3.30	2.84	2.40	2.08
Range	13.00	9.00	8.00	8.00
Minimum	6.00	11.00	8.00	9.00
Maximum	19.00	20.00	16.00	17.00

Table 4.13 provides the descriptive statistics for sentence-reading in the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-8.51$, $p=0.00<0.05$) and no significant difference for the control group ($t=-1.34$, $p=0.19>0.05$). Specifically, the experimental group advanced significantly after the treatment whereas the control group witnessed no significant progress. An independent samples t-test was performed to see whether the two groups of students' performances in reading these sentences were different or not after the treatment. It was found that there was a statistically significant difference between the experimental and control groups ($t=8.04$, $p=0.00<0.05$). That is, the experimental group was stronger than the control group in reading the same sentences after the experiment.

Part 3 Singing

Table 4.14 Descriptive statistics for singing in the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	11.70	13.60	10.95	10.58
Std. Deviation	3.78	4.14	1.63	2.08
Range	14.00	12.00	7.00	8.00
Minimum	5.00	7.00	7.00	6.00
Maximum	19.00	19.00	14.00	14.00

The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-3.26$, $p=0.00<0.05$) and no significant difference for the control group ($t=-0.86$, $p=0.39>0.05$). Specifically, the experimental group made significant progress in singing after the treatment whereas the control group went a little backward albeit in a non-significant way (mean in the pretest=10.95, mean in the posttest=10.58). The results of the independent samples t-test indicated that there was a statistically significant difference between the performances of the experimental and control groups ($t=3.89$, $p=0.00<0.05$). That is, the experimental group was better than the control group in singing after the experiment.

Part 4 Oral interview

The statistical analyses of the oral interview part in the pretest and posttest will be presented in the following sequence: overall proficiency score, and scores for vocabulary, grammar, pronunciation, fluency and comprehensibility.

Overall proficiency score

Table 4.15 provides the descriptive statistics for overall proficiency scores of the two groups of participants in the oral interview part of the pretest and posttest. A paired samples t-test was first employed to compare the means of the overall proficiency scores in the pretest and posttest for each group, and thus to check for any improvements in the oral interview part within groups. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-8.62$, $p=0.00<0.05$) and no significant difference for the control group ($t=-0.89$, $p=0.33>0.05$). Specifically, the experimental group advanced significantly after the treatment whereas the control group made no significant progress. The results of the independent samples t-test demonstrated that there was a statistically significant difference between the performances of the experimental and control groups ($t=7.08$, $p=0.00<0.05$). That is, the experimental group performed better than the control group in the oral interview of the posttest.

Table 4.15 Descriptive statistics for overall proficiency scores in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	16.40	22.10	15.05	15.58
Std. Deviation	4.59	5.19	3.43	3.58
Range	19.00	29.00	14.00	17.00
Minimum	5.00	9.00	9.00	8.00
Maximum	24.00	38.00	23.00	25.00

The results of the five sub-skills are provided as follows:

Vocabulary

Table 4.16 Descriptive statistics for vocabulary in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	13.32	16.25	12.12	12.42
Std. Deviation	4.24	3.60	3.92	3.85
Range	16.00	23.00	17.00	17.00
Minimum	3.00	5.00	2.00	3.00
Maximum	19.00	28.00	19.00	20.00

Table 4.16 provides the descriptive statistics for vocabulary in the oral interview part of the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for both the experimental ($t=-5.02$, $p=0.00<0.05$) and control groups ($t=-3.22$, $p=0.00<0.05$). Specifically, the two groups both performed better in terms of vocabulary after the treatment. The results of the independent samples t-test indicated that there was a statistically significant difference between the performances of the experimental and control groups ($t=4.58$, $p=0.00<0.05$). That is, the experimental group significantly outperformed the experimental group in vocabulary use in the posttest.

Grammar

Table 4.17 Descriptive statistics for grammar in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	11.57	15.55	11.60	11.77
Std. Deviation	4.40	4.09	4.10	3.51
Range	17.00	19.00	16.00	17.00
Minimum	2.00	8.00	2.00	3.00
Maximum	19.00	27.00	18.00	20.00

Table 4.17 provides the descriptive statistics for grammar in the oral interview part of the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-5.46$, $p=0.00<0.05$) and no significant difference for the control group ($t=-1.36$, $p=0.18>0.05$). Specifically, the experimental group improved significantly after the treatment whereas the control group remained almost the same. The results of the independent samples t-test showed that there was a statistically significant difference between the performances of the experimental and control groups ($t=4.42$, $p=0.00<0.05$). That is, the experimental group went far ahead of the control group in terms of grammar after the experiment.

Pronunciation

Table 4.18 Descriptive statistics for pronunciation in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	17.70	21.77	18.00	18.15
Std. Deviation	3.29	3.12	2.56	2.44
Range	14.00	12.00	11.00	11.00
Minimum	9.00	15.00	12.00	13.00
Maximum	23.00	27.00	23.00	24.00

Table 4.18 provides the descriptive statistics for pronunciation in the oral interview part of the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-8.06$, $p=0.00<0.05$) and no significant difference for the control group ($t=-1.89$, $p=0.09>0.05$). Specifically, the experimental group progressed significantly after the treatment whereas the control group's progress in pronunciation was non-significant. The results of the independent samples t-test demonstrated that there was a statistically significant difference between the performances of the experimental and control groups ($t=5.78$, $p=0.00<0.05$). Specifically, the experimental group performed better than the control group in terms of pronunciation after the experiment.

Fluency

Table 4.19 provides the descriptive statistics for fluency in the oral interview part of the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group (t=-9.24, p=0.00<0.05) and no significant difference for the control group (t=-1.00, p=0.32>0.05). Specifically, the experimental group made significant progress after the treatment whereas the control group became less fluent though in a non-significant way (mean in the pretest=12.70, mean in the posttest=12.20). The results of the independent samples t-test showed that there was a statistically significant difference between the performances of the experimental and control groups (t=5.70, p=0.00<0.05). That is, the experimental group (mean=17.62) was better than the control group (mean=12.20) in terms of fluency after the experiment.

Table 4.19 Descriptive statistics for fluency in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	11.60	17.62	12.70	12.20
Std. Deviation	4.41	4.20	4.28	4.30
Range	17.00	21.00	17.00	17.00
Minimum	2.00	8.00	2.00	2.00
Maximum	19.00	29.00	19.00	19.00

Comprehensibility

Table 4.20 provides the descriptive statistics for comprehensibility in the oral interview part of the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group (t=-9.28, p=0.00<0.05) and no significant difference for the control group (t=0.90, p=0.37>0.05). Specifically, the experimental group improved significantly in terms of comprehensibility after the treatment whereas the performance of the control group did not change. The results of the independent samples t-test indicated that there was a statistically significant difference between the performances of the experimental and control groups (t=6.13, p=0.00<0.05). That is, the experimental group significantly outperformed the control group for comprehensibility after the experiment.

Table 4.20 Descriptive statistics for comprehensibility in the oral interview part of the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	18.35	23.07	19.12	19.15
Std. Deviation	3.64	3.22	2.78	2.61
Range	15.00	14.00	12.00	11.00
Minimum	9.00	15.00	12.00	13.00
Maximum	24.00	29.00	24.00	24.00

4.2 Assessment of phonological working memory

Again, since the ability to repeat an unfamiliar phonological form - in a sense, the first step in learning a new word - greatly relies on the capacity of phonological working memory (PWM), the nonword repetition (NWR) performance has been increasingly acknowledged as a relatively effective measure of the capacity of PWM (Baddeley et al., 1998). Thus, the present study adopted the most influential NWR test - The Children's Test of Nonword Repetition established by Gathercole and colleagues in 1994 to assess participants' PWM. Students' repetition attempts were audio recorded for blind rating by two Chinese expert raters. The number of nonwords correctly repeated by each student was calculated as his/her PWM performance. The results of the statistical analysis of students' PWM capacity in the pretest and posttest are provided as follows.

4.2.1 Pretest results

Table 4.21 Descriptive statistics for PWM in the pretest

Item	Experimental group	Control group
Number	40	40
Mean	14.10	14.37
Std. Deviation	3.12	2.21
Range	10.00	9.00
Minimum	9.00	9.00
Maximum	19.00	18.00

Table 4.21 presents the results of the statistical analysis of participants' phonological working memory capacity before the experimentation. An independent

samples t-test was performed to see whether the mean differences between the two groups were significant or not. It was obtained that there was no statistically significant difference between the performances of the experimental and control groups (t=0.45, p=0.65>0.05). Specifically, the PWM capacity of both the experimental and control groups were at the same level before the experiment.

Besides, the two groups of participants' performances in repeating nonwords with different word length and wordlikeness in the pretest are presented in Table 4.22 and Table 4.23. It can be observed from the two tables that both groups performed better in repeating nonwords with 2 syllables (mean=4.35, 4.37) and with high wordlikeness (mean=8.13, 8.23) than nonwords with 4 (mean=3.03, 3.15) and 5 syllables (mean=2.71, 2.82) and low wordlikeness (mean=5.97, 6.14). The results of the independent samples t-test indicated that there were no significant differences between the two groups in repeating nonwords with 2 ((t=-0.16, p=0.89>0.05), 3(t=-0.17, p=0.89>0.05), 4 (t=-1.14, p=0.23>0.05) and 5 (t=-1.40, p=0.16>0.05) syllables as well as nonwords with high (t=-1.43, p=0.12>0.05) and low wordlikeness (t=-1.09, p=0.35>0.05). More specifically, both the experimental and control groups performed almost at the same level in repeating nonwords with different word length and wordlikeness before the intervention.

Table 4.22 Descriptive statistics for nonwords with different word length and wordlikeness in the pretest (the experimental group)

Item	Mean	Std. Deviation	Range	Minimum	Maximum
2-syllable	4.35	1.63	3	2	5
3-syllable	4.01	1.48	3	2	5
4-syllable	3.03	1.57	4	1	5
5-syllable	2.71	2.32	4	0	4
Low wordlikeness	5.97	2.75	8	1	9
High wordlikeness	8.13	1.97	8	2	10

Table 4.23 Descriptive statistics for nonwords with different word length and wordlikeness in the pretest (the control group)

Item	Mean	Std. Deviation	Range	Minimum	Maximum
2-syllable	4.37	1.59	3	2	5
3-syllable	4.03	1.48	3	2	5
4-syllable	3.15	1.67	2	2	4
5-syllable	2.82	2.21	2	2	4
Low wordlikeness	6.14	2.67	5	3	8
High wordlikeness	8.23	1.86	7	3	10

4.2.2 Posttest results

Table 4.24 provides the descriptive statistics for PWM in the pretest and posttest. The results of the paired samples t-test showed that there was a statistically significant difference between the pretest and posttest for the experimental group ($t=-5.64$, $p=0.00<0.05$) and no significant difference for the control group ($t=-0.32$, $p=0.74>0.05$). Specifically, the experimental group improved substantially after the treatment whereas

the control group remained the same. The results of the independent samples t-test indicated that there was a statistically significant difference between the performances of the experimental and control groups (t=2.06, p=0.04<0.05). That is, the experimental group significantly outperformed the control group for phonological working memory capacity after the experiment.

Table 4.24 Descriptive statistics for PWM in the pretest and posttest

Item	Experimental group		Control group	
	Pretest	Posttest	Pretest	Posttest
Number	40	40	40	40
Mean	14.10	15.90	14.37	14.50
Std. Deviation	3.12	3.05	2.21	3.01
Range	10.00	11.00	9.00	9.00
Minimum	9.00	9.00	9.00	10.00
Maximum	19.00	20.00	18.00	19.00

Additionally, the two groups of participants' performances in repeating nonwords with different word length and wordlikeness in the posttest are presented in Table 4.25 and Table 4.26. Similar to that in the pretest, both groups performed better in repeating nonwords with 2 syllables (mean=4.41, 4.39) and with high wordlikeness (mean=8.35, 8.30) than nonwords with 4 (mean=3.73, 3.17) and 5 syllables (mean=3.69, 2.85) and low wordlikeness (mean=7.55, 6.20). The results of the paired samples t-test showed that there were statistically significant differences between the pretest and posttest for the experimental group in repeating nonwords with 4 (t=-5.21, p=0.03<0.05) and 5 syllables (t=-0.76, p=0.02<0.05), and with low wordlikeness (t=-8.24,

p=0.00<005) after the experiment while no such significant differences were found for the control group.

The results of the independent samples t-test demonstrated that there were significant differences in repeating nonwords with 5 (t=2.96, p=0.02<0.05)) syllables and low wordlikeness (t=2.74, p=0.01<0.05) between the two groups. That is to say, the experimental group was stronger than the control group in repeating nonwords with 5 syllables and low wordlikeness after the intervention.

Table 4.25 Descriptive statistics for nonwords with different word length and wordlikeness in the posttest (the experimental group)

Item	Mean	Std. Deviation	Range	Minimum	Maximum
2-syllable	4.41	1.60	3	2	5
3-syllable	4.07	1.42	3	2	5
4-syllable	3.73	1.64	2	3	5
5-syllable	3.69	1.94	1	4	5
Low wordlikeness	7.55	2.65	8	2	10
High wordlikeness	8.35	2.01	5	5	10

Table 4.26 Descriptive statistics for nonwords with different word length and wordlikeness in the posttest (the control group)

Item	Mean	Std. Deviation	Range	Minimum	Maximum
2-syllable	4.39	1.59	3	2	5
3-syllable	4.09	1.48	3	2	5
4-syllable	3.17	1.67	3	2	5
5-syllable	2.85	2.21	2	2	4
Low wordlikeness	6.20	2.67	6	3	9
High wordlikeness	8.30	1.86	7	3	10

4.3 Relationship between speaking skills and PWM

This section seeks to unravel the relationship between phonological working memory and speaking skills. The results of the correlation analysis between the two variables are presented in the following subsections.

4.3.1 Pretest results

The Pearson correlation analysis was conducted to disclose the relationship between speaking skills and phonological working memory in the pretest. It can be seen from Table 4.27 that there was a statistically positive correlation between speaking skills and phonological working memory in the pretest *(r*= 0.73, p=0.00). That is to say, increases in speaking skills were correlated with increases in phonological working memory, or vice versa. Figure 4.3 is the scatter plot diagram in which the line slopes upward from zero, demonstrating again a positive correlation between the two variables.

Table 4.27 Correlation matrix for speaking skills and PWM in the pretest

		Speaking	PWM
Speaking	Pearson Correlation	1	.725**
	Sig. (2-tailed)		.000
	N	80	80
PWM	Pearson Correlation	.725**	1
	Sig. (2-tailed)	.000	
	N	80	80

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

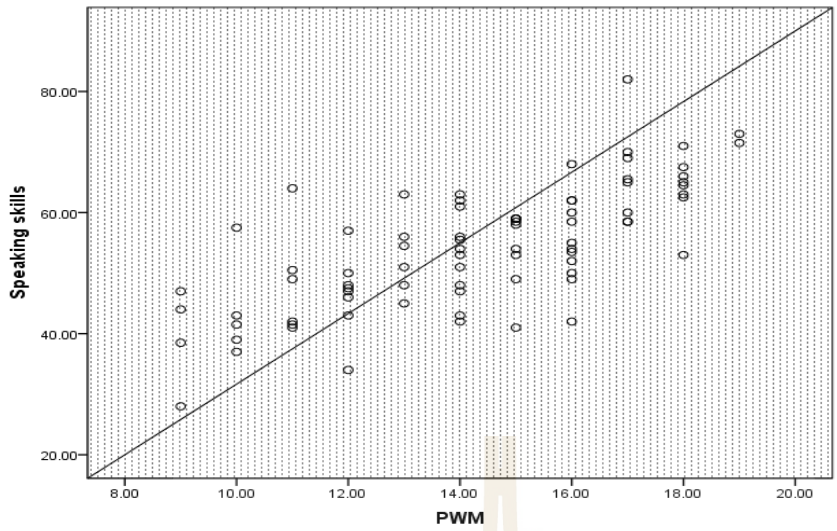


Figure 4.3 Scatter plot of speaking skills and PWM (pretest)

The Pearson correlation analysis was then applied to the four individual parts of the pretest to see their relationship with PWM. As can be obtained from Table 4.28, there were statistically positive correlations between phonological working memory and word-reading (*r*=0. 57, p=0.00), sentence-reading (*r*=0. 60, p=0.00), singing (*r*=0. 63, p=0.00) and oral interview (*r*=0. 57, p=0.00) in the pretest.

Table 4.28 Correlation matrix for individual tests and PWM in the pretest

		PWM	Word reading	Sentence reading	Singing	Oral interview
PWM	Pearson Correlation	1	.566**	.601**	.638**	.568**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	80	80	80	80	80

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

The Pearson correlation analysis was also applied to the five subskills in the pretest to see their connection with PWM. Table 4.29 reveals that there were statistically positive correlations between PWM and vocabulary (*r*=0.60, p=0.00), grammar (*r*=0.67,

p=0.00), pronunciation (*r*=0.63, p=0.00), fluency (*r*=0.70, p=0.00), and comprehensibility (*r*=0.67, p=0.00) in the pretest.

Table 4.29 Correlation matrix for subskills and PWM in the pretest

		PWM	Vocabulary	Grammar	Pronunciati-on	Fluency	Compreh-ensibility
PWM	Pearson Correlation	1	.596**	.665**	.625**	.700**	.674**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	80	80	80	80	80	80

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

4.3.2 Posttest results

Table 4.30 indicates that there was a statistically positive correlation between speaking skills and phonological working memory in the posttest (*r*= 0.79, p=0.00<0.05). The results enhanced the findings of the pretest in which a strong association between speaking skills and PWM was discovered *(r*= 0.73, p=0.00<0.05). The scatter plot diagram (see Figure 4.4) also displays a positive correlation between the two variables.

Table 4.30 Correlation matrix for speaking skills and PWM in the posttest

		PWM	Speaking skills
PWM	Pearson Correlation	1	.786**
	Sig. (2-tailed)		.000
	N	80	80
Speaking skills	Pearson Correlation	.786**	1
	Sig. (2-tailed)	.000	
	N	80	80

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

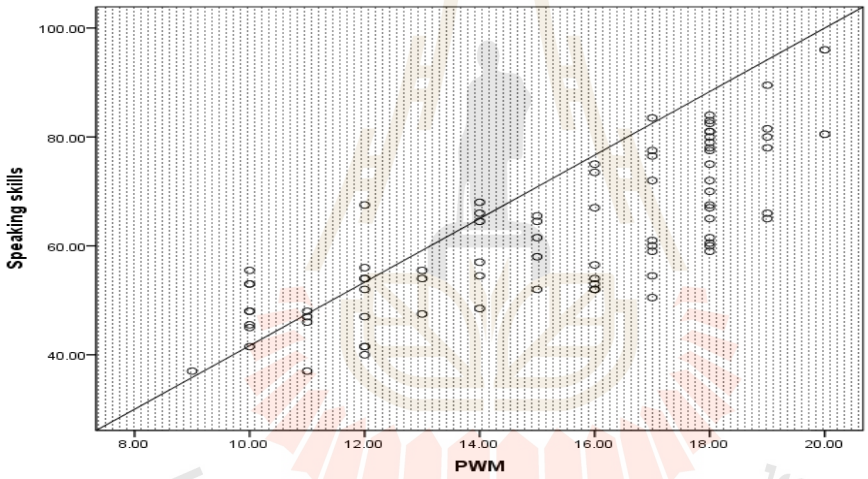


Figure 4.4 Scatter plot of speaking skills and PWM (posttest)

The Pearson correlation analysis was then applied to the four individual parts of the posttest to see their relationship with PWM. It can be obtained from Table 4.31 that there were statistically positive correlations between phonological working memory and word-reading (r=0.59, p=0.00), sentence-reading (r=0.59, p=0.00), singing (r=0.73, p=0.00) and oral interview (r=0.74, p=0.00) in the posttest.

Table 4.31 Correlation matrix for individual tests and PWM in the posttest

		PWM	Word reading	Sentence reading	Singing	Oral interview
PWM	Pearson Correlation	1	.593**	.589**	.737**	.744**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	80	80	80	80	80

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

The Pearson correlation analysis was also applied to the five subskills in the posttest to discern their relationship with PWM. It can be found from Table 4.32 that there were statistically positive correlations between PWM and vocabulary (r=0.76, p=0.00), grammar (r=0.74, p=0.00), pronunciation (r=0.70, p=0.00), fluency (r=0.73, p=0.00), and comprehensibility (r=0.70, p=0.00) in posttest. These findings are consistent with those in the pretest where PWM had proven to be positively correlated with the five subskills.

Table 4.32 Correlation matrix for subskills and PWM in the posttest

	PWM	Vocabulary	Grammar	Pronunciation	Fluency	Comprehensi-bility
PW Pearson Correlation	1	.761**	.738**	.703**	.731**	.700**
M Sig. (2-tailed)		.000	.000	.000	.000	.000
	80	80	80	80	80	80

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

The above results of the statistical analysis established that VTA was able to improve both students' speaking skills and PWM. In addition, although started almost at the same level, the experimental group outperformed the control group by a large

margin in every aspect that was being tested after the pedagogical treatment. In relation to the association between PWM and peaking skills, it was obtained that there was a significantly positive correlation between PWM and speaking skills.

4.4 Students' opinions of VTA

This section deals with students' understandings of VTA, their perceptions on the effectiveness of the treatment, and the problems they encountered during the experiment. A closed written questionnaire and semi-structured interviews were employed to gather data. All the obtained data were analyzed quantitatively or qualitatively. The results are presented as follows.

4.4.1 Results from questionnaires

As mentioned in the previous chapter, a questionnaire (see Appendix G) was constructed and distributed to all the participants from the experimental group to gather their opinions of VTA after the treatment. The questionnaire comprised two parts: 1) demographic and background information, such as age, gender, and the amount of time spent in learning to speak English each week after class, and 2) students' perceptions and attitudes towards VTA. There were 12 statements involving students' perceived effects of learning to speak English and enhancing PWM via VTA. The content validity of the questionnaire was checked using IOC (see Appendix I) and its reliability was analyzed through Cronbach's Alpha Coefficient ($\alpha = 0.95$). The questionnaire proved to be a valid tool to elicit data needed.

The questionnaire adopted a 5-point Likert scale which required the students to respond to the 12 statements by ticking the choice from “strongly agree”, “agree”, “undecided”, “disagree” and “strongly disagree”. “Strongly agree” received a weight of 5 points, agree 4, undecided 3, disagree 2, and strongly disagree 1. Hence, a higher score represents higher agreement on each item. The scale ranges: high (3.5-5.0), medium (2.5-3.4), and low (1.0-2.4) developed by Oxford (1990) were applied in the current study.

The demographic and background information collected indicated that most of the students in the experimental group were 9 years old (38 out of 40), and they all had only half-a-year English learning experience. In addition, all of them spent around 1.5 hours a week to learn to speak English outside of class.

Table 4.33 shows the quantitative data of students’ opinions of VTA as a manner of learning to speak in English and training phonological working memory. What is worth noting is that the mean score of each item was greater than 3.5. That is to say, students’ agreement on each statement was generally high according to the scale standards established by Oxford (1990). The majority of them thought VTA was effective (mean=4.83) and interesting (mean=4.92) as a way of learning to speak English. Most of them also thought that VTA improved their confidence (mean=4.52) and sparked their interest in learning to speak English (mean=4.87). They felt happy with the approach (mean=4.89). They tended to agree that they were able to remember the sound of words more quickly (mean=4.76) and accurately (mean=4.79) than before

with the help of the new learning approach. There were more students who would like to learn English speaking via VTA in class (mean=4.52) than on their own (mean=3.61). Besides, an overwhelmingly large number of students preferred to learn via VTA (mean=4.94) than via the traditional approaches. With regard to learner autonomy, most students thought VTA boosted their learner autonomy (mean=4.34).

Table 4. 33 Descriptive statistics for students' opinions on VTA

	N	Mean	S.D.
13. VTA is effective in the learning of speaking.	40	4.83	0.59
14. I believe that my speaking skills are improved by using VTA.	40	4.67	0.63
15. My confidence in speaking English is improved via VTA.	40	4.52	0.71
16. My interest in learning to speak English is improved via VTA.	40	4.87	0.60
17. I am happy with VTA in the learning of speaking.	40	4.89	0.57
18. VTA is interesting in the learning of speaking.	40	4.92	0.53
19. I prefer VTA to traditional approaches in the learning of speaking.	40	4.94	0.51
20. I would like to learn to speak English via VTA in class.	40	4.52	0.72
21. I would like to learn to speak English via VTA on my own.	40	3.61	1.07
22. My learner autonomy is improved via VTA.	40	4.34	0.81
23. I am able to remember the sound of words more quickly than before via VTA.	40	4.76	0.67
24. I am able to remember the sound of words more accurately than before via VTA.	40	4.79	0.68
Valid N	40		

4.4.2 Results from semi-structured interview s

In order to probe in-depth qualitative information of the effects of VTA as a manner of learning speaking skills and training PWM on the part of the students, semi-structured interviews were conducted by the researcher with 6 students from the experimental group. Ten open-ended questions guided the interviews (see Appendix J). The data gathered from the semi-structured interviews were first transcribed verbatim. Then, the researcher and an English teacher who had been teaching English in YAPS for 3 years, coded and analyzed the data.

All the interviewees held a positive attitude towards VTA. They thought the new approach was very effective in learning English speaking skills and improving PWM. As reported by the students, VTA had quite a few advantages over the traditional approaches. First, VTA made their learning experience motivating, enjoyable and fulfilling. Second, the program involved the experimental group in multimodal, collaborative and interactive learning. This way of learning speaking skills developed their learner autonomy and made them feel an intrinsic need for learning and using English. Third, their confidence in speaking English was boosted despite the fact that most considered themselves as failures and unworthy of success in speaking English because of their perceived limitations in using the language at the onset of the pedagogical treatment. Fourth, VTA enabled them to be faster and more accurate in remembering the pronunciation of words than before. Most important of all, their ability to store and retrieve chunks or streams of speech seemed to be strengthened.

4.5 Teacher's opinions of VTA

To gather the teacher's understandings and opinions of the pedagogical treatment, a semi-structured interview was developed (see Appendix M) and administered to the teacher who was responsible for both the experimental and control groups.

The teacher set a high value on the new approach. According to her, VTA was *"unusual," "effective"* and *"useful"*, and was capable of improving not only speaking but also other language skills such as listening and reading. She thought that this specific approach had achieved *"extremely"* positive results in the experimental group.

She told the researcher that she preferred VTA to traditional approaches in the teaching of speaking skills. The teacher also expressed that her self-efficacy which refers to *"beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments"* in teaching English speaking skills was relatively low (Bandura, 1997, p. 3) because her specialized field at college was mathematics although she had been recognized as an experienced English teacher in the district. For her, the greatest strength of VTA lay in the empowerment for the students to resolve problems and conflicts springing up in the course of learning, thereby interesting them to engage in experiential and autonomous learning. As students were encouraged to take the initiative to direct their own learning, the pressure on the part of the teacher to model in class was alleviated.

At the end of the interview, she hoped that VTA could be known by a wider population and flourish in rural primary schools in China.

4.6 Information gained from observations

Naturalistic observations in the real-world teaching context were also utilized to capture data that cannot be reached by other data collection instruments in the present study. To minimize the impact on the participants being studied, the researcher adopted an "**observer as participant**" role. Meanwhile, in an attempt to reduce the Hawthorne effect which refers to the phenomenon that the participants tend to perform better in response to their awareness of being observed, and to allow the participants time to become less aware of the presence of the researcher, a total of 12 observations (2 hours for each) had been carried out. Field notes were utilized to record both descriptive and reflective accounts of the situation being observed. To make the observations rigorous and systematic, an observation protocol (see Table 4.34) was predesigned to guide the observations. Analytic induction was applied to make sense of the data from the field notes.

Table 4.34 Observation protocol

Observation notes
Note the time of the class, the number of students
Describe the teaching settings
Describe the teacher's role in each pedagogical procedure and the problems facing her
Record students' performances in each step and the problems they encounter
Describe the class atmosphere
Record the interactions between the teacher and students
Record the interactions between the students
Add reflective accounts

According to the researcher's observation, VTA classroom was characterized by a positive and optimistic learning climate. The teacher did not teach but play a supportive role in the whole learning process. There was not any form of imposition on the students. The students were given an opportunity to be in charge of their own learning. They were empowered to negotiate and construct meaning by themselves. This empowerment made learning a joyful and rewarding experience, full of self-motivation, great enthusiasm and positive engagement. It was observed that students were willing to be checked by their teacher or the researcher for their pronunciation and speaking performances. They enjoyed practicing speaking skills both in and out of class.

4.7 Answers to research questions

- **Research questions 1**: Is VTA effective for learning to speak in English? If yes, in what ways?

After 18-week experimentaion, the experimental group improved significantly in both overall speaking proficiency and individual tests: word-reading, sentence-reading, singing, and oral interview as well as the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility. In addition, the experimental group outperformed the control group in all aspects that were tested. These findings suggest that the verbotonal approach was effective in improving the English-speaking skills of the experimental group.

- **Research questions 2**: Is VTA effective for enhancing PWM? If yes, in what ways?

The phonological working memory capacity of students in the experimental group improved substantially after the treatment. Results also indicated that the experimental group improved significantly in repeating nonwords with 4 and 5 syllables, and with low wordlikeness after the pedagogical intervention. Moreover, the experimental group significantly outperformed the control group in terms of phonological working memory performance after the experiment. These findings indicate that the verbotonal approach was of enormous use in enhancing the phonologcial working memory of the participants.

- **Research questions 3**: Is there a relationship between students' PWM performance and speaking skills? If yes, in what ways?

There was a statistically positive correlation between speaking skills and phonological working memory both in the pretest and in the posttest. In addition, statistically positive correlations between phonological working memory and individual tests including word-reading, sentence-reading, singing, and oral interview, and subskills involving vocabulary, grammar, pronunciation, fluency and comprehensibility were also discovered in the pretest and posttest.

- **Research questions 4**: What are the students' opinions of VTA as a way of learning to speak English and enhancing PWM?

All the students in the experimental group held that the verbotonal approach

was of considerable benefit to the learning of English speaking skills and improvement of phonological working memory. According to them, this approach enabled them to learn to speak English in an unusually enjoyable, interesting and motivating way. They also thought that VTA was capable of helping them to memorize the pronunciation and spelling of words more quickly and accurately than before.

- **Research questions 5**: What are the teacher's opinions of the implementation and outcomes of VTA as a way of learning to speak English and enhancing PWM?

The teacher who took the responsibility for both the experimental and control groups held a considerable positive attitude towards VTA. Specifically, she believed that VTA was effective in teaching and learning English speaking skills and enhancing PWM. She anticipated that this approach would be widely accepted by teachers teaching English in rural primary schools in China.

4.8 Summary

This chapter presented the results of the present study. The statistical analyses indicated that the experimental group improved tremendously both in speaking skills and phonological working memory while the control group made no significant progress after one semester's (from March 1 to July 25, 2016) study. Further, the experimental group outperformed the control group in every aspect that was tested. In regards to the relationship between speaking skills and phonological working memory,

a significantly positive correlation was spotted between them. Besides, both the teacher and students viewed VTA as a very effective approach to learn speaking skills and enhance PWM. Insights gained from the observations also served as evidence for the effectiveness of the intervention. The forthcoming chapter will focus on discussing the major findings of the current research.

CHAPTER 5
DISCUSSION

As described in the preceding chapter, the findings of the current research reveal the effectiveness of the verbotonal system in improving speaking skills as well as in enhancing phonological working memory (PWM) capacity. This chapter presents a discussion on these findings. The discussion is structured around the five research questions proposed in Chapter 1. First and foremost, the results of the assessment of speaking skills are discussed. Then comes the discussion on the enhancement of PWM via VTA. After that, a discussion on the correlation between PWM and speaking skills is provided. Finally, both students' and teacher's opinions of VTA are explored.

5.1 Discussion on the results of speaking skills

This section discusses students' speaking performances before and after the intervention. Students' overall performances will be discussed first. Then, findings on all individual tests: word-reading, sentence-reading, singing, and oral interview along with the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility will be discussed.

5.1.1. Overall performances

As reported in Section 4.1.1，no significant difference was found between the two groups of participants in terms of overall speaking performances in the pretest. That is, the English-speaking skills of the experimental (mean=54.01) and the control (mean=53.81) groups were very similar in the pretest. This was because the two groups of students had similar English learning backgrounds before the experiment. More specifically, the experimental and control groups were drawn from two intact classes that were taught by the same teacher. According to the teacher, the two classes were "parallel classes," which meant that their academic performances on all subjects including English were at the same level. All the students began to learn English from the first semester of Grade 3, and no single student had been receiving any form of out-of-class training in English language skills except the homework assigned by the teacher before and during the experimentation. Moreover, the parents of these students were villagers in Anjiaqiao Village. They made a living mainly from growing and selling vegetables. Their English proficiency level was found to be too inadequate to assist their children to learn English. It is therefore not surprising to find that the overall speaking performances of the two groups of participants were almost the same in the pretest. Although all the students had more or less the same learning backgrounds, the individuality in English learning cannot be underestimated. As presented in 4.1.1, there was a greater variation in students' speaking skills in the experimental group (std. deviation=10.72, maximum=83.00, minimum=24.50) than in the control group (std.

deviation=8.14, maximum=69.00, minimum=37.00) in the pretest. This was confirmed by the teacher who reported that there was a big gap between students' speaking performances in the experimental group as some students seemed to have a facility for the acquisition of a second language while others seemed to have great difficulty in learning English.

What is worth reiterating here is that the control group was taught to pronounce segmental elements and apply grammatical rules and translation skills deliberately. In contrast, the experimental group was never exposed to the specific teaching for any of these items. However, although starting at the same level, after 18 weeks of experimentation, the experimental group advanced significantly in spoken English proficiency, while no significant improvement was found in the control group. The mean of the total scores of the experimental group increased by 16.22 (from 54.01 to 70.23), a rise of 30%, indicating a significant difference between the pretest and posttest (p=0.00) whereas no significant difference between the pretest and posttest scores (p=0.86) were found in the control group whose mean increased only by 0.28 (from 53.81to 54.09). Further, the results of the independent samples t-test suggested that the experimental group significantly outperformed the control group in terms of overall speaking performances after the intervention (p=0.00). These findings indicate that VTA was of substantial advantage to the learning of speaking skills in the experimental group. Like that in the pretest, the variability of students' speaking skills in the experimental group (std. deviation=13.27, maximum=96.00, minimum=37.00) was

greater than that in the control group (std. deviation=8.46, maximum=70.00, minimum=39.00). This phenomenon reveals that, though being exposed to the same approach, learning tended to be an individualized and idiosyncratic experience. It can be concluded that VTA was able to improve students' speaking skills and learning with VTA could promote more significant outcomes than learning with the traditional approach as the verbotonal approach proposed a holistic, multi-sensory and multi-modal manner of learning to speak English. VTA held that an individual speaker was a perceiver, and at the same time, a producer of speech. If we want to improve his/her speech, we need first to improve his/her perception (Guberina & Asp, 1981). The playful activities such as listening to and repeating the filtered materials and original sentences in chorus, utilizing body movements and visual resources, and role-playing the dialogues in the pedagogic procedures all contributed significantly to changing students' perceptions and improving their production of speech.

5.1.2 Word-reading

The first part of both the pretest and posttest consisted of word-reading wherein 40 words with 2 to 4 syllables required by the syllabus were listed for students to read aloud. The overall maximum score of this part amounted to 20 points.

As can be obtained in 4.1.1, in the pretest, the control group (mean=15.19) performed better than the experimental group (mean=13.64) in reading the 40 words (p=0.01). In the posttest, both the experimental (mean=17.25, p=0.00) and control (mean=15.76, p=0.00) groups advanced substantially. However, the mean score of the

experimental group witnessed an increase of 26% against 3% for that of the control group. It means that the experimental group made up the differences and overtook the control group by a large margin in terms of word-reading after the treatment (p=0.01).

Moreover, by analyzing the recordings of students' performances, it can be found that, in the pretest, there were mainly three problems that faced students in the word-reading part. The biggest problem for both groups lied in word recognition. As stated before, all 40 words were selected from the textbook. And they were specified by the syllabus as the keywords that should be mastered by Grade 3 students after completion of one semester of English. However, many words in the word-reading part appeared to be unfamiliar for students from the two groups. Most of the students (87%) fell short of recognizing words such as *pencil, mouth, white, brother, black, bread, school, bird* and *seven.* Another problem that hindered students' performances in this part was their inability to pronounce certain vowels and consonants accurately (segmental features). To illustrate, some students mispronounced [fɪʃ] as [fɪs], [pɪg] as [pɪ:g], [mʌm] as [mamu], ['sev(ə)n] as ['sew(ə)n], [bæg] as [beg], etc. When it came to the posttest, both groups made significant progress (mean=17.25, p=0.00; mean=15.76, p=0.00), but the experimental group performed better than the control group in word recognition, segmental pronunciation and word stress identification. The control group had more problems than the experimental group in recognizing and pronouncing words such as *Canada, eighteen, grape, strawberry, fruit* and *watermelon.*

The control group's improvement in word-reading in the posttest was expected, as they had plenty of practice in reading words. As reported by the teacher, in the traditional instruction setting, she spent most of the class time teaching pronunciation, spelling and translation skills. She served as the model and source of English language input. The students were supposed to follow the model of their teacher. On the contrary, the students in the experimental group undertook VTA, and never had any specific training for pronouncing segments and spelling words as the concern of the approach was on suprasegmental features and entire sentences. So, in this sense, the experimental group's considerable progress in the word-reading part was intriguing and thought-provoking.

There are several language processing skills operating concurrently in reading words, for instance, word recognition and segmentation, segmental pronunciation, sound discrimination as well as word stress identification (Carter, 2012; Gervain & Werker, 2013). In modern phonetics, segments are defined as *"any discrete unit that can be identified, either physically or auditorily, in the stream of speech"* such as phones and phonemes or vowels and consonants depending on different categorizing criteria (Crystal, 2011, p.408-409). Suprasegmentals are most often used to refer to those abstract properties of speech which coexist with and cannot be derived from multiple segments, for instance, tone, stress, rhythm, intonation, pitch and loudness, in other words, the prosody or prosodic features. Although suprasegmental features are not included in normal orthography or in conventional segmental phonetic transcriptions, they have been shown to be an integral part of human speech and

facilitate the acquisition of both native and foreign languages (Nooteboom, 1997). Studies indicate that newborns are sensitive and able to discriminate the prosodic patterns between different languages, and these prosodic cues are of great help to the processing of word-level information (Gonzalez- Gome & Nazzi, 2012; Ragó et al., 2014). For example, Nazzi and Ramus (2003) argued that infants, though starting with very little lexical and language-specific phonological knowledge, were capable of using prosodic features to group syllables, segment and recognize words, and identify word boundaries as early as the second half of the first year of their life. Also, according to Gonzalez-Gomez and Nazzi (2012), the prosodic properties of language play a central role in phonetic/phonotactic acquisition, especially in stress pattern and phoneme information processing.

As reviewed in Section 2.3.2, He Bi (2015) implemented a verbotonal-based approach, which she called as *"the CALL-VT system"*, to teach EFL pronunciation to Chinese university students. In her study, after the pedagogical treatment, the experimental group who learned English pronunciation with the CALL-VT program went far ahead of the control group who were taught with the traditional articulatory approach in reading the same 50 words. The findings of the current research are in concordance with the results of her study, confirming the effectiveness of the verbotonal approach in improving students' ability in reading words, the assessment of which relies heavily on segmental pronunciation, without intentional teaching for segmental production. Emphasizing suprasegmental features was just one aspect of the verbotonal

approach. Clearly, the effectiveness of this approach in improving word-reading performance cannot only be attributed to this one single aspect. There must be mechanisms operating in synchrony to contribute to improving both segmental and suprasegmental pronunciation in the absence of the specific training for the former. So, following the general discussion of results, major characteristics of the verbotonal approach will be discussed to account for its effectiveness in improving word-reading and other language skills.

5.1.3 Sentence-reading

The part following word-reading in both pretest and posttest consisted of sentence-reading. Again, to make the task commensurate with students' levels of ability and knowledge of the English language, all the 10 sentences in the pretest were selected from the textbook E*nglish (Grade 3) (Vol.1),* and those in the posttest were drawn from the textbook E*nglish (Grade 3) (Vol.2).* Different sentence types including declarative, imperative, and interrogative sentences were covered. This part was intended to test students' pronunciation of both segmentals and suprasegmentals at the sentence level. The overall maximum score of this part added up to 20 points.

As shown in 4.1.1, there was no statistically significant difference between the experimental (mean=12.50) and control groups (mean=12.70) in the sentence-reading part in the pretest (p=0.57). This means the performances of the two groups of students in reading the same 10 sentences before the pedagogical treatment were at the same level. However, in the posttest, it was found that the mean score of the experimental

group rose to 16.25, an increase of 30%, demonstrating a significant improvement (p=0.00). On the other hand, the mean of the control group only increased from 12.70 to 12.85, a rise of 1%, 30 times smaller than that of its counterpart, showing non-significant progress in sentence-reading in the posttest (p=0.19). Further, the statistical results indicated that the experimental group was ahead of the control group in terms of reading sentences (p=0.00).

The assessment of the sentence-reading part focuses not only on segmental elements but also on suprasegmentals such as fluency, rhythm, intonation, duration, tempo, pitch, and loudness. After the treatment, the majority of students in the experimental group were faster and more precise in dealing with both segmental and suprasegmental features at the sentence level. They performed especially well in mastering the intonation patterns of the *wh*-question (e.g., Where is my pencil box?) and the *yes/no* question (e.g., Do you like oranges?).

As discussed earlier, in the pretest, the control group was better than the experimental group in reading words, but when it came to reading sentences, the performances of the two groups were observed to be the same. It reveals that better performances in reading words does not necessarily guarantee better results in reading sentences because the sentence-reading task involves broader linguistic, cognitive and processing skills including speech perception, vocabulary knowledge, syntactic parsing, semantic integration, and language segmentation and production than the word-reading task does (Gonzalez- Gome & Nazzi, 2012; Ragó, 2014).

A sentence takes the form of a linear string of words. That linear string does not occur randomly but is organized by a series of established phonological, morphological and syntactic rules. A large pool of evidence indicates that suprasegmental features, the rhythmic and melodic aspects of language, play a vital role in processing the acoustic, phonetic, phonotactic, semantic and syntactic information at the sentence level (Nazzi & Ramus, 2003; Hawthorne & Gerken, 2014; Ragó et al., 2014). For instance, Pell and Baum (1997) found that pitch (F0 frequency) had an impact on the correct classification of spoken words in languages such as German and English. Prosody was also shown to cue word order in 7-month-old bilingual infants (Gervain & Werker, 2013). Langus et al. (2012) demonstrated that prosody helped adults with segmenting speech into phrases that were embedded in sentences. In addition, Hawthorne and Gerken (2014) found that prosody could facilitate the learning of syntactic constituency.

The better performance of the experimental group served as valuable evidence of the great value of the verbotonal approach in increasing sentence-reading performance. As VTA attaches greater importance to the prosodic properties of language than the traditional approach, this finding supports the above-mentioned research, revealing again the facilitative role of prosody in acquiring segmental elements at the sentence level.

5.1.4 Singing

The third part of the pretest and posttest for speaking skills consisted of singing an English song. The songs in the pretest and posttest were selected from the textbooks used by the third-graders in YAPS in their first and second semesters respectively.

As presented in Section 4.1.1, the results of the independent samples t-test analysis demonstrated that there was no statistically significant difference between the performances of singing the English song for the experimental (mean=11.70) and control (mean=10.95) groups in the pretest (p=0.25). This was confirmed by the teacher who reported that the two groups of students showed substantial enthusiasm and interest in learning to sing the English song in each unit of their textbook and their ability to sing these songs were approximately at the same level before the pedagogical intervention. For the teacher, singing English songs was the most interesting activity for the two groups of students before the experiment. She thought that teaching speaking skills through singing was easy-going, motivating and enjoyable. There is indeed a large body of evidence showing the merits of using singing in the teaching and learning of foreign languages. For example, Good et al. (2015) asserts that group singing in the classroom is both hedonically and cognitively stimulating. Singing is able to create a positive and agreeable atmosphere that is conducive to learning (Paquette & Rieg, 2008). This risk-free and joyful environment created by learning through singing can also help to develop confidence, encourage participation and promote pro-social behavior on the part of the learners (Kirschner & Tomasello, 2010). Besides, learning via singing is capable of focusing attention and enhancing levels of concentration (Bird, 2008; Sandberg, 2009). There are also numerous empirical studies that demonstrate the effectiveness of singing in facilitating foreign language learning. For example, Ludke et al. (2014) carried out an experiment to study the effectiveness of singing in short-

term paired-associate phrase learning in Hungarian. Sixty adults (non-native speakers of Hungarian) participated in their study. All the participants were randomly assigned to one of three learning conditions: speaking, rhythmic speaking, and singing conditions. Results showed that participants who were exposed to the singing condition outperformed the participants in the speaking and rhythmic speaking conditions in recalling and producing spoken Hungarian phrases after a 15-minute learning period. Similarly, Good et al. (2015) investigated the effect of singing in EFL learning. In their study, two groups of Spanish-speaking Ecuadorian children learned a novel English passage for two weeks. One group conducted their learning in a singing condition in which the passage was learned as a song while the other group was assigned to the spoken condition wherein the passage was learned as an oral poem. After the experimentation, the children's performances in recalling the passage verbatim, pronouncing English vowel sounds, and translating target terms from English to Spanish were tested. It was shown that the group in the singing condition outperformed the group in the spoken condition in all aspects that were being tested.

As discussed above, singing plays a facilitative role in learning various aspects of language. Both the teacher and students in the current research held a positive attitude towards the singing activity in their English class. However, in the posttest, only the experimental group improved significantly in singing (mean=13.60, p=0.00). The control group's singing performance was found to become worse albeit not in a significant way (mean=10.58, p=0.39). In addition, the experimental group outperformed the control

group in terms of singing in the posttest (p=0.00). In VTA, students learned to sing an English song by listening to and humming the filtered materials, and repeating the original sentences in the recordings while in the traditional approach, the teacher usually taught the students to read the lyrics first, and then required the students to follow her model of singing. In the posttest, the experimental group was found to be more accurate than the control group in combining the tune with the melody and lyrics of the song. It can be observed that VTA was demonstrably more effective and beneficial than the traditional approach in improving students' ability to sing English songs.

5.1.5 Oral interview

A face-to-face oral interview in which an interviewee is questioned on a series of specified topics has been considered as a robust technique for assessing speaking skills (Brown, 2000). The popularity of this technique is mainly derived from the belief that it is an appropriate means of testing not only linguistic factors such as lexical richness, grammatical accuracy, and pronunciation etc., but also aspects which are known as communicative and interactional skills (Brown, 2003). In an oral interview, candidates are encouraged to demonstrate their ability to produce a variety of eliciting functions, for example, to take the initiative, to seek information, to express ideas and attitudes with precision, to switch register, and to use non-verbal features and affective components, etc. (Ingram & Wylie,1993; Brown, 2000). Given all the strengths of an oral interview in assessing speaking proficiency, in this study, oral interviews were adopted both in the pretest and posttest to elicit students' authentic language in natural

settings. The oral interview for each student ran for approximately 5 minutes. It was intended to assess students' speaking skills in terms of pronunciation, grammar, vocabulary, fluency and comprehensibility. As mentioned before, three Chinese expert raters were required to give two sets of scores. The first set was the overall proficiency score, and the second set was the scores for the five subskills: pronunciation, grammar, vocabulary, fluency and comprehensibility. The overall maximum score for this part added up to 40, and the score for each of the subskill was 30.

In the pretest, the students were required to use 4 to 6 sentences to describe different parts of their body, and then answer a question "*Which part of your body do you like most? Why*? " As reported in Section 4.1.1, the results of the statistical analysis showed that the experimental (mean=16.40) and control groups (mean=15.05) performed the same in the oral interview part before the pedagogical treatment (p=0.14). In comparison to their performances in other individual tests, namely, word-reading (14.40 out of 20, 72%), sentence-reading (12.60 out of 20, 63%) and singing (11.32 out of 20, 57%), the two groups of students' performances in the oral interview were poorer (14.83 out of 40, 37%). This poorer performance can be attributed to the fact that, unlike the reading and singing tasks where students did not need to improvise speech, the oral interview focused more on the students' ability to utilize linguistic and extra-linguistic features to produce a flow of meaningful speech without preparation in advance, which is the principal goal of learning speaking skills.

In the posttest, the students were expected to talk briefly about their family members, and then answer the question "*Do you like your family? Why?"* Results showed that the experimental group made great progress in the oral interview (p=0.00), the mean of the overall proficiency scores increasing from 16.40 to 22.10, a 35% increase. However, no significant improvement was found in the control group (p=0.33), the mean of the overall proficiency scores rising from 15.05 to 15.58. On top of that, the experimental group went ahead of the control group by a large margin in the overall assessment of the oral interview after the pedagogical intervention. This finding was in concordance with He Bi's (2015) study, revealing that the verbotonal approach was a better alternative for improving students' overall performance in an oral interview than its counterpart.

The following space will be set aside to discuss the findings on the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility.

Vocabulary

Vocabulary knowledge is a complex construct that involves various language-related aspects, for example, form and meaning, grammatical characteristics, and collocations (Nation, 2001, Van Zeeland & Schmitt, 2013). In the pretest, the experimental group (mean=13.32) and the control group (mean=12.12) started at the same level in terms of vocabulary as judged by the three expert raters (p=0.19). Both groups were not good at retrieving appropriate words to express their ideas. Most students in the two groups were unable to answer the question *"Which part of your*

body do you like most? Why? *"* Again, this finding was unsurprising because the two groups of participants had almost the same English learning history. They acquired vocabulary mainly through the in-class formal teaching from their teacher. Normally, the teacher taught the students to read and spell words specifically, and then familiarized them with word meaning and usage knowledge. During the experimentation, the control group continued to learn lexicons this way whereas the experimental group was assigned to VTA in which participants never received the explicit teaching for vocabulary given to the control group. VTA stressed the importance of acquiring speaking skills from the sentence level rather than the word level. All the sentences in this study were filtered and repeated in their entirety.

In the posttest, both the experimental ((p=0.00) and control (p=0.00) groups improved significantly in lexical range and accuracy. The mean of the experimental group rose from 13.32 to 16.25, an increase of 22 percentage points, 11 times greater than that of the control group whose mean increased from 12.12 to 12.42, a rise of 2 percentage points. An independent samples t-test was performed to compare the performances of the two groups. Results showed that there was a statistically significant difference between the performances of the experimental and control groups (t=4.58, p=0.00<0.05). Specifically, the experimental group performed significantly better than the control group in the posttest. Evidence from the recordings also demonstrated that the lexical repertoire, i.e., depth and breadth of lexical knowledge, of the students from the experimental group was greater than that of the students from the control group.

Their vocabulary size was larger than that of their counterparts. When interviewed, they even used some words which did not appear in their textbook such as *"friendly"*, *"kind"*, *"farmer"*, *"vegetable"*, *"market"*, etc.

The improvement of the control group was within expectation because they had intensive practice in reading, spelling and memorizing words. However, the positive outcomes of the experimental group were unexpected and thus worth noting. As is widely known, vocabulary is crucial to proficiency in any language (Coady & Huckin, 1997; Hosoda et al., 2013). Being a foundation of FL/L2 learning, vocabulary acquisition is key to both spoken and written language production, verbal fluency, and reading comprehension ability (Nation, 1993; Laufer & Nation, 1995; Qian, 2002; Luo et al., 2010; Pu et al., 2016). The investigation of vocabulary acquisition has a history of more than a hundred years (Laufer, 2009). It has risen as a prominent specialist field of research over the past few decades with many established sectors examining different aspects of explicit and implicit instruction and acquisition of vocabulary (Laufer, 2009; Nation, 2001, Eckerth & Tavakoli, 2012). There is research into second/foreign language vocabulary acquisition indicating that intentional and direct word learning is more effective than incidental word learning where learners acquire lexicons without being focused on doing so (De la Fuente, 2002; Laufer, 2005). For example, incidental word learning through listening and reading has shown to result in non-significant learning gains (Barcroft, 2009; Sonbul & Schmitt, 2010).

In the current research, the control group received the explicit teaching for vocabulary while the experimental group engaged in incidental word learning. However, evidence showed that the experimental group was ahead of the control group after completing the experiment. This finding is at odds with the above-mentioned literature, demonstrating the potential exemplary merits of incidental vocabulary acquisition via VTA.

Grammar

Grammar has been recognized as a base for all language-related skills: listening, speaking, reading and writing, and it functions as building blocks of second/foreign language acquisition (Azar, 2014). According to Canale and Swain (1980), grammatical competence is also a critical component of communicative competence. In this study, the participants in the experimental and control groups performed the same in terms of grammatical accuracy in the pretest (p=0.98), the mean of the experimental group being 11.57 against 11.60 for that of the control group. However, the experimental group enjoyed substantial progress in the posttest, with the mean increasing from 11.57 to 15.55, a rise of 34 percentage points (p=0.00). On the other hand, the performance of the control group did not change significantly (p=0.18), with the mean increasing from 11.60 to 11.77. Evidence from the recordings showed that in the pretest, both groups could barely use isolated words to communicate ideas. For instance, when asked to introduce the different parts of their body, most students could not produce grammatically correct sentences but only come up with isolated words such as *"eye,"*

"ear," "mouth," etc. However, in the posttest, students in the experimental group were found to be able to use more grammatically correct sentences in their speech than students from the control group.

Central to the instruction of grammar is the debate that whether grammar should be taught explicitly (specific teaching for grammatical rules) or implicitly (exposure to comprehensible and meaningful input without resources to rules) (Loewen et al., 2009). In fact, looking back at the history of second/foreign language teaching, explicit grammar instruction has long been recognized as the organizing principle of many teaching methods, especially the well-known grammar-translation method (Christison, 2015). The traditional model of explicit grammar instruction highlights conscious presentation and manipulation of forms through intensive drills and practice (Nazari, 2013). It has been widely criticized for teaching grammar in an isolated and decontextualized way (Christison, 2015). An increasing number of researchers who emphasize the need for explicit teaching have directed their attention to explore more effective approaches to grammar instruction (Ranalli, 2001). Amongst the approaches, *"*Focus on Form,*"* which refers to drawing the learners' attention to grammatical forms in a communicative context, is the most popular representative (Burgess & Etherington, 2002). According to Nassaji and Fotos (2011), a suitable explicit teaching approach ought to enable learners to "*encounter, process and use*" the target forms in different ways so as to internalize the form and integrate it into their interlanguage (p.130). Nazari (2013) claims that if learners receive communicative exposure to grammatical knowledge that has already been presented explicitly, they will be able to develop a

longer-lasting awareness of the knowledge and their grammatical accuracy will improve. Azar (2007) created a grammar-based teaching approach (GBT) which incorporated grammar instruction into communicative teaching methodology. In contrast to explicit instruction, the implicit instruction of grammar focuses on acquisition of grammatical features without awareness. Krashen (1982) holds that explicit instruction cannot contribute to grammatical gains and instead, he advocates natural and unconscious acquisition of grammar. Similarly, Ellis (2009) claims that implicit instruction is capable of providing learners with contexts wherein they can infer grammatical rules without awareness and the rules inferred this way will be internalized without having learners' attention focused on them.

Many studies have been carried out to demonstrate which type of instruction is more effective, the results of most show that explicit instruction tends to be more useful than the implicit type (Norris & Ortega, 2000, Jean & Simard, 2013).

In this study, the control group learned grammar in an explicit way. The teacher raised students' awareness of the grammatical points and then illustrated the rules. Students needed to memorize the rules learned. The medium of instruction was Chinese. In contrast, the experimental group acquired grammar implicitly. The teacher did nothing to raise students' awareness of grammatical features. Grammar rules were never taught to the students in the experimental group specifically. The better performance of the experimental group was significant because it demonstrated that VTA had created an effective implicit learning environment for grammar.

Pronunciation

As described in 4.1.1, the two groups of participants performed almost the same for pronunciation in the pretest (p=0.65). However, after the treatment, the experimental group saw a significant improvement (p=0.00), whose mean rose from 17.70 to 21.77, whereas the control group's performance showed little change (p=0.09), with the mean increasing from 18.00 to 18.15. When the performances of the two groups were compared, it was found that the experimental group outperformed the control group by a large margin (p=0.00).

Pronunciation covers both segmental features which phonetics and phonemics deal with, and suprasegmental elements which operate on longer stretches of utterances than individual sounds or phonemes (Cruttenden, 2014). When examining the recordings of the two groups, it was found that the experimental group had fewer segmental and suprasegmental errors than the control group.

In the course of the past five decades, the role of pronunciation in second/foreign language teaching has undergone substantial changes. In the grammar-translation method, pronunciation was mostly irrelevant. While in some other approaches such as the direct method and audiolingualism, it was elevated to a pinnacle of importance (Levis, 2005). Recent decades witness the heydays of the communication-oriented approach. With the ascendency of the communicative teaching approach, pronunciation has been seen more *"in light of how it can help achieve communicative competence without setting unrealistic standards of attainment"*

than a primary focus of speaking proficiency (Jones, 1997, p.69). Many language teaching specialists have sought to integrate pronunciation and oral communication. To illustrate, Pennington (2014) posits that to become a competent language perceiver and producer, language learners should pay attention to not only the acoustic and articulatory features of pronunciation, but also to the meaningful correlations of those features in the immediate linguistic context as well as the real communicative context.

The past few decades also saw a fundamental change in the content of pronunciation teaching (Jones, 1997, Koike, 2014). To be more detailed, there was a change in focus from segments (i.e., vowel and consonant sounds) to suprasegmentals (i.e., rhythm, intonation and stress) due to the increasing awareness of the communicative function of suprasegmental features in spoken discourse (Brazil et al., 1980; Brown & Yule, 1983; Zarifi & Sayyadi, 2015). VTA in the current research highlighted the role of suprasegmentals in raising participants' phonological awareness. The experimental group never undertook the training for segmental pronunciation given to the control group. However, in the posttest, the experimental group improved considerably and outperformed the control group. This finding was consistent with previous research manifestly demonstrating that suprasegmental features contribute significantly to improving learners' pronunciation (Avery & Ehrlich, 1992; McNerney & Mendelsohn, 1992; Celce-Murcia et al., 1996).

Fluency

Speaking a language fluently is the chief aim of learning a new language (Kormos & Dénes, 2004). Although fluency is a commonly used concept in the field of second/foreign language teaching, it is difficult to reach a precise definition of it. (Chambers, 1997; O'Brien, 2014). It is common to find that even among language teachers and learners, there is more or less confusion as to what fluency exactly means (Chambers, 1997). Crystal (1987) defined the term as *"smooth, rapid and effortless use of language"* (p. 421). With the rise of the communicative language teaching (CLT) approach, the definition of fluency extends to encompass both linguistic knowledge and communicative competence as well. Fluency in the CLT context means learners' effective use of knowledge to achieve linguistic and communicative purpose (Chambers, 1997; Canale & Swain, 1980). Lennon (1990) proposed both broad and narrow definitions of the term. According to him, fluency in a broad sense is synonymous with overall oral proficiency. Fluency in a narrow sense, on the other hand, denotes fluidity of speech and is composed of temporal measures including speech rate, hesitations, and pausing. The present study adopted the narrow definition of fluency put forward by Lennon because it restricts fluency in the speech stream to temporal properties that are empirically quantifiable and identifiable.

Students' fluency in the oral interview was assessed by three expert raters independently. In the pretest, the mean for fluency of the experimental group was 11.60, and that for the control group was 12.70. The results of the statistical analysis

demonstrated that there was no significant difference between the two groups. That is, both the experimental group and the control group started at the same level in terms of fluency in the pretest (p=0.26). After the pedagogical intervention, the two groups were given a posttest for fluency in their spoken production. The results of the paired samples t-test showed that there was a significant difference between the pretest and posttest for fluency in the experimental group (p=0.00), whose mean increased from 11.60 to 17.62, a sharp rise of 52 percentage points. On the contrary, the mean for fluency in the control group decreased from 12.70 to 12.20, albeit in a non-significant way (p=0.32). When the performances of the two groups were compared, it was found that the experimental group was far better than the control group.

These above findings were in conformity with the teacher's impression. According to the teacher, before the experiment, both groups showed intermediate fluency when speaking English. Generally speaking, they exhibited a slow rate of speech. It was not uncommon to find unnatural hesitations and pauses in their utterances. In other words, students had word-finding pauses and problems when speaking English. However, the teacher reported that, after the treatment, the experimental group appeared to be able to speak English much more fluently and naturally than before. Their spontaneous oral language turned out to be smooth, coherent and effortless. That is, they were capable of speaking in an automatic and reflex-like way after completing VTA. The experimental evidence suggests that VTA also acts on fluency in spontaneous spoken production.

Comprehensibility

One of the chief goals in learning a second/foreign language is to be understood by listeners (Munro & Derwing, 1999; Murphy, 2014). *"The ease or difficulty with which a listener understands L2 accented speech"* is termed as comprehensibility, which has been acknowledged as a very important facet of pronunciation (Derwing et al., 2008, p. 360). Comprehensibility has shown to play a more important role in facilitating communication in an L2/FL context than accentedness or nativelikeness (Derwing & Munro, 2009). As comprehensibility is listeners' perceived difficulties of understanding L2/FL speech, fundamental to comprehensibility is how it is assessed by listeners with different linguistic backgrounds (O'Brien, 2014). Some studies compared the ratings for comprehensibility provided by listeners with different language backgrounds. For example, in their study, Munro and colleagues (2006) invited listeners from native Cantonese, Japanese, Mandarin, and English backgrounds to evaluate the same set of English utterances of a total of 48 native speakers of Cantonese, Japanese, Polish, and Spanish. They found that, in spite of varied native language backgrounds, the comprehensibility ratings provided by all listeners were positively correlated. Despite somewhat between-group differences, there was general agreement on which of the 48 speakers were the easiest and most difficult to understand.

In the present study, non-native expert raters evaluated students' EFL comprehensibility. In the pretest, the experimental (mean=18.35) and control (mean=19.12) groups started at the same level in terms of comprehensibility (p=0.29).

It was observed in the posttest that there was a significant increase of comprehensibility in the experimental group (mean=23.07, p=0.00), but the performance of the control group did not change (mean=19.15, p=0.37). Moreover, the experimental group was far more ahead of its counterpart in terms of comprehensibility after the treatment. This finding, again, was in accordance with the teacher's perception of comprehensibility in students' oral performances. The teacher stated that students in the experimental group tended to produce more comprehensible speech than the control group after the completion of VTA. When asked what factors she thought had contributed to comprehensibility, the teacher replied: *"intonation, stress, rhythm and accent."* Her opinion was in line with some studies examining raters' perceived factors that affect comprehensibility. Their findings revealed that L2 speakers' better control of suprasegmental elements led to more comprehensible speech (Derwing & Munro, 1997; Wennerstrom, 2001; Kang et al., 2010; Murphy, 2014). For example, Kang (2010) investigated the relative weights of individual temporal and prosodic features for listeners' evaluations on comprehensibility and accentedness in an L2 context. In his study, five minutes of continuous in-class lectures given by 11 international teaching assistants were acoustically analyzed for measures of speech rate, stress, pauses and pitch using the PRAAT computer program. Then, 58 American undergraduate students rated the speech production of the international teaching assistants. Results indicated that prosody independently contributed to listeners' ratings for comprehensibility.

To sum up, VTA has shown to act both on overall oral proficiency and its constituent parts. The positive outcomes of the present study echoed previous research carried out by Lian (1980), Hu and Uno (2005), Zhang (2006), Alazard et al. (2011), and He (2015), confirming the effectiveness of the verbotonal-based approach in improving the teaching and learning of a foreign language.

5.2 Discussion on the results of phonological working memory

This section discusses the results of the evaluation of students' phonological working memory (PWM) capacity before and after the experiment. A discussion on students' overall performances will be provided first. Then, discussions on students' performances in repeating nonwords of different word length as well as of high and low wordlikeness are given.

5.2.1 Overall performances

Being the most investigated modular system of the multi-component working memory model proposed by Baddeley (2000), phonological working memory (i.e. the phonological loop) is specialized for the maintenance and manipulation of verbal-acoustic information over a matter of seconds. It comprises two related but separable subcomponents: a brief store that retains incoming inputs in phonological form together with a subvocal rehearsal system which acts to register both auditory and nonauditory information into the store through subvocalization. Nonword repetition tasks are assumed to assess the capacity of the phonological loop in that the phonological form

of a nonword is necessarily unfamiliar, so the listener has to rely heavily on the capacity of their phonological memory system to encode and maintain the novel phonological sequence for sufficiently long (Gathercole & Baddeley, 2014). Moreover, compared with other measures of PWM, nonsense word repetition test is more natural because hearing and repeating novel lexical items is an inevitable step in the learning of any language (Gathercole & Baddeley, 1993). Therefore, as noted before, the present study adopted the Children's Test of Nonword Repetition created by Gathercole and colleagues (1994) to assess students' PWM.

In the pretest, there was no statistically significant difference between the experimental (mean=14.10) and control (mean=14.37) groups (p=0.65). To be more specific, the PWM capacity of the two groups belonged to the same level before the experiment. However, the experimental group improved significantly (mean=15.90, p=0.00) and outperformed the control group (m=14.50) in terms of PWM in the posttest (p=0.04). As stated before, the verbotonal approach was primarily implemented for language therapy and teaching. It has never been utilized for training phonological working memory in the literature. However, the basic principles and practices of VTA were found to be conducive to the improvement of phonological working memory. So, VTA was applied in the present study to improve students' speaking skills, and at the same time, to strengthen their PWM. The experimental evidence indicated that VTA was of great use not only in increasing students' oral proficiency gain but also in enhancing their PWM. As regarding working memory training, there are generally two

approaches. One is termed as explicit or strategy training (Klingberg, 2010; Morrison & Chein, 2011). The explicit training approach is intended to teach domain-specific strategies such as rehearsal and chunking to increase memory capacity. And the other is implicit or core training which focuses on repetition of WM tasks engaging general WM mechanisms. In this sense, VTA is more an implicit training approach because it did not teach strategies explicitly to enlarge memory capacity. The basic elements within the approach involving prosody, body movements and situational learning have been proven to target general WM mechanisms. They all help particularly at the encoding stages of memory, thereby increasing performances on memory tasks (Thaut et al., 2005; Sammler et al., 2010; Adams, 2013; Shintel, 2014).

5.2.2 Nonwords of different word length

As noted in Section 3.2.5, both the pretest and posttest for PWM consisted of 20 nonwords, 5 each containing two, three, four and five syllables. All these nonwords were coined to be in conformity to the phonotactic and prosodic rules of English. The two syllable nonwords follow the strong-weak stress pattern, and the three syllable follow the strong-weak-weak pattern. There is no single dominant stress structure for four and five syllable English words.

Students' correct repetitions of nonwords of different syllable length were calculated. In the pretest, both the experimental and the control groups performed almost the same in repeating nonwords with two to five syllables ($p=0.89$, $p=0.89$, $p=0.23$, $p=0.16$). After the experiment, the control group showed no progress in

repeating nonwords with different syllable length. Whereas the experimental group improved significantly in repeating nonwords with four ($t=-5.21$, $p=0.03<0.05$) and five syllables ($t=-0.76$, $p=0.02<0.05$). An independent samples t-test was employed to compare the performances of the two groups. Results indicated that there was a significant difference between the experimental and control groups in repeating nonwords with five syllables ($p=0.02$). Specifically, the experimental group was better than the control group in repeating nonwords with five syllables in the posttest.

Additionally, it was found that the two groups had better performances in repeating nonwords with two syllables (mean in the pretest for the experimental and control groups: 4.35, 4.37; mean in posttest: 4.41, 4.39) than in repeating nonwords with four (mean in the pretest=3.03, 3.15; mean in the posttest: 3.73, 3.17) and five syllables (mean in the pretest: 2.71, 2.82; mean in the posttest: 3.69, 2.85) both in the pretest and posttest.

The finding that both of the two groups performed better in recalling shorter nonwords is hardly surprising because it is supported by a range of studies examining the impact of word length on PWM performance. According to Baddeley (2000), forgetting results from simple trace decay within the store of the phonological loop. It is assumed that longer sequences take longer to articulate during rehearsal and output, allowing more trace decay and thus leading to poorer performance (Gathercole & Baddeley, 2014). For instance, in their experiment, Baddeley and colleagues (1975) demonstrated that participants repeated approximately 80% monosyllabic words

correctly, but repetition accuracy dropped to about 50% when recalling multisyllabic words.

There is another finding which is promising. That is, VTA increased participants' performances in recalling lengthy nonwords. This implies that after exposed to VTA, the participants were able to perceive, maintain and articulate longer novel phonological representations than before, which, undoubtedly, gave convincing evidence of the effectiveness of VTA in increasing participants' PWM capacity.

5.2.3 Nonwords of high and low wordlikeness

Studies evidence that repetition accuracy is strongly influenced by the rated wordlikeness of nonwords (Gathercole et al., 1991; Gathercole, 1995; Gathercole, 2006). Both the pretest and posttest in this study comprised 20 nonwords, 10 being high-wordlike and 10 being low-wordlike. The wordlikeness ratings were obtained by Gathercole and colleagues (1991). In their study, they invited a total of 20 adults to rate all the 40 nonwords on a 5-point scale ranging from 1 (very unlike a real word) to 5 (very like a real word). The rating was *"based on the extent to which the sound structure of a spoken nonword would pass for a real word"* (Gathercole, 1995, p.86). The mean for each nonword was kept as its final wordlikeness rating.

In the pretest, both groups had better performances in repeating nonwords with high (mean for the experimental group: 8.13, mean for the control group: 8.23) rather than low wordlikeness (mean for the experimental group: 5.97, mean for the control group: 6.14). The same trend was observed in the posttest where the two groups still

performed better in repeating nonwords of high wordlikenss (mean for the experimental group: 8.35, mean for the control group: 8.30) than of low wordlikeness (mean for the experimental group: 7.55, mean for the control group: 6.20). Besides, it was found that, in the pretest, both the experimental and control groups performed at the same level for repeating nonwords of high ($t=-1.43$, $p=0.12>0.05$) and low wordlikeness ($t=-1.09$, $p=0.35>0.05$). However, in the posttest, the experimental group significantly outperformed the control group in repeating nonwords of low wordlikeness ($t=2.74$, $p=0.01<0.05$).

There is much supportive evidence for the finding that both groups repeated nonwords of high wordlikeness more accurately than they repeated nonwords of low wordlikeness. Gathercole (1991, 2006) found that language knowledge exerted a strong impact on the nonword repetition performance. Participants tend to have higher level of performances in repetition of nonwords of high wordlikeness. This is because for repeating nonwords of high wordlikeness such as “trumpetine” and “defermication”, they are more likely to retrieve similar phonological structures from long-term lexical memory. However, for repeating stimuli of low wordlikeness such as “perplisteronk” and “loddernapish”, there is less long-term lexical knowledge to be dependent on, thereby resulting in lower level of performance (Gathercole, 1995).

Consistent with this, Dollaghan et al. (1993) tested the impact of the lexical status (word or nonword) of the stressed syllables of multisyllabic nonwords on repetition performance. Their study revealed that participants repeated nonwords with

lexical stressed syllables more accurately than nonwords with nonlexical stressed syllables. Additionally, increases in phonotactic probability which refers to *"the frequency with which phonological segments and sequences of phonological segments occur in words in a given language,"* have been found to facilitate nonword repetition performance (Vitevitch & Luce, 2005, p.193; Munson et al., 2005).

It can be concluded from the above discussion that repetition of low-wordlike stimuli is largely dependent on phonological memory because there is less available long-term lexical knowledge to resort to. In the current research, the experimental group who undertook VTA improved greatly and outperformed the control group in repeating nonwords of low wordlikeness. This established the evidence that the verbotonal approach led considerable value to the enhancement of PWM.

5.3 Discussion on the relationship between speaking skills and PWM

In this section, the relationship between speaking skills and phonological working memory is discussed. The discussion on the relationship between overall speaking proficiency and PWM is presented first. Then, the relationship between PWM and individual tests as well as subskills is discussed.

5.3.1 The relationship between overall speaking proficiency and PWM

As reported in Chapter 4, there is a statistically positive correlation between speaking skills and phonological working memory both in the pretest (*r*= 0.73, p=0.00) and in the posttest (*r*= 0.79, p=0.00<0.05). This means that increases in speaking skills

lead to increases in phonological working memory capacity, or vice versa. This finding is within expectations. Phonological working memory (PWM) is considered a language learning device (Baddeley et al., 1998), or a key component of language learning aptitude (Carroll, 1962; Skehan, 1998). The review of available evidence supports the claim that there is a close link between PWM and both native and second/foreign language development. For example, PWM has proved to predict reading comprehension ability (Gathercole & Baddeley 1993; Gathercole, 1995; Muter & Snowling, 1998; Seigneuric et al., 2000; Xu et al., 2015). Considerable experimental studies also support a strong association between PWM and listening and reading skills (Swanson, 1996; McInnis, 2003; Kellogg, 2007; Kormos, 2008). Besides, several studies investigating the relationship between PWM and speech production suggest that PWM is uniquely linked to speaking proficiency (Adams & Gathercole, 1995, 2000; Alloway et al., 2005; O'Brien et al., 2006, 2007), which was confirmed by the present study. It should be pointed out that correlation alone does not suggest causation. More empirical studies need to be carried out to examine the intricate relationship between speaking proficiency and PWM.

5.3.2 The relationship between PWM and individual tests

A Pearson correlation analysis was performed to examine the links between PWM and the four individual tests in the pretest. It was found that there was a statistically positive correlation between phonological working memory and word-reading (r=0. 57, p=0.00), sentence-reading (r=0. 60, p=0.00), singing (r=0. 63, p=0.00)

and oral interview (*r*=0. 57, p=0.00). This was confirmed by the results obtained in the posttest wherein direct links also persisted between PWM and word-reading (r=0.59, p=0.00), sentence-reading (r=0.59, p=0.00), singing (r=0.73, p=0.00) and oral interview (r=0.74, p=0.00).

Interpretation of the above findings calls for a sound theoretical understanding of the cognitive systems underlying the phonological loop. One cognitive skill which has been considered strongly linked to both PWM and learners' capacities to acquire language skills is phonological awareness which refers to the ability to perceive, encode, and manipulate the sounds of language (Alloway, 2004). The first and foremost step towards phonological awareness is identification of syllables and rhymes which are characterized by *"slow, temporal rhythmic modulations, including changes in vowel intensity or amplitude rise time (ART)"* (Ugolini, 2016, p.17). Experimental evidence validated the connection between phonological awareness, language and literacy abilities. In this sense, it is not unexpected to find that PWM was closely related to all the individual tests of speaking skills both in the pretest and posttest of the study.

5.3.3 The relationship between PWM and subskills

A Pearson correlation analysis was also applied to discern the relationship between PWM and the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility, both in the pretest and posttest. Results showed that there were statistically positive correlations between PWM and vocabulary (*r*=0.60, p=0.00), grammar (*r*=0.67, p=0.00), pronunciation (*r*=0.63, p=0.00), fluency (*r*=0.70, p=0.00),

and comprehensibility (r=0.67, p=0.00) in the pretest. The posttest also demonstrated the same tendency. Specifically, positive correlations persisted between PWM and vocabulary (r=0.76, p=0.00), grammar (r=0.74, p=0.00), pronunciation (r=0.70, p=0.00), fluency (r=0.73, p=0.00), and comprehensibility (r=0.70, p=0.00) in the posttest. Again, most of these findings are in agreement with previous studies revealing the connection between PWM and measures of speaking skills such as lexical richness, grammatical accuracy, and fluency.

As reviewed in Chapter 2, Adams and Gathercole (1995) examined the connection between PWM and spoken language development in preschool children. Their findings indicate that, compared with that of children with poor phonological memory capacity, the speech production of children with good phonological memory capacity are characterized by a higher-level of grammatical accuracy, richer lexical density and longer utterances. Similarly, O'Brien et al. (2006, 2007) investigated the relationship between PWM and speech output by English-speaking adults learning Spanish. They pointed out that PWM was positively correlated with narrative skills at earlier stages of language learning and with grammatical competence at later stages. Moreover, PWM was found to contribute significantly to oral fluency gain.

Reviewing the literature, it can be found that there are few studies concerning the association between PWM and two important L2 speech performance measures, namely, pronunciation and comprehensibility. Thus, the finding that there is a direct link between PWM and pronunciation and comprehensibility in the current research is

of immense significance because it sheds light on the understanding of the operationalization of PWM and its role in oral proficiency gain, and even in language development in general.

5.4 Students' and teacher's perceptions of VTA

This section first comments briefly upon the students' and teacher's perceptions of the verbotonal approach as a manner of learning to speak English and training phonological working memory. Then, findings emerging from the observations conducted by the researcher are discussed.

5.4.1 Students' perceptions

All the students in the experimental group considered VTA more effective in gaining speaking proficiency than traditional approaches. According to them, this approach was helpful and interesting mainly in four ways.

First, this new way of learning speaking skills aroused their interest in learning. As one of the interviewee said, *"It's my first time to know that learning of speaking English can be as interesting, easy, and enjoyable as this. We listened to very soft music to relax ourselves at the beginning of the class. Then, we were exposed to the tunes of the sentences. After that, we hummed and danced to the prosody, mouthed the words, repeated the original sentences in chorus and acted out the dialogues. All the steps are fresh and fun to us."* Other interviewees all expressed similar ideas. One interviewee also said that *"when it was time for us to attend the English class, we all would become*

excited and happy. We could not wait to start our class. We learned English speaking in such an unusual, fascinating and impressive way instead of a dull, stereotyped, and boring way."

Second, VTA stimulated their motivation towards learning. Most of the interviewees reported that they had a stronger desire for learning to speak English than ever before. When interviewed, an interviewee said that *"We enjoyed participating the activities and became eager to speak English."* Some other interviewees articulated their opinions that in this new approach, they learned English the way they learned to sing and dance. As there were less pressure, stress, push, interruption and imposition in VTA class than in the traditional one, the students themselves felt a strong inclination to learn to speak in English both in and out of class. An interviewee told the researcher: *"I could not help humming the intonation patterns and uttering the original sentences on the way I headed for my home after school."*

Third, VTA increased their confidence in speaking English. One interviewee shared with the interviewer: *"When I spoke, I found my pronunciation is very much like that of a real foreigner. I felt excited and was quite proud of that. Besides, when I spoke, I could think of the word I needed more easily than before. I could understand others and also made myself understood. So I became very confident when I spoke English."*

Last but not least, the new approach cultivated their creativity and initiative in learning. The interviewees all expressed that in VTA class, they were given more opportunities to initiate learning by themselves, as one interviewee said: *"We could do*

whatever we favored to relax ourselves and dance to the sentences. This prompted us to create our own understandings of the 'music'." Another interviewee said, *"In the acting-out phase, our group always did our utmost to invent new plots and props."* All the interviewees held that compared with traditional approaches, they preferred to learn speaking skills via VTA.

When asked whether VTA was capable of helping them to memorize the pronunciation of words more quickly and accurately than before and improving their listening skills or not, all of them answered in the affirmative. They found that they could remember not only the sounds but also the spelling of words in a more effective and efficient way. Their listening skills had also been improved. According to them, under the influence of this new teaching approach, they could blurt out not only the tunes but also the original sentences effortlessly.

When faced the question *"Do you think that VTA can develop your learner autonomy and why*?*"* 3 out of 6 interviewees thought the new system developed their learner autonomy dramatically. They said that they *"formed a habit of engaging in learning by themselves," "did not rely too much on their teacher to perform corrections and give answers,"* and *"took pleasure in resolving problems and deciding learning materials."* Another 3 out of 6 interviewees thought their learner autonomy had been improved *"to some extent."* They found that despite the help of the new approach, they still have difficulties in *"setting goals for learning," "managing their own learning process,"* and *"addressing the problems emerging in learning."*

Most interviewees (4 out of 6) preferred to learn speaking skills via VTA in groups rather than on their own. They thought learning with others could produce positive learning outcomes and made learning a thing of joy, exploration and cooperation. As one interviewee said, *"I loved to learn in groups because I could discuss with other group members, and we could work together to overcome difficulties. This 'feel' and atmosphere of teamwork is great."*

As defined by Fishbein and Ajzen (1975), attitude is *"a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object"* (p. 6). A positive attitude is more likely to engender better learning achievements (Liu, 2014; Oetomo, 2014). So the extremely positive attitude held by the students toward VTA in this study formed one of the sources of the extremely fruitful results achieved by the approach.

5.4.2 The teacher's perceptions

A semi-structured interview was conducted with the teacher to obtain her perceptions of the effectiveness of VTA. The teacher firmly believed that VTA had improved the students' speaking skills. She saw the system as a very helpful way of teaching and learning English speaking skills mainly because of the following three reasons.

To begin with, it freed her from the huge pressure of modelling in class. As in VTA class, students were exposed to the recorded filtered materials and original sentences. So, the teacher did not need to model the pronunciation of words and

sentences. This lightened her workload, and more importantly, removed her worries about her own inaccurate pronunciation of certain words. The teacher told the researcher that she originally majored in mathematics at college. Upon graduation, she was assigned to teach mathematics in YAPS. Due to the lack of English teachers, she was then appointed to teach English to grades 3 and 4 students in the school. She had been teaching English for 4 years. Although being an experienced English teacher in the district, she considered her pronunciation as not *"standard."* But she did not know how to improve the situation. As soon as the researcher detailed the aims of the treatment and modeled a class for her, she was fascinated by the new teaching approach, and in her own words, *"fell in love with it immediately."*

Next, VTA enabled the students to take an active role and be involved in class almost all the time. The teacher reported that, as required by the experimental design, during the whole pedagogical procedures, she acted only as facilitator, supporter and organizer. She did not teach but oversee the whole learning process, and monitored students' performances and provided help when needed. As students were entitled to manage their own learning in each step, they became more concentrated in class.

Last, the new learning system improved students' confidence in speaking English and allowed students more opportunities to practice speaking skills both in and out of class than traditional teaching approaches. The teacher found that the listening, humming, dancing and role-playing activities made students' learning of speaking skills easy, joyful and effective. All the students grew in their willingness to speak

English than ever before. The teacher presented to the researcher, *"It was no surprise to hear the humming or repetition everywhere, say, in the classroom, on the playground or on their way back home."* Because of these advantages, the teacher preferred VTA to traditional approaches.

In the teacher's opinion, VTA had also improved students' pronunciation and their ability to memorize words. According to her, through the intensive training, students' "feel" towards English had been cultivated. The students are found to be more accurate both in pronouncing and spelling words. Additionally, the teacher held the position that VTA had fostered students' ability to adjust their own learning. She expressed that in all the pedagogical procedures, students could not rest on her to perform corrections and to carry out instructional duties. This drove the students to become an active constructor instead of a passive receiver of knowledge.

When asked what else she would like to say about VTA, the teacher stated *"I think VTA is of considerable practical value in the teaching of English in the schools of rural areas where there is a severe shortage of qualified English teachers. So I hope this approach can be known and implemented by more and more teachers teaching English in rural primary schools."*

The teacher's positive view towards VTA constituted another source of the success of the approach.

5.4.3 Findings from the observations

In order to triangulate the source of data, a series of 12 observations were conducted to obtain an in-depth understanding of VTA. Three main findings were identified as follows.

The first finding is about the **teacher's role**. After training by the researcher, the teacher mastered the principles and practices of the experiment. It was observed that the teacher conformed to all the pedagogic steps specified in the experimental design. What she actually did in class was to serve students, guide them to accomplish all the tasks and provide assistance when needs arose. Most of time, she would join the students in listening to the materials and repeating them. Sometimes, she would demonstrate her body movements in rhythm and hum along with the melody to facilitate students' understandings of the prosody of sentences. But she would emphasize that those were only her understandings of the tune of English and students did not need to follow her. If students made mistakes, instead of correcting them directly, she would say something like this *"Are you sure? Please think about it for a moment,"* or *"Let's listen to the recordings again and compare our pronunciation with the standard one,"* or she would raise such questions as *"Have you noticed something wrong with her answer?"* etc.

The second finding is about the **students' performances in class.** Overall**,** the students struck the observer by their strong passion for the new teaching approach and also by their active participation in class. At the very beginning, especially in the first

class, the students were somewhat confused. For example, they did not know what to do in the relaxation phase, felt curious about the filtered sentences, had difficulties in *"*dancing to the tune*"* and were at a loss in the role-play game that they had never played before. In response to students' reaction towards the new approach, the teacher made a full explanation of the pedagogical experiment, especially the filtered materials. Then, she guided the students to move on to each step. Gradually, students familiarized themselves with all these steps and became adept in fulfilling the tasks. In the relaxation phase, at the request of the students, Baroque music was played as the background music most of the time. Most students chose to rest on the desk with eyes closed for 5-10 minutes to relax and prepare themselves for learning. After the mind-calming exercise, they were asked to listen to the filtered materials 15 times through and then hummed the tune in chorus. It was observed that some students were capable of remembering and humming the tune by listening to the materials less than 15 times.

Inspired by their teacher, most students created a variety of body movements to express their understandings of the intonation patterns. Some groups demonstrated their ingenuity, creativity and imagination, particularly in the last step - acting out the situational dialogue. They invented many parodies and brainstormed a large number of words and expressions. They also produced various props using materials available, such as textbooks, water bottles, and their stationery. There were also some groups who could not improvise a parody but merely acted out the original dialogue in their textbook.

The third finding is regarding the **language practicing environment.** According to the observer, this approach nurtured a classroom environment which was favorable for practicing speaking skills. Students were found to be enthusiastic about humming the intonation patterns with body movements, and reading aloud and repeating the original sentences. Interestingly, they had a strong desire to show off their English to their teacher and the researcher. For example, they would ask their teacher or the observer to listen to them and check their pronunciation of certain words and recitation of some sentences during the break time of the class. Besides, in the preparatory stage of the role-play activity, more often than not, they did their utmost to speak more English than Chinese although it was no easy job for them to do so. They seemed to enjoy the "feel" of speaking English.

These findings added immensely to the effectiveness of VTA.

5.5 Major characteristics of VTA

This particular verbotonal-based approach was designed in the present study primarily for improving speaking skills, and at the same time, enhancing phonological working memory (PWM) as a byproduct. The extent of the experimental group's increased performance in both speaking skills and PWM was unexpected and exciting. Moreover, of relevance and interest but extending beyond the scope of the research questions, is the fact that, in the final examination held by Gucheng Town for all 403 third-graders (10 classes) in this district, the experimental group came out ahead of all

participating groups. The group claimed first place with a mean score of 93.75 out of 100 against 77.45 for the control group and 75 for all Grade 3 students in the district. The final examination set by the district consisted of listening, reading, spelling and translating. As mentioned earlier, the control group received intentional and explicit teaching for all these skills tested whereas the experimental group did not. The better performance of the experimental group strongly demonstrates that the verbotonal approach was able to improve not only the speaking skills of young learners but their other language skills as well and that it did so more effectively than the other approaches used. No individual student regressed in the experimental group whereas some students regressed in the control group. It seems that, quite literally, everybody wins with VTA, but not everybody wins with the traditional approach. This result is not limited to the present experiment (with young children) but is mirrored in the experiment with adults performed by He Bi (2015). Besides, there was also a significant improvement of the phonological working memory capacity in the experimental group. A summary of the major characteristics of VTA are drawn as follows to attempt to account for the effectiveness of the approach.

Above all, the verbotonal approach developed in the present study based its theoretical foundation on the principle that language learning involves a process of awareness-raising based on students' individual meaning-making mechanisms. Each individual speaker is a perceiver, and at the same time, a producer of speech. And the auditory and visual performances in one's production mirrors how he/she perceives

speech. Individuals' perceptions of the sounds of the foreign language are influenced by their "phonological sieve" resulting from the phonological system of their native languages: a mechanism for keeping sounds which are recognized and rejecting those which are not (Lian & Lian, 1997; Zhang, 2012). So, this approach was designed to raise learners' awareness of the new language, defeat their phonological sieve, and change their perceptions of the foreign sounds and sound sequences by creating an optimal field of hearing for them. Once their perceptions changed, their speech would also be able to change. And this is what seems to have happened.

Second, this specific approach engaged students with brain-based learning and focused on the activation of both right and left hemispheres. It first used low-pass filtering to act directly and preferentially on the right hemisphere of the brain and thus bypassed normal language-processing mechanisms. Studies investigating functional anatomy of language show that right brain is associated with the perception of changes of the frequency spectrum and speech prosody (Klein & Zatorre, 2011; Garratt, 2015). When students were exposed to the low-frequency rhythm and intonation patterns, right superior temporal activation occurred (Sammler, 2015). Further, the rhythm and intonation patterns facilitated the brain's perception of the high frequencies. In other words, the prosody of speech provided a foundation for the brain to encode and perceive the individual sounds (vowels and consonants), although they were actually removed through filtering. Then, students were required to listen to unfiltered sentences. Extensive exposure to the unfiltered sentences acted on the left hemisphere which has

been shown to be involved in normal speech processing such as auditory word recognition (the left temporoparietal cortex), word generation (the dorsolateral prefrontal cortex), word retrieval and sentence generation (the left anterior frontal areas) (Price, 2012). Thus, both hemispheres resumed parallel processing.

Third, VTA supported the notion that the perception and production of speech was a whole-body and multi-sensory experience. There is a strong association between cognition and action (Condon & Ogston, 1966). Physical activities such as stretching, walking, and moving promotes thinking and enhances the learning process (Diamond, 2000). Condon and Ogston (1966) also posit that there exists synchronization between body movements and speech. According to them, the body motion of speakers is synchronous with their speech rhythms. Meanwhile, the body movements of the listener also spontaneously move in rhythm with those of the speaker (Condon & Ogston, 1966; Shockley et al., 2003; Richardson et al., 2008). So, VTA integrated body movements into the learning of speaking skills. These embodied representations of speech engendered the enactment effect which facilitated verbatim memory for language and made learning outcomes positive and productive (Macedonia, 2012). More importantly, in the experiment, the students were not required to copy anyone else's models but to generate their own patterns of movement. In so doing, their meaning-making mechanisms were, in effect, respected. Models were not imposed but were allowed to develop according to students' understandings and preferences. Besides, this approach saw the perception and production of speech was a multisensory process. One goal of

the pedagogical procedures was to stimulate and orchestrate simultaneously students' visual, auditory, vibrotactile, vestibular and proprioceptive senses, in other words, to integrate senses. There is substantial evidence suggesting that sensory integration is able to increase the neuroplasticity of the brain by developing new neural pathways and thus increase memory capacity and expand learning potential (Ryugo et al., 2000).

Fourth, VTA stressed the importance of students' sensitivity to the suprasegmental features such as stress, rhythm, loudness, pitch, and intonation, the prosodic patterns of the target language. Different languages have different prosodic patterns. These prosodic patterns organize the vowel and consonant sounds of a particular language in a way which is unique to that language (Lian, 1980). Studies show that prosody assists learners in parsing continuous speech and recognizing words, and provides cues to syntactic structure, grammatical boundaries and sentence type (Nooteboom, 1997; Ragó et al., 2014). Moreover, McNerney and Mendelsohn (1992) argued that suprasegmentals could aid intelligibility and make learners' speech more easily comprehensible. Further, Philippe Martin (2016) maintained that suprasegmental features were essential to generate and understand speech and that they were always present and must be recovered even in silent reading. Though learners' brains use both the suprasegmental and segmental features to perceive speech, developmentally speaking, learners acquire the suprasegmental elements first (Asp, 2006). So, emphasizing the role of suprasegmentals and its impact on the fine-tuning of the segmental tonalities is a special feature of VTA.

Fifth, the particular VTA used here seemed to be teacher-proof as it required no explicit teaching to be performed. The teacher only acted as facilitator and catalyst. And, in the same way that students were empowered to generate their own personal gestures to interpret rhythmic and phonetic structures, students were empowered here to self-adjust their learning and be free from any intervention and teaching from the teacher. What learners did is to follow a set of pre-planned activities. The activities, although pre-planned, were designed in such a way that students' individual differences and different meaning-making mechanisms were respected, thus effectively creating an individualized personal learning environment for every student. In other words, students were not required to conform to any pre-determined models but constructed their own based on their operational history. VTA did not require any teaching or an experienced teacher. All that it required was exposure to the materials in the order presented. In this context, language teacher training did not seem to be a prerequisite for using this approach. It also did not require the setting of class size limits as students generated their own individual personal learning environments and it also protected students from an uneven quality of teaching. In this sense, VTA addressed, at least partially, the problems existing in the teaching and learning of English speaking skills in Chinese primary schools: large class sizes, inadequate teacher preparation and materials, and inappropriate teaching approaches.

Taken together, all the above-mentioned factors and mechanisms underlying the verbotonal approach developed in the current research operated in synchrony to act on students' improvement both in language skills and PWM capacity.

5.6 Summary

This chapter began with discussion on the results of the assessment of speaking skills and phonological working memory (PWM). Then, the correlation between speaking skills and PWM was laid out. Additionally, students' and teacher's perceptions of VTA were interpreted. The whole chapter was closed by a brief reflection on the major characteristics of VTA.

CHAPTER 6

CONCLUSION

This chapter concludes the dissertation. It is made up of three sections. Section One refocuses the aims of the research and reveals a synopsis of the findings. Section Two proposes the implications of the present study. Section Three expounds the limitations of the study and suggests major promising directions of further research.

6.1 Summary of the study

The current research conducted a quasi-experiment wherein a verbotonal-based approach (VTA) was devised and implemented to enhance the English-speaking skills and phonological working memory of Chinese primary EFL learners. Eighty Grade 3 students with an average age of 9 years and enrolled in two intact classes were sampled as subjects from the Yiliang Anjiaqiao Primary School (YAPS). The two classes were randomly assigned as the experimental group that underwent the verbotonal program, and the control group who were taught with a normal teaching approach. Both the experimental and control groups were taken care of by the same teacher. As a reminder, the research questions of the study are reiterated as follows.

- Is VTA effective for learning to speak in English? If yes, in what ways?
- Is VTA effective for enhancing PWM? If yes, in what ways?

- Is there a relationship between students' PWM performance and speaking skills? If yes, in what ways?
- What are the students' opinions of VTA as a way of learning to speak English and enhancing PWM?
- What are the teacher's opinions of the implementation and outcomes of VTA as a way of learning to speak English and enhancing PWM?

To resolve all these research questions, a mixed methods design was adopted. The data obtained comprised both quantitative and qualitative data. The quasi-experiment ran for 18 weeks. Both pretests and posttests for speaking skills and PWM were administered to check the merits of VTA and compare it with the traditional teaching approach. Three Chinese experts were invited to rate students' speaking skills and two of them were also requested to assess PWM capacity. To triangulate the source of data, questionnaires and semi-structured interviews were conducted with the students in the experimental group to elicit their opinions of VTA. Meanwhile, a semi-structured interview was also carried out with the teacher to obtain her perceptions of the effectiveness of the new approach. In addition, observations were utilized to gain a deeper understanding of the research problem.

After collecting data, both descriptive and inferential statistics including a paired samples t-test, an independent-samples t-test and Pearson correlation coefficient were applied using SPSS (Version 19.0) to analyze quantitative data. The qualitative data were first transcribed verbatim. Then, they were subject to content analysis. In

order to make the analysis of data rigorous, a statistician from the Faculty of Environmental Engineering, Kunming University of Science and Technology was invited to suggest and validate the data analysis techniques and procedures. The major findings of this study are outlined in the following part.

First, VTA was of considerable value to the learning of not only speaking but also other language skills. Although starting at the same level, after 18 weeks of treatment, the experimental group advanced significantly in speaking English, while no significant improvement was detected in the control group. In addition, when the performances of the two groups of participants were compared after the treatment, it was found that the experimental group was far in advance not only in total score but also in every individual component and subskill tested. There is another surprising finding which is of relevance and interest but extending beyond the scope of the research questions. That is, in the final examination held by Gucheng Town, Yiliang County, Kunming City, Yunnan Province, China, for all 403 third-graders in the district, the experimental group topped all participating groups with a mean score of 93.75 out of 100 against 77.45 for the control group and 75 for all Grade 3 students in the district. The final examination was organized into four parts: listening, reading, spelling and translating. Again, it should be noted here that the control group was given specific teaching for segmental pronunciation, spelling, grammar and translation. In contrast, the experimental group never received any intentional teaching for any of these items. These findings thus confirm the effectiveness of the verbotonal approach in the learning

of both speaking and other language-related skills.

Second, VTA was able to enhance PWM. After completing the experiment, the overall phonological working memory capacity of the students in the experimental group improved significantly. The experimental group also showed substantial progress in repeating polysyllabic nonwords and nonwords with low wordlikeness. By contrast, no significant improvement in overall PWM capacity and subtests was found in the control group. Moreover, the experimental group outperformed the control group for overall PWM capacity and also for repetition of lengthy nonwords and nonwords with low wordlikeness.

Third, a statistically positive correlation between speaking skills and phonological working memory was discovered in the present study. Besides, there was also a strong association between phonological working memory and individual tests for speaking skills including word-reading, sentence-reading, singing, and the oral interview, and subskills involving vocabulary, grammar, pronunciation, fluency and comprehensibility.

Fourth, almost all students in the experimental group held a positive attitude towards VTA. They considered VTA as a very helpful, useful and interesting way of learning English speaking skills and training PWM. They also thought that VTA elevated their confidence and boosted their motivation in speaking English. According to them, VTA enabled them to remember the sound of words more quickly and accurately than ever before.

Fifth, the teacher thought VTA was of considerable benefit to the learning of speaking skills and training of phonological working memory. She reported that VTA freed her from teaching and alleviated the pressure of modelling pronunciation and speaking in class by engaging the students with self-directed learning. When students were entrusted to manage their own learning, their meaning-making mechanisms and idiosyncrasies in learning were respected. Meanwhile, students' intrinsic motivation for learning was sparked, which tended to result in optimal and effective learning outcomes (Grolnick & Ryan, 1987; Taylor, 2015). Additionally, the teacher believed that VTA was also able to improve students' ability to memorize words and sentences. So, she preferred VTA to traditional teaching programs and hoped that the new approach could be expanded to include more rural primary schools.

Last, findings from observations conducted by the researcher also supported the value of VTA in improving English learning in general and not just the learning of speaking skills. It was observed that VTA fostered an enjoyable, meaningful and community-oriented learning climate. Students were delighted and active to participate in the classroom activities. They were found to be more willing to speak English both inside and outside of the classroom.

6.2 Implications of findings

The research results bring potential enlightenment not only to EFL teaching and learning in Chinese primary schools but also to the training of PWM (phonological

working memory) or the whole WM (working memory) system. To be more detailed, the following implications can be drawn from the findings.

First, this study brings implications to EFL/ESL learning and teaching theory. There are two principles that are critical to the development of language-learning and teaching theories. One is understanding how language learning or learning in general takes place. And the other is how to translate this understanding into teaching practices. When learners learn a foreign/second language, they are more likely to be insensitive to the new language system because of the influence of their first language. In this sense, language learning or teaching is no more than a matter of defeating learners' insensitivity to and raising their awareness of the new language. The primary goal of VTA is to improve the speaking skills of Chinese primary EFL learners by raising their awareness of English speech. Unlike traditional approaches, the verbotonal approach places emphasis on learners' auditory perception and speech production of the rhythm and intonation patterns of English. That is to say, language learning with VTA starts with prosody. Also, VTA sees learning as a complex, idiosyncratic and nonlinear process of meaning-making. Learners' meaning-making mechanisms are influenced by their personal operational history, and thus unique and different. Students' individual meaning-making mechanisms are respected in VTA. In addition, VTA engages students with brain-based learning and focuses on the activation of both hemispheres of the brain. The positive outcomes achieved shed light on the understanding of language learning

and brain functions, which is of essential importance to the development of learning theories.

Second, this study has implications for the practice of language learning and teaching. Results show that VTA is able to improve not only the speaking but also other language skills. These promising findings have significant implications for both EFL teachers and learners. Both teachers and learners need to reconsider the nature of language learning, including the roles of individual meaning-making, awareness-raising and brain functions, and adopt teaching/learning approaches able to capitalize on these. Meanwhile, language learning should be conducted in a holistic, multi-modal and multi-sensory manner which provides multiple memory patterns rather than operate in a fragmented and isolated way. Besides, students, even young students, can and should be empowered to manage their own learning. That is, they need to be seen as whole, capable and resourceful in the learning process rather than as incompetent and helpless.

Third, implications exist for language-learning policies at all levels. The present study provides support to the evaluation and development of language-learning policies in the Chinese context. The ultimate aim for any government, institution or organization to develop a language-learning policy is to improve learners' language proficiency and make them competent users of the target language. The teaching and learning of English as a foreign language (EFL) in China has a long history. But the learning outcomes seem to be disappointing despite the huge investment in teaching infrastructures. As for

teaching EFL in the primary schools, there is considerable inconsistency between the goals set by the curriculum standards and reality. The language skills, especially the speaking skills of Chinese primary EFL learners are regarded as poor. Thus, the specific VTA was designed for developing intelligible speech by using a range of pre-planned activities. Results indicate that the approach is also successful in improving other language-related skills such as listening, reading, spelling and translating. The theories and practices of the approach have significant implications for the designing and delivering of EFL programs at different levels.

Fourth, the findings of the research give implications for teacher education programs, which may require a revision of programs' prior assumptions. During the experimentation, the control group receives intentional and explicit teaching for vocabulary, grammar, spelling and translation skills whereas the experimental group does not. VTA encourages students to assume responsibility for and take charge of their own learning. The teacher oversaw the learning process and provided help when requested by the students. That is, VTA creates a learning environment that is devoid of teachers deciding how to teach it. In other words, this approach is teacher-proof as it requires no direct instruction from the teacher. In contrast, it engages students with pre-designed sensitization and reinforcement activities. So, in this sense, it does not place high demands on teachers' expertise. This is of significant importance to the teaching and learning of English in the rural primary schools in China where there is an enormous shortage of teachers who have the required English language proficiency for

teaching. The counter-intuitive findings of the study cause a rethink of the teacher training programs.

Fifth, implications are present for enhancing PWM or WM in general. VTA is primarily designed to improve students' speaking skills. However, the basic principles and practices of this approach have shown to be useful in enhancing PWM. Thus, it was also employed in the study as a way of increasing students' PWM. Results show that participants who underwent VTA had significant gains in PWM. VTA holds that learners, especially young learners process language through feeling, hearing, touching and seeing (Asp, 2006). It implements multisensory learning of speaking skills by engaging learners' visual, auditory, vestibular, tactile, and proprioceptive senses simultaneously. As mentioned earlier, this sensory integration is able to boost neuroplasticity of learners' brains by causing developmental changes in synaptic density and synaptic pruning and developing new neural pathways (Ryugo et al., 2000). Also, in VTA, young learners are repeatedly exposed to both filtered and unfiltered materials. Filtered material acts primarily on the right brain. Unfiltered material activates left brain areas which are associated with normal language processing. The concurrent activation of both hemispheres boosts neural firing thereby contributing to neuroplastic change of the brain (May et al., 2007). This learning-induced plasticity leads to improvements in performance in the non-trained task that relies on PWM, namely, the nonword repetition task. These findings provide valuable insights for the understanding and training of PWM or WM in general.

6.3 Strengths and limitations of the study

This study conducted a quasi-experiment to investigate the effectiveness of VTA as a way of learning English speaking skills and training PWM. The pedagogical procedures were designed on the basis of previous work carried out by Andrew Lian (1980), with modifications being made to meet the specific and local needs. Prior to performance of the main study, a pilot study was performed to test the feasibility and improve upon the research design. Besides, multiple methods including pretest, posttest, questionnaires, interviews and observations were employed to collect data. This mode of data triangulation boosted credibility of the findings by cross-verifying the same information. Meanwhile, all the instruments used for gathering data were tried out and validated. It is important to be noted here again that a statistician was invited to validate and suggest the data analysis techniques and procedures. These add weight to the reliability and validity of the research findings.

Despite the above-mentioned strengths, however, it is necessary to examine some of the limitations of the study. First, according to Asp (2006), the rhythm and intonation patterns and accent are better acquired by 4 to 12 years old children than by adults because children's brain is more neural plastic than that of adults. So the participants of the study were drawn from third-graders aged between 8 to 9 years. They had learned English for only one semester. In other words, they were beginner-level EFL students. So the generalization of the findings of the study to intermediate- and advanced-level EFL learners should be considered carefully.

Second, as explained before, YAPS is a rural primary school located in Anjiaqiao Village, Gucheng Town, Yiliang County, Kunming City, Yunnan Province, China. Since there are regional disparities in the English learning backgrounds and conditions of primary students from different parts of China, the ability to generalize and apply the findings of the present study to rural primary students in other places or regions or to urban primary school students needs to be treated with caution.

Third, the students did not have access to the Internet in the in-class sensitization stage. This represented a huge limitation. Again, VTA freed the students from the teacher's teaching and intervention, and empowered them to resolve the problems emerging in learning. We surmise that more thought-provoking and promising findings could be obtained if the networked information resources were accessible to the students to provide learner support.

6.4 Recommendations for further study

Nevertheless, within the limits set by this study and other related studies, there is good cause for optimism. Suggestions for future research to verify and improve the verbotonal approach are provided as follows.

First, as the participants of this study were only limited to Grade 3 primary school students who had just learned English for one semester and belonged to beginner-level EFL learners, it is strongly suggested that more empirical studies be carried out with students from other levels (i.e. intermediate- or advanced-level EFL

learners) to fully recognize the potential value of the verbotonal system in improving EFL teaching and learning.

Second, the participants in this study were sampled from a rural primary school in Yunnan Province, China. Future replication studies could be conducted in the primary schools of urban areas. Also, there could be a contrastive study of the performances of students in rural and urban areas to gain more insight into VTA as a manner of learning to speak English in Chinese primary schools.

Third, as stated previously, the verbotonal approach was initially developed as a language therapy to treat hard-of-hearing/hearing-impaired children or adults. Also, this approach has been implemented by most verbotonalists worldwide to treat patients with speech or communication problems. However, in the Chinese context, there is no implementation of VTA as a speech and language therapy. It is hoped that future studies could bridge the gap and apply the approach to special education.

Fourth, the new era boasts an unprecedented speed of change and explosion of information. The traditional view of learning wherein a specified and predesigned body of knowledge was taught just in case it was needed some time seems to be out of place (Lian, 2014). The phenomena of life cannot be exhaustively accounted for by a pool of stable knowledge. In other words, with an ever-more rapid speed of development, the new era requires a new form of learning which is flexible and enables learners to decide what to learn so as to cope with unpredicted and unpredictable real life problems occasioned by changing circumstances. However, to conform to the regulations

specified by the government, participants in this study were not totally free to decide what to learn although they could use other learning materials out of class. It is therefore recommended that further research focus on a learning environment where students can decide and access individualized instructional material.

Fifth, this specific VTA fostered a holistic and multi-modal way of learning a foreign language. Specifically, it utilized filtered and unfiltered materials and incorporated body-movements into learning. Importantly, it emphasized sensory integration and allowed parallel processing of both hemispheres of the brain. Results showed that VTA was able to improve young learners' language skills and phonological working memory (PWM) capacity. These findings raise important issues concerning brain areas activated when learning with VTA. Consequently, there could be studies to fully investigate the functional imaging of learners' brain when they are listening to filtered and unfiltered materials or engaging in dichotic listening using EEG or MRI.

Finally, the study used nonword repetition tasks to test students' PWM because it is an effective measure of PWM, and more importantly, it is particularly appropriate for use with young learners (Gathercole et al., 1994). Also, the experimental evidence substantiates the significant positive correlation between nonword repetition test and other PWM measures such as auditory digit span and immediate serial recall. However, new insights might be gained if other measures were used to assess PWM.

The present study serves as a point of departure for exploring a field of investigation that can add considerably to the improvement of learning of a general

nature and not just foreign/second language learning in this most intellectually and socially challenging century. The above-mentioned unresolved questions arising out of the study unearthed new directions for future research, and also conclude this dissertation.

REFERENCES

Abikoff, H., & Gittelman, R. (1985). Hyperactive children treated with stimulants: Is cognitive training a useful adjunct? *Archives of general Psychiatry*, *42*(10), 953-961.

ACTFL Proficiency Guidelines. (1985). New York: American Council on the Teaching of Foreign Languages.

ACTFL Proficiency Guidelines. (1999). New York: American Council on the Teaching of Foreign Languages.

Adams, A. M. & Gathercole, S. E. (2000). Limitations in working memory: implications for language development. *International Journal of Language & Communication Disorders, 35* (1), 95–116.

Adams, L. J. (2013). *Movement and memory* (Unpublished PhD thesis).

Adams, Marianne L. (1980). Five co-occurring factors in speaking proficiency. In James R. Firth (Eds.), *Measuring spoken language proficiency* (pp. 1–6). Washington, DC: Georgetown University Press.

Adamson, B. (2001). English with Chinese characteristics: China's new curriculum. *Asia Pacific Journal of Education*, *21*(2), 19-33.

Adamson, B. and P Morris (2000). Changing Hong Kong's schools. In B. Adamson, T. Kwan and K. K. Chan (Eds). *Changing the curriculum: The impact of reform on primary schooling in Hong Kong* (pp. 71-91). Hong Kong: Hong Kong University Press.

Alazard, C., Astésano, C., and Billières. M. (2010). “The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills”. Proceedings of the 5th Speech Prosody 2010 (Chicago, USA).

Alazard, C., Astesano, C., BilliÈres, M., Alazard, C., Astesano, C., & BilliÈres, M. (2011). Multiphonia: A multimodal database of phonetics teaching methods in classroom Inter Actions. Retrieved from http://www.lreccnf.org /proceedings/lrec2012/pdf/254_Paper.pdf.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of experimental child psychology*, 87(2), 85-106.

Alonso, R. A. (2014). Teaching speaking: an exploratory study in two academic contexts. *Porta Linguarum*, *145*, 160-165.

Anderson, O.R. (1992). Some interrelationships between constructivist models of learning and current neurobiological theory, with implications for science education. *Journal of Research in Science Teaching*, *29* (10), 1037-1058.

Anggraeni, P. (2007). *Audio lingual teaching as an alternative method in teaching speaking*. Semarang: Semarang State University.

Ariza, E. N. (2002). *Fundamentals of teaching English to speakers of other languages in k-12 mainstream classrooms*. Dubuque: Kendall Hunt Publishing Co.

Arnold. Higgs, T. V. (1987). Oral proficiency testing and its significance for practice. *Theory into Practice*, *26* (4), 282-287.

Asher, J. J. (1969). The total physical response approach to second language learning. *The modern language journal*, *53*(1), 3-17.

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

Asp, C. W., & Kline (2012), M. Timeless verbotonal corrective body movements. Retrieved from http://verbotonal.utk.edu/welcome.html.

Asp, C. W., Kline, M. W. & Koike, K. J. (2012b). Verbotonal worldwide. Retrieved from http://verbotonal.utk.edu/welcome.html.

Asp, C. W., Kline, M. W. & Koike, K. J. (2012c). Verbotonal body movements. Retrieved from http://verbotonal.utk.edu/welcome.html.

Asp, C. W., Koike, K. J., & Kline, M. W. (2012a). Breaking news: Verbotonal rehabilitation: Are we doing enough? *The Hearing Journal*, *65*(01), 28-30.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The psychology of learning and motivation*, *2*, 89-195.

Avery, P., & Ehrlich, S. (1992). Common pronunciation problems. *Teaching American English Pronunciation*, 95-109.

Avons, S. E., Wragg, C. A., Cupples, W. L., & Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics*, *19* (4), 583-601.

Azar, B. (2007). Grammar-Based Teaching: A Practitioner's Perspective. *Tesl-ej*, 11(2), n2.

Azar, B.S. (2014). A conversation with Betty Azar and her new co-author Stacey Hagan. Retrieved from http://azargrammar.com/assets/authorsCorner/

Azzam, A. M. (2009). Why creativity now? A conversation with Sir Ken Robinson. *Educational Leadership*, *67*(1), 22-26.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests (Vol. 1)*. Oxford: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of communication disorders*, *36*(3), 189-208.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience*, *4*(10), 829-839.

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

Baddeley, A. D. (1997). *Human memory: Theory and practice*. Hove: Psychology Press.

Baddeley, A. D. (2013). Levels of working memory. In Naveh-Benjamin, M., Moscovitch, M., & Roediger III, H. L. (Eds), *Perspectives on Human Memory and Cognitive Aging: Essays in Honor of Fergus Craik* (pp.111). Hove: Psychology Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, *8*, 47-89.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *The Quarterly Journal of Experimental Psychology*, *60*(4), 497-504.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior*, 14(6), 575-589.

Baddeley, A., & Wilson, B. A. (1993). A developmental deficit in short-term phonological memory: Implications for language and reading. *Memory*, *1*(1), 65-78.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, *105*(1), 158.

Baddeley, A., Papagno, C., & Norris, D. (2014). Phonological memory. In Hockley, W. E., & Lewandowsky, S. (Eds), *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock* (pp.175). Hove: Psychology Press

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of memory and language*, *27*(5), 586-595.

Baddeley, A.D. (2012). Working memory: theories, models, and controversies. *Annual review of psychology*, *63*, 1-29.

Bagnall, R. G. (1999). *Discovering Radical Contingency: Building a Postmodern Agenda in Adult Education. Counterpoints: Studies in the Postmodern Theory of Education, (Vol. 81)*. New York: Peter Lang Publishing.

Bai, R. W., Zhou, R., & Zhang, W. L. (2013). The strategies for the implementation of E-schoolbag in the teaching of primary English. *Modern Education Technology, 23*(8), 50-54.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(04), 723-742.

Bancroft, W. J. (2005). *Suggestopedia and Language*. London: Routledge.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, 79-103.

Barrows, H.S. (1986). A taxonomy of problem based learning methods. *Medical Education 20,* 481—86.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and brain sciences*, *22*(4), 637-660.

Barton, J., Collins, A. (1997). *Portfolio assessment: A handbook for education.* California: Addison-Wesley Publishing Company.

Batool, N., Anosh, M., Batool, A., & Iqbal, N. (2015). The direct method: A good start to teach oral language. *Journal of Literature, Languages and Linguistics*, *5*, 53-55.

Bauman, Z. (1987) *Legislators and Interpreters,* Cambridge: Polity Press.

Bauman, Z. (1992). *Intimations of postmodernity*. London: Routledge.

Berns, A., Gonzalez-Pardo, A., & Camacho, D. (2013). Game-like language learning in 3-D virtual environments. *Computers & Education, 60*(1), 210-220.

Best, S., & Kellner, D. (1991). *Postmodern Theory: Critical Interrogations.* New York: Guilford.

Bird, A. A. (2007). *The role of music in the English language development of Latino prekindergarten English learners*. (Unpublished EdD thesis).

Blumer, H. (1969). *Symbolic interactionism*. New Jersey: Prentice-Hall.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*(3), 441-469.

Boyne, R. and Rattansi, A. (1990). *Postmodernism and society,* London: Macmillan.

Brazil, D., Coulthard, M., Johns, C., & Johns, C. (1980). *Discourse intonation and language teaching* (pp. 1-82). London: Longman.

Briton, D. (1996). *The modern practice of adult education: A postmodern critique.* New York: State University of New York Press.

Brooks, R. L. (2013). Comparing native and non-native raters of US Federal Government speaking tests. (Unpublished PhD thesis).

Brown, A. (2000). An investigation of the rating process in the IELTS oral interview. *International English Language Testing System (IELTS) Research Reports 2000*: Volume 3, 49.

Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language testing*, 20(1), 1-25.

Brown, G., & Yule, G. (1983). *Teaching the spoken language (Vol. 2)*. Cambridge University Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10* (1), 12-21.

Brownlee, P. P. (1980). Suggestopedia and its application to the education of children with learning disabilities. *Journal of the Society for Accelerative Learning and Teaching*, *6*(2), 137-144.

Burgess, J., & Etherington, S. (2002). Focus on grammatical form: explicit or implicit?. *System*, *30*(4), 433-458.

Burgess, R. G. (1984). Methods of field research 2: Interviews as conversations. *In the Field: An Introduction to Field Research*, 101-122.

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience*, *2*, 167-179.

Butterfield, E. C., Wambold, C., & Belmont, J. M. (1973). On the theory and practice of improving short-term memory. *American journal of mental deficiency*, *77* (5), 654-69.

Bygate, M. (2009). 15. Teaching the spoken foreign language. *Handbook of foreign language communication and learning*, *6*, 401-409.

Caddick, A. (1986). Feminism and the body. *Arena*, *74*, 60-88.

Campbell, R., & Wales, R. (1970). The study of language acquisition. *New horizons in linguistics*, 242-260.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics*, *1*(1), 1-47.

Carroll, J. B. (1964). The Prediction of Success in Intensive Foreign Language Training. Pennsylvania: University of Pittsburgh.

Carter, R. (2012). *Vocabulary: Applied linguistic perspectives.* London: Routledge.

Chambers, F. (1997). What do we mean by fluency?. *System*, *25*(4), 535-544.

Chang, S. C. (2011). A contrastive study of grammar translation method and communicative approach in teaching English grammar. *English Language Teaching*, *4*(2), 13.

Chen, L. (2008). Primary school English education in China. Paper presented at the International Conference on Primary English, Beijing. Retrieved from http://www.chinabfl e.org/ylp/newsdetails.asp?icntno=2317.

Christison, M., Christian, D., Duff, P. A., & Spada, N. (Eds.) (2015). *Teaching and learning English grammar: Research findings and future directions*. London: Routledge.

Clark, J. L. (1972). *Foreign language testing: Theory and practice (Vol. 15)*. Boston: Heinle & Heinle Publishers.

Clark, J. L. (1975). Theoretical and technical considerations in oral proficiency testing. In Jones, R. & Spolsky, B. (Eds), *Testing Language Proficiency* (pp. 10-28). Arlington, VA : Center for Applied Linguistics.

Coady, J., & Huckin, T. (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.

Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *The Journal of Nervous and Mental Disease*, *143*(4), 338-347.

Condon, W. S., & Ogston, W. D. (1971). Speech and body movement synchrony of the speaker-hearer. *Perception of language*, 98-120.

Connor, S. (1989). *Postmodernist culture: An introduction to theories of the contemporary*. Amsterdam: John Benjamins Publishing Company.

Cortazzi, M., & Jin, L. (2003). English Language Teaching in China: A Bridge to the Future. *Asia-Pacific Journal of Education*, *22*(2), 53–64.

Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment.* Cambridge: Cambridge University Press.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Crook, S., Pakulski, J., & Waters, M. (1992). *Postmodernization change in advanced society.* London: Sage Publications.

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

Cruttenden, A. (2014). *Gimson's pronunciation of English*. London: Routledge.

Crystal, D. (1987). *The encyclopedia of language*. Cambridge: CUP.

Crystal, D. (2011). *A dictionary of linguistics and phonetics* (5th ed.). Malden, MA: Blackwell.

Crystal, D. (2011). *Dictionary of linguistics and phonetics* (Vol. 30). New Jersey: John Wiley & Sons.

Cubillos, J., Flynn, D., Vincent, F. M., Azeem, M., Ashfaq, M., Van Lengen, K., & Faizi, M. (2010). Computer-mediated oral proficiency assessments: Validity, reliability and washback. *International Journal of Technology, Knowledge and Society, 6*(6) 34-45.

Curtis, C. E., & D'sposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in cognitive sciences, 7*(9), 415-423.

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320* (5882), 1510-1512.

De Jong, J. H., & Van Ginkel, L. W. (1992). 15 Dimensions in oral foreign language proficiency. Verhoeven, L., & de Jong, J. H. (eds.), *The construct of language proficiency: Applications of psychological models to language assessment*, 187.

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition, 34*(01), 5-34.

De La Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary. *Studies in second language acquisition*, 24(01), 81-112.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition*, *19*(01), 1-16.

Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review*, *66*(2), 181-202.

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.

Dey, I. (2003). *Qualitative data analysis: A user friendly guide for social scientists*. London and New York: Routledge.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child development*, *71*(1), 44-56.

Doll, W. (1993). *A post-modern perspective on curriculum*. New York: Teachers

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136-1146.

Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech, Language, and Hearing Research*, *36*(5), 1051-1054.

Dong, Y. (2003). Are we ready for an early start in foreign language learning-A survey of primary school English education in Guangdong Province. *Modern Foreign Languages, 26* (1), 39–47.

Douglas, D. (1994). Quantity and quality in speaking test performance. *Language Testing, 11*(2), 125-144.

Du, X. (2010). Primary school English teaching in Guangzhou: Current situation and some analyses. *Education Journal, 1*, 36–39.

Dufva, M., & Voeten, M. J. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied psycholinguistics, 20*(03), 329-348.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*(2), 227-252.

Edin, F., Macoveanu, J., Olesen, P., Tegnér, J., & Klingberg, T. (2007). Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *Journal of cognitive neuroscience, 19* (5), 750-760.

Efrizal, D. (2012). Improving students' speaking through communicative language teaching method at Mts Ja-alhaq, Sentot Ali Basa Islamic Boarding School of Bengkulu, Indonesia. *International Journal of Humanities and Social Science, 2*(20), 127-134.

Efthymiou, G. (2012). Portfolio Assessment of Speaking Skills in English as a Foreign Language in Primary Education. *Research Papers in Language Teaching and Learning*, *3*(1), 200-205.

Elliot, C., Murray, D., & Pearson, L. (1983). *British abilities scales. National Foundation for Educational Research.* Nelson: Windsor, UK.

Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, *46*(3), 533-558.

Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). New York: Springer.

Ericcson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*(4448), 1181-1182.

Evans, S. (1996). The context of English language education: The case of Hong Kong. *Relc Journal*, 30-55.

Fabbro, F., Muzur, A., Bellen, R., Calacione, R., & Bava, A. (1999). Effects of praying and a working memory task in participants trained in meditation and controls on the occurrence of spontaneous thoughts. *Perceptual and motor skills*, *88*(3), 765-770.

Fang, W., Shao, H., & Tang, L. (2004). An investigation on qualifications of primary school English teachers. *Basic English teaching and research, 10*, 19-22.

Farnia, F., & Geva, E. (2011). Cognitive correlates of vocabulary growth in English language learners. *Applied Psycholinguistics*, *32*(4), 711-719.

Featherstone, M. (1991). *Postmodernism and consumer culture*. London: Sage Publications.

Felix, U. (1993). The contribution of background music to the enhancement of learning in suggestopedia: A critical review of the literature. *Journal of the Society for Accelerative Learning and Teaching*, *18*(3-4), 277-303.

Ferrara, S. (2008). Design and psychometric considerations for assessments of speaking proficiency: The English Language Development Assessment (ELDA) as illustration. *Educational Assessment, 13*(2), 132-169.

Foster, H. (1985). *Postmodern Culture.* London: Pluto Press.

Fotos, S. (2005). Traditional and grammar translation methods for second language teaching. *Handbook of research in second language teaching and learning*, *1*, 653-670.

Foucault, M. (1979). *Discipline and Punish: The Birth of the Prison.* Harmondsworth: Penguin Books

Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings*. Brighton: Harvester Press.

Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press.

French, L. M., & O'brien, I. R. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*(3), 463-487.

Fu, H. (2006). An investigation on English curriculum in some primary schools in Zhejiang province. *Basic English teaching and research, 3*, 34-37.

Fulcher, G. (2003). *Testing second language speaking*. London: Pearson Education.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. London and New York: Routledge.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of neurophysiology*, *61*(2), 331-349.

Galaczi, E. (2005). Upper Main Suite Speaking Assessment: Towards an Understanding of Assessment Criteria and Oral Examiner Behaviour. *Research Notes. 20,* 16-19.

Galloway, V. (1987). From defining to developing proficiency. A new look at the decision. H. Byrnes (eds): *Defining and developing proficiency: Guidelines, implementations, and concepts*, 25–73.

Gathercole S E, Baddeley A D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of memory and language, 28*(2): 200-213.

Gathercole, S. E. (1990). *The structure and functioning of phonological short-term memory*. England: University of Bristol.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*(1), 83-94.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(04), 513-543.

Gathercole, S. E. and Adams, A. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language*, *33*, 672-688.

Gathercole, S. E., & Baddeley, A. D. (1990a). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of memory and language*, *29*(3), 336-360.

Gathercole, S. E., & Baddeley, A. D. (1990b). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*(4), 439-454.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition?. *European Journal of Psychology of Education*, *8*(3), 259-272.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition?. *European Journal of Psychology of Education*, *8*(3), 259-272.

Gathercole, S. E., & Baddeley, A. D. (2014). *Working memory and language*. Hove: Psychology Press.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84.

Gathercole, S. E., Hitch, G. J., Service, E. and Martin, A. J. (1997). Short-term memory and new word learning in children. *Developmental Psychology, 33*, 966-979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied cognitive psychology*, *13*(1), 65-77.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103-127.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*, *28*(5), 887.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology*, *82*(3), 387-406.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*(03), 349-367.

George, E. M., & Coch, D. (2011). Music training and working memory: an ERP study. *Neuropsychologia*, *49*(5), 1083-1094.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature communications*, *4*, 1490.

Gold, R. L. (1958). Roles in sociological field observation. *Social Forces*, 36, 217–223.

Gonzalez- Gomez, N., & Nazzi, T. (2012). Phonotactic acquisition in healthy preterm infants. *Developmental science*, *15*(6), 885-894.

Good, A. J., Russo, F. A., & Sullivan, J. (2015). The efficacy of singing in foreign-language learning. *Psychology of Music*, *43*(5), 627-640.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science*, *337*(6102), 1623-1627.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science*, *337*(6102), 1623-1627.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove: Lawrence Erlbaum.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences*, *99*(16), 10911-10916.

Gray, J. (2002). *An investigation into the oral English language proficiency gain of pupils taught by native English-speaking teachers in Hong Kong secondary schools* (Unpublished PhD thesis).

Griffin, D. R. (1997). *Parapsychology, philosophy, and spirituality: A postmodern exploration*. New York: Suny Press.

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: an experimental and individual difference investigation. *Journal of personality and social psychology*, *52*(5), 890.

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

Habermas, J. (1970). Towards a theory of communicative competence. *Inquiry, 13*(1-4), 360-375.

Hall, H. M., & Austin, T. (2004). *Content-based second language teaching and learning: An interactive approach*. Boston, MA: Pearson Education Inc.

Halliday, M. A. K. (1981). *A course in spoken English: Intonation*. Oxford: Oxford University Press.

Hand, J. D. (1984). Split brain theory and recent results in brain research: Implications for the design of instruction. In R. K. Bass & C. R. Dills (Eds.), *Instructional development: The state of the art, II* (pp. 23-27). Dubuque: Kendall/Hun.

Hang, J. (2012). Implementation of the Verbotonal method in teaching Japanese pronunciation. *Journal of Changchun Education Institute*, 28(11), 124-125.

Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. *The development of second language proficiency*, 7-25.

Harmer, J. (1998). *How to teach English: an introduction to the practice of English language teaching*. London: Longman.

Hart, L. (1975). *How the Brain Works.* New York: Basic Books.

Harvey, D. (1991). *The Condition of Postmodernity,* Oxford: Basil Blackwell.

Hawthorne, K., & Gerken, L. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, *133*(2), 420-428.

He, B. (2015). *Improving the English pronunciation of Chinese EFL learners through the integration of CALL and Verbotonalism.* (Unpublished PhD thesis)

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory.* Hove: Psychology Press.

Hempel, A., Giesel, F. L., Garcia Caraballo, N. M., Amann, M., Meyer, H., Wüstenberg, T., & Schröder, J. (2004). Plasticity of cortical activation related to working memory during training. *American Journal of Psychiatry, 161*(4), 745-747.

Henner-Stanchina, C., & Riley, P. (1978). Aspects of autonomous learning. *ELT Documents*, *103*, 75-97.

Hiemstra, R. (1994). Self-directed learning. *The sourcebook for self-directed learning*, 9-20.

Higgs, T. V., & Clifford, R. (1982). The push toward communication. In Higgs, T.V. (Eds), *Curriculum, Competence, and the foreign Language Teacher* (pp. 57-79), Chicago: National Textbook Company.

Ho, J., & Crookall, D. (1995). Breaking with Chinese cultural traditions: Learner autonomy in English language teaching. *System*, *23*(2), 235-243.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental science*, *12*(4), F9-F15.

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary

acquisition: A multimodal imaging study. *The Journal of Neuroscience*, *33*(34), 13663-13672.

Hu, G. (2002a). Recent important developments in secondary English-language teaching in the People's Republic of China. *Language Culture and Curriculum*, *15*(1), 30-49.

Hu, G. (2002b). Recent important developments in secondary English-language teaching in the People's Republic of China. *Language Culture and Curriculum*, *15*(1), 30-49.

Hu, G. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development*, *24*(4), 290-318.

Hu, G. (2005). English language education in China: Policies, progress, and problems. *Language Policy*, *4*(1), 5-24.

Hu, G. W. (2004). Building a strong contingent of secondary EFL teachers in China: Problems and policies. *International Journal of Educational Reform*, *13*(4), 20-23.

Hu, R. (2013). Task-Based Language Teaching: Responses from Chinese Teachers of English. *TESL-EJ*, 16(4), n4.

Hu, X. (2009). Investigating qualifications of English teachers in primary schools in Heilongjiang province. *Chinese Adult Education*, *18*, 89–90.

Hu, Y. (2007). China's foreign language policy on primary English education: What's behind it? *Language Policy, 6,* 359–376.

Hu, Y., & Uno, S. (2005). Effectiveness of a new teaching method based on the verbo-tonal method for Japanese beginners' learning of chinese voice tones. *Japanese Journal of Educational Psychology*, 53(4), 541-550.

Hu, Y.H., & Nuo, S.B. (2005). Effectiveness of a new teaching method based on the Verbo-Tonal method for Japanese beginners' learning of Chinese voice tones. *The Japanese Journal of Educational Psychology, 53*(4), 541-550.

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of memory and language*, *30*(6), 685-701.

Hymes, D. (1972). On communicative competence. *Sociolinguistics*, *4,* 269-293.

Illeris, K. (2003). Towards a contemporary and comprehensive theory of learning. *International journal of lifelong education*, *22*(4), 396-406.

Illeris, K. (2007). *How we learn: Learning and non-learning in school and beyond*. London: Routledge.

Illeris, K. (2009). *Contemporary theories of learning: learning theorists...in their own words*. London: Routledge.

Ingram, D. E., & Wylie, E. (1993). Assessing speaking proficiency in the international English language testing system. *A New Decade of Language Testing Research. TESOL, Alexandria, VA,* 220-234.

Isidor, C. H. (1976). An introduction to sampling. In Selltiz, C., Wrightsman, L.S. & Cook, S.W. (Eds.), *Research Methods in Social Relations* (pp. 21- 29). New York: Holt.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied linguistics, 29*(1), 24-49.

Jarvis, P. (2009). Learning to be a person in society: Learning to be me. In Knud Illeris (Eds.), *Contemporary Theories of Learning: Learning Theorists -- In Their Own Words* (pp. 21-34). London: Routledge.

Jean, G., & Simard, D. (2013). Deductive versus inductive grammar instruction: Investigating possible relationships between gains, preferences and learning styles. *System, 41*(4), 1023-1042.

Jones, J. D. (2006). *The effects of music training and selective attention on working memory during bimodal processing of auditory and visual stimuli.* (Unpublished PhD thesis)

Jones, R. H. (1997). Beyond "listen and repeat": Pronunciation teaching materials and theories of second language acquisition. *System, 25*(1), 103-112.

Juntunen, M. L., & Hyvönen, L. (2004). Embodiment in musical knowing: how body movement facilitates learning within Dalcroze Eurhythmics. *British Journal of Music Education, 21*(2), 199-214.

Kang, D. J. (2007). Rhizoactivity: Toward a postmodern theory of lifelong learning. *Adult Education Quarterly*, *57*(3), 205-220.

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

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

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental science*, *12*(6), 978-990.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica*, *124*(3), 382-397.

Kilgore, D. W. (2001). Critical and postmodern perspectives on adult learning. *New directions for adult and continuing education*, *2001*(89), 53-62.

Kim, Y. H. (2009). A G-theory analysis of rater effect in ESL speaking assessment. *Applied linguistics*, *30*(3), 435-440.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354-364.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in cognitive sciences*, *14*(7), 317-324.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of cognitive neuroscience*, *14*(1), 1-10.

Klingberg, T., Kawashima, R., & Roland, P. E. (1996). Activation of Multi- modal Cortical Areas Underlies Short-term Memory. *European Journal of Neuroscience*, *8*(9), 1965-1971.

Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, *16*(02), 155-172.

Koike, Y. (2014). Explicit Pronunciation Instruction: Teaching Suprasegmentals to Japanese Learners of English. In N. Sonda, & A. Krause (Eds.), *JALT 2013 Conference Proceedings* (pp. 361-374).

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. New Jersey: Financial Times Press.

Kormos, J. (2014). *Speech production and second language acquisition*. London: Routledge.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145-164.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, *11*(02), 261-271.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and cognition*, *11*(02), 261-271.

Krashen, S. (1982). Principles and practice in second language acquisition. *Modern Language Journal*, 67(2).

Kvale, S. (1996). *Interviews. An introduction to qualitative research writing*. London: Sage Publications.

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. London: Sage Publications.

Laban, R. (1974). *The Language of Movement: A Guidebook to Choreutics*. Boston, MA: Play.

Lachman, M. E., Neupert, S. D., Bertrand, R., & Jette, A. M. (2006). The effects of strength training on memory in older adults. *Journal of aging and physical activity*, *14*(1), 59.

Land, G. & Jarman, B. (1992). *Breakpoint and beyond: Mastering the future today*. New York: HarperCollins.

Langer, N., von Bastian, C. C., Wirz, H., Oberauer, K., & Jäncke, L. (2013). The effects of working memory training on functional brain network efficiency. *Cortex*, *49*(9), 2424-2438.

Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech?. *Journal of Memory and Language*, *66*(1), 285-306.

Larsen-Freeman, D., & Anderson, M. (1986). *Techniques and Principles in Language Teaching.* Oxford : Oxford University Press.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics*, *27*(4), 590-619.

Lash, S. (1990) *Sociology of Postmodernism*, London: Routledge.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *Eurosla yearbook*, *5*(1), 223-250.

Laufer, B. (2009). Second language vocabulary acquisition from language input and from form-focused activities. *Language Teaching*, *42*(03), 341-354.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, *16*(3), 307-322.

Lave, J. & Wenger, E. (1991). *Situated learning.* New York: Cambridge University Press.

Leemann, A., Kolly, M. J., & Dellwo, V. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic science international*, 238, 59-67.

Lehiste, I. (1976). Suprasegmental features of speech. *Contemporary issues in experimental phonetics*, *3*, 236-239.

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

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

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

Levis, J. M., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *Tesol Journal*, *12*(2), 13-19.

Li, J. P. (2001). MOE briefing on primary ELT in China. *Chinese Education Journal*, p. 3.

Li, M. (2007). Foreign language education in primary schools in People's Republic of China. *Current Issues in Language Planning*, *8*, 148–161.

Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychology and aging*, *23*(4), 731.

Li, S. S. (2016). A study on the approaches to teach English speaking skills in Chinese primary schools. *Examinations Weekly*, (16), 96.

Lian, A. P. (1980). *Intonation patterns of French*. Sydney: River Seine.

Lian, A. P. (2000). From first principles: Constructing language learning and teaching environments. Keynote address to the 9th International Symposium and Bookfair, English Teachers' Association of the Republic of China, Taipei. In Selected Papers from the Ninth International Symposium on English Teaching, 49-62. Taipei: Crane Publishing. Retrieved from http://www.andrewlian.com/andrewlian/prowww/first_principles/index.html.

Lian, A. P. (2004). Technology-enhanced language-learning environments: A rhizomatic approach. *Computer-assisted language learning: Concepts, contexts and practices*, 1-20.

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

Lian, A., & 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. (2011). Reflections on language-learning in the 21st century: The rhizome at work. *Rangsit Journal of Arts and Sciences*, 1(1), pp. 5-17.

Lian, A-P. (2001). Imagination in language teaching and learning. Keynote address to the English Language Teaching and Knowledge Transformation Conference, Chaoyang University of Technology, Taichung, Taiwan. In Proceedings of the Conference. Retrieved from http://www.andrewlian.com/andrewlian/prowww/imagination/index.html.

Liang, W. X., & He, K. K. (2011). The effectiveness of the application of multimedia resources in improving students' listening and speaking skills and phonetic perception, *Audio-visual Education*, *10*, 20.

Lier, L. (1989). Reeling, writhing, drawling, stretching, and fainting in coils: Oral proficiency interviews as conversation. *TESOl Quarterly*, *23*(3), 489-508.

Lin, H. (2005). An investigation on qualifications of primary school English teachers in Zhejiang province. *Journal of Zhejiang Normal University, 30* (3), 82–85.

Linden, D. E. (2007). The working memory networks of the human brain. *The Neuroscientist, 13*(3), 257-267.

Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge: Cambridge University Press.

Liu, D. Y. (2001). The characteristics and challenges of Primary English Teaching. *Teaching Practice and Research*, *7*(7), 2012.

Liu, J. L. (2006). Factors affecting speaking teaching in primary schools. *Teaching research and practice*, *9,* 34-35.

Liu, Q. X., & Shi, J. F. (2007). An Analysis of Language Teaching Approaches and Methods-Effectiveness and Weakness. *Online Submission*, *4*(1), 69-71.

Liu, Y. (2013). Multimedia and primary English teaching. *Curriculum Research, 2* (28) 20-21.

Liu, Y. (2014). Motivation and Attitude: Two Important Non-Intelligence Factors to Arouse Students' Potentialities in Learning English. *Creative Education*, 5(14):1249-1253.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: From theory to practice.* New Jersey: John Wiley & Sons.

Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukasa, K., Ahn, S., & Chen, X. (2009). Second language learners' beliefs about grammar instruction and error correction. *The Modern Language Journal*, *93*(1), 91-104.

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory & cognition*, *42*(1), 41-52.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29-41.

Luoma, S. (2004). *Assessing Speaking.* Cambridge: Cambridge University Press.

Luria, A. R. (1976). *The cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Lyotard, J. F. (1984). *The postmodern condition: A report on knowledge* (Vol. 10). U of Minnesota Press.

Lyotard, J. F. (1984). *The postmodern condition: A report on knowledge* (Vol. 10). Minneapolis: University of Minnesota Press.

MacFarlane, A., Albrair, A., Marshall, C. R., & Buchanan, G. (2012). Phonological working memory impacts on information searching: An investigation of dyslexia. In *Proceedings of the 4th Information Interaction in Context Symposium* (pp. 27-34). ACM.

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental science*, *14*(3), 582-590.

Mackie, C. J., Dockrell, J., & Lindsay, G. (2013). An evaluation of the written texts of children with SLI: The contributions of oral language, reading and phonological short-term memory. *Reading and Writing*, *26*(6), 865-888.

Magnan, S. S. (1988). Grammar and the ACTFL oral proficiency interview: Discussion and data. *The Modern Language Journal*, *72*(3), 266-276.

Malerstein, A.J. (1986). *The conscious mind*. New York: Human Sciences Press

Maley, A. (1987). Poetry and song as effective language-learning activities. *Interactive language teaching*, 93-109.

Marshall, B.K. (1992) *Teaching the Postmodern.* London: Routledge

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379-413.

Martin, p. (2016). *The structure of spoken language*. Cambridge: Cambridge University Press.

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(4), 377-384.

Masoura, E. V., & Gathercole, S. E. (1999). Phonological short- term memory and foreign language learning. *International Journal of Psychology*, *34*(5), 383-388.

Masoura, E. V., & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, *13*(3), 422-429.

Maturana, H.R. & Varela (1992). *The tree of knowledge. The biological roots of human understanding.* Boston, MA: Shambhala.

May, A., Hajak, G., Gänssbauer, S., Steffens, T., Langguth, B., Kleinjung, T., & Eichhammer, P. (2007). Structural brain alterations following 5 days of intervention: dynamic aspects of neuroplasticity. *Cerebral Cortex*, *17*(1), 205-210.

McInnes, A., Humphries, T., Hogg-Johnson, S., & Tannock, R. (2003). Listening comprehension and working memory are impaired in attention-deficit hyperactivity disorder irrespective of language impairment. *Journal of abnormal child psychology*, *31*(4), 427-443.

McNamara, T. (1996). *Measuring Second Language Performance.* London and New York: Longman

McNamara, T. F. (1990). Item response theory and the validation of an ESP test for health professionals. *Language Testing*, *7*(1), 52-76.

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

Mehrabian, A. (1969). Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, *71*(5), 359.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*(2), 270.

Michas, I. C., & Henry, L. A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology*, *12*(2), 147-163.

Miller, J. L. (1978). Interactions in processing segmental and suprasegmental features of speech. *Perception & Psychophysics*, *24*(2), 175-180.

Miller, L. A., Collins, R. L., & Kent, T. A. (2008). Language and the modulation of impulsive aggression. *Language*, *20*(3), 261-273.

MOE (1978). *English syllabus for ten-year full-time primary and secondary schools*. Beijing: MOE.

MOE (2001a). *Guidelines for promoting primary English language teaching.* Beijing: MOE.

MOE (2001b). *English curriculum standards for compulsory education and senior secondary education*. Beijing: MOE.

Morales-Jones, C. A. (2011). Methods/Approaches of Teaching ESOL: A Historical Overview. In H. Zainuddin, N, Yahya, CA Morales-Jones & EN Whelan Ariza (Eds.), *Fundamentals of Teaching English to Speakers of Other Languages in K-12 Mainstream Classrooms* (pp. 63-74). Dubuque: Kendall Hunt.

Morgan, G. A., Griego, O. V., & Gloeckner, G. (2000). *SPSS for Windows: An introduction to use and interpretation in research*. Hove: L. Erlbaum Associates Inc.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic bulletin & review*, *18*(1), 46-60.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, *49*(s1), 285-310.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in second language acquisition*, *28*(01), 111-131.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research*, *48*(1), 61-78.

Murcia, M. C., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation.* Cambridge University Press.

Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, *42*, 258-269.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short- term memory skills. *Reading Research Quarterly*, *33*(3), 320-337.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short- term memory skills. *Reading Research Quarterly*, *33*(3), 320-337.

Nagaraj, G. (2003). *English language teaching: approaches methods techniques*. Telangana: Orient Longman.

Nassaji, H., & Fotos, S. S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. London: Routledge.

Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. P. (1993). Vocabulary size, growth, and use. *The bilingual lexicon*, 115-134.

Nation, I. S. P., & Newton, J. (2008). *Teaching ESL/EFL listening and speaking*. London: Routledge.

Nazari, N. (2013). The effect of implicit and explicit grammar instruction on learners' achievements in receptive and productive modes. *Procedia-Social and Behavioral Sciences*, *70*, 156-162.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*(1), 233-243.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. *The handbook of phonetic sciences*, *5*, 640-673.

Nooteboom, S.G. and J.G. Kruyt (1987). Accents, focus distribution, and the perceived distribution of given and new Information: an experiment. *Journal of the Acoustical Society of America, 82,* 1512-1524.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta- analysis. *Language learning*, *50*(3), 417-528.

Nudo, R. J., Milliken, G. W., Jenkins, W. M., & Merzenich, M. M. (1996). Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *Journal of Neuroscience*, *16*(2), 785-807.

Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. New Jersey: Prentice Hall.

Oakley, D.A. (1985). Animal awareness, consciousness, and self-image. In D.A. Oakley (Eds.), *Brain and mind* (pp. 132-151), London: Methuen.

O'brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, *27*(3), 377-402.

O'brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, *29*(4), 557-581.

O'Brien, M. G. (2014). L2 Learners' Assessments of Accentedness, Fluency, and Comprehensibility of Native and Nonnative German Speech. *Language Learning*, *64*(4), 715-748.

Oetomo, S. (2014). *The correlation between the attitude towards english and the english achievement of grade eight students* (Unpublished PhD thesis).

Oppenheim, A. N. (2000). *Questionnaire design, interviewing and attitude measurement*. London: Bloomsbury Publishing.

Ornstein, R. E., & Thompson, R. F. (1984). *The amazing brain*. Boston: Houghton Mifflin Harcourt.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Massachusetts: Newbury House Publishers.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331-347.

Paquette, K. R., & Rieg, S. A. (2008). Using music to support the literacy development of young English language learners. *Early Childhood Education Journal*, *36*(3), 227-232.

Pell, M. D., & Baum, S. R. (1997). Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain and Language*, *57*(2), 195-214.

Pennington, M. C. (2014). *Phonology in English language teaching: An international approach*. London: Routledge.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology*, *58*(3), 193.

Piaget, J. (1973). *The child and reality.* New York: Grossman.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production (Vol. 20).* Amsterdam: John Benjamins Publishing.

Pu, H., Holcomb, P. J., & Midgley, K. J. (2016). Neural changes underlying early stages of L2 vocabulary acquisition. *Journal of Neurolinguistics*, *40*, 55-65.

Punch, K. F. (2013). *Introduction to social research: Quantitative and qualitative approaches*. London: Sage Publications.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language learning*, *52*(3), 513-536.

Qiang, W. (2002). Primary school English teaching in China-New developments. *English Language Teacher Education and Development*, *7*, 99-108.

Racle, G. (1976). The key principles of Suggestopedia. *Journal of Suggestive Accelerative Learning and Teaching*, *1*(3), 149-163.

Ragó, A., Honbolygó, F., Róna, Z., Beke, A., & Csépe, V. (2014). Effect of maturation on suprasegmental speech processing in full-and preterm infants: a mismatch negativity study. *Research in developmental disabilities*, *35*(1), 192-202.

Rainer, G., & Miller, E. K. (2000). Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron*, *27*(1), 179-189.

Ranalli, J. M. (2001). Consciousness-raising versus deductive approaches to language instruction: A study of learner preferences. Retrieved from http://www.cels,bham,ac.uk/resources/essays/Ranalli1.pdf.

Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, K. A., & Dinse, H. R. (1992). Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *Journal of neurophysiology*, *67*(5), 1031-1056.

Reed, D. J., & Cohen, A. D. (2001). Revisiting raters and ratings in oral language assessment. *Experimenting with uncertainty: Essays in honor of Alan Davies*, *11*, 82-96.

Renard, R., & Van Vlasselaer, J. J. (1976). *Foreign Language teaching with an integrated methodology: the SGAV (Structuro-Global Audio-Visual) methodology*. Paris: Didier.

Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge university press.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Richardson, D. C., Dale, R., & Shockley, K. (2008). Synchrony and swing in conversation: coordination, temporal dynamics, and communication. *Embodied communication*, 75-94.

Richardson, G. (1983). An introductory overview of methods and methodology. *Teaching Modern Languages, Beckenham, Kent: Croom Helm*, 19-37.

Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*, *55*(3), 683-694.

Rosenau, P.M. (1992) *Postmodernism and the Social Sciences*. Princeton: Princeton University Press.

Ryugo, D. K., Limb, C. J., & Redd, E. E. (2000). Brain plasticity: The impact of the environment on the brain as it relates to hearing and deafness. *Cochlear implants: Principles and practices*, 33-56.

Sammler, D., Anwander, A., Bestelmeyer, P. E., & Belin, P. (2015). Dual streams for prosody in the right hemisphere. Retrieved from http://pubman.mpdl.mpg.de.

Sammler, D., Baird, A., Valabrègue, R., Clément, S., Dupont, S., Belin, P., & Samson, S. (2010). The relationship of lyrics and tunes in the processing of unfamiliar songs: a functional magnetic resonance adaptation study. *The Journal of Neuroscience*, *30*(10), 3572-3578.

Sandberg, H. F. (2009). *Get Moving! The Effects of Music and Movement on Student Attention and Engagement*. ProQuest.

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational technology, 35*(5), 31-38.

Savignon, S. J. (1972). *Communicative competence: An experiment in foreign-language teaching (Vol. 12).* Philadelphia: Center for Curriculum Development.

Schifter, D., & Fosnot, C.T. (1993). *Reconstructing mathematics education: Stories of teachers meeting the challenge of reform.* New York: Teachers College Press.

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a Fluent Reader: Reading Skill and Prosodic Features in the Oral Reading of Young Readers. *Journal of educational psychology, 96(*1), 119.

Seidman, S. & Wagner, D.G. (1992). *Postmodernism and Social Theory,* Oxford: Basil Blackwell.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and writing*, *13*(1-2), 81-103.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and writing*, *13*(1-2), 81-103.

Seliger, H. W., & Shohamy, E. G. (1989). *Second language research methods*. Oxford: Oxford University Press.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology*, *45*(1), 21-50.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly journal of experimental psychology*, *22*(2), 261-273.

Sharan, Y., & Sharan, S. (1992). *Expanding Cooperative Learning through Group Investigation*. New York: Teachers College Press.

Shintel, H., Anderson, N. L., & Fenn, K. M. (2014). Talk this way: The effect of prosodically conveyed semantic information on memory for novel words. *Journal of Experimental Psychology: General*, *143*(4), 1437.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological bulletin*, *138*(4), 628.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326.

Shu, D., Li, Z., & Zhang, Y. (2003). Primary school English education in Shanghai: A survey and some thoughts. *Foreign Language World, 3*, 54–62.

Silverman, D. (2000). *Doing qualitative research. A practical guide*. London: Sage.

Simon, H. (1991) Bounded rationality and organizational learning, *Organization Science*, *2*(1), 125-34.

Simon, M. (2011). *Assumptions, limitations and delimitations. Dissertation and scholarly research: Recipes for success.* Seattle, WA: Dissertation Success, LLC. Retrieved from www. dissertationrecipes.com.

Singleton, D. M. (1999). *Exploring the second language mental lexicon*. Stuttgart: Ernst Klett Sprachen.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Smart, B. (1994) *Modern Conditions, Postmodern Controversies*, London: Routledge.

Snowling, M. J., & Hulme, C. (2008). *The science of reading: A handbook* (Vol. 9). New Jersey: John Wiley & Sons.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*(1), 121-159.

Springer, S. P., & Deutsch, G. (1998). *Left brain, right brain: Perspectives from cognitive neuroscience.* New York: WH Freeman/Times Books/Henry Holt & Co.

Srikaew, D., Tangdhanakanond, K., & Kanjanawasee, S. (2015). English Speaking Skills Assessment for Grade 6 Thai Students: An Application of Multivariate Generalizability Theory. *International Journal of Psychology: A Biopsychosocial Approach*, *16*, 47-66.

Stern, H. H. (1983). *Fundamental concepts of language teaching: Historical and interdisciplinary perspectives on applied Linguistic research*. Oxford: Oxford University Press.

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the national Academy of Sciences*, *105*(19), 6791-6792.

Swanson, H. L., & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of experimental child psychology*, *63*(2), 358-385.

Swanson, H. L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology*, *93*(4), 720.

Swanson, H. L., Sáez, L., Gerber, M., & Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology, 96*(1), 3.

Takeuchi, H., Taki, Y., Nouchi, R., Hashizume, H., Sekiguchi, A., Kotozaki, Y., & Kawashima, R. (2013). Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex, 49*(8), 2106-2125.

Tang, W. W. (2003). Creating situations to improve speaking and listening of primary school students. *Shanghai Education and Research, 4*, 19.

Teng, Y. (2010). The comparison ad implications of oral teaching evaluation in primary schools between China and foreign countries. *Curriculum, Textbooks, Methodology, 3*, 108-112.

Thaut, M. H., Peterson, D. A., & McIntosh, G. C. (2005). Temporal entrainment of cognitive functions. *Annals of the New York Academy of Sciences*, *1060*(1), 243-254.

Thorn, A. S. (2013). Phonological short-term memory and foreign language learning. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 141–158). Mahwah, NJ: Erlbaum.

Tian, L. L. (2014). Collaborative learning of English in primary schools. *Journal of Small Writers election: Teaching Experience Exchange, 1*, 453-454.

Tong, X., McBride, C., Lee, C. Y., Zhang, J., Shuai, L., Maurer, U., & Chung, K. K. (2014). Segmental and suprasegmental features in speech perception in Cantonese-speaking second graders: An ERP study. *Psychophysiology, 51*(11), 1158-1168.

Ugolini, M., Wagley, N., Ip, K., Hsu, L. S. J., Arredondo, M. M., & Kovelman, I. (2016). In young readers, the left hemisphere supports the link between temporal processing and phonological awareness. *Speech, Language and Hearing, 19*(1), 17-26.

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques.* Cambridge: Cambridge University Press.

Ur, P. (1996). *Course in language teaching: Practice and theory.* Cambridge: Cambridge University Press.

Usher, R. (1989). Qualifications, paradigms and experiential learning in higher education. *Access and institutional change, 4*, 64-81.

Usher, R., & Edwards, R. (1994). *Postmodernism and education.* London and New York: Routledge.

Usher, R., Bryant, I., & Johnston, R. (1997). *Adult Education and the Postmodern Challenge. Learning beyond the Limits*. London: Routledge.

Van Vugt, M. K., & Jha, A. P. (2011). Investigating the impact of mindfulness meditation training on working memory: a mathematical modeling approach. *Cognitive, Affective, & Behavioral Neuroscience, 11*(3), 344-353.

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, 41(3), 609-624.

Varela, M. L. R., & Martínez, I. M. P. (2013). How are spoken skills assessed in proficiency tests of general English as a Foreign Language? A preliminary survey. *International Journal of English Studies, 13*(2), 53-68.

Vint, L. (2005). Fresh Thinking Drives Creativity & Innovation. *Quick*, 94, 22-24.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of memory and language*, *52*(2), 193-204.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748-751.

Von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, *80*(1), 121-140.

Wakefield, N. (1990) *Postmodernism-The Twilight of the Real*. London: Pluto Press.

Wang, N. (2007). Report on a survey of English education in 12 primary schools in some villages. *Modern Education Forum, 4*, 103–104.

Wang, R. (2009). Primary school English teaching in Gansu province: Current status, problems and solutions. *Northwest Adult Education Journal, 1,* 73–75.

Wechsler, D. (1974). *WISC-R, Wechsler intelligence scale for children*. New York: Psychological Corporation.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised.* New York: Psychological Corporation.

Wenden, A. (1987). Conceptual background and utility. *Learner strategies in language learning, 5,* 3-13.

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.

Westerveld, M. F., & Moran, C. A. (2011). Expository language skills of young school-age children. *Language, speech, and hearing services in schools, 42*(2), 182-193.

Widodo, H. P. (2005). Teaching children using a Total Physical Response (TPR) method: Rethinking. *Bahasa Dan Seni, Tahum, 33*(2), 34-38.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning, 53*(1), 67-121.

Wolf, M. K., Farnsworth, T., & Herman, J. (2008). Validity issues in assessing English language learners' language proficiency. *Educational Assessment, 13*(2-3), 80-107.

Wu, X. (2011). Reviewing the reform of English teaching at primary schools in China. *Curriculum, Textbooks, Methodology,* 4.

Wu, X., Swartz, L. & Levy, B. (2012). *English (Grade 3) (Vol.1) (Vol.2),* Beijing: People's Education Press.

Xi, X. (2008). Methods of Test Validation. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education (Vol. 7)* (pp. 177-196). New York: Springer.

Xu, M., Yang, J., Siok, W. T., & Tan, L. H. (2015). Atypical lateralization of phonological working memory in developmental dyslexia. *Journal of Neurolinguistics*, *33*, 67-77.

Yi. Y. (2009). The application of communicative approach in English teaching in Primary schools. *Shangdong Education*, *7*, 99-100.

Zarifi, A., & Sayyadi, A. (2015). How English Suprasegmental Features of Pronunciation Are Viewed and Treated by Instructors in Iranian Private Language Centers. *Theory and Practice in Language Studies*, *5*(6), 1166.

Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: evidence of brief mental training. *Consciousness and cognition*, *19*(2), 597-605.

Zhang, D. (2012). Chinese primary school English curriculum reform. *Perspectives on Teaching and Learning English Literacy in China*, *4*, 67-83.

Zhang, F. (2012). Teaching Mandarin prosody: a Somatically-Enhanced Approach for second language learners. *Chinese Teaching and Research in Taiwan, 4(*1), 95-130.

Zhang, X. H. (2014). Situational pedagogy in primary English instruction. *Learning Weekly, 5*, 62-63.

Zhang, X. Y. (2016). Strategies for improving the teaching of English speaking skills in Chines primary schools. *China Extracurricular Education*, (z1), 440-442.

Zhang, Z.Y. (2011). A reflection of basic English teaching in China. *Curriculum, Textbooks, Teaching approaches, 1*, 95-99.

APPENDICES

APPENDIX A

A Semi-structured Interview on the Learning of English Speaking (For Students)

Dear students,

This interview is to elicit your opinions on the learning of English speaking. There is no right or wrong answer. Please feel free to say what you think about the learning of speaking. Your response will be kept confidential. Thank you for your participation!

Part 1 Personal Information

Gender ☐ male ☐ female

Age ☐ below 7 ☐ 8 ☐ 9 ☐ 1 0 up

Years of EFL study ☐ 0 ☐ 1 year ☐ 2 years ☐ 3 years and more

Part 2 Interview Questions

1. What do you think of your English speaking ability? (e.g., fluent, can only express basic information using isolated words or short formulae, can hardly speak, etc.)
2. Does your speaking performance influence the result of your English exam? (e.g., have a great influence, a small influence, no influence), why?

3. Can you understand your teacher's talk in English in class? (e.g., always, almost sometimes, never)
4. What (who) is your model when you learn to speak English? (e.g., your teacher, recordings, CD, etc.)
5. Do you usually practice speaking after class? If so, how much time do you spend in practicing speaking per week?
6. Do you meet any problems or difficulties in the learning of speaking? If so, what are they?
7. Is there any speaking test conducted in your school? If yes, what is your score?
8. What do you think of your English pronunciation?
9. What do you think of your vocabulary size?
10. What do you think of your grammatical knowledge?
11. Among the four skills: listening, speaking, reading and writing, which skill do you think is the most important one? Why?
12. Do you often practice speaking with others?
13. Do you like your teacher's teaching? What aspects do you expect him or her to improve?
14. Do you receive after-school English training?
15. What is your opinion on the ways to improve speaking ability?

APPENDIX B

The Chinese Version of a Semi-structured Interview on the Learning of English Speaking

(For Students)

关于小学英语口语学习的调查问卷

（学生部分）

亲爱的同学，你好！

本访谈是为了全面了解小学生的英语口语学习现状，访谈内容仅供研究使用。采取不记名方式，回答无对错之分，希望你不要有任何顾虑，对以下问题真实作答。谢谢你的配合！

第一部分　个人信息

性别　☐ 男　☐ 女

年龄　☐ 7 岁及以下　☐ 8 岁　☐ 9 岁　☐ 10 岁及以上

学习时间　☐ 0 年　☐ 1 年　☐ 2 年　☐ 3 年及以上

每周学习时间　☐ 1 个小时及以下　☐ 2-5 小时　☐ 6 小时及以上

第二部分　访谈问题

1. 你认为自己的英语口语能力怎样？（比如，流利、能基本自由会话、只能说简单单词、简单句子、几乎不能开口）
2. 你觉得口语说得好与坏对你的英语学习成绩有影响吗？（比如，影响大、影响小、基本无影响等）
3. 上英语课时，老师讲英语你能听懂吗？（比如，完全听不懂、 能听懂一点、能听懂大部分、完全能听懂等）
4. 你练习口语时以什么作为模仿对象？（比如老师、录音磁带等）
5. 在课外，你每天另有时间练习口语吗？如果有，你通常花多少时间练习？
6. 对于英语口语学习，你现在面临的最大困难与问题是什么？（比如，不敢开口说、发音不准、知道的单词太少、不知道怎么说一个完整的句子等）
7. 学校有组织英语口语考试吗？如果有，你的分数是多少？
8. 你觉得自己的英语发音怎样？
9. 你觉得自己的词汇量怎样？
10. 你认为自己的语法知识怎样？（比如，不能完整地说出句子、能完整地说出一些句子等）
11. 在听，说，读，写中，你认为哪一项在英语学习中最重要？为什么？哪一项最不中要？为什么？
12. 你经常和同学一起练习口语吗？
13. 你喜欢老师的教学吗？如果可以，你希望老师在什么方面改进？
14. 你课后参加英语辅导班吗？
15. 你认为怎样才能学好口语呢？

APPENDIX C

IOC Analysis for a Semi-structured Interview on the Learning of English Speaking

(For Students)

Item	Experts					Result of analysis
	1	2	3	4	5	
Q1	+1	+1	+1	+1	+1	√
Q2	+1	+1	+1	+1	+1	√
Q3	+1	+1	+1	+1	+1	√
Q4	+1	0	+1	+1	+1	√
Q5	0	+1	+1	+1	0	√
Q6	+1	+1	+1	+1	+1	√
Q7	+1	+1	+1	+1	+1	√
Q8	+1	0	+1	+1	+1	√
Q9	-1	+1	+1	0	+1	√
Q10	+1	+1	+1	+1	+1	√
Q11	+1	+1	+1	+1	+1	√
Q12	+1	+1	+1	+1	+1	√
Q13	0	+1	+1	+1	+1	√
Q14	+1	0	+1	0	+1	√
Q15	+1	+1	+1	+1	+1	√
Total	11	12	15	13	14	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 15

R=11+12+15+13+14=65(Scores from experts)

N=5(Number of experts)

IOC=65/5=13.0

Percentage: 13.0/15×100%=86.7%

The table above shows that the analysis result of IOC is 13.0, and the percentage is 86.7% which is higher than 80%. Therefore, the items are suitable for adoption in an interview for students.

APPENDIX D

A Semi-structured Interview on the Teaching of English Speaking

(For Teachers)

Dear teachers,

This interview is to elicit your opinions on the teaching of speaking. There is no right or wrong answer. Please feel free to say what you think about the teaching of speaking. Your responses will be kept confidential. Thank you for your participation!

Part 1 Personal Information

Gender ☐ male ☐ female

Age ☐ 20-25 ☐ 26-31 ☐ 31 up

Years of EFL teaching ☐ 1-5 years ☐ 6-11 years ☐ 11years or more

Academic degree ☐ associate degree ☐ bachelor's degree ☐ MA or PhD

Major ☐ English majors ☐ non-English majors

Part 2 Interview Questions

1. How do you usually conduct your teaching of speaking? (e.g., ask students to read aloud, recite, play games, role-play, etc.)
2. Is your teaching conducted mainly in English or in Chinese? Why?

3. What do you think are the major problems existing in the teaching of speaking?
4. What do you think of your speaking ability?
5. What do you think should be done to improve one's English speaking ability?
6. Are there any extracurricular activities organized by the school for the students to practice English?
7. Are there multi-media resources available for teaching in your school?
8. What's your opinion of the effectiveness of your teaching method in improving students' speaking skills?
9. What do you usually do to stimulate students' interest in learning speaking?
10. What do you think are the important aspects of students' speaking skills?
11. What do you think is the chief goal of English (speaking) learning in primary school?
12. Do you have any training opportunities in English (speaking) teaching? Why?
13. Do you have any opportunities to take part in research projects about English (speaking) teaching? Why?
14. What do you think of your students' speaking ability? What are the major problems met by the students when they speak?
15. What (who) is the model when you teach pronunciation?
16. Is there an English speaking test conducted in your school?
17. Do you usually assign listening and speaking tasks to the students after class? If yes, please specify the types of tasks you normally assign?
18. How many academic periods of English teaching are there for Grade 3 students in your school? And how long does one period last?

APPENDIX E

The Chinese Version of a Semi-structured Interview on the Teaching of English Speaking

(For teachers)

关于小学英语口语教学的访谈问卷

（教师部分）

亲爱的老师，您好！

本访谈是为了全面了解您在英语口语教学方面的情况，访谈内容仅供研究使用。采取不记名方式，回答无对错之分，希望您不要有任何顾虑，对以下问题真实作答。谢谢您的配合！

第一部分　个人信息

性别	□ 男	□女	
年龄	□ 20-25	□ 26-31	□ 31 岁以上
教龄	□ 1-5 年	□ 6-11 年	□ 11 年以上
学历：	□ 专科	□ 本科	□ 研究生
专业：	□ 英语	□ 非英语	

第二部分　　访谈问题

1. 您在课堂上是否使用英语教学？（比如，很少使用、大部分时间使用、完全使用等）
2. 您认为英语口语教学所面临的最大的困难是什么？
3. 您认为自己的口语水平如何？（比如，不好、一般、良好、很好等）
4. 你认为应该怎样加强自己的口语能力呢？
5. 学校有为学生组织英语课外活动来练习口语吗？
6. 学校有多媒体设备供教学使用吗？
7. 您觉得自己的教学方法在帮助学生提高口语上的效果如何？
8. 您是如何激发学生练习口语的兴趣的？
9. 学生在练习口语时您最注重的是什么？
10. 您认为小学阶段英语学习最大的目的是什么？
11. 学校有提供进修机会给英语老师吗？
12. 您参加科研项目的情况如何？（比如，基本不参加、很少参加、经常参加等）
13. 您认为学生的口语水平怎样？他们说英语时最大的问题是什么？
14. 上课时您是以什么方式教学生读英语呢？
15. 期中或期末考试考口语吗？有没有系统的口语考试评价标准？
16. 目前使用的教材配套的音频资料只有录音磁带吗？还是有 MP3、CD 等其它音频或视频资料？
17. 课后您会布置听说作业给学生吗？如果有，一周大概多少时间？
18. 您所在的学校 3 年级每周有多少节英语课，每节课多长时间？

APPENDIX F

IOC Analysis for a Semi-structured Interview on the Teaching of English Speaking

(For teachers)

Item	Experts					Result of analysis
	1	2	3	4	5	
Q1	+1	+1	+1	+1	+1	√
Q2	+1	+1	+1	+1	+1	√
Q3	+1	+1	+1	+1	+1	√
Q4	+1	0	+1	+1	0	√
Q5	0	+1	+1	+1	+1	√
Q6	+1	+1	+1	+1	+1	√
Q7	+1	+1	+1	+1	+1	√
Q8	+1	+1	+1	+1	+1	√
Q9	0	-1	0	+1	+1	√
Q10	+1	+1	+1	0	+1	√
Q11	+1	0	+1	+1	+1	√
Q12	+1	+1	+1	+1	+1	√
Q13	+1	+1	+1	+1	+1	√
Q14	0	+1	+1	0	+1	√
Q15	+1	+1	+1	+1	0	√
Q16	+1	+1	+1	+1	+1	√
Q17	+1	+1	+1	0	+1	√
Q18	+1	+1	+1	+1	+1	√
Total	15	14	17	15	16	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 18

R=15+14+17+15+16=77(Scores from experts)

N=5(Number of experts)

IOC=77/5=15.4

Percentage: 15.4/18×100%=85.6%

The table above shows that the analysis result of IOC is 15.4, and the percentage is 85.6% which is higher than 80%. Therefore, the items are suitable for adoption in an interview for teachers.

APPENDIX G

A Questionnaire on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

Part 1 Personal Information

Gender ☐ male ☐ female

Age ☐ below 7 ☐ 8 ☐ 9 ☐ 10 up

Years of EFL study ☐ 0 ☐ 1 year ☐ 2 years ☐ 3 years and more

Time spent
Per week after class ☐ 1 hour or less ☐ 2-5 hours ☐ 6 hours or more

Part 2 Students' Perceptions of VTA

Directions:

This questionnaire is designed to gather information about your opinions on learning speaking and training PWM via VTA. VTA stands for a Verbotonal-based approach for learning speaking and training PWM (phonological working memory). Please read each statement carefully and tick (√) the response which represents your opinion.

No.	Statements of students' perceptions	Strongly agree	agree	undecided	disagree	Strongly disagree
1	VTA is effective in the learning of speaking.					
2	I believe that my speaking skills are improved by using VTA.					
3	I gain confidence in speaking English via VTA.					
4	My interest in learning to speak English is improved via VTA.					
5	I am happy with VTA in the learning of speaking.					
6	VTA is interesting in the learning of speaking.					
7	I prefer VTA to traditional approaches in the learning of speaking.					
8	I would like to learn to speak English via VTA in class.					
9	I would like to learn to speak English via VTA on my own.					
10	My learner autonomy is improved via VTA.					
11	I am able to remember the sound of words more quickly than before via VTA.					
12	I am able to remember the sound of words more accurately than before via VTA.					

APPENDIX H

The Chinese Version of a Questionnaire on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

学生关于用 VTA 学习口语以及提高语音工作记忆能力的观点调查问卷

第一部分　个人信息

性别　☐ 男　☐ 女

年龄　☐7 岁及以下　☐ 8 岁　☐ 9 岁　☐ 10 岁及以上

学习时间　☐ 0 年　☐ 1 年　☐ 2 年　☐ 3 年及以上

课后每周学习时间　☐ 1 个小时及以下　☐2-5 小时　☐ 6 小时及以上

第二部分　学生对于利用 VTA 学习口语以及提高语音工作记忆能力的观点

提示：本问卷旨在收集你对于语调听觉法（VTA）用于口语学习的观点。请你根据自身情况，在代表你观点的选项下面打勾（√）。

题号	内容	很同意	同意	一般	不同意	很不同意
1	VTA 对口语学习有效。					
2	我相信我能利用 VTA 提高英语口语。					
3	通过 VTA 学习方式，我敢开口说英语了。					
4	通过 VTA，我的听力水平提高了。					
5	我对 VTA 用于英语学习很满意。					
6	VTA 很有趣。					
7	比较传统的学习方式，我更喜欢 VTA。					
8	我更愿意在课堂上利用 VTA 学习。					
9	我更愿意单独利用 VTA 学习。					
10	通过 VTA，我的自主学习能力提高了。					
11	通过 VTA，我能更快地记住单词的发音了。					
12	通过 VTA，我能更准确地记住单词的发音了。					

APPENDIX I

IOC Analysis for a Questionnaire on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

Item	Experts					Result of analysis
	1	2	3	4	5	
Q1	+1	+1	+1	+1	+1	√
Q2	+1	+1	+1	+1	+1	√
Q3	+1	0	+1	+1	+1	√
Q4	+1	+1	+1	+1	+1	√
Q5	+1	+1	+1	+1	+1	√
Q6	0	+1	+1	+1	0	√
Q7	+1	0	-1	+1	+1	√
Q8	+1	+1	0	+1	+1	√
Q9	+1	+1	+1	+1	+1	√
Q10	+1	+1	+1	+1	+1	√
Q11	-1	+1	+1	+1	+1	√
Q12	+1	+1	+1	+1	+1	√
Total	9	10	9	12	11	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 12

R=9+10+9+12+11=51(Scores from experts)

N=5(Number of experts)

IOC=51/5=10.2

Percentage: 10.2/12×100%=85.0%

The table above shows that the analysis result of IOC is 10.2, and the percentage is 85.0% which is higher than 80%. Therefore, the items are suitable for adoption in a questionnaire.

APPENDIX J

A Semi-structured Interview on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

Part 1 Personal Information

Gender ☐ male ☐ female

Age ☐ below 7 ☐ 8 ☐ 9 ☐ 10 up

Years of EFL study ☐ 0 ☐ 1 year ☐ 2 years ☐ 3 years and more

Time spent
Per week after class ☐ 1 hour or less ☐ 2-5 hours ☐ 6 hours or more

Part 2 Interview Questions

1. Is VTA effective in your learning of speaking? If yes, in what ways? If not, why?
2. Do you like to learn speaking via VTA in groups or on your own? Why?
3. Do you think that VTA can help you memorize the pronunciation of words more quickly than before? Why?
4. Do you think that VTA can help you memorize the pronunciation of words more accurately than before? Why?
5. Do you believe that VTA can improve your listening skill? Why?
6. Do you prefer VTA to traditional approaches? Why?
7. Is VTA interesting? If so, in what ways?
8. Do you think that VTA can improve your confidence in speaking English? Why?
9. Do you think that VTA can develop your learner autonomy?
10. What else would you like to say about learning to speak via VTA

APPENDIX K

The Chinese Version of a Semi-structured Interview on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

学生关于用 VTA 方式学习口语和提高语音工作记忆的观点访谈问卷

第一部分　个人信息

性别　□ 男　□ 女

年龄　□ 7 岁及以下　□ 8 岁　□ 9 岁　□ 10 岁及以上

学习时间　□ 0 年　□ 1 年　□ 2 年　□ 3 年及以上

每周学习时间　□ 1 个小时及以下　□ 2-5 小时　□ 6 小时及以上

第二部分 访谈问题

1. 你觉得 VTA 对你的英语口语学习有效吗？如果有，在哪些方面有？如果没有，为什么？
2. 你更愿意利用 VTA 进行小组合作学习呢还是单独学习？为什么？
3. 你觉得 VTA 能帮助你更快地记住单词的发音吗？为什么？
4. 你觉得 VTA 能帮助你更准确地记住单词的发音吗？为什么？
5. 通过 VTA 学习方式，你对英语的学习兴趣提高了吗？为什么？
6. 相比传统的学习方式，你更喜欢 VTA 吗？为什么?
7. VTA 有趣吗？如果有趣，怎么有趣？
8. 通过 VTA 学习方式，你讲英语时更自信了吗？为什么？
9. 课后你愿意用 VTA 方式学习英语吗？为什么？
10. 对于 VTA 用于英语口语学习，你还有什么其他要说的？

APPENDIX L

IOC Analysis for a Semi-structured Interview on Students' Perceptions of Learning to Speak English and Improving PWM via VTA

Item	Experts					Result of analysis
	1	2	3	4	5	
Q1	+1	+1	+1	+1	+1	√
Q2	+1	+1	+1	+1	+1	√
Q3	+1	+1	+1	+1	+1	√
Q4	+1	0	+1	0	+1	√
Q5	-1	+1	+1	+1	+1	√
Q6	+1	0	-1	+1	+1	√
Q7	+1	+1	+1	+1	+1	√
Q8	+1	+1	0	+1	+1	√
Q9	+1	+1	+1	+1	+1	√
Q10	+1	+1	+1	+1	+1	√
Total	8	8	8	9	10	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 10

R=8+8+8+9+10=43(Scores from experts)

N=5(Number of experts)

IOC=43/5=8.6

Percentage: 8.6/10×100%=86.0%

The table above shows that the analysis result of IOC is 8.6, and the percentage is 86.0% which is higher than 80%. Therefore, the items are suitable for adoption in an interview.

APPENDIX M

A Semi-structured Interview on the Teacher's Perceptions of Teaching English speaking and Improving PWM via VTA

Dear teacher,

This semi-structured interview is designed to gather information about your opinions on teaching speaking and training PWM via VTA. VTA stands for a Verbotonal-based approach for learning speaking and training PWM (phonological working memory). Please answer the questions below. Thank you for your cooperation!

Part 1 Personal Information

Gender ☐ male ☐ female

Age ☐ 20-25 ☐ 26-31 ☐ 31 up

Years of EFL teaching ☐ 1-5 years ☐ 6-11 years ☐ 11 years or more

Academic degree ☐ associate degree ☐ bachelor's degree ☐ MA or PhD

Major ☐ English majors ☐ non-English majors

Part 2 Interview Questions

1. Is VTA helpful to your teaching of speaking? If yes, in what ways? If no, why?
2. Do you believe that VTA will improve students' speaking skills? Why?
3. Do you believe that VTA will improve students' pronunciation? Why?
4. Do you believe that VTA will improve students' ability in memorizing words? Why?
5. Do you think that VTA can develop students' learner autonomy? Why?
6. Do you believe that VTA can stimulate students' interest in English learning? Why?
7. Do you prefer VTA to traditional approaches? Why?
8. Do you like students to learn to speak via VTA in group or on their own? Why?
9. Please describe your opinions of VTA.
10. What else would you like to say about teaching speaking via VTA?

APPENDIX N

The Chinese Version of a Semi-structured Interview on the Teacher's Perceptions of Teaching English speaking and Improving PWM via VTA

教师关于用 VTA 方式进行口语教学以及训练语音工作记忆的观点访谈问卷

（教师部分）

亲爱的老师，您好！

本访谈是为了全面了解您在运用语调听觉法 VTA 进行英语口语教学以及语音工作记忆训练方面的情况，访谈内容仅供研究使用。采取不记名方式，回答无对错之分，希望您不要有任何顾虑，对以下问题真实作答。谢谢您的配合！

第一部分　个人信息

性别　☐ 男　☐ 女

年龄　☐ 20-25　☐ 26-31　☐ 31 岁以上

教龄　☐ 1-5 年　☐ 6-11 年　☐ 11 年以上

学历：　☐ 专科　☐ 本科　☐ 研究生

专业：　☐ 英语　☐ 非英语

第二部分　　访谈问题

1. VTA 对于你的口语教学有帮助吗？如果有，在哪些方面？如果没有，为什么？
2. 你认为 VTA 能提高学生的英语口语水平吗？为什么？
3. 你认为 VTA 能提高学生的语音水平吗？为什么？
4. 你认为 VTA 能提高学生记忆单词的能力吗？为什么？
5. 你认为 VTA 能提高学生的自主学习能力吗？为什么？
6. 你认为 VTA 能提高学生对英语学习的兴趣吗？为什么？
7. 你更喜欢 VTA 还是传统的教学方法？为什么？
8. 你喜欢学生利用 VTA 进行单独学习呢还是与别人一起合作学习？
9. 请描述你对 VTA 用于口语教学的态度。
10. 关于 VTA 用于英语口语教学你还有其它想要表达的吗？

APPENDIX O

IOC Analysis for a Semi-structured Interview on the Teacher's Perceptions of Teaching English speaking and Improving PWM via VTA

Item	Experts					Result of analysis
	1	2	3	4	5	
Q1	+1	+1	+1	+1	+1	√
Q2	+1	+1	+1	-1	+1	√
Q3	+1	+1	+1	0	+1	√
Q4	-1	+1	0	+1	+1	√
Q5	+1	+1	+1	+1	+1	√
Q6	+1	+1	-1	+1	0	√
Q7	+1	+1	+1	+1	+1	√
Q8	+1	+1	0	+1	+1	√
Q9	+1	+1	+1	+1	+1	√
Q10	+1	+1	+1	+1	+1	√
Total	8	10	8	7	9	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 10

R=8+8+8+9+10=42(Scores from experts)

N=5(Number of experts)

IOC=42/5=8.4

Percentage: 8.4/10×100%=84.0%

The table above shows that the analysis result of IOC is 8.4, and the percentage is 84.0% which is higher than 80%. Therefore, the items are suitable for adoption in an interview.

APPENDIX P

Speaking Test Paper (Pretest)

Part 1 Word-reading (20%). There are 40 words in this part. Please read them aloud.

1	cat	11	face	21	mum	31	brother
2	fish	12	pig	22	two	32	black
3	pen	13	red	23	water	33	zoo
4	five	14	rice	24	ten	34	bread
5	foot	15	bear	25	white	35	egg
6	pencil	16	arm	26	eye	36	body
7	ruler	17	bag	27	nose	37	school
8	mouth	18	milk	28	no	38	bird
9	panda	19	monkey	29	nine	39	dog
10	three	20	tiger	30	hand	40	seven

Part 2 Sentence-reading. (20%) There are 10 sentences in this part. Please read them out.

1. What's your name?
2. Let's go to school!

3. Can I have some water, please?
4. Nice to meet you.
5. How are you?
6. I’m fine, thank you.
7. Good morning, Miss Green.
8. Happy birthday!
9. How old are you?
10. I’m six years old.

Part 3 Singing (20%). Please sing the following song out.

Hello, Mike. Hello, Chen Jie. Hello, hello! How are you?

This is Mary. This is Willy. Nice to meet you. How are you?

I’m fine. I’m fine. Thank you, thank you very much.

I’m fine. I’m fine. Thank you, thank you very much.

Part 4 Oral interview (introduce your body) (40%). First, please introduce the different parts of your body using at least 4 sentences. For example, look at me! This is my face…. Then, answer the following question about the topic using 1-3 sentences.

Question: Which part of your body do you like most? Why?

APPENDIX Q

IOC Analysis for Pretest Paper

Part 1 Word-reading

Item	Content	Experts					Result of analysis
		1	2	3	4	5	
1	cat	+1	+1	+1	+1	+1	√
2	fish	+1	+1	+1	+1	+1	√
3	pen	+1	+1	+1	+1	+1	√
4	five	+1	+1	+1	+1	+1	√
5	foot	+1	+1	+1	+1	+1	√
6	pencil	+1	+1	+1	+1	+1	√
7	ruler	+1	+1	+1	+1	+1	√
8	mouth	+1	+1	+1	0	+1	√
9	panda	+1	+1	+1	+1	+1	√
10	three	+1	+1	+1	+1	+1	√
11	face	+1	+1	+1	+1	+1	√
12	pig	+1	+1	+1	+1	+1	√
13	red	+1	+1	+1	+1	+1	√
14	rice	+1	+1	+1	+1	+1	√
15	bear	+1	+1	+1	+1	+1	√
16	arm	+1	+1	+1	+1	+1	√
17	bag	+1	+1	+1	+1	+1	√
18	milk	+1	+1	+1	+1	+1	√
19	monkey	+1	+1	+1	+1	+1	√
20	tiger	+1	+1	+1	+1	+1	√
21	mum	+1	+1	+1	+1	+1	√
22	two	+1	+1	+1	+1	+1	√
23	water	+1	+1	+1	+1	+1	√
24	ten	+1	+1	+1	+1	+1	√
25	white	+1	+1	+1	+1	+1	√
26	eye	+1	+1	0	+1	+1	√
27	nose	+1	+1	+1	+1	+1	√
28	no	+1	+1	+1	+1	+1	√
29	nine	+1	+1	+1	+1	+1	√

30	hand	+1	+1	+1	+1	+1	√
31	brother	+1	+1	+1	0	+1	√
32	black	+1	+1	+1	+1	+1	√
33	zoo	+1	+1	+1	+1	+1	√
34	bread	+1	+1	+1	+1	+1	√
35	egg	+1	+1	+1	+1	+1	√
36	body	+1	+1	+1	+1	+1	√
37	school	+1	+1	0	+1	+1	√
38	bird	+1	+1	+1	+1	+1	√
39	dog	+1	+1	+1	+1	+1	√
40	seven	0	+1	0	+1	0	√
Total		39	40	37	38	39	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 40

R=39+40+37+38+39=193(Scores from experts)

N=5(Number of experts)

IOC=193/5=38.6

Percentage: 38.6/40×100%=96.5%

The table above shows that the analysis result of IOC is 38.6, and the percentage is 96.5% which is higher than 80%. Therefore, the items are suitable for adoption in a test for Grade 3 students at the end of their first semester.

Part 2 Sentence-reading

Item	Experts					Result of analysis
	1	2	3	4	5	
S1	+1	+1	+1	+1	+1	√
S2	+1	+1	+1	0	+1	√
S3	+1	+1	+1	+1	+1	√
S4	0	+1	+1	+1	+1	√
S5	+1	+1	+1	+1	+1	√
S6	+1	+1	0	+1	0	√
S7	+1	+1	+1	+1	+1	√
S8	+1	+1	+1	+1	+1	√

S9	+1	+1	+1	+1	+1	√
S10	+1	+1	+1	+1	+1	√
Total	9	10	9	9	9	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 10

R=9+10+9+9+9=46 (Scores from experts)

N=5 (Number of experts)

IOC=46/5=9.2

Percentage: 9.2/10×100%=92%

The table above shows that the analysis result of IOC is 9.2, and the percentage is 92.0% which is higher than 80%. Therefore, the items are suitable for adoption in a test for Grade 3 students at the end of their first semester.

APPENDIX R

Speaking Test Paper (Posttest)

Part 1 Word-reading (20%). There are 40 words in this part. Please read them aloud.

1	boat	11	grape	21	twelve	31	sister
2	Canada	12	long	22	small	32	short
3	desk	13	map	23	pear	33	friend
4	father	14	nineteen	24	she	34	children
5	eighteen	15	orange	25	man	35	new
6	banana	16	pupil	26	car	36	toy
7	chair	17	under	27	dad	37	girl
8	grandfather	18	watermelon	28	big	38	buy
9	eleven	19	woman	29	tall	39	fruit
10	fat	20	seventeen	30	strawberry	40	and

Part 2 Sentence-reading. (20%) There are 10 sentences in this part. Please read them out.

1. Where are you from?
2. Come here, Children!
3. Is she your mother?
4. Who's that woman?
5. I'm from the UK.
6. It has small eyes and big ears.
7. Where is my pencil box?
8. Honey, let's buy some fruit.
9. Do you like oranges?
10. How many kites do you see?

Part 3 Singing (20%). Please sing the following song.

One, two, three, four, five.

Once I caught a fish alive.

Six, seven, eight, nine, ten.

Then, I let it go again.

Why did you let it go?

Cause it bit my finger so.

Which finger did it bite?

This little finger on the right.

Part 4 Oral interview (introduce your family) (40%). First, please introduce your family members using at least 4 sentences. For example, I am Wei Hua. This is my family…. Then, answer the following question about the topic using 1-3 sentences.

Question: Do you like your family? Why?

APPENDIX S

IOC Analysis for Posttest Paper

Part 1 Word-reading

Item	Content	Experts					Result of analysis
		1	2	3	4	5	
1	boat	+1	+1	+1	+1	+1	√
2	Canada	+1	+1	+1	+1	+1	√
3	desk	+1	+1	+1	+1	+1	√
4	father	+1	+1	+1	+1	+1	√
5	eighteen	+1	+1	+1	+1	+1	√
6	banana	+1	+1	+1	+1	+1	√
7	chair	+1	+1	+1	+1	+1	√
8	grandfather	+1	+1	0	+1	+1	√
9	eleven	+1	+1	+1	+1	+1	√
10	fat	+1	+1	+1	+1	+1	√
11	grape	+1	+1	+1	+1	+1	√
12	long	+1	+1	+1	+1	+1	√
13	map	+1	+1	+1	+1	+1	√
14	nineteen	+1	+1	+1	+1	+1	√
15	orange	+1	+1	+1	+1	+1	√
16	pupil	+1	+1	+1	+1	+1	√
17	under	+1	+1	+1	+1	+1	√
18	watermelon	0	+1	+1	0	+1	√
19	woman	+1	+1	+1	+1	+1	√
20	seventeen	+1	+1	+1	+1	+1	√
21	twelve	0	+1	+1	+1	+1	√
22	small	+1	+1	+1	+1	+1	√
23	pear	+1	+1	+1	+1	+1	√
24	she	+1	+1	+1	+1	+1	√
25	man	+1	+1	+1	+1	+1	√
26	car	+1	+1	+1	+1	+1	√
27	dad	+1	+1	+1	+1	+1	√
28	big	+1	+1	+1	+1	+1	√
29	tall	+1	+1	+1	+1	+1	√

30	strawberry	0	+1	0	+1	0	√
31	sister	+1	+1	+1	+1	+1	√
32	short	+1	+1	+1	+1	+1	√
33	friend	+1	+1	+1	+1	+1	√
34	children	+1	+1	+1	0	+1	√
35	new	+1	+1	+1	+1	+1	√
36	toy	+1	+1	+1	+1	+1	√
37	girl	+1	+1	+1	+1	+1	√
38	buy	+1	+1	+1	+1	+1	√
39	fruit	+1	+1	+1	+1	+1	√
40	and	+1	+1	+1	+1	+1	√
Total		37	40	38	38	39	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 40

R=37+40+38+38+39=192(Scores from experts)

N=5(Number of experts)

IOC=192/5=38.4

Percentage: 38.4/40×100%=96.0%

The table above shows that the analysis result of IOC is 38.4, and the percentage is 96.0% which is higher than 80%. Therefore, the item are suitable for adoption in a speaking test for Grade 3 students at the end of their second semester.

Part 2 Sentence-reading

Item	Experts					Result of analysis
	1	2	3	4	5	
S1	+1	+1	+1	+1	+1	√
S2	+1	+1	+1	0	+1	√
S3	+1	+1	+1	+1	+1	√
S4	0	+1	+1	+1	+1	√
S5	+1	+1	+1	+1	+1	√
S6	+1	+1	+1	+1	0	√
S7	+1	+1	+1	+1	+1	√
S8	+1	+1	+1	+1	+1	√

S9	+1	+1	+1	+1	+1	√
S10	+1	+1	+1	+1	+1	√
Total	9	10	10	9	9	

Notes:

1. +1= the item is congruent with the objective
2. -1= the item is not congruent with the objective
3. 0=uncertain about this item

Result of IOC:

(IOC=∑R/N)

Item number: 10

R=9+10+10+9+9=47 (Scores from experts)

N=5 (Number of experts)

IOC=47/5=9.4

Percentage: 9.4/10×100%=94%

The table above shows that the analysis result of IOC is 9.4, and the percentage is 94.0% which is higher than 80%. Therefore, the items are suitable for adoption in a test for Grade 3 students at the end of their second semester.

APPENDIX T

Children's Test of Nonword Repetition

with Wordlikeness Ratings (pretest)

Low-wordlike set	High-wordlike set
Loddernapish 1.1	Prindle 3.9
Ballop 2.4	Confrantually 3.4
Woogalamic 1.6	Pennel 3.4
Tafflest 1.8	Sladding 2.7
Detratapillic 2.0	Thickery 3.1
Brasterer 2.1	Barrazon 2.9
Altupatory 2.5	Stopograttic 3.5
Underbrantuand 2.2	Glistering 3.0
Blonterstaping 1.6	Fennerizer 3.0
Frescovent 1.8	Voltularity 3.0

APPENDIX U

Children's Test of Nonword Repetition with Wordlikeness Ratings (posttest)

Low-wordlike set	High-wordlike set
Perplisteronk 1.1	Hampent 2.7
Skiticult 1.7	Versatrationist 2.8
Bannow 1.9	Commerine 3.0
Contramponist 2.1	Bannifer 2.7
Empliforvent 2.2	Doppelate 3.0
Glistow 2.3	Trumpetine 3.4
Pristoractional 2.3	Rubid 3.8
Penerriful 2.4	Defermication 3.9
Diller 2.3	Sepretennial 3.3
Reutterpation 2.4	Commeecitate 3.2

APPENDIX V

Rubrics for Nonword Repetition Test

Dear colleague,

Thank you for your cooperation. You are invited to score students' recordings of nonword repetition test. There are altogether 20 nonwords in the test. Please listen carefully and thoroughly to the recordings and then judge whether the student's repetition attempt is phonologically incorrect or correct, with allowance being made for any regional accent or other cases where a student is known consistently to pronounce one phoneme as another. The total test score is provided by the number of correct repetition attempts made by the student. (Maximum=20).

Rubrics for the nonword repetition test

Criteria	Score
Phonologically correct	1
Phonologically incorrect	0

APPENDIX W

Children's Test of Nonword Repetition with Wordlikeness Ratings and Phonetic Transcriptions (pretest)

Low-wordlike set

Loddernapish 1.1 [lɔdə'neipiʃ]

Ballop 2.4 ['bæləp]

Woogalamic 1.6 [wugə'læmik]

Tafflest 1.8['təfləst]

Detratapillic 2.0 [ditrətə'pilik]

Brasterer 2.1 ['bræstərə]

Altupatory 2.5 [æltʃu'peitəri]

Underbrantuand 2.2 [undə'brantʃuənd]

Blonterstaping 1.6['blɔntə'steipiŋ]

Frescovent 1.8 ['freskəvent]

High-wordlike set

Prindle 3.9 ['prindl]

Confrantually 3.4 [kən'frantʃuli]

Pennel 3.4 ['penl]

Sladding 2.7 ['slædiŋ]

Thickery 3.1 ['θikəri]

Barrazon 2.9 ['bærəzən]

Stopograttic 3.5 [stɔpə'grætik]

Glistering 3.0 ['glistəriŋ]

Fennerizer 3.0 [fenə'raizə]

Voltularity 3.0 [vɔltʃu' læriti]

APPENDIX X

Children's Test of Nonword Repetition with Wordlikeness Ratings and Phonetic Transcriptions (posttest)

Low-wordlike set	**High-wordlike set**
Perplisteronk 1.1 [pəˈplistərɔŋk]	Hampent 2.7 [ˈhæmpənt]
Skiticult 1.7 [ˈskitikʌlt]	Versatrationist 2.8 [vəsəˈtreiʃənist]
Bannow 1.9[ˈbænau]	Commerine 3.0 [ˈkɔməriŋ]
Contramponist 2.1[ˈkənˈtrampənist]	Bannifer 2.7 [ˈbænifə]
Empliforvent 2.2 [empliˈfɔvənt]	Doppelate 3.0 [ˈdɔpəleit]
Glistow 2.3 [ˈglistəu]	Trumpetine 3.4 [trʌmpətin]
Pristoractional 2.3 [pristəˈrækʃənl]	Rubid 3.8 [ˈrubid]
Penerriful 2.4 [pənˈerifl]	Defermication 3.9 [difəmiˈkeiʃən]
Diller 2.3 [ˈdilə]	Sepretennial 3.3 [seprəˈtenjul]
Reutterpation 2.4 [riʌtəˈpeiʃən]	Commeecitate 3.2 [kəmˈisəteit]

APPENDIX Y

Rubrics for Speaking Test (pretest and posttest)

Dear colleague,

Thank you for your cooperation. You are invited to score students' recordings of the speaking test. There are altogether 4 parts in the test: (1) Word-reading, (2) Sentence-reading, (3) Singing, and (4) Oral interview. The rubrics for scoring these parts are outlined as follows.

Part 1. Word-reading: This part is to assess students' pronunciation of segments, i.e. vowels and consonants. Students are required to read aloud the 40 words. The maximum score of this part is 20. This part will be scored dichotomously (phonologically correct or incorrect).

Description	**Score**
Phonologically correct	0.5
Phonologically incorrect	0

Part 2. Sentence-reading: This part is to assess students' pronunciation of both segmentals and suprasegmentals. There are 10 sentences to be read out. And the maximum score of this part is 20.

Description	Score
No Fluency; insufficient information to score or incomprehensible	0-5
Intermediate fluency; numerous problems with pronunciation of individual sounds, speech rate, intonation, and rhythm; slightly comprehensible	6-10
Good fluency; some problems with pronunciation of individual sounds, speech rate, intonation, and rhythm; no serious problems in comprehensibility	11-15
Advanced Fluency; accurate pronunciation of individual sounds; no or very few minor phonemic/phonetic errors; natural speech rate, intonation, and rhythm; fully comprehensible	16-20

Part 3. Singing: Students are required to sing a song in this part. This singing part is to assess students' sense of the rhythm of English songs. And the maximum score of this part is 20.

Description	Score
Insufficient information to score or incomprehensible	0-5
Fairly good sense of rhythm; numerous problems with pronunciation of individual sounds, tempo, intonation, and rhythm; slightly comprehensible	6-10
Good sense of rhythm; some problems with pronunciation of individual sounds, tempo, intonation, and rhythm; no serious problems in	11-15

intelligibility	
Very good sense of rhythm; accurate pronunciation of individual sounds; no or very few minor phonemic/phonetic errors; appropriate tempo, intonation, and rhythm; fully intelligible	16-20

Part. 4 Oral interview: In this part, students are required to talk on a given topic using 4-6 sentences, then answer a question about the topic using 1-3 sentences. This part is to assess students' speaking skills in terms of pronunciation, grammar, vocabulary, fluency and comprehensibility. You are required to give two scores. The first is the overall proficiency score, and the second is the score for the five subskills: pronunciation, grammar, vocabulary, fluency and comprehensibility. The overall maximum score for this part is 40, and the score for each subskill is 30.

Rubrics for the overall proficiency score

Description	**Score**
Insufficient information to score or incomprehensible; no fluency	0-5
Very little correct pronunciation, grammar, and vocabulary; single words instead of complete thoughts; slightly comprehensible; intermediate fluency	6-13
Numerous errors in pronunciation, grammar, and vocabulary; heavy reliance on simple sentence structures, with almost no use of idiomatic	14-22

expressions; significantly comprehensible; good fluency	
Some Errors (mostly minor) in pronunciation and grammar; some variety in sentence structure; generally appropriate and varied vocabulary including some use of idiomatic expressions; mostly comprehensible; advanced fluency	23-31
Very few or no errors in pronunciation and grammar; use complex sentence structures; precise and varied vocabulary including idiomatic expressions; fully comprehensible; native-like fluency	32-40

Rubrics for pronunciation

Description	**Score**
Insufficient information to score or incomprehensible	0-5
Serious problems with pronunciation of individual sounds, speech rate, intonation, and rhythm	6-11
Numerous problems with pronunciation of individual sounds, speech rate, intonation, and rhythm	12-17
Some problems with pronunciation of individual sounds, speech rate, intonation, and rhythm, but do not cause serious problems in intelligibility	18-23
Accurate pronunciation of individual sounds; no or very few minor phonemic/phonetic errors; natural speech rate, intonation, and rhythm	24-30

Rubrics for grammar

Description	**Score**
Insufficient information to score	0-5
Severely limited in the range and control of grammar; used isolated words or short utterances to communicate ideas	6-11
Limited in the range and control of grammar; used some complex structures but typically contain errors	12-17
Fairly automatic and effective use of grammar; some imprecise or inaccurate use of grammatical structures; somewhat limited in the range of structures	18-23
Good control of basic and complex grammatical structures that allow for coherent, efficient (automatic) expression of relevant ideas	24-30

Rubrics for vocabulary

Description	**Score**
Very limited ability in using vocabulary	0-5
Limited ability in using vocabulary; used repeated vocabulary in every or almost every sentence; used incorrect or inappropriate vocabulary	6-11
Moderate ability in using vocabulary; no variety of vocabulary used; often used repeated vocabulary; often used incorrect or inappropriate vocabulary	12-17
Used variety of vocabulary; sometimes used repeated vocabulary; sometimes used incorrect or inappropriate vocabulary	18-23
Used variety of vocabulary; used correct and appropriate vocabulary most of the time; very few or no mistakes	24-30

Rubrics for fluency

Description	**Score**
Intrusive	0-5
Intermediate	6-11
Good	12-17
Advanced	18-23
Native-like	24-30

Rubrics for comprehensibility

Description	**Score**
Incomprehensible	0-5
Slightly comprehensible	6-11
Significantly comprehensible	12-17
Mostly comprehensible	18-23
fully comprehensible	24-30

Example:

Student number	Overall proficiency score of Part 4	Vocabulary	Grammar	Pronunciation	Fluency	Comprehensibility
20130512	35	23	26	28	27	29

APPENDIX Z

Sample lesson plan (two academic periods/90 minutes)

Lesson: Unit 1 Welcome back to school (Part A)

Learning objective: Students learn to introduce themselves and their friends.

Key sentences/expressions:

Good morning.

Hi, I'm…(your name)

I'm from … (UK, Canada, China and USA).

Nice to meet you.

Where are you from?

This is my friend…(name of the friend)

We have a new friend today.

We are friends.

We like to play.

Welcome!

Materials: Filtered and unfiltered sentences in Part A, a video clip (Part A)

Procedure:

8:00–8:10 →Relaxation phase. The teacher explains to the students the pedagogical procedures they will undertake and the filtered materials they will encounter. Students

engage in about 10 minutes of mind-calming activities to nurture a sensation of well-being conducive to learning. The classroom is darkened, and Baroque music is played. With music in the background, students are free to do any activity which can relax them physically and/or spiritually such as resting on the desk or leaning against the chair with eyes closed.

8:10–8:20 → The teacher plays the recordings of the filtered sentences of Part A for the students. Students listen to the filtered sentences. The filtered materials are played 15 times consecutively. The filtered materials are played using good quality amplifiers and loudspeakers with a good bass response and the volume is as loud as possible in order to enhance students' perceptions of the language input through the body's natural sensitivity to low frequencies.

8:20–8:30 →Students continue to listen to the filtered sentences for another 10 times. In this phase, students repeat the filtered patterns by "humming" them in chorus. At the same time, they are required to use body movements and gestures to express their personal perceptions of the rhythm of the filtered sentences. In order for the students to understand the requirements well, the teacher can present her version of body movements and gestures to them in a non-prescriptive way. That is, the teacher can not impose the model on the students and does not expect them to follow it. That is made clear to everybody. Students may draw the shape of the intonation curve in the air, beat out or clap out the rhythm of the sentences, come down on every syllable or at the end

of every rhythm group with feet, stretch arms at each rise in pitch, or even dance to the rhythm.

8:30–8:45→Following the humming and movement exercises, students are encouraged to guess the meaning of the prosodic patterns and to discriminate between them on the basis of what they have previously studied. For example, the teacher plays the filtered sentences (where are you from? Nice to meet you! Welcome back to school! Boys and girls, we have two new friends today) and then let the students guess the sentence type and the original unfiltered sentences. In addition, the key words specified by the Curriculum Standards for Grade 3 students in the second semester are written down on the blackboard by the teacher: **the UK, Canada, China, the USA, friend, student, she, boy, girl and new**. The students are then guided to guess the stress patterns of the keywords. Students are told to refer to the ***Word List*** part to check the meaning of these words in Chinese.

8:45–8:55→ break time

8:55–9:05 →Mouthing the words. The unfiltered sentences are played 10 times continuously, and the students are required to only mouth the words without producing any sound. Then, the teacher leads the students to discover whether their previous "guesses" were correct or not. The teacher raises students' awareness of the different intonation patterns between a statement and an interrogative.

9:05–9:15→Articulation of words to the intonation patterns. The students hum the

intonation patterns first, and then begin to add words to the “tune” produced by themselves. Unfiltered sentences continue to be played 10 times and students repeat the sentences in chorus.

9:15-9:20→The video in which all the original sentences are acted out in a cartoon is played to the students 3 times. After watching the video, students refer to the ***Word List*** part to check for the meaning of these sentences in Chinese.

9:20-9:40→Situational dialogue. Students are divided into groups of 3 or 4. First, each group member chooses to play a role or more in the speech dialogue that they have watched in the video. Then, all groups are encouraged to recreate and make fun of the original dialogue, in other words, to create a parody. The teacher comes to each group to offer help when need arises. Finally, the teacher comments upon students’ performances.

Homework: 1. Listen to and hum the filtered sentences (filtered sentences from the textbook and/or other materials students have access to) (around10 minutes).

2. Listen to and repeat the unfiltered sentences (the same as the above) (around10 minutes).

3. Watch the video clip and role-play the situation with/without the help from their parents (around 10 minutes).

CURRICULUM VITAE

Yan Yang was born in Sichuan Province, China in August 1981. She started to teach English to Chinese university EFL learners after she obtained her master's degree in linguistics and applied linguistics in Southwest Jiaotong University, Chengdu City, Sichuan Province, China in 2008. She is currently working in the Faculty of Foreign Languages and cultures of Kunming University of Science and Technology, Kunming City, Yunnan Province, China. She was promoted as an associate professor in October 2016. From 2014 to 2017, she pursued her Ph.D. in English language studies in the School of Foreign Languages, Institute of Social Science Technology, Suranaree University of Technology, Thailand. Her research interests mainly revolve around pragmatics and language education.