Available online at http://rjas.rsu.ac.th

Rangsit Journal of Arts and Sciences, July-December 2017 RJAS Vol. 7 No. 2, pp. 141-156

ISSN 2229-063X (Print)/ISSN 2392-554X (Online)

Improving the English-speaking skills of Chinese primary EFL learners with a verbotonal approach

Yan Yang[1*], Anchalee Wannaruk[2], and Andrew-Peter Lian[3]

School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

*Corresponding author; E-mail: yangyanhch@163.com

Submitted 30 December 2016; accepted in final form 29 March 2017
Available online 22 December 2017

Abstract

Among the four language skills, speaking was identified as the poorest skill for Chinese primary EFL learners. Although many teaching approaches from the mainstream have been adopted to redress the problem, teaching and learning outcomes for speaking are still disappointing. 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. To bridge the gap, the present study devised a verbotonal-based approach for improving the speaking skills of Chinese primary EFL learners. Eighty Grade 3 students from a rural primary school in China participated in the study. The experimental group undertook the intervention while the control group followed the traditional way of learning to speak English. Time on task was the same for both groups. After a period of 18 weeks, the experimental group improved significantly in both overall speaking proficiency and individual tests: word-reading, sentence-reading, singing and oral interview, as well as five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility. Additionally, the experimental group outperformed the control group in all aspects tested. Finally, the experimental group topped the district examinations whereas the control group obtained an average score. These findings suggest that the verbotonal approach was of considerable benefit to the learning of English speaking skills in the participating group.

Keywords: *Chinese, EFL learners, English speaking skills, improving, primary, verbotonal approah*

1. Introduction

Stepping into the new era, China is beginning to enjoy substantial economic growth, and simultaneously, faces the challenges of internationalization and globalization. To meet the demands resulting from social development, the Chinese Ministry of Education (hereinafter referred to as MOE) issued a directive in January 2001 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 (M. Li, 2016; MOE, 2001). Fourteen years have elapsed since the implementation of the universal provision of English in primary schools. However, the outcomes of primary English education appear to be unsatisfactory and discouraging. Studies indicate that among the four macro-skills (listening, speaking, reading and writing), speaking is identified as the most difficult for Chinese primary school students (S. S. Li, 2016; Zhang, 2011). The major challenges hindering English-speaking instruction in primary schools in China are: a significant shortage of qualified teachers and a severe lack of suitable learning approaches and materials (Bing, 2016; Hu, 2005; Teng, 2010; Wang, 2007; Zhao, Joshi, Dixon, & Huang, 2016). In addition, the overly large class size, inadequate facilities for instruction and insufficient coordination between primary and secondary English courses also pose serious challenges to the teaching and learning of speaking skills in Chinese primary schools (Hu, 2005; D. Zhang, 2012). In order to address these significant problems, researchers, educators and teachers in English language education in China are beginning to invest more effort in exploring progressive teaching approaches (Hu, 2013; D. Zhang, 2012; Zhang, 2016). Nevertheless, the fact remains that the problems in the teaching and learning of English-speaking skills in primary schools have not been resolved successfully despite the adoption of

the latest approaches from the mainstream. Clearly, there is an urgent need for more developed learning approaches.

To develop a suitable learning approach to improve students' English-speaking skills, it seems valuable to engage in a fundamental rethink of what speaking skills are and the nature of learning to speak a foreign language. Although it is difficult to arrive at a precise definition of speaking skills, there is general agreement that speaking skills include learners' accurate and effective use of linguistic and sociolinguistic features to achieve a wide range of communicative functions. 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 (De Jong & Van Ginkel, 1992). The learning of speaking skills, like the learning of other language skills or learning in general, is essentially a process of personal meaning-making (Lian, 2004, 2011). When learners begin to learn a foreign language, they are liable to be insensitive and "deaf" to the new language signals because their perceptual and productive systems are attuned to perceive and produce their native language or other languages that they know well. So, the first and foremost step in learning is to raise learners' awareness and change their perceptions of the foreign language. The verbotonal approach was originally a language therapy whose principles and practices are conducive to the learning of a foreign language. Thus, the objective of the study was to develop a verbotonal-based approach (VTA) in an attempt to improve the English speaking skills of Chinese primary EFL learners. It was hypothesized that, after treatment, VTA would increase participants' speaking performances as defined above and that there would be a statistically significant difference between the experimental group and the control group that engaged with a traditional learning program for developing speaking skills. Two research questions were derived from these hypotheses:

1) Is the VTA effective for learning to speak English? If yes, in what ways?
2) Is there a significant difference in speaking skills improvement between the experimental and control groups? If so, in what ways?

2. Literature review

2.1 General introduction to Verbotonalism

The verbotonal approach was conceived in 1939 by Petar Guberina, a linguist specializing in "the linguistics of speech" (Asp, Kline, & Koike, 2012). It originally applied auditory-based strategies to treat hard of hearing children and adults by maximizing their listening skills, and at the same time, developing intelligible speech through binaural listening (Asp, 2006). This approach holds that hearing impairment or even deafness is not so much a physical deficit as a different way of organizing audition (Renard & van Vlasselaer, 1976). So, the central preoccupation of the approach is to create an optimal field of hearing in order to restructure, i.e., reorganize, the way participants perceive language signals.

Since its advent, the verbotonal approach has been studied and applied mainly in the following areas: rehabilitation of children and adults with hearing impairment, 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 noted here 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 article. 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 system has been thought of by some as an effective approach to phonetic correction. To illustrate, as early as 1980, Andrew-Peter Lian, in his book *Intonation Patterns of French (Teacher's Book)*, introduced Verbotonalism and adopted it in the teaching of French pronunciation. He designed sensitization and reinforcement sessions to attempt to create optimal conditions to reinforce students' perceptual and articulatory abilities. Thirty-one years later, still in the teaching of French pronunciation, Alazard, Astesano, and Billières (2011) compared two approaches of phonetic

correction: the traditional articulatory approach, and the verbotonal approach. Their study provides good evidence of the benefits of Verbotonalism in the teaching of second language pronunciation over the articulatory approach.

Besides, Hu and Uno (2005) developed a teaching method based on Verbotonalism to teach Japanese beginners learning the tones of Chinese. Thirty-five university students in a basic Chinese class were selected as subjects. After 7 weeks of training, the students were given a test, including the task of dictating and pronouncing the tones of monosyllabic and two-syllable words. Results showed that this approach was useful for distinguishing and learning to pronounce 4 different kinds of tones, and students' performances on almost all tasks had improved substantially. Hang (2012) adopted the verbotonal system in the teaching of Japanese to Chinese learners. Her study also confirmed the usefulness of the approach. In the Chinese EFL context, He (2015) conducted a teaching experiment to study the effectiveness of the verbotonal system in the teaching of pronunciation to Chinese university EFL learners. Her findings again suggest that the verbotonal program was highly effective in the teaching of pronunciation of a foreign language.

As Alazard and colleagues (2011) 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 relatively long and respected history within the limits of the research programs undertaken.

2.2 Studies on the teaching of English speaking skills in primary schools in 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 (Chen, 2008; Dong, 2003; Hu 2007; Zhang, 2014). Although there are a large number of studies on primary English instruction, only a few focus deliberately and exclusively on the teaching of speaking skills.

Browsing the existing literature, it can be concluded that research on the teaching and learning of English-speaking skills in Chinese primary schools concerns itself mainly with the identification of the challenges and problems faced by teachers and learners, and the exploration of teaching methods for improving students' speaking skills.

In terms of challenges and problems, studies show that the speaking proficiency of students is generally poor. There are mainly three problems existing in the teaching and learning of speaking skills (Hu, 2007; D. Zhang, 2012). The first is the severe scarcity of qualified teachers. The second is the adoption of inappropriate teaching methods by the teachers. Some teachers, especially those in rural primary schools are inclined to accord speaking only secondary priority in teaching and continue to adopt very traditional pedagogy which attaches importance to the instruction of grammar and translation (Liu, 2006). The third problem lies in the large class size and lack of environment for students to practice speaking.

In respect of studies on the development of appropriate methods for teaching speaking skills, since the introduction of the communication-oriented teaching approach from the west, a growing number of linguists, educators and language teachers have paid attention to the search for teaching methods suitable for Chinese primary school students. Some researchers have suggested that teachers should abandon the traditional methods which focus only on delivery of linguistic knowledge, and adopt situational, communicative and task-based pedagogies (Tian, 2014; Zhang, 2014). Others realize the significance of resorting to technology to teach speaking skills. For example, Liu (2013) advocated that teachers should take advantage of multimedia resources to make their teaching more attractive, interesting and exciting. Bai, Zhou, and Zhang (2013) proposed the application of an "electronic schoolbag" in teaching. Liang and He (2011) conducted an experiment to examine the efficacy of the use of multimedia resources in improving students' listening and speaking skills.

Reviewing what has been done so far in the research on speaking skills instruction in Chinese primary schools, it can be concluded that there is a severe shortage of empirical studies aiming to improve students' speaking skills. More importantly, the verbotonal approach has never been applied to the teaching of EFL speaking in primary schools in China. Thus, the present study devised an approach basaed on Verbotonalism, to

attempt to solve, at least partly, problems in the learning/teaching of English speaking for young children.

3. **Methods and materials**

The present study was conducted using a quasi-experimental design. This section describes the methods and materials of the research which include participants, learning materials, pedagogic procedures, instruments, and data collection and analysis.

3.1 Participants

The participants were sampled from the Yiliang Anjiaqiao Primary School (YAPS). YAPS is a primary school and is located in Anjiaqiao Village, Gucheng Town, Yiliang County, Kunming City, Yunnan Province, China. As of September, 1st, 2015, there were altogether 10 classes in the school from Grade 1 to Grade 6, the full-time students amounting to 400. As mandated by the directive issued by the MOE in 2001, English classes are provided to students from Grade 3 and above. The sample of the present study was drawn from Grade 3 students who had learned English for only one semester. Eighty third-grade students with an average age of 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, and received the VTA intervention. All 40 students from Class 201302 constituted the control group, and followed the traditional way of learning speaking skills. 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. And the 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. And classes for both groups were mostly scheduled in the morning.

3.2 Learning materials

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.

3.2.1 Filtered sentences

In the present study, the phrase "filtered sentences" refers to sentences that were digitally filtered through a low-pass filter using *Audacity (Version 2.1.0),* a sound-editing software. The cutoff frequency was set at 320 Hz. A low pass filter removes frequencies above its cutoff frequency. It allows the low frequencies to pass, hence its name. What is left behind sounds a little like humming. Such a procedure removes vowel and consonant sounds and makes the intonation and rhythm patterns salient. According to Asp et al. (2012, p. 323), auditory training with low-pass filters can facilitate listeners' acquisition of rhythm and intonation patterns, which are "the foundation of both listening and spoken language".

3.3 Pedagogic procedures

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

3.3.1 In-class sensitization session

The chief aim of the session was to change students' perceptions of and cultivate their sensitivity to the language. The 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. The classroom was darkened, music which could help students 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. It was observed that mind-calming exercises adopted by the students on the basis of their own choices involved Yoga moves, Yoga

breathing, meditation, resting on the desk, leaning against the chair with eyes closed, deep breathing, free dance or even walking slowly in the classroom.

Step 2: The students listened to the filtered sentences. 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 (Lian, 1980). The filtered materials in this 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 signals through the body's natural sensitivity to low frequencies.

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. 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.

Step 4: Following the humming and movement exercises, students were encouraged to guess the meanings of the prosodic patterns (e.g. is it a statement? is it a question?) and to discriminate between them on the basis of what they had previously studied. In addition, the keywords 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.

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, and the students were required only to mouth the words without producing any sound. Besides, 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. By this stage, after hearing both the filtered and unfiltered sentences for many times, the intonation patterns had been internalized. It was therefore the right time for the students to articulate words. Unfiltered sentences continued to be played 10 times and students repeated the original 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 settings.

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 create a parody of the original dialogue. 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. This 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.

3.3.2 Out-of-class reinforcement session

Both the control and the experimental groups were required 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. 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 from Anjiaqiao Village. They received training from the researcher to filter materials chosen by their children. 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. In this study, parents were requested to cooperate with the students to role-play the situational dialogue by acting out some of the 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. And their children played all the roles themselves.

3.3.3 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. In the pedagogic procedures of the experimental group, the teacher acted only as facilitator, supporter and guide. More specifically, the teacher did not teach but oversaw the learning environment, monitored students' performances and provided help when needed. Thus, students in the experimental group were empowered to self-adjust their 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.

3.4 Instruments

As in most other primary schools in China, an English-speaking test is not required in YAPS. So, a pretest and posttest for speaking skills were constructed by the researcher based on a number of sources: 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 (Teng, 2010). The tests in the present study were designed primarily to test students' speaking skills in terms of vocabulary, grammar, pronunciation, fluency and comprehensibility. Initially, four speaking tasks were formulated in both the pretest and posttest: Word- reading, Sentence-reading, Singing, and Role-play. The two tests were pre-piloted among 8 students (2 groups) in YAPS. Revisions were made as a result of inappropriate difficulty levels of the role-play task. Both the students and the teacher commented that the role-play task was too difficult for the students to accomplish and was unable to elicit their speaking performances effectively. Therefore, the role-play task was replaced by the oral interview task. The finalized pretest and posttest both consisted of 4 parts respectively: Word-reading (20%), Sentence-reading (20%), Singing (20%) and Oral interview (40%). The content validity of the pretest and posttest was checked by 5 experts who specialized in English education using the Item Objective Congruence (IOC) approach. The results of the IOC analysis indicated that the two tests were highly valid (pretest: IOC= 38.6, percentage=96.5%; posttest: IOC=38.4, percentage=96.0%). Then, the pretest was tried out among 40 Grade 3 students twice within a time interval of 2 weeks, and the posttest was trialed twice among 40 Grade 4 students in YAPS, to check the test-retest reliability. The Pearson correlation coefficients of the test and retest of the pretest (r= 0.84, p=0.00) and posttest (r= 0.81, p=0.00) indicated that the two tests were highly reliable. Students' speaking performances were audio recorded.

3.4.1 Raters

Three Chinese expert raters were invited to rate students' speaking skills in both the pretest and posttest. Rating was blind. One rater 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 developed rubrics for the speaking tests and familiarized the raters with the objectives of the assessment, the rating procedures, and the assessment criteria. Results of the inter-rater reliability analysis are presented in Table 1. It can be seen that all correlations are statistically significant, serving as solid evidence of the high reliability of the assessment scores used in the study.

Table 1 Correlation matrix on inter-rater reliability

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

*****p=<0.05

3.5 Data collection and data analysis

To check the speaking proficiency level of the experimental and control groups before the experiment and to compare the performances between them, the two intact groups received a pretest for speaking skills. After that, VTA was implemented in the experimental group for a period of 18 weeks while the control group underwent the normal teaching program. After the experimentation, a posttest for speaking skills was administered to both the control and experimental groups to determine if there had been any improvements in the performances of the two groups, and at the same time, to check whether there was a significant difference between the two groups after the intervention. 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. Moreover, the confidentiality and anonymity of the participants were protected.

The data collected were then subject to statistical analysis. Descriptive statistics were first applied using SPSS (Version 19) to get basic information about the data. Then, inferential statistics including a paired samples t-test and an independent-samples t-test were utilized to conduct subsequent analyses.

4. Results

4.1 Total score

Table 2 displays the descriptive statistics for the overall speaking performances and individual tests of the two groups of students in the pretest and posttest. An independent samples t-test was applied to check whether there was any difference between the speaking performances of the two groups of students **in the pretest**. Results showed that there was no significant difference between the experimental and control groups. Both were at the same level as judged by the overall scores of the four individual tests mentioned above ($t=0.15$, $p=0.88>0.05$).

A paired samples t-test was employed to compare the means of the total scores in the pretest and posttest of 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 ($t=-13.43$, $p=0.00<0.05$) but 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 performed to check whether there was any significant difference between the overall speaking performances of the two groups of students **in the posttest**. Results indicated that there was a statistically significant difference between the performances of the experimental and control groups ($t=6.49$, $p=0.00<0.05$). Specifically, the experimental group outperformed the control group in the posttest.

4.2 Word-reading

The results of the independent samples t-test showed that, **in the pretest**, there was a statistically significant difference between the two groups ($t=-2.69$, $p=0.01<0.05$). Specifically, the control group was stronger than the experimental group in reading the 40 words.

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 ($t=-5.31$, $p=0.00<0.05$) groups. Specifically, both groups progressed significantly in word-reading after their respective pedagogical interventions. An independent samples t-test was applied to see whether the two groups of students' performances in reading these words were different or not **in the posttest**. There was a statistically significant difference between the performances of the experimental and control groups ($t=2.74$, $p=0.01<0.05$). Specifically, the experimental group caught up the difference and went ahead of the control group in word-reading after the experiment.

Table 2 Descriptive statistics for overall speaking skills and individual tests

Item	Group	Test	Number	Mean	S.D.
Total score	Experimental group	Pretest	40	54.01	10.72
		Posttest	40	70.23	13.27
	Control Group	Pretest	40	53.81	8.14
		Posttest	40	54.09	8.46
Word-reading	Experimental group	Pretest	40	13.64	2.98
		Posttest	40	17.25	2.72
	Control Group	Pretest	40	15.19	2.07
		Posttest	40	15.76	2.11
Sentence-reading	Experimental group	Pretest	40	12.50	3.30
		Posttest	40	16.25	2.84
	Control Group	Pretest	40	12.70	2.40
		Posttest	40	12.85	2.08
Singing	Experimental group	Pretest	40	11.70	3.78
		Posttest	40	13.60	4.14
	Control Group	Pretest	40	10.95	1.63
		Posttest	40	10.58	2.08
Oral interview	Experimental group	Pretest	40	16.40	4.59
		Posttest	40	22.10	5.19
	Control Group	Pretest	40	15.05	3.43
		Posttest	40	15.58	3.58

4.3 Sentence-reading

The results of the independent samples t-test indicated that, **in the pretest**, there was no statistically significant difference between the two groups of students ($t=-0.58$, $p=0.57>0.05$). In other words, the ability of the two groups of students to read the 10 sentences were at almost the same level before the pedagogical intervention.

The results of the paired samples t-test indicated that there was a statistically significant difference **between the pretest and posttest** for the experimental group ($t=-8.51$, $p=0.00<0.05$) but no significant difference for the control group ($t=-1.34$, $p=0.19>0.05$). More specifically, the experimental group advanced considerably after the treatment whereas the control group witnessed no significant progress. 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 ($t=8.04$, $p=0.00<0.05$) **in the posttest**. That is, the experimental group outperformed the control group in reading the 10 sentences after the experiment.

4.4 Singing

The results of the independent samples t-test demonstrated that, **in the pretest**, there was no statistically significant difference between the two groups of students ($t=1.15$, $p=0.25>0.05$). The performances of the experimental and control groups in singing the English song were similar in the pretest.

The results of the paired samples t-test demonstrated that there was a statistically significant difference **between the pretest and posttest** for the experimental group ($t=-3.26$, $p=0.00<0.05$) but no significant difference for the control group ($t=-0.86$, $p=0.39>0.05$). This means that the experimental group improved significantly 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 showed that there was a statistically significant difference between the performances of the experimental and control groups ($t=3.89$, $p=0.00<0.05$). To be more specific, the experimental group outperformed the control group in singing **in the posttest.**

4.5 Oral interview

The oral interview component was meant to assess students' speaking skills in terms of pronunciation, grammar, vocabulary, fluency and comprehensibility. So, the three Chinese expert raters were required to give two sets of scores. The first was the overall proficiency score, and the second were the scores for each of the five subskills: vocabulary, grammar, pronunciation, fluency and comprehensibility. The overall maximum score for this part aggregated to 40, and the score for each subskill was 30. The statistical analyses of the oral interview are provided in the following sequence: overall proficiency score (See Table 2), and then scores for vocabulary, grammar,

pronunciation, fluency and comprehensibility (see Table 3).

4.5.1 Overall proficiency score

The results of the independent samples t-test demonstrated that, **in the pretest**, there was no statistically significant difference between the experimental and control groups (t=1.49, p=0.14>0.05).

The results of the paired samples t-test indicated that there was a statistically significant difference **between the pretest and posttest** for the experimental group (t=-8.62, p=0.00<0.05) but no significant difference was found in the control group (t=-0.89, p=0.33>0.05). This means, the experimental group improved significantly after the treatment whereas the control group made no significant progress. An independent samples t-test was performed to see whether the two groups of students' performances in the oral interview were different or not after the treatment. It was found 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 was far ahead of the control group in the oral interview part **in the posttest.**

Table 3 Descriptive statistics for subskills in the oral interview

Item	Group	Test	Number	Mean	S.D.
Vocabulary	Experimental	Pretest	40	13.32	4.24
	group	Posttest	40	16.25	3.60
	Control Group	Pretest	40	12.12	3.92
		Posttest	40	12.42	3.85
Grammar	Experimental	Pretest	40	11.57	4.40
	group	Posttest	40	15.55	4.09
	Control Group	Pretest	40	11.60	4.10
		Posttest	40	11.77	3.51
Pronunciation	Experimental	Pretest	40	17.70	3.29
	group	Posttest	40	21.77	3.12
	Control Group	Pretest	40	18.00	2.56
		Posttest	40	18.15	2.44
Fluency	Experimental	Pretest	40	11.60	4.41
	group	Posttest	40	17.62	4.28
	Control Group	Pretest	40	12.70	4.30
		Posttest	40	12.20	2.08
Comprehensibility	Experimental	Pretest	40	18.35	3.63
	group	Posttest	40	23.07	3.22
	Control Group	Pretest	40	19.12	2.78
		Posttest	40	19.15	2.61

4.5.2 Vocabulary

Table 3 presents the statistics for all five subskills. The results of the independent samples t-test demonstrated that, **in the pretest**, there was no statistically significant difference between the two groups (t=1.31, p=0.19>0.05). That is, the lexical richness of the two groups of students started at the same level before the treatment. 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 progressed significantly in terms of lexical richness after the treatment. The results of the independent samples t-test indicated that there was a statistically significant difference between the groups (t=4.58, p=0.00<0.05) **in the posttest**. To be more specific, the experimental group was better than the control group in terms of lexical richness after the treatment.

4.5.3 Grammar

An independent samples t-test analysis showed that, **in the pretest**, 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 grammatical accuracy before the treatment.

The results of the paired samples t-test showed that there was no 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 substantially after the treatment whereas the performance of the control group remained almost the same. The results of the independent samples t-test indicated that there was a statistically significant difference between the two groups (t=4.42, p=0.00<0.05) **in the posttest**. That is, the experimental group significantly outperformed the control group in terms of grammar after completing the verbotonal program.

4.5.4 Pronunciation

An independent samples t-test analysis demonstrated that, **in the pretest**, there was no statistically significant difference between the two groups (t=-0.46, p=0.65>0.05). More specifically, the two groups of students performed the same in terms of pronunciation before the treatment.

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 improved considerably after the treatment whereas the control group's progress in pronunciation was not significant. The results of the independent samples t-test demonstrated that there was a statistically significant difference between the experimental and control groups (t=5.78, p=0.00<0.05) **in the posttest**. That is, the experimental group performed better than the control group in terms of pronunciation after the experiment.

4.5.5 Fluency

An independent samples t-test analysis indicated that, **in the pretest**, there was no statistically significant difference between the two groups (t=-1.13, p=0.26>0.05). Specifically, the two groups of students started at the same level in terms of fluency before the treatment.

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). It means that the experimental group made significant progress after the treatment whereas the control group remained at almost the same level (mean in pretest=12.70, mean in 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) **in the posttest.** That is, the experimental group (mean=17.62) significantly outperformed the control group (mean=12.20) in terms of fluency after the treatment.

4.5.6 Comprehensibility

The results of the independent samples t-test demonstrated that, **in the pretest**, there was no statistically significant difference between the two groups (t=-1.07, p=0.29>0.05). That is, the two groups of students performed at the same level for comprehensibility before the treatment.

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 advanced considerably in terms of comprehensibility after the treatment whereas the control group's performance did not change. The results of the independent samples t-test indicated that there was a statistically significant difference between the experimental and control groups (t=6.13, p=0.00<0.05) **in the posttest**. Specifically, the experimental group significantly outperformed the control group in terms of comprehensibility after the treatment.

5. Discussion

This section discusses the aforementioned findings on students' speaking performances before and after the intervention. In terms of the overall assessment of speaking skills, as reported above, no significant difference was found between the two groups of participants in the pretest. That is, the speaking skills of both groups were much alike before the pedagogical intervention. This finding was expected because the two groups had similar English learning histories before the experiment. However, in the posttest, 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 was found in the control group (p=0.86) whose mean increased by only 0.28 (from 53.81 to 54.09). It can be concluded that the experimental group improved greatly in overall speaking proficiency

while the control group showed no significant progress.

With regard to the individual tests, the experimental group progressed significantly in word-reading, sentence-reading, singing, and oral interview. The control group also improved significantly in word-reading, but made no significant progress in other components. It should be noticed that, in the pretest, the control group had performed significantly better than the experimental group in word-reading. However, in the posttest, the experimental group made up the difference and overtook the control group by a large margin. As for the five subskills tested, in the pretest, both the experimental and control groups started at the same level in terms of vocabulary, grammar, pronunciation, fluency and comprehensibility. However, in the posttest, the experimental group advanced significantly in all the subskills, while the control group was found to make significant progress only in vocabulary. Besides, the experimental group outperformed the control group in all of the aspects tested: overall proficiency, word-reading, sentence-reading, singing, oral interview, vocabulary, grammar, pronunciation, fluency and comprehensibility.

In the traditional instruction settings, most of the class time was occupied teaching, specifically, segmental pronunciation (individual language sounds), spelling, grammar and translation skills. In contrast, the students in the experimental group were exposed to VTA, and never received any of the kind of teaching given to the control group except the instructions for doing the various tasks required of them and occasional provision of help when requested by the students. So, in this sense, the extent of the experimental group's increased performance in every area tested was both unexpected and intriguing. 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 the district, the experimental group came out ahead of all participating groups. It 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 translation. As stated above, the control group received intentional teaching for these skills 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. Thus, the verbotonal approach yielded fruitful results. 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 (2015). The effectiveness of the approach is explored from the following perspectives.

To begin with, 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 information 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 (A.P. Lian & B. Lian, 1997; F. 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, Anwander, Bestelmeyer, & Belin, 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).

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 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, Santana, & Fowler, 2003; Richardson, Dale, & Shockley, 2008). So, VTA integrated body movements into the learning of speaking skills. These embodied representations of speech engendered the enactment effect which made learning outcomes positive and productive (Macedonia & Von Kriegstein, 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 expand learning potential (Ryugo, Limb, & Redd, 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, Honbolygó, Róna, Beke, & Csépe, 2014). Besides, 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' brain 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 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 experienced teacher. All that it required was exposure to the materials in the order they

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 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 speaking and other language skills.

6. Conclusion

The present study performed a quasi-experiment wherein a verbotonal-based approach was developed and implemented to enhance the English-speaking skills of Chinese primary EFL learners. 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 were compared after 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. And, literally, no student in the experimental group regressed unlike the situation in the control group. The control group was given specific teaching on how to produce/pronounce segmental elements, and apply grammatical rules and translation skills whereas the experimental group never received any specific teaching for any of these items. These counter-intuitive findings seem to support the effectiveness of the verbotonal approach in the learning of foreign language skills but also point to the need for further investigations.

Pedagogical implications can be drawn from the findings of the present study. First, language teachers 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. Then, 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, isolated, way. Last but not least, 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. Thus, the study has potential implications not only for the practice of language learning and teaching but also for language-learning policy at all levels, for language-learning and teaching theory and also for teacher education programs which may need to revise some of their assumptions.

Unavoidably, there remain limitations in the current research. To start with, the present study focused on Grade 3 primary students who had learned English for only one semester, and those from other grades were excluded. Moreover, as mentioned previously, YAPS is a rural primary school situated in a village of Yunnan Province. Since there are regional disparities in the learning backgrounds and conditions of primary students from different parts of China, the ability to generalize and apply the present study's findings needs to be treated with caution. It is strongly suggested that further empirical studies be carried out on a larger population and in a wider context to assess fully the potential merits of this form of implementation of the verbotonal approach. Nevertheless, within the limits set by this and other related studies, there is room for cautious optimism.

7. References

Alazard, C., Astesano, C., & Billières, M. (2011). MULTIPHONIA: A MULTImodal database of PHONetics teaching methods in classroom Inter Actions. (pages 2578-2583). Retrieved from http://www.lrec-conf.org/proceedings/lrec2012/pdf/254_Paper.pdf

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

Asp, C. W., Kline, M. W., & Koike, K. J. (2012). Verbotonal worldwide. In R. Goldfarb (Ed), *Translational speech-language pathology and audiology* (Essay 42, pp. 319-326). Plural Publishing. ISBN-10: 1597564451; ISBN-13: 978-1597564458. Retrieved from http://verbotonal.utk.edu/Verbotonal/Resources_files/Essay%2042.pdf

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. 白若微, 周榕, 张文兰 (2013). 电子书包在小学英语教学中的应用策略研究[J].《现代教育技术》, *23(*8), 50-54.

Bing, W. (2016). The English education in primary schools in minor ethnic areas in Western China-Taking Leshan City as an example. *English Language Teaching*, *9*(8), 152-155. DOI: 10.5539/elt.v9n8p152

Chen, L. (2008). Primary school English education in China. Paper presented at the International Conference on Primary English Education, Beijing. Retrieved from https://paper.i21st.cn/story/46180.html.

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. http://dx.doi.org/10.1097/00005053-196610000-00005

De Jong, J. H., & Van Ginkel, L. W. (1992). 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*, Chapter 15, pp. 187-212.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development*, *71*(1), 44-56. DOI: 10.1111/1467-8624.00117

Dong, Y. P. (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. 董燕萍(2003). 从广东省小学英语教育现状看"外语要从小学起"的问题[J].《现代外语》, *26*(1), 39-47.

Garratt, G. M. (2015). *Perception of prosodic and linguistic material in depression*. Master's thesis, California State University, San Marcos, USA. Retrieved and download from http://csusm-dspace.calstate.edu/handle/10211.3/149656?show=full

Hang, J. (2012). Implementation of the Verbotonal method in teaching Japanese pronunciation. *Journal of Changchun Education Institute*, *28*(11), 124-125. Retrieved and get full text from http://caod.oriprobe.com/articles/31428223/yu_diao_ting_jue_fa_zai_ri_yu_sheng_diao_zhi_dao_z.htm

He, B. (2015). Improving the English pronunciation of Chinese EFL learners through the integration of CALL and Verbotonalism. Proceedings of the Seventeenth International CALL Conference, 6-8 July 2015, at the Universitat Rovira i Virgili in Tarragona., Spain, pp. 276-285. Retrieved from http://wwwa.fundacio.urv.cat/congressos/public/usr_docs/call_2015_conference_proceedings.pdf

Hu, G. W. (2005). Building a strong contingent of secondary EFL teachers in China: Problems and policies. *International Journal of Educational Reform*, *14*(4), 458-485. https://books.google.co.th/books?hl=en&lr=&id=ybl_CgAAQBAJ&oi=fnd&pg=PA454&dq=Hu,+G.+W.+(2004).+Building+a+strong+contingent+of+secondary+EFL+teachers+in+China:+Problems+and+policies&ots=sJVBQn_Bd7&sig=QwPUqCucpyBoxpSwEAY5pYnWq2A&redir_esc=y#v=onepage&q&f=false

Hu, R. (2013). Task-based language teaching: Responses from Chinese teachers of English. *The Electronic Journal for English as a Second Language (TESL-EJ)*, *16*(4). Retrieved from http://www.tesl-ej.org/wordpress/issues/volume16/ej64/ej64a1/

Hu, Y. (2007). China's foreign language policy on primary English education: What's behind it? *Language Policy*, *6*(3-4), 359-376. DOI: 10.1007/s10993-007-9052-9. Retrieved from https://link.springer.com/article/10.1007/s10993-007-9052-9

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. http://doi.org/10.5926/jjep1953.53.4_541

Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia*, *49*(5), 878-887. DOI: 10.1016/j.neuropsychologia.2011.01.008

Lian, A. P. (1980). *Intonation patterns of French (Teacher's Book)*. Sydney, Australia: River Seine Publications.

Lian, A. P. (2004). Technology-enhanced language-learning environments: A rhizomatic approach. In. Sun, J.-B. (Ed.), *Computer-assisted Language Learning: Concepts, Contexts and Practices* (Chapter 1, pp. 1-20).

Lian, A. P. (2011). Reflections on language learning in the 21st Century: the Rhizome at work. *Rangsit Journal of Arts and Sciences*, *1*(1), 3-15. DOI: 10.14456/rjas.2011.2

Lian, A. P., & Lian, A. B. (1997). The secret of the Shao-Lin Monk: Contribution to an intellectual framework for language-learning. *On-CALL*, *11*(2), 2-19. Retrieved from http://www.andrewlian.com/andrewlian/prowww/shaolin/psupres2.htm

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*, 96-102. 梁文鑫, 何克抗 (2011). 多媒体资源的语音辨析支持对于小学生英语听说学习的影响研究[J].《电化教育研究,10, 96-102.

Li, M. (2016). Power relations in the enactment of English language education policy for Chinese schools. *Discourse: Studies in the Cultural Politics of Education*, *38*(5), 713-726. https://doi.org/10.1080/01596306.2016.1141177

Li, S. S. (2016). A study on the approaches to teach English speaking skills in Chinese primary schools. *Examinations Weekly*, *16*, 96-96. 李松松 (2016). 小学英语口语教学方法与策略浅析[J]. 《考试周刊》,16 :96-96.

Liu, J. L. (2006). Factors affecting speaking teaching in primary schools. *Teaching Research and Practice*, *9*, 34-35. 刘金兰 (2006). 试论影响英语口语教学效果的相关因素[J].《教育实践与研究》, (9b):34-35.

Liu, Y. F. (2013). Multimedia and primary English teaching. *Curriculum Research*. *35*,213-214. 刘颖范(2013). 巧用多媒体优化小学英语课堂教学[J].《课程教育研究:新教师教学》, 35, 213-214.

Macedonia, M., & Von Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biolinguistics*, *6*(3-4), 393-416.

Martin, P. (2016). *The structure of spoken language*. Cambridge, UK: Cambridge University Press.

McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English Pronunciation* (pp. 185-196). Oxford, UK: Oxford University Press.

MOE. (2001). *Guidelines for promoting primary English language teaching*. Beijing, China: The Chinese Ministry of Education (MOE). http://en.moe.gov.cn/

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. In W. J. Hardcastle & J. Laver (Eds.) (Oxford, USA: Blackwell), *The Handbook of Phonetic Sciences*, *5*, 640-673.

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*(2), 816-847. DOI: 10.1016/j.neuroimage.2012.04.062

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.

Renard, R., & Van Vlasselaer, J. J. (1976). *Foreign language teaching with an integrated methodology: the SGAV (Structuro-Global Audio-Visual) methodology*. Paris, France: Didier.

Richardson, D. C., Dale, R., & Shockley, K. (2008). Synchrony and swing in conversation: coordination, temporal dynamics, and communication. *Embodied Communication in Humans and Machines*, 75-94. DOI: 10.1093/acprof:oso/9780199231751.003.0004

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. In J. K. Niparko (Ed.), *Cochlear implants: Principles and practices*, 33-56. Philadelphia, PA, USA: Lippincott Williams and Williams.

Sammler, D., Anwander, A., Bestelmeyer, P. E., & Belin, P. (2015). Dual streams for prosody in the right hemisphere. In *Proceedings of the 7th Annual Meeting of the Society for the Neurobiology of Language (SNL 2015)*, held in Chicago, Illinois at the legendary Drake Hotel, USA. 15-17 October, 2015. Retrieved from http://www.neurolang.org/programs/SNL2015_Abstracts.pdf

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), 326-332.

Teng, Y. (2010). The comparison and implications of oral teaching evaluation in primary schools between China and foreign countries. *Curriculum, Teaching Method and Material*, 30 (3), 108-112. 滕云 (2010). 中外小学英语口语教学评价比较及启示[J].《课程•教材•教法》, *30* (3), 108-112.

Tian, L. L. (2014). Collaborative learning of English in primary schools. *Journal of Small Writers Election: Teaching Experience Exchange*, *1*, 453-454. 田丽丽 (2014). 合作学习在小学英语教学中的应用[J].《小作家选刊杂志》, *1*, 453-454.

Wang, N. N. (2007). Report on a survey of English education in 12 primary schools in some villages. *Modern Education Forum*, 4, 103-104. 王宁宁 (2007). 12 所农村小学英语教育调查报告[J].《当代教育论坛:学科教育研究》, 4, 103-104.

Zhang, D. (2012). Chinese primary school English curriculum reform. In J. Ruan and C.B. Leung (Eds.), *Perspectives on Teaching and Learning English Literacy in China*, Multilingual Education 3, Chapter 5, pp. 67-83. DOI: 10.1007/978-94-007-4994-8_5

Zhang, F. (2012). Teaching Mandarin prosody: a somatically-enhanced approach for second language learners. *Taiwan Journal of Chinese As A Second Language*, *4*(1), 95-130.

Zhang, X. H. (2014). Situational pedagogy in primary English instruction. *Learning Weekly*, *15*, 62-63. 张新华(2014). 情景教学法在小学英语教学中的应用[J].《学周刊》, *15*, 62-63.

Zhang, X. Y. (2016). Strategies for improving the teaching of English speaking skills in Chinese primary schools. *China Extracurricular Education*, (S1), 440-442. 张晓宇 (2016). 小学英语口语教学改进策略研究[J].《中国校外教育》, (S1), 440-442.

Zhang, Z. Y. (2011). A reflection on basic English teaching in China. *Curriculum, Teaching Method and Material*, *31*(1), 95-99. 张志远 (2011). 中国基础英语教学:在反思中前行[J].《课程•教材•教法》, *31*(1), 95-99.

Zhao, J., Joshi, R. M., Dixon, L. Q., & Huang, L. (2016). Chinese EFL teachers' knowledge of basic language constructs and their self-perceived teaching abilities. *Annals of dyslexia*, *66*(1), 127-146. DOI: 10.1007/s11881-015-0110-2