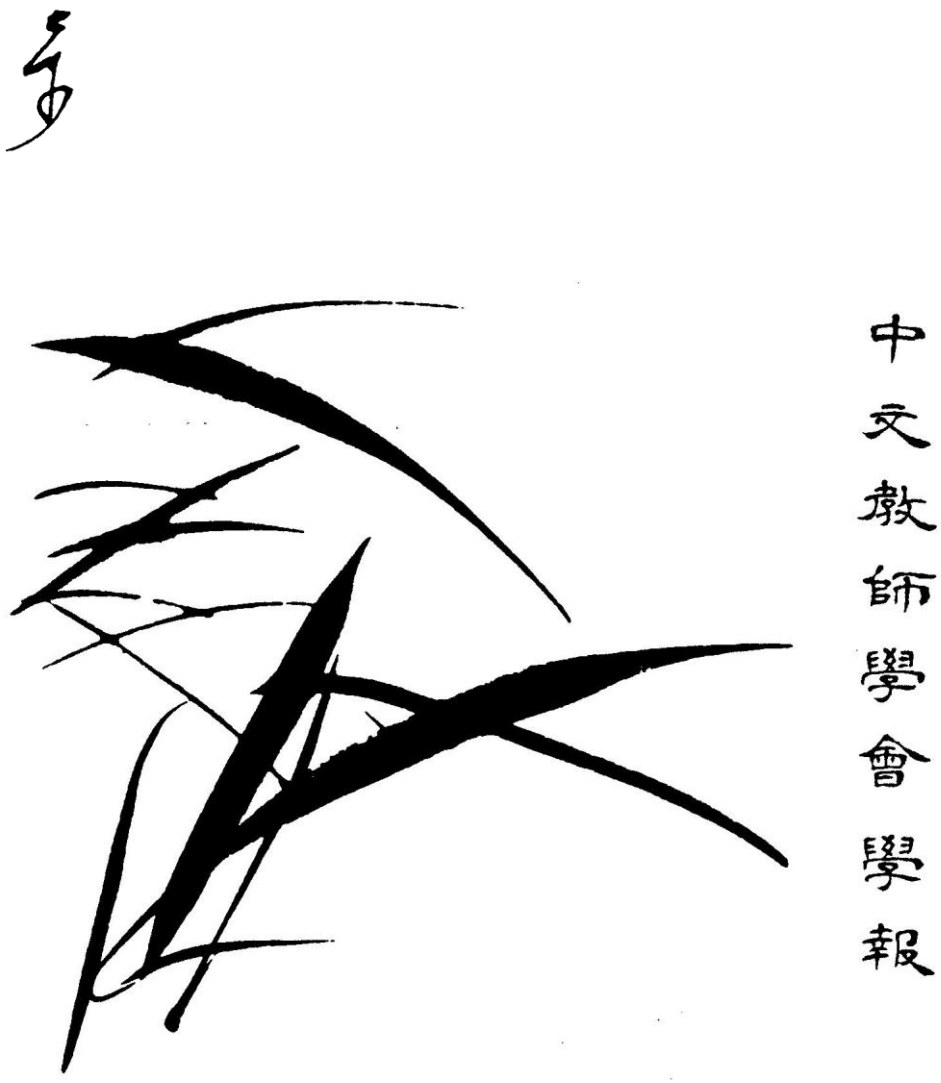
Wu, Yongan. "Using External Text Vocalization to Enhance Reading Development among Beginning Level Chinese Learners." Journal of Chinese Language Teachers Association 47.1 (2012): 1-22.

JOURNAL

of the

CHINESE LANGUAGE TEACHERS ASSOCIATION

Volume 47 : 1 February 2012



中文教師學會學報

Using External Text Vocalization to Enhance Reading Development Among Beginning Level Chinese Learners[1]

使用文本录音提高中文初学者的阅读能力

Yongan Wu
University of Northern Florida

Abstract Phonological awareness plays a key role in reading but Chinese language learners' self-initiated efforts to connect sounds with writing are oftentimes inadequate. To bridge the gap, this study examines the effect of externally supplying text vocalization on reading development in beginning level Chinese learners. The author presented the text and its vocalization to participants in reading exercises for four weeks and then measured participants' performance in four areas related to reading (when text vocalization was removed): word recognition, word recall, reading speed and reading comprehension. The results suggest that, compared with the traditional visual presentation without any aural stimuli, supplying text vocalization along with the text helped participants become better overall at retaining meanings and the general graphic features of Chinese characters; in particular, it helped struggling readers to complete reading in a quicker and better fashion. The pedagogical implications of these findings are suggested.

Key Word: Text vocalization, Reading, CFL

摘要 音系意识是阅读的重要一环，但汉语学习者在连接声音和文本方面往往力不从心。为了跨越这道障碍，本文考察外部文本录音输入对初级中文学习者阅读能力发展的影响。作者将文本录音和文本同时呈现给被试，为期四周，然后移走文本录音，测量被试在四个方面的情况：单词认读，单词记忆，阅读速度和阅读水平。结果显示，将文本录音和文本同时输入的方法，相比于传统的只有文本而无语音的方法而言，能够促进单词记忆和对汉字的一般形状的认读。此外该方法能帮助后进，让他们更快更好地完成阅读任务。相关教学建议也一并讨论。

关键词：文本录音，阅读，对外汉语

[1] I would like to express my gratitude to the editor, Zheng-sheng Zhang, and the two anonymous reviewers, for their insightful comments and valuable suggestions throughout the revision and editing process.

1. Introduction

In alphabetically written languages, phonology and orthography generally are sufficiently closely related that readers have the luxury of recalling the meaning of a word by just looking at its orthographic features as cues and then sub-vocalizing it (Lee, 2009). By comparison, however, Chinese characters' phonology has diverged from the orthography, at least to people who are less familiar with the language, and the text does not have natural boundaries (spaces between strings of letters) between characters to indicate words (Chu, 2006). Such features create two extra hurdles that many CFL (Chinese as a Foreign Language) learners have to overcome to achieve fluency in reading. First, readers at all levels and novice readers in particular, who are accustomed to using orthographic cues to produce sound and recall meaning, crave a way to put the missing phonological elements back into the reading process(*c.f.* Lee-Thompson, 2008). Second, readers have to constantly make word decisions; erroneous combinations of adjacent characters will certainly interrupt the flow of meaning (Ke, 1996; Shen, 2008). Facing the challenges created by these two factors combined, readers are oftentimes observed using phonological remediation, including writing phonetic notes (e.g., Pinyin, a system using the Roman alphabet to transliterate sounds in the Chinese language) next to the text or guessing at the sound of a character from its component(s), sounding out characters or transliterating the whole text in their heads (Everson & Ke, 1997; Lee-Thompson, 2008; Pavlidis, 1992). Remedial methods like these are not only laborious but also misleading and counterproductive due to readers' lack of a firm understanding of the relationship between the sound, meaning and writing of Chinese characters and words.

This issue can now be addressed by technology from a new perspective. In the last decade, the advancement of technology and the proliferation of online platforms enable teachers to easily place the text next to its vocalized recordings (Xie & Yao, 2009; Zheng, 2005). It has fundamentally changed the way in which the text can be presented. Unlike the vocalization produced by readers themselves as they read aloud or under their breath, the audio recordings are recorded in advance by native speakers in a loud and clear voice. The technological development thus opens up opportunities to explore how reading can be taught in this new context, i.e., if students are supplied the vocalized text while reading Chinese, what will be the effect on reading, especially in such areas as word recognition, reading speed and reading comprehension? Two studies thus far have looked at the effects of presenting sound along with the graphic and phonetic representations of characters (Chung, 2008; Jin, 2006), yet no study has directly

tackled the effect of coupling text and its vocalized sound on reading in the CFL area. This deficiency becomes more pronounced due to U.S. colleges experiencing a sharp increase in the number of CFL students (Furman, Goldberg, & Lusin, 2007) who usually speak an alphabetically written languages at home and often use phonetic notes (usually Pinyin) as their main aid to vocalize words in Chinese (Wang, 2008; Xing, 2006).

This study aims to investigate the effect of coupling text with its vocalized recording on beginning level college CFL learners' reading comprehension. Following Everson (1994), the author first surveyed existing studies to examine the role played by phonology or phonological knowledge in the reading process at the character/word level and the sentential level as well. A couple of relevant studies using Pinyin to aid learning characters were also reviewed. Two experiments were then designed to examine whether providing vocalized recordings of the text would enhance reading proficiency, with the hope that the findings provide empirical evidence on the potential benefits of incorporating mixed-modal information (text along with its audio vocalization) to expedite reading development. In the age when educational technology can easily be utilized to juxtapose text and its audio form, it is natural and important to examine the significance of this change brought on by technology.

2. Review of Related Studies

Researchers have spent great efforts to understand unique features of reading in a second language and CFL reading is no exception. Early models of second language reading hypothesized that readers, regardless of any specific language they were learning, had to go through a set of hierarchical, developmental stages to gain proficiency (Bernhardt, 1986; Everson, 2007). Word recognition and phonemic/graphemic decoding, aided by phonological prompts embedded in orthographic features, were considered lower-level skills, the "basics", to be mastered relatively early and easily (Bernhardt, 1991; Bernhardt & Kamil, 1995). For reading alphabetically written languages in which spelling and sounds are so closely intertwined that word-recognition can be assisted when readers sound out words in their mind during reading, this model works relatively well, but it quickly falls short when applied to reading Chinese, a language whose orthography and phonology are not directly related (Wu & Liu, 1996). In the absence of phonologic prompts, it is indeed a challenging and time-consuming task, as many CFL learners have found, to proficiently perform the so-called lower-level tasks, including recognizing characters, associating them with meaning, and segmenting adjacent characters into words, just to name a few. Such difficulties are cited

across proficiency levels (Everson, 1988; Ke, 1998; Shen, 2005b, 2008).

When few natural phonological prompts can be found in a text, how can CFL readers adjust their strategies to learn characters, recognize them and make word decisions? More importantly, will readers use their knowledge on Chinese phonology to identify or re-create the missing phonological prompts? Shen (2005a, 2008) examined these two questions. In the first study, Shen used an open-ended questionnaire and gathered 176 character learning strategies in total, then condensed these down to 30 after deleting duplicates and removing less-common ones. A careful look at the finalists suggests that phonologically related items can explain roughly a total 30% of the variance (9 out of 30). The factor analysis later performed by Shen also revealed the important and almost omnipresent role phonological knowledge plays in the cognitive process of learning characters, whenever participants "create mental linkages among sound, shape, and meaning of the character…use both aural-oral cues and writing in receiving and encoding information…[and] emphasize the use of sound as cues to make connections to meaning and shape within a character" (p. 59). The second study found that both beginning and advanced learners predominantly relied on the strategy of matching the target item to their existing mental lexicon as a way to parse sentences into meaningful words. In the follow-up discussion, Shen speculated that only when three linguistic values of a word (phonological, graphemic, and semantic representations) from readers' mental lexicon matched the printed words, can readers make correct word decisions. If phonologic and semantic representations are somehow disconnected or mis-connected, readers have to resort to other methods. For beginning level learners, such attempts often degenerate into aimless guessing.

At the sentential level, there are also a number of studies focusing on the role of phonology. Hayes (1988) asked participants (native Chinese speakers or proficient CFL readers) to do a sentence validity test after reading a number of sentences, each containing one of three types of distracters (semantically, phonetically or graphically oriented) to examine and measure how character recognition was achieved within a sentence. The results indicate that CFL readers are not able to quickly activate the phonological dimension of the text. They have to allocate a disproportionate amount of attention to the graphic nature of the characters, leaving only a limited amount of mental capacity for actual comprehension. Hayes' classic study convincingly illustrates how phonological inputs can alleviate the burden placed on readers' working memory, so they can read more accurately and faster, not just individual words but also sentences.

In a related study, Everson and Ke (1997) again encountered the same issue. They asked participants to silently read an authentic short newspaper passage that

was at or just above the intermediate level, perform a verbal report after reading, re-read the passage in silence and then complete a recall-protocol. After reviewing the verbal report data of the participants at both levels, Everson and Ke found that participants from both levels demonstrated a clear trend of consistently using sound remediation during their reading. They noticed that "[t]he protocols were filled with countless instances of reading aloud, reading under their breath, and other forms of sound remediation, sub-vocalization and pronunciation" (p. 11).

In Lee-Thompson's (2008) study of reading strategies developed and adopted by CFL students, eight college CFL students at the intermediate level of proficiency were included to examine readers' approaches to narrative and argumentative texts. Lee-Thompson was intrigued to observe that a large amount of Pinyin was written by participants alongside the Chinese text. This use of Pinyin, as it turned out, was used as a phonological prompt to help readers re-connect phonology and semantics. A follow-up interview with the participants faithfully documented CFL learners' desires and methods of using phonological mediation to aid reading. According to participants, "they did not feel comfortable with just recognizing and knowing the meaning of a character or a word without knowing how to pronounce it, and that knowing the pronunciation (i.e., with the aid of Pinyin) helped them when encountering difficulties in reading" (p. 711).

Since phonological prompts play such an important role in every aspect of the CFL reading process and cannot be easily extracted from the text, it is then natural and important to ask the question: if phonological prompts are supplied externally in the learning process, how will reading be affected? Practically speaking, there are two ways to do so, by using either Pinyin scripts or text vocalization. Pinyin has been used as a simple remedy but often leads to more problems that it can solve. CFL teachers have reached the consensus that supplying Pinyin along with the text will lure students to virtually overlook characters (Chung, 2003). Moreover, it has the side-effect of permanently trapping CFL learners in the difficult situation that they always have to painstakingly match Pinyin spelling with the writing of the character, not to mention they also have to remember its meaning in their first language. Such effort considerably overloads learners' cognitive capacity and working memory (Xing, 2006). Two studies, Jin (2006) and Chung (2007), have shown that presenting Pinyin along with characters does little good in helping learn characters. Jin's study demonstrated that Pinyin presentation (plus a sound playback system) was the least effective in aiding character recognition, compared to radical presentation and stroke presentation. Chung advised teachers to present characters by following a mixed visual-audio method (aurally presenting the sound of the characters through headphones, besides Pinyin script and English at the same time), rather than presenting Pinyin script and

English alone without the sound, in order to better enhance beginning CFL learners' ability to recognize characters and recall their associated meaning.

The other reason to discourage students from using Pinyin in their reading is that Pinyin's orthography does not always faithfully represent the sound. Bassetti (2006) demonstrated how the orthographic input of Pinyin is affected by CFL learners' first language, resulting in inaccurate and sometimes erroneous pronunciations on a number of Pinyin sounds. When adult learners' first language is a language of letters, Bassetti suggested that adding an extra layer of Pinyin scripts bring up problematic phonological representations and can negatively impact readers' phonological reconstruction of the text, especially among beginning level learners.

Considering the role and the value of phonological prompts in CFL reading and the ineffectiveness of Pinyin remediation, this study attempts to investigate the issue from a new perspective. The aim of the study is to examine the effect of externally-supplied text vocalization on CFL learners' reading comprehension, at both the character level and the sentential level. Specifically, the following two questions will be addressed: 1) How will supplying text vocalization to beginning CFL learners in reading exercises affect their performance in word recognition and word recall? 2) How will this affect reading speed and overall comprehension?

3. Method

3.1. Participants

A total of 73 students enrolled in beginning level Chinese classes at one public university in the United States participated in this study (44 males and 29 females, mean age = 21). They were taught by instructors using the same textbooks and same curriculum. According to the questionnaire administrated at the beginning of the study, most of the participants are native speakers of English and had no or little previous knowledge of Chinese. By the time they participated in this study, they had learned around 400 characters within the previous 5 months of study (interrupted by a winter break). When learning Chinese, participants learned Pinyin first, before any characters were introduced, and were often found using Pinyin in their everyday learning activities. To avoid any complications brought on by participants' previous knowledge of the Chinese language, students with any background in Asian languages were include in the study but their answers were removed from the data analysis. This resulted in 65 participants (39 males and 26 females, mean age = 20.7).

3.2. Instruments

The author prepared a vocabulary and a comprehension test. The vocabulary test was paper-based and contained two types of questions: 30 items for recognition (Recog) and 30 items for recall (Recal). In the recognition test, drawing on Hayes (1988), the author used two types of distracters (phonological and graphic)[2], one distracter per sentence, if there was one at all. Ten sentences contained phonological distracters, ten contained graphic, and another ten had none (see Appendix A). Participants needed to judge whether a sentence was distracter-free by choosing "yes" or "no". To discourage random guessing, participants must mark the distracter out if they believed there was one. In the recall test, the target words were presented in Chinese and participants needed to provide the meaning in English. Grading was done by two graders (the author himself and another native Chinese speaker) and the inter-rater agreement was 100%. Sentences in the Recog test were written by the author containing the grammar and vocabulary that participants had already studied, as were the words in the Recal test.

The comprehension test (Comp) was hosted online and done in a computer laboratory to better record participants' reading speed (see Appendix B). The online platform featured a built-in function to record the amount of time during which the test remained active, from the moment when the test was first accessed and displayed on the screen to the moment when the "submit" bottom was hit. The time was originally recorded in an hour-minute-second format and was later converted to seconds for data analysis. The author presented 30 sentences in Chinese (similar to those used in the Recog task, but without distracters) and asked participants to faithfully translate them into English to pinpoint the participants' abilities and issues when reading Chinese.

Sentences involved in both tests were 15 to 24 characters long, containing only narrations/statements to avoid further stratification of the data. All items were equally weighted. The scores of the vocabulary test were calculated based on the number of errors participants made whereas the scores of the comprehension test were based on the number of answers correct. In the vocabulary test, an error was marked and one point was generated when participants mistakenly perceived a distracter in a distracter-free sentence, or failed to identify the distracter in a sentence that contained one. In the comprehension test, participants were asked to translate sentences from Chinese into English. A complete and accurate translation earned two points. One point was deducted if minor error(s) were found and two points were deducted if the translation deviated from the original significantly. As noted by Shen (2005b), translation is a time-consuming

[2]Compared to the target character, a phonological distractor is an exact homophone in both sound and tone; a graphic distractor closely resembles in graphic features.

task compared to multiple questions or cloze test and thus is less favored by researchers when measuring reading comprehension, yet it became an ideal choice for this study due to one of its unique features, namely, the capability to document and exhibit details in how readers construct meaning from the text by accessing and retrieving pertinent knowledge at both the word and sentence level.

Participants were given opportunities to do mock tests to avoid unwanted test-anxiety or confusion, so scores could better reflect their actual reading level. To ensure reliability, identical tests were administrated in parallel sessions that used the same textbook, followed the same curriculum and produced CFL learners whose Chinese level was similar to that of the participants'. Cronbach's alpha, based on scores of students from parallel sessions, was calculated to measure the internal reliability of the two tests respectively. Reasonable Cronbach's alphas were obtained, 0.77 in Recog, 0.86 in Recal and 0.82 in Comp.

3.3. Procedure

The study lasted for four weeks. All participants were randomly assigned to the experimental group or the control group at the beginning of the study and given the first round of tests as a pretest. Testing was done under the supervision of the author and no reference of any kind was allowed. No answers or feedback was provided after the test. After taking the pretest, every two weeks both groups were asked to read and translate 60 Chinese sentences into English as part of their homework, which was assigned regularly with their scores counted toward their final grades. These sentences (without distracters) closely resembled those in the pretest. Over a four-week period, a total of 120 sentences were assigned. Both groups completed the reading assignment using an online platform where the experimental group was provided the text and a link to its vocalization so participants could click the link to hear the text, whereas the control group could only read the text. After the study ended, participants in the treatment group were contacted by the author and asked about their habits regarding the use of audio links next to the text. Their responses showed that they always clicked the links to listen to the recordings. The data provided by the web platform was carefully analyzed in order to confirm this claim and found that, when doing the reading assignment online, the treatment group spent considerably more time listening, almost twice as much as the control group did. This suggests that participants in the treatment group made use of audio input while reading. The change of the reading manner then led to longer reading time. The vocalization of the text was recorded by the author in collaboration with other native Chinese speakers. No broken links or faulty recordings were reported. The author promptly graded both groups' homework and provided detailed feedbacks.

At the end of this study (week 4), the second round of tests (post-test) was conducted in the same manner as the first round (no reference, no feedback, etc). Individual questions on the test remained the same while the order of questions was randomized to avoid the unwanted test effects.

4. Results

How will supplying text vocalization to beginning CFL learners in reading exercises affect their performance in word recognition and word recall?

A 2×2 one-between-one-within subjects ANOVA was conducted to assess whether there were differences in word recognition scores for the different conditions, with Recog scores as the dependent variable, audio inputs (with or without audio links adjacent to texts) as between groups factor and time (week0 and week4) as the within group factor (see Table 1). A significant main effect was found for audio input ($F = 13.55$, $p < .05$). Two follow-up independent T-tests revealed that the treatment and control group did not score differently upon entering the study ($t = 1.26$, $p > .213$), while scoring significantly differently by the end of the study ($t = -6.46$, $p < .05$) with a very large effect size ($d = 1.60$), and those who received audio input ($M = 10.34$, $SD = 2.13$) made significantly fewer errors than those in the control group ($M = 14.09$, $SD = 2.52$). Also, the main effect for time was found to be significant ($F = 67.34$, $p < .05$), with both a large effect size (partial $\eta^2 = .52$) and a high observed power at .99. The follow-up contrast analysis suggests that participants made significantly fewer mistakes at the end of the study, compared with themselves than when they entered the study ($t = 6.43$, $p < .05$), with a large effect size ($d = 0.80$), regardless of whether they had received the treatment or not. The interaction of audio input and time on Recog was significant ($F = 38.82$, $p < .05$), with a large effect size (partial $\eta^2 = .38$) and a high observed power at .99.

The significant main effect of audio input positively answered the first part of the research question. However, such effects were overshadowed by the fact that both groups were more capable of telling the differences between target words and distracters in the post-test. It is natural to hypothesize that since audio input helped participants in the treatment group in scrutinizing the graphic features of a character during the reading process, consequently it helped them outperform those in the control group, even after the effect of time (the main factor leading to the study effect) was controlled. To examine this hypothesis, the author conducted independent t-tests on the pretest data and post-test data, one dataset at a time, in order to compare participants' performance under three conditions (phonological or graphic or zero distracter) to test whether there were any be-

tween-group differences in identifying characters that sound alike or look alike. For example, the phonological distracter of "前" (early) was set as "钱" (money); the graphic distracter of 字 (character) was set as 学 (study). When entering the study, there was no between-group difference in any of the conditions. In the fourth week, however, those who had received audio input made significantly fewer errors on phonological distracters than those who did not ($t = 4.45$, $p < .05$), with a large effect size ($d = 1.10$). Similar results were found under zero distracter conditions ($t = 6.94$, $p < .05$) with a large effect size ($d = 1.72$). The graphic distracters, on the other hand, misled both groups and the difference was insignificant ($t = .38$, $p = .70$).

Being fully aware of the study effect on word-recognition, the author conducted a 2×2 one-between-one-within subjects ANOVA test on scores of word recall and received a result similar to that of word recognition (see Table 2). Both main effects and the interaction were significant. Again, two independent t-tests were conducted to compare the treatment and the control group at the beginning and at the end of the study. The author found both groups' ability to recall words were significantly different at the end of the study ($t = 4.45$, $p < .05$), with a large effect size ($d = 1.10$), while such differences did not exist at the beginning of the study, ($t = .40$, $p = .70$). Furthermore, two paired t-tests on the score differences between week0 and week4 showed that the treatment group had a bigger mean difference ($MD = 1.88$ compared to $MD = 0.91$) and a larger effect size ($d = 1.10$ compared to $d = 0.60$) than that of the control group.

In sum, audio input supplied during the reading process positively affected participants' performance on word recognition and word recall in a significant way. It was true that those abilities could be improved by the activity of reading alone, yet the audio input gave learners an edge and made them do even better.

How will supplying text vocalization to beginning CFL learners in reading exercises affect reading speed and overall comprehension?

The focus now shifted from the word level to the sentential level to examine how audio input could affect reading speed and comprehension. A 2×2 one-between-one-within subjects ANOVA test on reading speed revealed that the main effect of neither time nor audio input was significant (see Table 3). Their interaction was insignificant as well. This insignificance must be approached with caution as the plot generated along with the ANOVA test depicted a visual trajectory of how groups differed over time, suggesting that the insignificance of the ANOVA test may result from the control group's unchanging performance on reading speed, i.e., the time they spent on reading remained by and large unchanged between week0 and week4. The treatment group clearly follows a different pattern, which was examined by follow-up tests. The first comparison be-

tween the treatment group and control group in week4 showed a significant difference in reading speed ($t = 2.46$, $p < .05$), with a medium effect size ($d = 0.61$), while this between-group difference was not observed in week0 ($t = .09$, $p = .93$). The second comparison of within-group difference (week0 vs. week4) on the control group yielded no significant difference ($t = .04$, $p = .97$), suggesting the reading speed of the control group remained the same over time. On the other hand, the treatment group made significant progress on reading speed at the end of the study, ($t = 2.77$, $p < .05$), with a small to medium effect size ($d = 0.49$).

In the same manner, a 2×2 one-between-one-within subjects ANOVA test was conducted on reading comprehension (see Table 4). The results show that the main effect of time on reading comprehension was significant ($F = 43.90$, $p < .05$) while the main effect of audio input was not significant ($F = 1.24$, $p = .27$). After a close examination of the data, the author observed the limit of comparing the test results without differentiating and acknowledging participants' reading condition, i.e., there was more space for struggling readers to improve, while scores of proficient readers had limited space to rise. It was possible that participants who scored low in week0 made more progress in week4, in comparison to those who already scored high in week0 and had little room for improvement by week4. To test this hypothesis, the author first ran correlation tests between scores of week0 and week4 of the experimental and control group respectively. For both groups, there was a strong correlation: $r(30)=0.83$ for the experimental group and $r(31)=0.98$ for the control group. He then used each group's average score as the benchmark to identify those who scored below the average in week0 and compared their scores in week4. It turned out that the participants who scored below average in both groups were almost at the same level in week0 ($t = .77$, $p = .45$). By the fourth week, the audio input made a significant difference, ($t = 3.96$, $p < .05$), with large effect size ($d = 1.40$). Compared with themselves, control group participants who scored below average in week0 did not make much progress in week4 ($t = 1.09$, $p = .29$), whereas participants who scored below average in week0 in the experimental group scored differently in week4, ($t = 5.19$, $p < .05$), with large effect size ($d = 1.26$).

The mixed result on the sentential level suggests that the speed and competency of reading are relatively stable attributes. A four-week long period is too short to produce substantial changes. Adding the factor of audio input, however, can lead to a moderate level of progress, which was more evident among challenged readers.

5. Discussions

This study investigated the effects of presenting the text and its vocalization in reading exercises to non-native learners of Chinese at the beginning level with an intent to probe the benefits readers can harness in areas relevant to reading, i.e., to recognize characters, to recall meaning, and to comprehend the text in a speedy manner. It also examined whether audio input could help participants focus on general as well as detailed features of characters.

The first observation obtained after conducting the data analysis on the word recognition and word recall tests was, in addition to the benefits of audio input, the large study effect. Comparing their relative progress, participants all benefited from a few more weeks of reading and learning the language as they ended up being able to detect more distracters and recall more words. Knowing of such an effect beforehand, the author intentionally designed this study with repeated measures to capture the reality of CFL reading, namely, learners often go through a long process to learn the language and phonological awareness takes time to form. Supplying text vocalization on one or two occasions may not be enough to incubate any significant change. Meanwhile, as Xing (2003:62) stated, there are "a number of factors [which] should be taken into consideration in the process of teaching and learning the language". By randomly assigning participants into two groups and comparing their performance at the beginning and the end of a four-week study, the author was able to balance influences of factors irrelevant to the experiment so the results faithfully demonstrate the role played by the audio input over a reasonable amount of time.

As it turned out, hearing the text vocalization while reading the text made participants perform better and they continued to do so even after the study effect was accounted for. The author believes there are at least two primary mechanisms involved. First, when participants in the experimental group read and heard the sound simultaneously, they had the privilege to search their lexicon to match the sound and the meaning, in addition to processing graphic features of characters to recall meaning. Reading exercises conducted in this form helped reinforce the connection between sound and meaning, consequently leaving participants with a deeper, stronger impression of the meaning of the text. Participants soon realized they could establish a shortcut between phonology and semantics in the mental lexicon, without having to go through the loop of writing-sound-meaning. The shortcut could have circumvented an underlying conceptual system that many researchers regarded as counterproductive because this system oftentimes diverted the effort and attention from directly comprehending the second language to self-initiated attempts to translate the second language into the first language

for comprehension (Kroll, 1993; Pavlenko, 2009). Furthermore, without relying on processing graphic features of characters to access meaning and segment words, participants could allocate more attention and cognitive resources for higher-level text comprehension and could employ more holistic reading strategies to construct meaning from the top down.

The second mechanism may come from the pattern in which grammar and words in a second language are retrieved and understood. According to Ullman's (2001) model of procedural memory and declarative memory, adult second language learners predominantly relied on the latter for both grammatical computations and lexical storage. Compared to their native speaking peers, they can still quickly identify requirements of a language task yet often suffer some delays in associating a preceding word with an array of possible succeeding expressions as they painstakingly weigh the correctness and appropriateness of different combinations. For example, a common expression such as "你那里" (the situation in your place, "你那里的天气怎么样?" means "how is the weather in your place?") is indeed puzzling because the association between "you(你)" and "there(那里)" in English is considerably different from that of Chinese. Instead, adult learners favor the solution of storing meaningful, manageable chunks of words in the declarative memory, instead of retrieving grammatical rules from procedural memory and computing them in real time while generating or comprehending the target language. The externally supplied vocalization in this study, under Ullman's theorization, accommodated participants' strategies for encoding and processing information by activating the mental lexicon and the mental grammar to produce meaning (Pavlenko, 2009). It expedited both levels of information processing, i.e., lexical access and text comprehension, by reinforcing the arbitrary association between the sound and the graphic representation. Reading became less laborious after the phonological stimulus lifted participants beyond the word-by-word decoding task, allowing them to process meaning in larger chunks. The data analysis on sentence comprehension provided empirical evidence to validate Ullman's hypothesis.

Before the discussion could continue, it was imperative to answer a crucial question: if using audio input helped participants pay more attention to meaning by establishing a sound-meaning shortcut and graphic features of characters were somehow downplayed by such treatment, then how do we explain their better performance in word recognition tasks at the end of the study? This question can be answered by paying closer attention to the task of recognition itself. By design, graphic distracters examine participants' knowledge of details of characters, while phonologic distracters probe the memory on general features. Differentiating these two types of distracters would help produce a more comprehensive

picture of the issue. For graphic distracters, participants in both groups made significantly fewer errors in the fourth week, compared with the first week. The one-within-one-between ANOVA showed this was due to the effect of study, rather than the effect of treatment, and participants in the experimental group did not score significantly higher than their peers in the post-test. In contrast, for phonological distracters, only those in the experimental group made significantly fewer errors in the fourth week, whereas those in the control group ended up making more errors. Furthermore, under zero-distracter conditions, the experimental group scored the same as the control group in the pretest but outperformed their peers in the post-test.

The data analysis on vocabulary shows that audio input can substantially enhance participants' abilities to remember and recognize general features of characters, confirming the earlier argument that audio input helped establish a short path between the sound and the meaning, so participants could better focus on the general features of the orthography while doing reading exercises and maintain this awareness even when the audio input was removed. As a matter of fact, the benefit of audio input in this regard was in line with other studies (Chung, 2008; Jin, 2006), where both audio and characters were presented as stimuli in the learning process, and an enhancement of, rather than impairment to, vocabulary learning was observed.

Meanwhile, it is important to notice that coupling recordings and texts could do little to help participants remember the fine and minute details of a character. Those who received audio input had a better understanding of the general visual image of a character, but when a graphic distracter was only one or two strokes different from the target character, such knowledge became shallow and insufficient. According to Shen (2004) and other researchers (Li & Lee, 2006; Taft & Chung, 1999), instructor-guided elaborations on characters' graphic features and components are highly valuable because such knowledge can help beginning level CFL learners better understand how sub-morphemic parts form characters and suggest meanings. The effect of audio input obviously does not touch this domain.

5.1. Pedagogical Implication

This study demonstrated the benefits of supplying text vocalization in reading exercises. It suggests that the potential to integrate aural stimulus along with the traditional visual stimuli to enhance reading deserves attention as much as, if not more than, the endeavor of teaching vocabulary with multimedia programs. Now since Computer Assisted Language Learning (CALL) has reached almost every level and aspect of language acquisition, the outlook of CFL teaching and learn-

ing is gradually changing to keep pace with the ramification of the technology (Yao, 2009). The author believes a number of actions need to be taken. Creating appropriate and manageable reading materials should be the first item on the list. Bounded by very limited vocabulary and grammar, beginning level adult CFL readers are susceptible to giving up on reading when materials are beyond their level. They also get bored quickly when text becomes too mundane and childish. Teachers need to come up with texts that readers consider rewarding because it covers grammar and vocabulary they have learned so far, and is challenging enough to keep them going. More importantly, the text must be embedded in a specific, meaningful context to give readers a sense of what they can do with the language, instead of what they know about the language. Teachers can utilize the recording for multiple purposes including listening comprehension and dictation, in addition to fostering reading development.

Since a fairly amount of labor and technology is involved in both recording the text and delivering the recording, teachers need to have the necessary means and timely support at their disposal to smoothly integrate the sound and text. This may be beyond the reach of an individual or a single program. Working in groups and collaborating with technicians is a viable solution to create, upload and maintain the text and its vocalization. Also, the author urges textbook writers and publishers to invest more effort accommodating lower level students' need to receive meaningful, engaging language input in more diverse forms, both aural and written.

5.2. Limitations

In this study participants' performances were only measured and compared before and after text vocalization was supplied in reading exercises. Due to the limitations of the experiment, however, a longitudinal design was not done to trace how participants did long after the audio input was removed and reading went back to the previous situation where no vocalization of the text was heard. A follow-up study will further explore the role played by audio input over an extended period of time, providing a comprehensive examination of related issues to better inform researchers of the advantages and disadvantages of introducing audio input in lower level reading. Meanwhile, the beginning level CFL learners and their reading materials have unique features that may not be applicable to higher levels. The result of this study, in this regard, must be taken with caution when applied in a different context.

Table 1

Two-way Analysis of variance for test scores of Word Recognition as a Function of Time and Audio Input

Variable and source	*Df*	MS	*F*	η^2
scores				
Time	1	272.39	67.34*	.52
Audio Input	1	77.94	13.55*	.18
Time X Audio	1	157.03	38.82*	.38
Error	63	5.75		

*$P < .05$

Table 2

Two-way Analysis of variance for test scores of Word Recall as a Function of Time and Audio Input

Variable and source	*Df*	MS	*F*	η^2
scores				
Time	1	182.18	76.86*	.55
Audio Input	1	37.44	20.95*	.25
Time X Audio	1	27.11	11.44*	.15
Error	63	1.79		

*$P < .05$

Table 3

Two-way Analysis of variance for test scores of Reading Speed as a Function of Time and Audio Input

Variable and source	*Df*	MS	*F*	η^2
Time in seconds				
Time	1	6231042.14	3.37	.05
Audio Input	1	5892786.52	3.04	.05
Time X Audio	1	6231042.14	2.91	.04
Error	63	1731052.91		

*$P < .05$

Table 4

Two-way Analysis of variance for test scores of Reading Comprehension as a Function of Time and Audio Input

Variable and source	*Df*	MS	*F*	η^2
Score				
Time	1	454.39	43.90*	.41
Audio Input	1	265.92	1.24	.02
Time X Audio	1	306.82	29.64*	.32
Error	63	214.54		

*$P < .05$

REFERENCES

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language and Literacy, 9*(1), 95-114.

Bernhardt, E. B. (1986). Reading in the foreign language. In B. Wing (Ed.), *Listening, reading, and writing: analysis and application.* (pp. 93-115). Middlebury, VT: Northeast Conference.

Bernhardt, E. B. (1991). *Reading development in a second language: theoretical, empirical & classroom perspectives*. Norwood, NJ: Ablex Publishing Corporation.

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics, 16*(1), 15-34.

Chu, C. Z. (2006). A reflection on traditional approaches to Chinese character teaching and learning. In T. Yao, C. Chu, Y. Wang, H. Xu & J. Hayden (Eds.), *Studies on Chinese instructional materials and pedagogy* (pp. 240-279). Beijing: Beijing Language University Press.

Chung, K. K. H. (2003). Effects of pinyin and first language words in learning of Chinese characters as a second language. *Journal of Behavioral Education, 12*(3), 207-223.

Chung, K. K. H. (2007). Presentation factors in the learning of Chinese characters: the order and position of Hanyu pinyin and English translations. *Educational Psychology Review, 27*(1), 1-20.

Chung, K. K. H. (2008). What effect do mixed sensory mode instructional formats have on both novice and experienced learners of Chinese characters? *Learning and Instruction, 18*(1), 96-108.

Everson, M. E. (1988). Speed and comprehension in reading Chinese: Romanization vs. characters revisited. *Journal of the Chinese Language Teachers Association, 23*(2), 1-20.

Everson, M. E. (1994). Toward a process view of teaching reading in the second language Chinese curriculum. *Theory into Practice, 33*(1), 4-9.

Everson, M. E. (2007). Viewing diversity of subject matter: The case of reading in Chinese. In S. McGinnis & M. H. Haley (Eds.), *The many views of diversity: understanding multiple realities.* Middlebury, VT: Northeast Foreign Language Conference Report.

Everson, M. E., & Ke, C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 32*(1), 1-20.

Furman, N., Goldberg, D., & Lusin, N. (2007). Enrollments in languages other than English in United States institutions of higher education, Fall 2006. Retrieved September 20, 2010, from http://www.mla.org/enroll_survey06_fin

Hayes, E. (1988). Encoding strategies used by native and non-native readers of Chinese Mandarin. *Modern Language Journal, 72*(2), 188-195.

Jin, H. (2006). Multimedia effects and Chinese character processing: An empirical study of CFL learners from three different orthographic backgrounds. *Journal of the Chinese Language Teachers Association, 41*(3), 35-56.

Ke, C. (1996). An empirical study on the relationship between Chinese character recognition and production. *The Modern Language Journal, 80*(3), 340-349.

Ke, C. (1998). Effects of strategies on the learning of Chinese characters among foreign language students. *Journal of the Chinese Language Teachers Association, 33*(2), 93-112.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 63-81). Amsterdam: Benjamins.

Lee-Thompson, L. (2008). An investigation of reading strategies applied by American learners of Chinese as a foreign language. *Foreign Language Annals, 41*(4), 702-721.

Lee, C. H. (2009). Testing the role of phonology in reading: Focus on sentence processing. *Journal of Psycholinguistic Research, 38*(4), 333-344.

Li, J., & Lee, K. (2006). The graphic factor in the teaching and learning of Chinese characters. *Journal of the Chinese Language Teachers Association, 41*(1), 79-92.

Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: interdisciplinary approaches* (pp. 125-160). Buffalo, NY: Multilingual Matters.

Pavlidis, M. (1992). Reading strategies of second-language learners of Chinese. *Melbourne Papers in Applied Linguistics, 4*, 22-56.

Shen, H. H. (2004). Level of cognitive processing: Effects on character learning among non-native learners of Chinese as a foreign language. *Language and Education, 18*(2), 167-182.

Shen, H. H. (2005a). An investigation of Chinese-character learning strategies among non-native speakers of Chinese. *System: An International Journal of Educational Technology and Applied Linguistics., 33*(1), 49-68.

Shen, H. H. (2005b). Linguistic complexity and beginning-level L2 Chinese

reading. *Journal of the Chinese Language Teachers Association, 40*(3), 1-28.

Shen, H. H. (2008). An analysis of word decision strategies among learners of Chinese. *Foreign Language Annals, 41*(3), 501-524.

Taft, M., & Chung, K. (1999). Using radicals in teaching Chinese characters to second language learners. *Psychologia, 42*, 234-251.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*(2), 105-122.

Wang, S. C. (2008). The ecology of the Chinese language in the united states. In A. Creese, P. Martin & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed., Vol. 9, pp. 169-181). New York: Springer Science.

Wu, J.-T., & Liu, I.-M. (1996). Chinese lexical access. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 31-42). Hong Kong: Oxford University.

Xie, T., & Yao, T. (2009). Technology in Chinese language teaching and learning. In M. E. Everson & Y. Xiao (Eds.), *Teaching Chinese as a foreign language* (pp. 151-172). Boston: Cheng & Tsui.

Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar*. Hong Kong: Hong Kong University Press.

Yao, T. (2009). The current status of Chinese CALL in the United States. *Journal of the Chinese Language Teachers Association, 44*(1), 1-23.

Zheng, T. (2005). Designing on-line Chinese language courses: New roles for educators. *Journal of Information Technology Education 4*, 275-285.

Appendix A

The following sentences may contain one "wrong" character, which looks like or sound like the right character. If you think there is one, then choose "YES" and mark it out. If not, choose "no". Example:

我足中国人 Yes no

我是钟国人 Yes no

我是中国人 Yes no

1.	这个老师是中国人,他的中文和英文都很不错。	YES	no
2.	对不起，这个学我觉得我好像写错了。	YES	no
3.	你明天上午十点能不能送我到机场？	YES	no
4.	我很喜欢吃牛内,可是最近太贵了。	YES	no
5.	她在医院上班，可是她不是医生。	YES	no
6.	王朋的宿舍离我的宿舍很近，离中国城也不远	YES	no
7.	我以钱不会用中文写信 我现在会用中文写信了。	YES	no
8.	我去过小高家，他的家又书服又漂亮。	YES	no
9.	我周末常常和我的朋友一起去会园。	YES	no
10.	我坐车咸者自己开车去学校上学。	YES	no
11.	我听说兆京春天很冷，常常下雪，对吗？	YES	no
12.	我昨天吃了很多凉拌黄爪, 现在很想上厕所。	YES	no
13.	小张吃了昨天晚上的剩菜，他的肚子疼死了。	YES	no
14.	我想用电脑练习发因，我可以用你的电脑吗？	YES	no
15.	小张不但说英文说得很好， 面且会用英文写信。	YES	no
16.	师博，对不起，这碗酸辣汤怎么这么酸？	YES	no
17.	我听说你喜欢吃中国菜，可是不喜欢吃辣，对吗?	YES	no
18.	我下个星期过生日，你送我什么声日礼物？	YES	no
19.	我昨天和小王去中国城，我们买了很多冬西。	YES	no
20.	我昨天有一个中文考事，我觉得我考得不好。	YES	no
21.	我去年在暑期班认识了一个美国女孩。	YES	no
22.	到第三个路口往左拐不远，就能看见我家。	YES	no
23.	我很喜欢吃饺子，可是不知道怎么作。	YES	no
24.	你渴不渴？我们有啤酒和可乐，你想要什么？	YES	no
25.	九点上课,你为什么十点半才来？	YES	no
26.	你一直往难,过三个路口，就会看到一个中国饭馆儿	YES	no
27.	你知道中国有什么好玩的第方吗?	YES	no
28.	明天有中文听力课,小白想跟朋友夏习语法。	YES	no
29.	我们学校有一个咖啡电，在餐厅和宿舍的中间。	YES	no
30.	她常常练习中文，她的中文一定不错。	YES	no

Appendix B

Comprehension test

Translate the sentences into English. Two points for each sentence.

1. 往右拐,过一个路口，你就能看到我的学校。
2. 我的三百块钱你什么时候能还给我?
3. 学校不准我们在宿舍吃饭，都得在餐厅吃。
4. 吃了那盘家常豆腐以后,我肚子不舒服。
5. 欢迎来到中国!你现在一定累了吧。
6. 请问，从北京到上海的飞机什么时候到？
7. 请问，这个包里有没有水？水不能带上飞机。
8. 那家饭馆上菜上得快不快，如果慢我们就不去了。
9. 服务员，请问二十个饺子多少钱?
10. 这个书架是谁送给你的，虽然不大但是很漂亮。
11. 旅行的时候,别喝太多啤酒，要不然你就会睡过头。
12. 我没有拿地图,可是我上个月常去那儿，知道怎么走。
13. 往前开,再往西拐，看到一个中国饭馆就到了。
14. 这家饭馆的黄瓜好吃极了,你想不想来一个？
15. 这些是我高中照片，我觉得这个人很像张老师。
16. 我没费很大的力气就找到了小张。
17. 我觉得北京的秋天很舒服，夏天太热。
18. 对不起,我没听清楚，再说一遍好吗？
19. 我觉得同学送给我的生日礼物很有意思。
20. 这张饭桌真大。你怎么搬进来的？
21. 我晚上七点请小王吃饭，可是他八点才来。
22. 在报纸上看到了四条广告，现在想打电话问问。
23. 我昨天晚上上了四次厕所。你以前吃坏过肚子吗?
24. 网上有很多电影，可是我的电脑太慢，不能看。
25. 他是我的叔叔,在学校的餐厅工作。
26. 他喜欢喝啤酒,常常喝很多，所以身体不好。
27. 这个网球拍多少钱?请问能不能刷卡？
28. 今天的报纸不好看，我看了一遍，没意思。
29. 北京上个星期下雪了，下了三天。
30. 你那里一个月电费多少钱，比这儿便宜还是贵?