Repeated-reading-based instructional strategy and vocabulary acquisition: A case study of a heritage speaker of Chinese

ZhaoHong Han and Cheng-ling Alice Chen
Teachers College, Columbia University
United States

Abstract

Repeated reading, a procedure involving repetition of the same text, has received copious attention from first language reading research providing highly converging evidence of its potency for reading fluency, accuracy, and comprehension. In contrast, second language research on repeated reading has been scarce. The very few studies extant have, nevertheless, shown similar, albeit inconclusive, findings. The present study was an attempt to foray into a hitherto uncharted area in both first and second language research, by investigating vocabulary gains from implementing a set of repeated-reading-based pedagogical and learning procedures. Using one heritage speaker of Chinese as its subject, the study administered 20 sessions of assisted repeated reading over three weeks. Results indicated both intentional and incidental vocabulary gains that would not otherwise have been possible through conventional reading or vocabulary instruction.

Keywords: repeated reading, second language vocabulary acquisition, Chinese, heritage speaker

It is trivial to point out that reading and vocabulary are closely related. In fact, they are mutually constraining and complimentary: On the one hand, reading ability depends on vocabulary knowledge. “Semantic processing is central to reading comprehension ... Ultimately, it is vocabulary that largely controls semantic processing” (Koda, 1994, p. 10). On the other hand, reading is a critical source of vocabulary growth (Krashen, 1989; Zahar, Cobb, & Spada, 2001). It therefore is no surprise that vocabulary instruction has taken a central place in all existing approaches to the development of reading ability (e.g., phonic, linguistic, sight-word, and language experience), nor is it that reading has served as a major scaffold for vocabulary instruction (e.g., Peters, Hulstijn, Sercu, & Lutjeharms, 2009).

The present study explores a particular type of reading, repeated reading, as a fulcrum for vocabulary acquisition. In the sections that follow, we will first introduce and discuss the theoretical background of the research, and then report the study. We will end with a discussion of the main findings and their implications for future research.
Repeated Reading

Repeated reading, initially known as multiple oral reading, involves multiple, successive encounters with the same visual material, the key being repetition—whether of the same words, sentences, or connected discourse. An instructional technique designed originally for improving reading fluency in learners with reading disabilities, repeated reading has been practiced with both disabled and non-disabled students in a variety of fashions, ranging from having the learner read aloud (Samuels, 1979), to listening to and simultaneously or subsequently reading aloud (Chomsky, 1978), and to silently reading (Anderson, 1993, 1999, 2008, 2009), the same material multiple times.

Despite the procedural divergence, research has shown that the technique benefits fluency development—defined as improved accuracy of word recognition and reading speed—and comprehension in slow readers. Chomsky (1978), for example, reported that the procedure increased the fluency of slow and halting readers and instilled in them a heightened sense of confidence, motivation, and willingness to undertake reading new material independently. Similar findings were reported by Samuels (1979) claiming:

The fact that starting rates were faster with each new selection and fewer rereadings were necessary to reach goals indicates transfer of training and a general improvement in reading fluency. (p. 404)

Reviewing the early research, Moyer (1982) concluded that “repeated reading practice can facilitate general reading fluency for some unskilled readers, for normal readers given difficult text, and in regular classroom instruction” (p. 620).

Recent research syntheses have corroborated this conclusion (Meyer & Felton, 1999; National Institute of Child Health and Human Development [NICHD], 2000; Therrien, 2004). Therrien (2004) stated:

Repeated reading can be used effectively with nondisabled students and students with learning disabilities to increase reading fluency and comprehension on a particular passage and as an intervention to increase overall fluency and comprehension ability. (p. 252)

The highly converging support from by now an extensive research base for repeated reading raises the question: What is it about repeated reading that facilitates reading fluency?

Researchers have suggested that difficulty in word recognition is a major obstacle to fluent reading, while maintaining that reading is a complex process involving multiple levels of processing from word decoding to deriving meaning from sentences, paragraphs, and the text as a whole and that fluency can be jeopardized by a breakdown at any of these levels (Logan, 1997). LaBerge and Samuels (1974) argued that slow decoding creates a “bottleneck” that impedes the flow of thought and hampers comprehension. Poor readers often spend a great deal of their cognitive resources on decoding and have little left for comprehension. Conversely, good readers decode words quickly and accurately, thus conserving more resources for comprehension. This is, however, one (albeit main) explanation for reading dysfluency. A
contending perspective posits that written language is devoid of prosodic cues, and this makes transfer from oral language to written language difficult for some readers. According to Schreiber (1980), readers who are unable to generate appropriate prosodic markings cannot divide sentences into meaningful phrases and therefore have difficulty comprehending written text, regardless of their ability to identify individual words. Thus, a lack of ability to appropriately segment meaningful units is deemed a co-factor for the lack of reading fluency, alongside failure to decode words.

Both difficulties can be surmounted by repeated reading—importantly, of connected discourse rather than isolated words (Gonzalez, 1974; Samuels, 1979). First, connected discourse affords macro (i.e., encyclopedic) and micro (i.e., linguistic) contexts, which, coupled with the learner’s repeated exposure to the same text leading to increased familiarity with the contextual constraints, may not only help accurate decoding of words but also its speed (Moyer, 1982). More specifically, experimental research has shown that repeated reading practice on words in isolation—such as repeated reading of word lists—does not enable transfer of the acquired speed to word identification in connected discourse (Fleisher, Jenkins, & Pany, 1979); that good and poor readers alike show the same general facilitation of context for rapid word identification (West & Stanovich, 1978); and yet, that skilled readers are faster in the naming of words in isolation, while poor readers are more reliant on context (Perfetti, Goldman, & Hogaboam, 1979). These findings combine to suggest that repeated reading practice of connected discourse is a more effective way of developing fluency, though an ultimate index of reading fluency can be rapid context-free word identification as well as word identification in context.

In addition, repeated reading, as practiced in many settings, often offers guided phonological practice which helps segmentation and parsing of written text, thus facilitating word and phrase decoding, chunking of larger meaningful units, and, ultimately, reading fluency and comprehension. Schreiber (1980) noted that there can be ‘epiphenomenal’ gains from repeated reading, such as the discovery of the appropriate syntactic phrasing including intonation, stress, and duration, which are not explicitly represented in the written form of language.

At the heart of repeated reading is repetition (Biemiller, 1978; Perfetti & Roth, 1981). A concept from information processing theory, repetition or redundancy may lead to an increase in familiarity and corresponding decrease in the amount of information to be processed while reading (Hyönä & Niemi, 1990). Drawing on insights from automaticity theory (LaBerge & Samuels, 1974), Samuels (1979) outlined three possible developmental stages of word recognition in relation to comprehension. The first is called the non-accurate stage, wherein the developing reader experiences great difficulty in recognizing words, even when given ample time, and achieves little comprehension. The second is called the accuracy stage, wherein the developing reader is able to recognize printed words with accuracy but with much cognitive effort, and as a result, comprehension is slow and halting. The third stage is called the automatic stage, wherein the developing reader is able to recognize words accurately and with ease; correspondingly, comprehension is fast and smooth. An important function of repeated reading, as Samuels summed it up, is that it provides the practice needed to become automatic, thereby contributing to fluency and comprehension.

In a meta-analytical attempt to identify essential components of repeated reading, Therrien
(2004) found that repeated reading differentially improves reading fluency and comprehension ($ES = .83$, $SE = .066$ for fluency increase vs. $ES = .67$, $SE = .080$ for comprehension for repeated reading of the same passage; $ES = .50$, $SE = .058$ for fluency increase vs. $ES = .71$, $SE = .177$ for comprehension for repeated reading of different passages). Moreover, cueing students’ attention to fluency, comprehension, or both at the outset may differentially impact the improvement in the specific areas. Furthermore, reading aloud to an adult produces greater improvement than reading to a peer. Additionally, provision of corrective feedback on word errors has a greater impact on fluency and comprehension than absence of corrective feedback. Lastly, it has been found that using a performance criterion to regulate repeated reading has a greater effect for fluency and comprehension than using a fixed number of readings. Aside from these insights, research has indirectly suggested that the effects of repeated reading are greater for difficult texts than for easy texts (Hyônä & Niemi, 1990). In other words, with repeated reading, students should be able to process texts far beyond their current reading level.

What has been said above about repeated reading comes from research on learners who grow up learning to speak and read a language, that is, on first language (L1) reading. How does repeated reading fare with other more unique populations such as post-pubescent second language (L2) learners, and even heritage speakers? Little research, as a matter of fact, exists on that front, as is briefly discussed below.

**L2 Research on Repeated Reading**

L2 research on repeated reading has by far been rare (Grabe, 1991, 2009; Koda, 1994). The lack of interest mirrors a lack of interest in researching word recognition in general, which, according to Koda (1994), is attributable to two factors. One is the dominance of schema theory in L2 reading research, and the other, the prevailing fallacy that there is a strong correlation between high levels of oral proficiency and better word recognition skills. Consequently, the available results on repeated reading in L2 learners are limited and inconclusive, as revealed, for example, by Taguchi (1997) and Taguchi and Gorsuch (2002).

Taguchi (1997) conducted repeated readings with 15 students enrolled in a university in Eastern Japan. The participants, all native speakers of Japanese and beginning-to-intermediate learners of English, were assigned to practice repeated readings from English textbooks that matched their reading levels, three times a week over a ten-week period, for a total of 28 sessions of 30 minutes. In each session, they read a passage seven times (the first and the last reading serving as pretest and posttest), with the first time unassisted, the next three times assisted by an audio-taped model of the passage, and the remaining three times again unassisted, the reading time being recorded throughout. Results showed, among other things, a correlational increase of reading rate with the number of readings:

Three repeated readings assisted by an audio-taped model of reading significantly improved students’ silent reading rates. Moreover, the results indicate that the three more unassisted repeated readings, in which no model of reading was provided for the students, also significantly increased their silent reading rates. (p. 110)
Hence, “in this study, repeated readings are found to be equally effective in developing word recognition skills of (foreign language) readers within practiced passages as it is for L1 readers” (p. 112). The study, however, yields little evidence for transfer of improved reading rate to new, unpracticed passages.

In a follow-up study, Taguchi and Gorsuch (2002) measured comprehension and transfer as well as rate. This time around, an experimental group and a control group were constituted from 18 first-year Japanese students at a university in Japan who at the time of the study had studied English for a total of six years in junior and senior high schools. The experimental group read level-appropriate passages, while the control group read Power Builder cards from the SRA Reading Laboratory 2c (Parker, 1989). As in Taguchi (1997), the study had a pretest and posttest design, enacted respectively through the first and last session. Altogether 28 sessions were administered over 10 weeks, and their procedures replicated those of Taguchi (1997):

1. Students (in the experimental group) read the previous passage to remember what they had read in the last session.
2. Students timed their first reading of a new passage with a stopwatch.
3. Students read the passage three times while listening to the exact taped version with headphones.
4. Students read the passage silently three more times and timed each of their readings with a stopwatch. (pp. 51–52)

Reading rate was measured by words per minute (WPM), comprehension by comprehension questions, and transfer by rate and comprehension in relation to practiced and unpracticed passages. Results were intriguing:

1. The reading rate of the experimental group improved significantly from the initial reading of the pretest passage to that of the posttest passage. But the experimental group did not perform significantly better than the control group.
2. With regard to reading comprehension, the reading performances by the experimental group were not significantly different from those by the control group.
3. Transfer effects from repeated reading were not apparent.
   (Taguchi & Gorsuch, 2002, p. 58)

The inconclusive evidence from this study for L2 repeated reading can be accounted for in part by methodological weaknesses such as the pretest and posttests being non-equivalent and both being more difficult than the treatment passages (see, however, Gorsuch & Taguchi, 2008). But more importantly, this study underscores the need for further research with L2 readers, who, by nature, are fundamentally different from L1 readers (Koda, 1994; Shiotsu, 2009) but who nevertheless still face the same issues: phonemic awareness, phonics, fluency, vocabulary, and comprehension (Hu, 2008).

In summary, the extant research, more on first than on second language readers, has established repeated reading as a valid method for developing reading fluency. Three categorical findings are particularly worthy of note for their conceptual relevance to the present study. First, assisted
repeated reading, namely, reading while listening to a model, aids fluency more than unassisted reading. Second, repeated reading with feedback produces greater improvement on fluency and comprehension than reading without. Third, repeated reading without a fixed number of reading times produces greater improvement on fluency and comprehension than with fixed reading times. These findings informed the design of the present study, as will be described shortly. Next, let us turn briefly to relevant L2 research on reading and vocabulary acquisition.

L2 Reading and Vocabulary Acquisition

Reading, as noted earlier, is a critical source of vocabulary acquisition, particularly for adolescent and adult L2 learners, and particularly in a foreign language context, where reading is a primary vehicle for target language input (Han & Anderson, 2009). Whereas so far in neither first language nor second language research has there been direct research on repeated reading and vocabulary acquisition, there has been plenty on normal reading and vocabulary development in both realms, with some offering direct insights for the present study. For example, in L2 research, a recent study by Brown, Waring, and Donkaewbua (2008) showed that reading-while-listening is more effective than either reading or listening alone for incidental vocabulary acquisition, in which case, vocabulary acquisition is a byproduct of reading comprehension rather than a focus in its own right, the latter being intentional vocabulary acquisition. Laufer (2006) demonstrated that in a foreign language context, intentional vocabulary acquisition facilitated by explicit instruction is more effective than incidental vocabulary acquisition via comprehension-oriented reading alone (cf. de la Fuente, 2006). She argued that in a foreign language context, which does not offer massive exposure to L2 input, meaning-focused instruction alone is not adequate for vocabulary development and recommends supplementing it with explicit form-oriented instruction. There was compelling evidence in several studies (e.g., Laufer, 2003; Min, 2008; Paribakht & Wesche, 1999) that reading supplemented with vocabulary-enhancement activities is more effective for vocabulary gains and retention. Further, Lee and Muncie’s (2006) study showed that post-reading explicit instruction in conjunction with contextualized writing practice enables transfer of receptive to productive vocabulary knowledge. The study also suggested that using a combination of instruction strategies benefits not only vocabulary development but also retention. In addition, most studies have documented the positive effect of frequency in vocabulary acquisition (e.g., Grabe & Stoller, 1997; Pigada & Schmitt, 2006).

In an attempt to develop a unified theoretical construct of lexical competence, Henriksen (1999) proposes three (possibly successive) dimensions: (a) partial/precise knowledge, (b) depth of knowledge, and (c) receptive and productive knowledge. From this framework it follows that empirical investigations of vocabulary acquisition should seek to measure these aspects in order to achieve descriptive adequacy. In the present study, as described in the next section, a combination of vocabulary tasks were designed and used to tap these dimensions, through testing recognition and production of words in context and in isolation, intentionally and incidentally.

The Study
In the context of repeated reading, the present study investigated vocabulary acquisition—defined as knowledge of form and meaning mappings of individual words and operationalized as the ability to recognize and produce words in context and isolation. The study attempted to examine the scope and depth of vocabulary acquisition through repeated reading, and, therefore, both intentional and incidental learning were probed and measured. Following Hulstijn (2005), here “intentional learning refers to the learning mode in which participants are informed, prior to their engagement in a learning task, that they will be tested afterward on their retention of a particular type of information. Incidental learning refers to the mode in which participants are not forewarned of an upcoming retention test for a particular type of information” (p. 132).

The study was guided by three interlocking questions:

1. Can repeated reading lead to vocabulary acquisition?
2. If so, does it lead to intentional, incidental acquisition, or both?
3. What is the depth of acquisition? Is it limited to receptive knowledge or does it encompass both receptive and productive knowledge?

Subject

The subject, pseudo-named Anna, was a heritage speaker of Chinese living in the United States. At the time of the study, she was a college senior. Anna was a fluent speaker of Mandarin Chinese but her literacy skills in Chinese lagged far behind. Since the age of six, she had received, intermittently, instruction in Chinese, the aggregate amount of instructional time being approximately 200 hours over a period of 15 years. At the time of the study, she had acquired a written receptive vocabulary of approximately 1000 high-frequency words, assessed according to Dew (1999), of which her written productive knowledge was estimated at approximately 200 words. She had mastered the pinyin system, a system of romanization (phonetic notation and transliteration to roman script) for Mandarin Chinese, and was therefore able to mark the pronunciation of unknown words with ease.

Materials

The reading materials used in the present study consisted of two types: (a) theme-related passages and (b) an independent article. Both types were authentic, taken from a then current issue of Ren Min Hua Bao (人民画报). Type A comprised 10 short passages ($M = 249$ characters; $Range = 202–296$ characters), and Type B an extended narrative broken down into 10 segments ($M = 227$ characters; $Range = 170–295$ characters). There was no significant difference in length between Type A and Type B ($p = .19$) readings, according to a one-way ANOVA.

Both Type A and Type B were chosen in consultation with Anna, based primarily on her interests. The Type A passages were vignettes about individual gold medalists in the 2008 Olympic Games in Beijing, and the Type B article was about how economic changes in China over the last thirty years had changed ordinary people’s lives.

The average comprehensibility of Type A readings was 79% ($Range = 71%–85%$), with an
unknown word density of 21%. Hence, the texts were largely incomprehensible to Anna. According to research, “reading becomes arduous and comprehension suffers when unknown word densities exceed 5%” (Cobb, 2008, p. 113; cf. Laufer, 1989). The average comprehensibility of Type B readings was 71% (Range = 62%–77%), with an average unknown word density of 29%. Type B was therefore of even lower comprehensibility than Type A. A one-way ANOVA indicates a significant difference in the average comprehensibility between Type A and Type B ($p < .05$). Table 1 summarizes the length and comprehensibility of Type A and Type B readings.

<table>
<thead>
<tr>
<th>Type</th>
<th>Length (characters)</th>
<th>Comprehensibility (%)</th>
<th>Unknown word density (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A</td>
<td>249</td>
<td>79</td>
<td>21</td>
</tr>
<tr>
<td>Type B</td>
<td>227</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td>Mean total</td>
<td>238</td>
<td>75</td>
<td>25</td>
</tr>
</tbody>
</table>

It is worth mentioning that in the past, Anna had complained about textbook readings, which were primarily what she had experienced in her instructed learning of Chinese, commenting that they were boring and irrelevant. Therefore, in the present study, the decision to involve her in choosing readings was in part based on considerations of how to satisfy her cognitive needs and interests, and in part on insights from extant research on repeated reading suggesting, in particular, that the repeated reading method allows slow readers to process materials that are way beyond their current capacity. As it turned out, the materials chosen did meet her cognitive needs and motivated her to read.

An audio recording was made of each text read at normal pace by one of the researchers using a digital voice recorder. The recording served as a model for Anna’s oral reading later on. Each recording took less than five minutes to prepare. The recorded files were then sent via email to Anna, who also had a hard copy of the texts available to read from.

**Target Words**

For each passage/segment, the subject was asked to identify 4–5 words for memorization, for a total of 96 target words (197 characters) for the entire study. These self-selected words ran the gamut from nouns to verbs, adjectives, adverbs, and conjunctions. For example, Anna picked the following four words as the target words for Session 7: ‘连续’ (continuously), ‘失败’ (failure), ‘痛苦’ (agony), and ‘弥补’ (compensate). Typically, the words chosen were from the ones Anna could understand when heard orally, but occasionally, she also chose words that she could not understand initially, for example, ‘领军’ (leading).

**Procedure**

The study spanned 23 days, involving 20 treatment sessions and 3 testing sessions. Table 2 displays the schedule of the study. Each treatment session consisted of two phases, each phase lasting an hour. The first phase involved self-directed repeated reading of a text, aided by the audio-recording pre-prepared by one of the researchers, and the second phase entailed direct
interaction with the researcher. Specifically, in the first phase, Anna worked alone, performing the steps summarized below:

1. Reading while listening to the audio recording of the passage for as many times as deemed sufficient to be able to read alone fluently without listening
2. Singling out four to five words to focus on (i.e., target words) for understanding and memorization
3. Practicing writing the target words until they can be written with ease from memory

In the second phase, Anna met with the researcher who performed the following procedures:

1. Having Anna orally read the text, and providing corrective feedback on word recognition errors
2. Having Anna explain her interpretation of the text, and providing explanation if needed
3. Inviting Anna to ask questions about the text
4. Having Anna orally read the text again and providing feedback on word recognition errors as needed
5. Testing Anna’s memory of the target words via sentence dictation
6. Providing Anna with corrective feedback on the sentences that she wrote down
7. Concluding the session by having Anna orally read the text one more time

Throughout the treatment period, Anna was instructed to keep a reading journal focusing on (a) time spent on reading the text, (b) reflections on the reading experience, and (c) evaluation of the word-writing practice (e.g., hard or easy).

Table 2. Schedule of the study

<table>
<thead>
<tr>
<th>Day</th>
<th>Session</th>
<th>Task</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–20</td>
<td>1–20</td>
<td>Treatment</td>
<td>2 hours per session</td>
</tr>
<tr>
<td>21–23</td>
<td>21–23</td>
<td>Testing</td>
<td>10–20 minutes per session</td>
</tr>
</tbody>
</table>

Note. During the 3-day testing period, the reading materials were kept out of Anna’s reach.

Test Tasks

As noted earlier, the present study sought to examine both intentional and incidental vocabulary acquisition as a function of the treatment sessions. To that end, intentional vocabulary acquisition was operationalized as receptive and productive knowledge of the words that Anna had explicitly studied. On the other hand, incidental vocabulary acquisition was operationalized as receptive and productive knowledge of the words that Anna had not explicitly studied but had transliterated while doing her own oral reading in Phase 1. The intentional and the incidental acquisition were in turn measured via seven tasks, listed and described below in the order in which they were administered:

1. isolated production of intentional words (Task 1)
2. isolated recognition of intentional words (Task 2)
3. isolated production of incidental words (Task 3)
4. isolated recognition of incidental words (Task 4)
5. contextual comprehension of words (Task 5)
6. contextual production of intentional (Task 6a) and incidental words (Task 6b)
7. contextual recognition of intentional (Task 7a) and incidental words (Task 7b)

Task 1, isolated production of intentional words, was designed to contribute to an understanding of Anna’s intentional acquisition of vocabulary. Specifically, it measures Anna’s ability to produce the target words in isolation. The task was composed of 20 two-character words, randomly drawn from the database of target words. The instructions given were “Write out the corresponding words based on the pinyin given. 根据拼音写出相应的词语,” and five minutes were allotted to the task.

Task 2, isolated recognition of intentional words, designed to tap Anna’s recognition of intentional words in isolation, was composed of 20 two-character words, randomly drawn from the inventory of target words. The instructions were “Write out the corresponding pinyin based on the characters or words given. 根据所给词语写出相应的拼音,” and similarly, five minutes were allowed for task completion.

Tasks 3 and 4, isolated production of incidental words and isolated recognition of incidental words, were both designed to shed light on Anna’s incidental acquisition of vocabulary. Specifically, they measured respectively her ability to produce and recognize words in isolation. Both tasks had 20 items, constituted randomly of two-character words from the transliterations in pinyin (n = 1057 characters) that Anna herself had provided for words she did not know during the self-directed repeated reading phase of the study, (i.e., the first phase of the treatment sessions). The instructions she received for each task respectively were “Write out the corresponding words based on the pinyin given. 根据拼音写出相应的词语,” and “Write out the corresponding pinyin based on the words given. 根据所给词语写出相应的拼音.” For both tasks, five minutes were allowed.

Task 5, contextual comprehension of words, measured Anna’s comprehension of words whose meaning she had inquired about in the second phase of the treatment sessions in which she interacted with the researcher. The test consisted of 20 sentences, each containing an underlined word for interpretation. The words tested were all incidental words. Anna was asked to give their meanings (“解释以下划线词语的意思”) in five minutes.

Tasks 6 and 7, contextual production of words and contextual recognition of words, were intended to contribute to an understanding of Anna’s intentional and incidental acquisition of words in context. They measured respectively Anna’s ability to produce and recognize words in context. Task 6, measuring production, involved two cloze tests, adapted from two paragraphs of the texts that Anna had read in the treatment period, each containing 12 blanks, 6 of which targeted intentional words and 6 incidental words. Anna was asked to “Fill in the blanks with appropriate words. 请在以下空位里填进合适的词.” Task 7, measuring receptive knowledge in context, was similarly based in two paragraphs extracted from the texts that Anna had read, with 12 words underlined in each paragraph, 6 of which targeted intentional words and 6 incidental words.
words. The task required Anna to orally read the paragraphs (“朗读以下两段课文”). Ten minutes were allowed for completion of each task. Table 3 summarizes the test tasks.

**Table 3. Measurements tasks**

<table>
<thead>
<tr>
<th>Task</th>
<th>Intentional (words)</th>
<th>Incidental (words)</th>
<th>Comprehension checks</th>
<th>Time (minutes)</th>
<th>Item total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>20</td>
<td></td>
<td>5</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Task 2</td>
<td>20</td>
<td></td>
<td>5</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Task 3</td>
<td></td>
<td>20</td>
<td>5</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Task 4</td>
<td></td>
<td>20</td>
<td>5</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Task 5</td>
<td></td>
<td></td>
<td>20</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Task 6</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Task 7</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>64</td>
<td>20</td>
<td>45</td>
<td>148</td>
</tr>
</tbody>
</table>

The tests were purposely administered over three days, with tasks 1–2 on Day 1, tasks 3–5 on Day 2, and tasks 6–7 on Day 3, in order to minimize priming and fatigue effects.

**Data Coding and Analysis**

Task 1 was coded for productive knowledge of the target words in isolation. A correct response was awarded 2 points (e.g., ke3lian2/可怜), an incorrect response 0 point (e.g., jia4ge2/个), and a partial response 1 point (e.g., fa1zhan3/发). The maximum score was 40.

Similarly, Task 2 was coded for receptive knowledge of the target words in isolation. A correct response was awarded 2 points (e.g., jing1shen2/精神), a partial response 1 point (e.g., jiu4xi4/续), and an incorrect response 0 point. The maximum score was again 40.

Tasks 3 and 4 were coded similarly to tasks 1 and 2, for productive and receptive knowledge of the incidental words which Anna transliterated during the first, self-directed repeated reading phase of the treatment sessions. The maximum score was 40 respectively.

Task 5 was coded for comprehension of the words Anna had inquired about during the feedback phase (Phase 2) of the treatment sessions. A correct response was awarded 2 points, a partially correct response 1 point, and an incorrect response 0 point. The maximum score was 40. An example of a correct response is given below.

**Prompt:**
年底的世界杯总决赛，李小鹏再度在双杠比赛中折桂。

**Response:**
‘Again’

Tasks 6 and 7 were coded respectively for productive and receptive knowledge of the intentional and incidental words in context. For the sake of clarity, they will be reported in the next section as Task 6a for production of intentional words (target words) and as Task 6b for incidental words, on the one hand, and as Task 7a for recognition of intentional words and Task 7b for
incidental words, on the other. Consistent with the coding of the other tasks, here a correct response was awarded 2 points, a partially correct response 1 point, and an incorrect response 0 point. The maximum score was 24 points respectively for tasks 6a, 6b, 7a, and 7b. Figure 1 depicts schematically the tasks vis-à-vis their intention to measure intentional versus incidental words.

Figure 1. Tasks measuring intentional vs. incidental words

The data were coded independently by two coders, with 99% inter-coder agreement, and were subsequently analyzed statistically, using SPSS V.12. First, descriptive statistics were computed. Second, accuracy scores were calculated for each task. Third, one-way ANOVAs were carried out to compare means across the tasks. Fourth, the Friedman Test was conducted to rank-order the tasks.

Results

The descriptive statistics for the accuracy scores are displayed in Table 4, where Column 1 specifies the tasks, Column 2 the number of test items per task, Columns 3 and 4 the range of scores Anna earned on each task, Column 5 the mean accuracy scores, and Column 6 standard deviations. What is worth highlighting here is the different mean scores (Column 5) across the tasks, suggesting task-specific, differential accuracy. For example, Anna was more accurate on Task 2 (isolated recognition of intentional words; $Mean = 1.85$) than on Task 1 (isolated production of intentional words; $Mean = 1.35$).

Table 4. Descriptive statistics

<table>
<thead>
<tr>
<th>Tasks*</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>20</td>
<td>.00</td>
<td>2.00</td>
<td>1.35</td>
<td>.88</td>
</tr>
<tr>
<td>Task 2</td>
<td>20</td>
<td>1.00</td>
<td>2.00</td>
<td>1.85</td>
<td>.37</td>
</tr>
<tr>
<td>Task 3</td>
<td>20</td>
<td>.00</td>
<td>2.00</td>
<td>.90</td>
<td>.85</td>
</tr>
<tr>
<td>Task 4</td>
<td>20</td>
<td>.00</td>
<td>2.00</td>
<td>1.10</td>
<td>.85</td>
</tr>
<tr>
<td>Task 5</td>
<td>20</td>
<td>.00</td>
<td>2.00</td>
<td>1.60</td>
<td>.68</td>
</tr>
<tr>
<td>Task 6a</td>
<td>12</td>
<td>.00</td>
<td>2.00</td>
<td>1.50</td>
<td>.90</td>
</tr>
<tr>
<td>Task 6b</td>
<td>12</td>
<td>.00</td>
<td>2.00</td>
<td>.75</td>
<td>.97</td>
</tr>
<tr>
<td>Task 7a</td>
<td>12</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>.00</td>
</tr>
<tr>
<td>Task 7b</td>
<td>12</td>
<td>.00</td>
<td>2.00</td>
<td>1.25</td>
<td>.87</td>
</tr>
</tbody>
</table>

Note. *Refers to the task list on pp. 9–10.
The differential accuracy is amplified in Table 5, where both raw scores and accuracy percentages are given for each task. Thus, the following picture emerged as a result of the treatment. First, with respect to isolated production and recognition of intentional words, Anna’s accuracy rate was 68% (Task 1), which translates into 66 target words or 134 characters, and 93% (Task 2), which amounts to 90 target words or 183 characters. Where isolated production and recognition of incidental words are concerned, her accuracy rates were respectively 45% (Task 3), meaning that she was able to produce 476 characters, and 55% (Task 4), suggesting that she was able to recognize 581 characters in isolation. In respect of contextual comprehension of words, Anna achieved 80% accuracy (Task 5). In terms of contextual production of intentional and incidental words, her accuracy rates were 75% (Task 6a), amounting to 148 target words or 296 characters for intentional words, and 38% (Task 6b), equivalent to 402 characters for incidental words. Finally, her accuracy rates for contextual recognition of intentional and incidental words were respectively 100% (Task 7a), suggesting she was able to recognize all the target words (197 characters) in context, and 63% (Task 7b), which is equivalent to 666 incidental characters.

Table 5. Vocabulary gains  

<table>
<thead>
<tr>
<th>Tasks*</th>
<th>Raw score (Max.)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>27 (40)</td>
<td>0.68</td>
</tr>
<tr>
<td>Task 2</td>
<td>37 (40)</td>
<td>0.93</td>
</tr>
<tr>
<td>Task 3</td>
<td>18 (40)</td>
<td>0.45</td>
</tr>
<tr>
<td>Task 4</td>
<td>22 (40)</td>
<td>0.55</td>
</tr>
<tr>
<td>Task 5</td>
<td>32 (40)</td>
<td>0.80</td>
</tr>
<tr>
<td>Task 6a</td>
<td>18 (24)</td>
<td>0.75</td>
</tr>
<tr>
<td>Task 6b</td>
<td>9 (24)</td>
<td>0.38</td>
</tr>
<tr>
<td>Task 7a</td>
<td>24 (24)</td>
<td>1.00</td>
</tr>
<tr>
<td>Task 7b</td>
<td>15 (24)</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Note. *Refers to the task list on pp. 9–10.

Table 6 shows the hierarchical order of difficulty of the test tasks for Anna, according to the Friedman Test ($X^2(8) = 27.863, p = .001$).

Table 6. Hierarchical order of task difficulty  

<table>
<thead>
<tr>
<th>Tasks*</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 7a</td>
<td>6.88</td>
</tr>
<tr>
<td>Task 2</td>
<td>6.75</td>
</tr>
<tr>
<td>Task 6a</td>
<td>5.54</td>
</tr>
<tr>
<td>Task 5</td>
<td>5.50</td>
</tr>
<tr>
<td>Task 7b</td>
<td>4.71</td>
</tr>
<tr>
<td>Task 1</td>
<td>4.33</td>
</tr>
<tr>
<td>Task 4</td>
<td>3.96</td>
</tr>
<tr>
<td>Task 3</td>
<td>3.79</td>
</tr>
<tr>
<td>Task 6b</td>
<td>3.54</td>
</tr>
</tbody>
</table>

Note. *Refers to the task list on pp. 9–10.

Thus, as shown, Task 7a (contextual recognition of intentional words) was the easiest for Anna,
while Task 6b (contextual production of incidental words) was the most difficult. In between, Task 2 (isolated recognition of intentional words) was easier than Task 6a (contextual production of intentional words), which in turn was easier than Task 7b (contextual recognition of incidental words), and so forth. Similarly, Task 1 (isolated production of intentional words) was easier than Task 4 (isolated recognition of incidental words), which in turn was easier than Task 3 (isolated production of incidental words).

Table 7 shows ANOVA results from the tasks on which Anna’s performance was statistically different. Thus, for example, her performance was significantly different on Task 1 versus Task 2, with lower performance on Task 1 than on Task 2. Overall, the following patterns were observed. First, Anna was significantly more accurate in recognizing and producing target or intentional words/characters than incidental words/characters (e.g., Task 1 > Task 3; Task 2 > Task 4; Task 6a > Task 6b). Second, she was more accurate in recognizing than producing words/characters (e.g., Task 1 < Task 2). Third, she was more accurate in recognizing and producing words/characters in context than in isolation (e.g., Task 3 < Task 7a; Task 4 < Task 7b). Fourth, Anna was more accurate in comprehending words/characters that had been explained to her than in recognizing and producing incidental words/characters, in isolation and in context (e.g., Task 5 > Task 3; Task 5 > Task 6b).

Table 7. Tasks with significant differences

<table>
<thead>
<tr>
<th>Tasks*</th>
<th>Mean difference</th>
<th>SE</th>
<th>Sig.</th>
<th>95% confidence level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean difference</td>
<td>SE</td>
<td>Sig.</td>
<td>Lower bound</td>
</tr>
<tr>
<td>Task 1 &lt; Task 2</td>
<td>-.50</td>
<td>.23</td>
<td>.03</td>
<td>-.95</td>
</tr>
<tr>
<td>Task 1 &gt; Task 3</td>
<td>.45</td>
<td>.23</td>
<td>.05</td>
<td>-.00</td>
</tr>
<tr>
<td>Task 2 &gt; Task 3</td>
<td>.95</td>
<td>.23</td>
<td>.00</td>
<td>.50</td>
</tr>
<tr>
<td>Task 2 &gt; Task 4</td>
<td>.75</td>
<td>.23</td>
<td>.00</td>
<td>.30</td>
</tr>
<tr>
<td>Task 2 &gt; Task 6a</td>
<td>1.10</td>
<td>.26</td>
<td>.00</td>
<td>.58</td>
</tr>
<tr>
<td>Task 3 &lt; Task 6a</td>
<td>-.60</td>
<td>.26</td>
<td>.03</td>
<td>-1.12</td>
</tr>
<tr>
<td>Task 3 &lt; Task 7a</td>
<td>1.10</td>
<td>.26</td>
<td>.00</td>
<td>-1.62</td>
</tr>
<tr>
<td>Task 3 &lt; Task 7b</td>
<td>-1.10</td>
<td>.26</td>
<td>.00</td>
<td>-1.62</td>
</tr>
<tr>
<td>Task 4 &lt; Task 7a</td>
<td>-.90</td>
<td>.26</td>
<td>.00</td>
<td>-1.42</td>
</tr>
<tr>
<td>Task 4 &lt; Task 7b</td>
<td>-.90</td>
<td>.26</td>
<td>.00</td>
<td>-1.42</td>
</tr>
<tr>
<td>Task 5 &gt; Task 3</td>
<td>.70</td>
<td>.23</td>
<td>.00</td>
<td>.25</td>
</tr>
<tr>
<td>Task 5 &gt; Task 4</td>
<td>.50</td>
<td>.23</td>
<td>.03</td>
<td>.05</td>
</tr>
<tr>
<td>Task 5 &gt; Task 6a</td>
<td>.85</td>
<td>.26</td>
<td>.00</td>
<td>.33</td>
</tr>
<tr>
<td>Task 6a &gt; Task 6b</td>
<td>.75</td>
<td>.30</td>
<td>.01</td>
<td>.17</td>
</tr>
</tbody>
</table>

Note. *Refers to the task list on pp. 9–10. 
“<“ = “significantly worse than” 
“>“ = “significantly better than”

Discussion

The results from the study shed positive light on all three research questions, repeated below for convenience:

1. Can repeated reading lead to vocabulary acquisition?
2. If so, does it lead to intentional, incidental acquisition, or both?
3. What is the depth of acquisition? Is it limited to receptive knowledge or does it encompass both receptive and productive knowledge?

First, the results indicated that repeated reading did lead to vocabulary acquisition in Anna. Second, it led to gains in both intentional words and incidental words (see Table 5). Third, it led to acquisition of both receptive and productive knowledge (see, again, Table 5). However, the gains were differential: As shown in Table 6, the tasks were of unequal ease or difficulty for Anna, and as shown in Table 7, she demonstrated unequal gains from the pedagogical treatment. As reported in the previous section, she was significantly more accurate in recognizing and producing intentional words/characters than incidental words/characters; her recognition was more accurate than production; and she was better at recognizing and producing words in context than in isolation.

The gains were, in fact, quite impressive, made apparent by converting the accuracy rates into concrete numbers of characters, as shown in Table 8.

<table>
<thead>
<tr>
<th>Tasks*</th>
<th>Accuracy (%)</th>
<th>Number of characters acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>0.68</td>
<td>134 (intentional)</td>
</tr>
<tr>
<td>Task 2</td>
<td>0.93</td>
<td>183 (intentional)</td>
</tr>
<tr>
<td>Task 3</td>
<td>0.45</td>
<td>476 (incidental)</td>
</tr>
<tr>
<td>Task 4</td>
<td>0.55</td>
<td>581 (incidental)</td>
</tr>
<tr>
<td>Task 5</td>
<td>0.8</td>
<td>846 (incidental)</td>
</tr>
<tr>
<td>Task 6a</td>
<td>0.75</td>
<td>148 (intentional)</td>
</tr>
<tr>
<td>Task 6b</td>
<td>0.38</td>
<td>402 (incidental)</td>
</tr>
<tr>
<td>Task 7a</td>
<td>1.00</td>
<td>197 (intentional)</td>
</tr>
<tr>
<td>Task 7b</td>
<td>0.63</td>
<td>666 (incidental)</td>
</tr>
</tbody>
</table>

*Refers to the task list on pp. 9–10.

According to Table 8, after the treatment Anna was able, for instance, to produce 134 out of the 197 intentional (or target) characters in isolation (Task 1), suggesting that she had transferable productive knowledge of these characters; she was able to recognize 666 (Task 7b) out of the 1057 incidental characters, suggesting that she had context-specific receptive knowledge of these many characters, and so forth. These figures combined, in turn, would translate into productive and transferable knowledge of 15 characters per two-hour session or 300 characters for a total of 20 sessions or 40 instructional hours, on the one hand, and receptive and context-specific knowledge of 43 characters per session or 860 characters for the entire treatment period of 40 hours, on the other.

The study therefore showed that repeated encounter—with frequency and consistency—with the same text helps acquisition of vocabulary. Moreover, as reported in previous research, repeated practice with multiple texts rather than a single text, apparently, helps knowledge transfer (Moyer, 1982), as evident in Anna’s performance on the recognition/production-in-isolation tasks. Outside the repeated reading paradigm, a recent study by Bolger, Balass, Landen, and Perfetti (2008) provides compelling evidence of the superiority of exposure to variable contexts in word learning to exposure to a single context. The researchers subsequently claim that word...
learning results from abstraction across varied word encounters (for L2 research, see Cobb, 1999; Folse, 2004).

Another prominent finding from the present study is the fact that Anna was, overall, better at recognizing and producing intentional (target) words than incidental words. This suggests that greater and, importantly, self-motivated attention to words facilitates vocabulary acquisition. It also lends further support to the by now widely accepted Involvement Load Hypothesis (Hulstijn & Laufer, 2001), according to which words are better retained if the task demands greater investment from the learner in terms of need, search, and evaluation, and, in turn, to depth of processing as a key variable in retention (Craik & Lockhart, 1972; Leow, Hsieh, & Moreno, 2008).

Additionally, the study demonstrated the capability of repeated reading in aiding the processing of difficult texts, again as reported in the previous research (e.g., Hyönä & Niemi, 1990; Schreiber, 1980). In the present case, the difficult written input and, in particular, the unfamiliar words were made comprehensible through repeated reading while listening to the taped model. The ‘phonological code’ as embodied in the recording might have facilitated segmentation of meaningful units (Koda, 1994). More pertinently, when the visual (i.e., the written text) and the phonological (i.e., listening to the recording) encoding of information happened in tandem, thereby activating both the phonological loop and the visuo-spatial sketchpad of working memory (Baddeley, 1986, 2003), it may have doubled the chance of the words being accurately processed and retained.

Not all the reported gains in the present study can be attributed to repeated reading alone, however. It may be recalled that, as in many previous studies (see the earlier section on repeated reading), in the present study the treatment invoked repeated reading along with several other strategies, notably, corrective feedback and writing practice. Thus, it is likely that each of these strategies has, too, made a contribution to the reported gains. While corrective feedback in this case may have played a key role in forming the right mental representations of the form, meaning, and function of the new words, the repeated writing practice (mostly, in isolation) with self-selected target words may have helped their retention. According to Swain’s (1985, 1995) Output Hypothesis, output has the potential of performing a number of functions, such as stimulating metalinguistic reflection, promoting fluency, and triggering noticing, all of which may have occurred as a result of the character-writing process that Anna underwent. Moreover, though an empirical question, it is likely that given the dual fact of Chinese written language being ideographic and of its lack of sound-grapheme correspondence, writing practice could have played a greater role in word retention in the present study than would have with an alphabetic language such as English.

While teasing apart the contributions made by the various components of the pedagogical intervention would be a task for future research, it is important to recognize that in its current composite form, repeated reading proved to be motivating and capable of promoting both text comprehension as well as vocabulary acquisition. Throughout the study period, Anna exhibited high motivation to read, this notwithstanding the fact that the reading passages were far beyond her then capacity to decode (see Table 1) and that the challenge had persisted as a result of continued and increased density of unknown words from Type A to Type B readings. Her
comprehension was, in fact, outstanding; she was even able to critically point out and comment on logical inconsistencies in the extended narrative (Type B readings).

**Conclusion**

The present study explored repeated reading as a fulcrum for vocabulary acquisition, providing evidence that repeated reading, buttressed by corrective feedback and writing practice, can lead to dramatic gains in receptive and productive knowledge of intentional and incidental vocabulary. Additionally, the study corroborates a categorical finding from previous research with first language readers (e.g., Moyer, 1982), namely that repeated reading can allow the reader to process texts that are substantially beyond his or her current decoding capacity. This finding is particularly important for instructional settings involving heritage speakers, where choice of instructional materials often presents a seemingly insurmountable hurdle. Due to the chasm between their cognitive and linguistic abilities, heritage speakers are often discouraged by the reading materials matching their linguistic ability. With repeated reading, as demonstrated in the present study, this chasm can easily be bridged. Here cognitive ability, relevance, and personal interest are the overriding considerations in materials selection.

Focusing on one heritage speaker, the present study is, admittedly, limited of which findings should not be (over)generalized. Apart from the single subject concern, the study, due to logistic constraints, did not examine the issue of long-term retention. It therefore remains unclear as to whether or not the reported gains, in part or in total, were durable. While future research must look into this issue, it can be predicted that retention is contingent upon continued interaction. Thus, if Anna kept up her interaction with written Chinese, she would be able to retain and even multiply the gains from the pedagogical intervention; otherwise, she would lose, at least, some.

Future research, as alluded to in the previous section, must also employ experimental designs, controlling for various components of the repeated reading package as used in the present study, to isolate their contributions to, as well as assess their interactional effects on, vocabulary acquisition. While the current study came away with encouraging findings by examining one heritage speaker—a putatively unusual case of ‘second’ language learning (Montrul, 2008), future research can be enriched by expanding the research avenue to include typical second and foreign language learners.

**Acknowledgments**

We wish to thank the anonymous reviewers for their thoughtful comments and constructive suggestions on an earlier version of the manuscript. Any errors that remain are exclusively ours.

**References**


rationale for pedagogy (pp. 98–122). Cambridge: Cambridge University Press.

About the Authors

ZhaoHong Han is Associate Professor of Language and Education at Teachers College, Columbia University, where she teaches a variety of courses ranging from theory to methodology to pedagogy and mentors doctoral students. She has published a number of books and articles on second language acquisition and education, including Fossilization in Adult Second Language Acquisition (2004), Reading Research and Instruction: Crossing the Boundaries (with Neil Anderson, 2009), and Linguistic Relativity in SLA: Thinking for Speaking (with Teresa Cadierno, 2010). Email: han@tc.columbia.edu
Cheng-ling Alice Chen is a doctoral student in the TESOL program at Teachers College, Columbia University. Her research interests include second language acquisition, particularly the effect of the repeated reading approach on second language vocabulary instruction and learning. Email: cc2550@columbia.edu