

Vocabulary learning through reading: Does an ELT course book provide good opportunities?

Warren Matsuoka & David Hirsh
University of Sydney
Australia

Abstract

This study investigates the vocabulary learning opportunities in an ELT course book designed for upper-intermediate learners. All the words appearing in the 12 chapters of the text were analyzed. The results suggest that the text would provide opportunities to deepen knowledge of the second 1,000 most frequent words in English, and would provide a context for pre-teaching of academic words met in the text for learners on an academic pathway. The results also suggest that the text would provide minimal opportunities for learners to develop vocabulary knowledge beyond high frequency and academic words. The findings demonstrate a need to supplement use of such texts with an extensive reading program and other forms of language rich input to promote vocabulary development.

Keywords: vocabulary learning, ELT course book, vocabulary demands, lexical coverage, frequency, range, repetition, extensive reading

The scholarship of Paul Nation spans more than 30 years and has heightened our understanding and appreciation of the role of vocabulary learning in second language development. Paul's sustained research interest is an inspiration to others, and serves as a starting point for a broad range of research avenues related to vocabulary and reading.

Vocabulary Knowledge and Reading Comprehension

There is a strong link between vocabulary knowledge and reading comprehension. This is the case in first language (L1) contexts (Anderson & Freebody, 1981) and in second language (L2) contexts (Laufer, 1991). In studies of L2 learners, which is the focus of this paper, Koda (1989) revealed high correlations between vocabulary knowledge and two measures of reading, close tasks and paragraph comprehension, while Coady, Magoto, Hubbard, Graney and Mokhtari (1993) found an increase in L2 reading proficiency can be attributed to increased proficiency in vocabulary.

There has been a research interest in defining the vocabulary size required for second language reading. A study by Hwang and Nation (1995) showed that knowledge of the 2,000 most frequent word families enables L2 readers to recognize 84% of the words in various types of authentic texts. This falls well short of the 95% of words which Laufer (1991, 1992) indicated needs to be recognized for unassisted reading, and which would require a minimum vocabulary size of 3,000 word families. More recently, Laufer (1997) suggested that the 95% lexical threshold in L2 reading would be required to enable L2 readers to apply their L1 reading strategies.

While research suggests that comprehension of texts may be achieved with a vocabulary size of 3,000 word families, a vocabulary size of 5,000 word families is needed to attain 98% lexical coverage of texts, allowing for more pleasurable reading and more accurate guessing of unknown words in their context (Coady et al., 1993; Hirsh & Nation, 1992; Laufer, 1997). Hu and Nation (2000) investigated the effect of four different text coverage levels (80%, 90%, 95%, and 100%) on unassisted reading comprehension of a fiction text and found that reading comprehension increased as density of known words increased. They found that while some readers required 95% lexical coverage for adequate comprehension, most required 98%.

The current study examines the vocabulary demands and vocabulary learning opportunities in *New Headway Student's Book Upper-Intermediate* (Soars & Soars, 2005), an ELT course book written for upper-intermediate learners. While unassisted reading of authentic texts may require up to 98% lexical coverage, it is assumed that the lower lexical threshold of 95% could enable assisted comprehension of ELT course books which, unlike authentic texts, are designed specifically to be used in a classroom language learning environment with language teacher support including pre-teaching of target vocabulary items.

Repetition and Vocabulary Learning

In considering the use of textbooks for vocabulary development, two related features of vocabulary occurrence in texts can be highlighted in terms of their role in promoting incidental vocabulary acquisition.

The first feature is repetition. Word repetition is a favorable condition in vocabulary learning (Nation, 2001). Research on the effect of word repetition on vocabulary learning has focused on three variables: (1) the number of repetitions, (2) spacing of repetitions, and (3) types of repetitions. No set number of repetitions of a word guarantees its learning (Huckin & Coady, 1999). However, research (see Nation & Wang, 1999; Saragi, Nation, & Meister, 1978; Webb, 2007) suggests a target of 10 repetitions for learning of unknown words. Single long continuous texts such as novels (Hirsh & Nation, 1992), separate texts related through topic such as newspaper reports on the same event (Hwang & Nation, 1989), and separate expository texts on a similar narrow theme (Gardner, 2008) provide more favorable repetitions of unknown vocabulary than is the case in unrelated texts, and thus provide improved conditions for vocabulary acquisition. This is due to the effect of incremental acquisition of repeated vocabulary while reading.

The second feature is spaced repetition. Research in memory (Baddeley, 1990) and L2 vocabulary learning (Bloom & Shuell, 1981; Dempster, 1987) has found that spaced repetition (i.e., the spreading out or ‘spacing’ of repetitions of a word throughout a text) is more conducive to the learning of vocabulary than repetition that is massed (i.e., the concentration of repetitions of a word in only one part of a text). Massed repetition refers to repeated attention to a word over a continuous period of time; say 6 minutes, while spaced repetition refers to giving the same amount of attention to a word over a longer period of time, such as 2 minutes on three occasions over a 2-week period. The study time of the word is 6 minutes in total in both cases, but the repetitions are ‘massed’ or ‘spaced’ depending on the approach. The current study examines patterns of repetition of unknown words in an ELT course book in light of this understanding of what constitutes favorable conditions for vocabulary learning.

Depth and Breadth of Vocabulary Knowledge

Vocabulary learning can be viewed in two dimensions, breadth and depth (Qian, 2002; Read, 1988; Wesche & Paribakht, 1996). Breadth of vocabulary knowledge refers to the number of words one has at least superficial knowledge of (i.e., one’s vocabulary size) and can be calculated in terms of recognition, recall or production of vocabulary items. Depth of vocabulary knowledge, on the other hand, refers to how well one knows a word. Nation (2001) provides a framework for describing aspects of word knowledge associated with vocabulary learning, covering form, meaning and use in both receptive—reading and listening—and productive—writing and speaking—contexts. Although increasing both breadth and depth of vocabulary knowledge is regarded as important to improving reading comprehension (Qian, 1999, 2002; Read, 1993), breadth has attracted more attention than depth in L2 empirical studies to date. An aim of the current study is to investigate the opportunities an ELT course book provides for promoting both depth and breadth of vocabulary knowledge.

Text-Based Studies Using Word Frequency Lists

The current study builds primarily on three previous text-based studies that employed computer-generated data based on word frequency lists to examine vocabulary demands and vocabulary learning opportunities in texts.

In the first of these studies, Hwang and Nation (1989) examined running stories in newspapers (i.e., a news story and follow-up stories). They regarded proper nouns such as names of people and places as familiar vocabulary, because: (1) meanings of many proper nouns have already been learned in the L1 so can be easily inferred in the L2, and (2) explanations of proper nouns that are less commonly known are frequently given in context in news stories. In analyzing the 80 news articles containing accumulatively 26,722 running words, it was found that the 2,000 most frequent words provide 80% text coverage while proper nouns account for an additional 10% coverage. They also found a reduced vocabulary load—the density of new words—in subsequent related stories due to the effect of vocabulary learning from repetitions of new vocabulary.

Following on from this, Hirsh and Nation (1992) analyzed three short novels written for young L1 readers. They found that the 5,000 most frequent word families plus knowledge of proper nouns provided 97-98% text coverage enabling pleasurable reading of such texts.

More recently, Nation and Wang (1999) investigated the vocabulary learning opportunities in a series of graded readers. They found that (a) if a learner systematically read several graded readers at each of the different levels, most words in the series would be frequently met, and (b) learners should read one graded reader each week to ensure new words are met again within a short period of time. They suggested that graded reader series should go up to the 5,000 word level to help the learner make the successful transition from graded readers to authentic texts.

Other text-based studies have examined vocabulary occurrence in secondary school textbooks (Klinmanee & Sopprasong, 1997), writing samples of EFL and ESL students (Laufer & Nation, 1995; Morrow, 1997), information pamphlets (Muhammad, 1990), a variety of types of input including novels, newspaper text, a graded reader and a movie (Nation, 2006), a combination of press, academic, and fiction texts (Cobb, 2007), and television programs (Webb & Rodgers, 2009). To date, no published studies have investigated vocabulary occurrence in commercial ELT course books.

ELT course books are widely used in language learning classrooms (Hutchinson & Torres, 1994; Littlejohn, 1998). There is some concern, however, that ELT course books contradict rather than reflect contemporary developments in applied linguistics (Hutchinson & Torres, 1994; Sheldon, 1988). In addition, ELT course books may not always be supplemented with language input from other written sources such as graded readers, novels or newspapers. Tanaka and Stapleton (2007) report that Japanese EFL high school learners tend to engage in intensive reading using short texts rather than extensive reading. This is in spite of research suggesting a positive role for extensive reading in second language development, including vocabulary development (Horst, 2005; Lee, 2007), particularly when the extensive reading occurs in conjunction with other language-rich learning activities such as intensive reading (see Laufer, 2003).

ELT course books have been evaluated for their effectiveness in learner training (i.e., teaching learners how to learn) (Lake, 1997; Sinclair & Ellis, 1992), and their flexibility, design and teachability (Masuhara, Hann, Yi, & Tomlinson, 2008; Tomlinson, Dat, Masuhara, & Rubdy, 2001). The current study examines the realm of vocabulary use in ELT course books through two research questions, the first informing the second:

1. What are the vocabulary demands of reading an ELT course book?
2. What vocabulary learning opportunities are provided in an ELT course book?

Method

The current study provides an in-depth investigation of one ELT course book to provide a methodology and initial findings to inform future studies of the type. The text, *New Headway Student's Book Upper-Intermediate* (Soars & Soars, 2005), is an integrated skills student book designed for upper-intermediate level language learners. The series at six levels from beginner to

advanced includes at each level a teacher's book, student workbook, CD of the students' book listening exercises and CD for supplementary self-study listening practice. The selected text, the student's book, has 12 chapters which present short reading passages, vocabulary and grammar tasks, and information to learners on related skills work including listening and writing tasks. All the words appearing in each chapter in the students' book were analyzed. The total number of running words in the 12 chapters is 44,877.

Data Analysis

Research question 1 was examined by calculating the text coverage for the first 1,000 most frequent words, second 1,000 words, proper nouns, technical, textual and technology words, and academic words. This was carried out to examine what vocabulary size would provide close to 95% lexical coverage which could allow for guessing of unknown words in context.

Research question 2 was examined by counting repetitions of word families in the second 1,000 word list (West, 1953), the Academic Word List (Coxhead, 2000), and lower frequency words. Data were collected for 5, 7 and 10 occurrences of words in the text to investigate opportunities for (1) deepening the knowledge of the second 1,000 word list, and (2) learning new words beyond the 2,000 word list (i.e., academic words from the AWL and other words).

A case study was conducted of six of the second 1,000 words which occurred at least 5 times in the text to examine the type of repetitions a learner would meet in terms of word form and collocations. The number of word families from the list of the second 1,000 wordlist, which were repeated 5 or more times in the text was found to be 187. In order to choose six words randomly from these 187 word families, the words were sorted by increasing frequency of occurrence, and 187 was divided by 6 to get 31.2 resulting in the selection of every 31st word starting from the 187th word up. The following six words were thus randomly selected: *breakfast, invite, hotel, island, crash* and *conversation*.

The analysis for this study involved the use of a computer program, Range (Heatley, Nation, & Coxhead, 2002), which compares vocabulary in up to 32 texts against baseword lists (see following description) to generate data on word frequency (i.e., number of occurrences of a word in a text) and range (i.e., number of texts a word occurs in) in a series of texts. The program reported on word type frequency (i.e., the number of times a word type occurs in the input texts), and word family frequency (the number of times a word and its family members occur in the input text) for separate chapters and the combined text. Range also indicates the text coverage provided by baseword lists.

Baseword Lists

Baseword lists used with Range allow for the classification of words in a text according to word families. Three existing baseword lists were used. The first list is made up of the most frequent 1,000 words of English which mostly include words with a frequency higher than 332 occurrences per 5 million words as well as numbers, letters, months, days of the week, titles (e.g., *Mr., Ms.*) and common greetings (e.g., *hi, hello*) (Heatley, Nation, & Coxhead, 2002). The second baseword list represents the second 1,000 most frequent English words. Both lists are

sourced from West's (1953) General Service List and appear in the current study as baseword lists 1 and 2. The third existing baseword list used is Coxhead's (2000) Academic Word List (AWL), appearing in the current study as baseword list 7.

In the baseword lists, word families are categorized and listed under headwords. For example, the headword *need*, which is listed in baseword list 1, includes the family members *needed*, *needing*, *needless*, *needs* and *needn* (apostrophes are recognized by the program as a space so the family member *needn* is considered the contraction of 'need' and 'not'). A word family in the context of reading can be defined as a headword such as *need* and all of its derived (e.g., *needless*) and inflected (e.g., *needed*, *needing*, *needs* and *needn't*) forms that a learner is assumed to understand without having to learn all the word forms separately (Bauer & Nation, 1993). The baseword lists include British and American spellings of English words. For example, the headword *favour* includes *favor*, *favourite* and *favorite*.

Modification of the Text

All pages of each chapter of the text were scanned and changed to text files using ABBYY FineReader 5.0 Sprint Plus. It was found that the meaning of certain hyphenated compound words could be inferred by their word parts but would be mistakenly listed as unknown unlisted words by Range. To remedy this, a hyphenated compound word was separated to be read as individual tokens by Range if it met two criteria: (1) the meaning of the compound word could be correctly inferred by its word parts, and (2) all word parts of the compound words are in one of the baseword lists. Hyphenated words separated included *foreign-owned*, *prize-winning* and *year-old*. Hyphenated words not separated included *absent-minded* and *washing-up*. A simpler but less principled approach would be to separate all hyphenated words.

Additional Baseword Lists

Studies of this type typically rely on existing baseword lists. Use of the three existing baseword lists in this study, the first 1000, second 1000 and AWL lists, generated a list of other words which included many which the researchers considered could be familiar to upper-intermediate learners with a knowledge of the most frequent 2,000 words of English.

It was thus deemed necessary to create additional baseword lists for words considered to be known or easily understood by learners at upper-intermediate level of English. Decisions were made by one of the authorial team who has taught EFL extensively over the past 12 years. This resulted in four additional baseword lists to represent proper nouns, technical words, textual words and technology words. The baseword lists were numbered 1-7 based on their relevance to the focus of the main study (see Table 1). The first two lists are the first and second 1,000 most frequent words lists, respectively. Baseword list 3 is proper nouns, baseword list 4 is technical words, baseword list 5 is textual words, baseword list 6 is technology words, and baseword list 7 is the Academic Word List.

Table 1. *Order of baseword lists*

Baseword list	Type of words
1	First 1,000 most frequent words
2	Second 1,000 most frequent words
3	Proper nouns
4	Technical words
5	Textual words
6	Technology words
7	Academic Word List

The four baseword lists developed for this study are as follows:

Baseword list 3—proper nouns: A proper nouns baseword list was created in line with previous text-based studies (Hirsh & Nation, 1992; Hwang & Nation, 1989). A total of 504 proper nouns were found which were sorted into 465 word families to form baseword list 3. An example of a proper noun entry is *Korea*, *Korean* and *Koreans* listed under the headword *Korea*.

Baseword list 4—technical words: Metalinguistic terms such as *adverb* were considered to be technical words that the upper-intermediate reader should have already learned in previous stages of language learning. Twenty-three technical words were found in the unlisted words list and were arranged into 16 word families to create baseword list 4. This list included terms such as *adjective*, *adverb*, *comma*, *pronoun* and *vocabulary*.

Baseword list 5—textual words: A total of 10 words related to the text structure were categorized into 8 textual word families. The list included words such as *bracket*, *column*, *heading*, *italics* and *tapescript*.

Baseword list 6—technology words: A technology word list was created to group words related to technology such as *email* and *television* which are seen as likely to be part of most learners' daily routine and thus could be considered words known. As the first 2,000 words are sourced from the General Service List (West, 1953), recent frequently used additions to the language such as technology words may not be included in the first and second 1,000 word lists. Eighteen technology words were found and categorized into 13 word families to make baseword list 6. The list included *download*, *DVD*, *email*, *internet*, *online*, *PC*, *software*, *television*, *video*, *website*, and *WWW*.

Modification to Baseword Lists 1, 2, and 7

The original first 1,000, second 1,000 and AWL baseword list word families were developed according to levels of affixation and derivation from Bauer and Nation (1993). The output file using Range indicated a number of items which needed to be added to the baseword lists 1, 2, and 7 families to ensure consistency in defining word families. In all, 50 new family members appearing in the text were added to baseword list 1 (i.e., the first 1,000 word list), 26 to baseword list 2 (the second 1,000 word list), and 3 to baseword list 7 (the Academic Word List) following the principles for developing the original word families (Heatley, Nation, & Coxhead, 2002). Examples of these additions are *builder* added to the word family *build* (first 1,000 list), *impolite*

added to the word family *polite* (second 1,000 list), and *trendy* added to word family *trend* (Academic Word List).

Results and Discussion

Research Question 1: What are the Vocabulary Demands of Reading an ELT Course Book?

The text coverage was calculated for the first and second 1,000 word lists, plus proper nouns, technical, textual and technology words, to see if this vocabulary size would provide 95% text coverage (see Table 2). The combined text coverage for this vocabulary size is 93.4%.

Table 2. *Text coverage of the first and second 1,000 most frequent words and assumed known words*

Baseword lists	Percent text coverage (%)	Cumulative text coverage at key levels (%)
First 1,000 word list	83.1	83.1
Second 1,000 word list	6.9	90.0
Proper nouns	2.8	92.8
Technical words	0.3	93.1
Textual words	0.1	93.2
Technology words	0.2	93.4
Academic Word List	2.1	95.5

To illustrate the vocabulary burden on the learner with 93.4% text coverage, an excerpt of the text is displayed with unknown words replaced with XXXX.

It's so cold here that people don't use freezers. They hang their meat in XXXX bags on nails above their windows. Spring and summer, when they arrive in June, last a mere 8 weeks. The Bering Sea, one of four seas that wash against Chukotkan shores, freezes hard enough to support weights of up to 35 tons. There's no crime because it's just 'too XXXX cold'!

Familiarity with other words is required to reach 95% text coverage. This could be attained through pre-teaching of academic vocabulary which provide 2.1% text coverage. If we look at the text coverage on a chapter by chapter basis (see Table 3), it is evident that coverage figures vary across chapters. In Table 3, assumed known words are baseword lists 3-6 (proper nouns, technical, textual and technology words).

Table 3 shows the effect that pre-teaching of academic words from the AWL would have on text coverage of assumed known words. The text coverage of the AWL ranges from 0.9% to 3.6%, with a mean coverage across 12 chapters of 2.1%. Pre-teaching of the AWL could enable text coverage of above 95% in 9 out of 12 chapters. The average coverage across the 12 chapters with inclusion of knowledge of academic words in the count would be 95.5%.

An opportunity is provided in the text for learners to focus on academic words. This would improve reading comprehension and provide a good return for learning effort for students on an

academic pathway because academic words represent a specialized vocabulary in academic texts (Coxhead, 2000).

Table 3. *Text coverage of individual chapters for the 2,000 words and assumed known words, and academic words*

Chapter	Text coverage of 2,000 words plus assumed known words (%)	Text coverage of the AWL (%)	Text coverage of 2,000 words plus assumed known words plus the AWL (%)
1	94.5	1.7	96.2
2	93.6	2.1	95.7
3	91.8	2.0	93.8
4	89.9	2.8	92.7
5	95.1	1.9	97.0
6	91.9	3.6	95.5
7	96.3	0.9	97.2
8	91.2	2.5	93.7
9	93.7	2.1	95.8
10	94.0	2.0	96.0
11	94.1	2.1	96.2
12	94.4	1.9	96.3

Research Question 2: What Vocabulary Learning Opportunities are Provided in an ELT Course Book?

The vocabulary learning opportunities provided in the text were investigated in two ways: (1) examining how often words were repeated in the text, and (2) conducting a case study of six randomly selected second 1,000 words that occurred at least 5 times in the text.

Words in the second 1,000 most frequent words list, the Academic Word List (AWL) and the lower frequency unlisted words list that occurred at least 5, 7, and 10 times in the text were counted to see if there were any opportunities in the text for deepening the knowledge of partially known second 1,000 words, for learning academic words, or for learning other words appearing in the text.

Deepening knowledge of second 1,000 words: Table 4 indicates that a total of 603 second 1,000 word families appeared in the text. Of these 603 word families, 187 (or 31%) occurred at least 5 times, 128 (21.2%) at least 7 times, and 73 (12.1%) at least 10 times. In contrast, 201 (33.3%) words from the second 1,000 word list occurred only once.

Table 4. *Repetition of word families in the course book text*

Occurrences	Second 1,000 most frequent words (603 word families found)	Words from the AWL (271 word families found)	Unlisted words (1,005 word families found)
Only once	201 (33.3%)	116 (42.8%)	667 (66.4%)
5 times or more	187 (31.0%)	44 (16.2%)	72 (7.2%)
7 times or more	128 (21.2%)	24 (8.9%)	36 (3.6%)
10 times or more	73 (12.1%)	15 (5.5%)	13 (1.3%)

One-third (33.3%) of the second 1,000 words in the text occur only once (e.g., *shallow, cheat, mystery*), and thus offer poor opportunities for deepening word knowledge. However, significant numbers of words occur at least 5 times (e.g., *excuse, informal, argue*). There appear to good opportunities in this text for a deepening of knowledge of words from the second 1,000 word list based on frequency of occurrence.

Learning of academic words: As Table 4 shows, 271 academic word families from the AWL occur in the text. 42.8% of these occurred only once (e.g., *flexible, reluctance, accumulate*) while 16.2% occurred 5 times or more (e.g., *partner, similar, comment*). Although one encounter with a new word can lead to partial learning of word meaning (Nagy & Anderson, 1984) reflecting the incremental nature of vocabulary acquisition, the amount of exposure to many of the academic words in the text may be too small for incidental learning to occur due to the abstract nature of many academic words. For these words to be acquired, direct pre-teaching is suggested.

Learning of unlisted words: 1,005 unlisted words were found in the text. 66.4% occurred only once (e.g., *cope, groceries, stubborn*) while 7.2% occurred 5 times or more (e.g., *interview, conspiracy, tiny*). It appeared that there were few opportunities for incidental learning of unlisted words.

Case study of second 1,000 words: Building on data for repetitions of second 1,000 words, the current study also examines the nature of repetitions of a random sample of these words to further appreciate the scope for deepening knowledge of such words. A case study of six randomly selected second 1,000 words occurring 5 times or more in the text was conducted to see how many aspects of vocabulary knowledge depth are presented to the reader with each occurrence of the selected word. Table 5 summarizes the frequency and range of the case study words as well as what aspects of depth of word knowledge were presented in the text for each selected word.

Table 5. *Case study results*

Selected words (frequency in parentheses)	Number of chapters they occur in (frequency in parentheses if more than 1 per chapter listed)	Part of speech (frequency in parentheses)	Number of inflections and derivations	Number of collocations
BREAKFAST (5)	R = 3: 1, 8, 12(×3)	noun	1 inflection: breakfast	4
INVITE (7)	R = 4: 4(×2), 6, 7(×3), 12	verb (6), noun (1)	3 inflections: invite, invites, invited 1 derivation: invitation	5
HOTEL (8)	R = 2: 2(×7), 5	noun	2 inflections: hotel, hotels	9
ISLAND (10)	R = 3: 1(×2), 2(×7), 4	noun	3 inflections: island, islands, island's	11
CRASH (15)	R = 4: 3(×2), 4(×11), 5, 7	noun (12), verb (3)	2 inflections: crash, crashed	14
CONVERSATION (70)	R = 11: 1 (×13), 2(×5), (×3), 4(×7), 5(×9), 6(×2), 7(×10), 8(×6), 10(×5), 11(×4), 2(×6)	noun	2 inflections: conversation, conversations	19

The words were selected to examine the nature of word occurrence across the text, and the random sampling may not have identified the strongest examples for deepening word knowledge within the second 1,000 word list.

The entry for *invite* in Table 5 provides an example of data collected on depth of word knowledge in the case study. The word family *invite* occurred 7 times in 4 different chapters, 6 times as a verb (e.g., *invites*, *have invited*, *invite*, *is invited*, *invited*) and once as a noun (*invitation*). In all, 5 collocations of the word family *invite* were found: *invite someone to do something*, *invite someone somewhere*, *be invited*, *want to invite someone to do something*, *reply to someone's invitation*. 3 inflections of *invite* were presented (*invite*, *invites*, *invited*) in multiple tenses (simple present, simple past, present perfect, present passive) exposing the learner to variable uses and meanings of the word. One of its derivations (*invitation*), which was the only derivation of any of the six selected words found in the text for this study, was presented and collocated with the verb phrase *reply to*.

Overall, Table 5 indicates different aspects relating to the depth of word knowledge. Repetitions of five of the six randomly selected words occurred in 3 or more chapters (i.e., with a range of 3 or higher). This pattern represents spaced repetition (i.e., the spreading out of the repetitions of a word throughout the text) rather than massed repetition (i.e., the concentration of the repetitions of a word in only one part of the text). This pattern of spacing of repetitions may lead to longer retention of the word in the learner's memory (Baddeley, 1990; Bloom & Shuell, 1981; Dempster, 1987). For the word families *invite* and *crash*, two to three different inflections were presented giving the reader more opportunities to learn different forms and uses of these words. *Invitation* was the only derivation of any of the six selected words to be presented in the text. This is not surprising given that very few affixes could be attached to the six words (e.g., *islander*, *conversationalist*, *hotelier*). The low number of different inflections, derivations and parts of speech of the selected words presented in the text was partly due to the nature of the words examined. For all of the six randomly selected words except the word *conversation*, there was almost one new collocation presented with every occurrence of the word. This suggests that the text provides good opportunities to deepen the knowledge of these partially known words by introducing different and possibly new collocations of the selected words to the reader. Concerning the word *conversation*, there were 70 occurrences of this word (the highest frequency of the six selected words), with 19 collocations.

Overall, the case study suggests favorable conditions for deepening knowledge of frequently occurring words in the second 1,000 word list in terms of frequency, spaced repetitions and variable collocations provided with new meetings of the words.

Conclusions

This study has suggested that, in order to adequately comprehend the ELT course book and guess unknown words from context, the reader requires knowledge of the first and second 1,000 most frequent words, familiarity with or recognition of proper nouns and a small number of technical, textual and technology words, and pre-teaching of academic words occurring in the text. Direct teaching of academic vocabulary in this context would provide good return for learners on an

academic pathway. Determining the vocabulary readiness of learners for the vocabulary demands of the text could be achieved through use of a diagnostic test such as the Vocabulary Levels Test (Nation, 2001).

The text provides favorable opportunities for deepening knowledge of the second 1,000 word list due to frequent and spaced repetitions of many words in various forms and with a range of collocations. The second 1,000 words list is under-researched in the literature. Future research into the presence of the second 1,000 words in texts could determine to what extent these words assume general meanings independent of the subject matter or assume more technical meanings associated more closely with subject matter. Research could also investigate the extent to which groups of learners develop deep knowledge of these words at different stages of their learning.

The text was found to offer few opportunities to expand vocabulary knowledge beyond the first 2,000 words and academic words. The data reported here assumed complete use of the text. In reality, teachers may select and omit parts of the text. Thus, the vocabulary learning opportunities in a classroom setting may be lower than those reported here.

A significant outcome of this research is a call to return to the issue of the role of extensive reading in EFL classrooms. While extensive reading alone is not suggested as a model for EFL vocabulary development (Laufer, 2003), it has been shown to lead to significant gains in vocabulary learning beyond high frequency words (Horst, 2005; Lee, 2007). Research continues into the relative importance of language input and instructional techniques on vocabulary acquisition, pointing thus far to an important role for both in the language classroom (see Laufer, 2009). In light of this, there would appear to be a strong case, when selecting ELT course books for classroom use, to supplement use of the text with extensive reading of graded readers or other suitable reading material. This extensive reading program would ideally be in combination with direct vocabulary instruction (Nation & Waring, 1997) and instructionally-enhanced reading to highlight vocabulary use (Hulstijn, Hollander, & Greidanus, 1996; Zahar, Cobb, & Spada, 2001) in order to provide a rich context for vocabulary development to equip learners with the lexical resources required to understand and use the language effectively in a range of communicative contexts.

References

- Anderson, R. C., & Freebody, P. (1981). Vocabulary and knowledge. In J. T. Gutrie (Ed.), *Comprehension and teaching: Research review* (pp. 77–117). Newark, DE: International Reading Assn.
- Baddeley, A. (1990). *Human memory*. London: Lawrence Erlbaum Associates.
- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6, 253–279.
- Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74, 245–248.
- Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes & J. Coady

- (Eds.), *Second language reading and vocabulary learning* (pp. 217–228). Norwood, NJ: Ablex.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11, 38–63.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238.
- Dempster, F. N. (1987). Effects of variable encoding and spaced presentation on vocabulary learning. *Journal of Educational Psychology*, 79, 162–170.
- Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading. *Reading in a Foreign Language*, 20, 92–122.
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range computer program. http://www.vuw.ac.nz/lals/staff/Paul_Nation
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689–696.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review*, 61, 355–382.
- Hu, H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403–430.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21, 181–193.
- Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80, 327–339.
- Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal*, 48, 315–328.
- Hwang, K., & Nation, I. S. P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6, 323–335.
- Hwang, K., & Nation, I. S. P. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, 23, 35–41.
- Klinmanee, N., & Sopprasong, L. (1997). Bridging the EFL vocabulary gap between secondary school and university: A Thai case study. *Guidelines*, 19(1), 1–10.
- Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22, 529–540.
- Lake, N. (1997). Survey review: Learning training in EFL course books. *ELT Journal*, 51, 169–182.
- Laufer, B. (1991). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Basingstoke: Macmillan.
- Laufer, B. (1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading*, 15, 95–103.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge: Cambridge University Press.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59, 565–585.

- Laufer, B. (2009). Second language vocabulary acquisition from language input and form-focused activities. *Language Teaching*, 42, 341–354.
- Laufer, B., & Nation, I. S. P. (1995). Lexical richness in L2 written production: Can it be measured? *Applied Linguistics*, 16, 307–322.
- Lee, S.-Y. (2007). Revelations from three consecutive studies on extensive reading. *RELC Journal*, 38, 150–170.
- Littlejohn, A. (1998). The analysis of teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 190–216). Cambridge: Cambridge University Press.
- Masuhara, H., Hann, N., Yi, Y., & Tomlinson, B. (2008). Adult EFL courses. *ELT Journal*, 62, 294–312.
- Morrow, P. R. (1997). Patterns of lexis in the writing of Japanese EFL students: A comparative study. *Pasaa*, 27, 87–99.
- Muhammad, R. (1990). *The vocabulary demands of reading information pamphlets*. Unpublished MA Thesis, Victoria University of Wellington, New Zealand.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304–330.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and speaking? *Canadian Modern Language Review*, 63, 59–82.
- Nation, I. S. P., & Wang, M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 1999.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge: Cambridge University Press.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282–308.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, 513–536.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19(2), 12–25.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355–371.
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6, 72–78.
- Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, 42, 237–246.
- Sinclair, B., & Ellis, G. (1992). Survey: Learning training in EFL course books. *ELT Journal*, 46, 209–225.
- Soars, L., & Soars, J. (2005). *New headway student's book (Upper-Intermediate)*. UK: Oxford University Press.
- Tanaka, H., & Stapleton, P. (2007). Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading Matrix*, 7, 115–131.
- Tomlinson, B., Dat, B., Masuhara, H., & Rubdy, R. (2001). EFL courses for adults. *ELT Journal*, 55, 80–101.

- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46–65.
- Webb, S., & Rodgers, M.P.H. (2009). Vocabulary demands of television programs. *Language Learning*, 59, 335–366.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53, 13–40.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, 57, 541–572.

About the Authors

Warren Matsuoka recently commenced his PhD in the Faculty of Education and Social Work, University of Sydney. This paper is based on a dissertation completed as part of his Master of Education (TESOL) studies at the University of Sydney. Warren has previously taught EFL extensively in Japan and Taiwan. His current research interests include vocabulary studies, English language education in Asia, and language use in L2 classrooms.

E-mail: wmat7124@uni.sydney.edu.au

David Hirsh is a lecturer in TESOL at the University of Sydney. His research focuses on vocabulary studies, language assessment and academic acculturation. Recent publications are *Teaching academic writing* (University of Michigan Press, 2009, with Paltridge et al.), *Researching vocabulary* in the edited volume *Companion to research methods in applied linguistics* (Paltridge & Phakiti, 2010, Continuum) and articles in *French Review of Applied Linguistics* and *Asian EFL Journal* (forthcoming). E-mail: david.hirsh@sydney.edu.au