

Reading a whole book to learn vocabulary

Paul Nation

Victoria University of Wellington

This paper investigates whether it is a good idea to choose a book that interests you and read it through from the beginning to the end learning all the new words you meet. For the analysis, it is assumed that learners already know the most frequent 3,000 words of English. The criteria used to guide this investigation include the number of unknown words met, the usefulness of the unknown words, the density of the unknown words, and the number of repetitions of the unknown words. Reading a whole book intensively is not a good idea unless the book is a graded reader, a technical text in a relevant subject area, or a set text that would be examined as a part of assessment. Where learners need to do such reading of unsimplified texts, they should be strategic in dealing with unknown vocabulary.

Keywords: word learning, unknown words, word learning strategy

When I studied French reading knowledge at university, the course consisted of intensive reading of a single book, *L'Etranger* by Albert Camus. The teacher led us through the book page by page explaining the vocabulary and grammar points as we went. Later when I did independent study of the Indonesian language, one of my activities was to slowly work my way through *Revolusi di Nusa Damai* (Revolt in Paradise) by K'tut Tantri with the help of a dictionary. Later I read a linguistics book called *Morfologi* by M. Ramlan, which turned out to be much easier if less interesting. When I have heard others talk about their foreign language learning, it seems that this intensive reading of a whole book is a fairly common practice. It is possible to buy books that contain parallel texts, one text in English and the other in the original language of the book. One of Burling's (1968) "outlandish proposals" for learning a language involved reading a proposed specially designed book that started off being completely in English but which gradually and systematically transformed into the foreign language, so that by the end of the book the learner was reading completely in the foreign language.

This intensive reading approach to learning another language seemed like a good idea at the time to me (as many things do), but much later as I learned more about vocabulary I wondered if it was such a good idea.

1. Vocabulary learning from reading

There is a considerable amount of research on vocabulary learning from reading (see Nation (2013 Chapters 5 & 8) for a review of the literature). A distinction is often made between incidental learning and deliberate learning. Incidental learning through reading occurs when the reader's attention is focused on the ideas content of the book and vocabulary is picked up through guessing from context. Deliberate learning occurs when the reader looks up the meaning of a word or when some deliberate attempt is made to remember it such as putting it on a word card or meeting it later in a related vocabulary-focused exercise such as filling the blanks.

Incidental learning results in learning small numbers of strongly known words and reasonable numbers of words at various strengths of knowledge (Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Waring & Takaki, 2003). Repetition has a strong effect on learning from reading, and the small amounts of strongly known words learned can become large amounts with substantial quantities of reading (Nagy, Herman & Anderson, 1985; Nation, 2014a). Adding a deliberate element to the learning greatly increases learning (Fraser, 1999; Mondria, 2003; Paribakht & Wesche, 1993; Sonbul & Schmitt, 2009), with vocabulary exercises having a stronger effect than dictionary look-up (Laufer & Rozovski-Roitblatt, 2011). Reading-while-listening results in a little more vocabulary learning than reading by itself, and either of these two activities are much more effective than listening by itself (Brown, Waring & Donkaewbua, 2008).

Although reading can be a very useful means of vocabulary learning, especially if it is accompanied by deliberate vocabulary learning, reading your way through a long difficult book may not be an efficient way of doing such learning.

This article therefore explores the following question.

Is it a good idea to choose a book that interests you and read it through from the beginning to the end learning all the new words you meet?

2. Arguments against reading a whole book

Let us look at the reasons why it is not a good idea first. Generally, it is not a good idea if the book is not written in a controlled vocabulary, or if the book is not on a topic that you already know a lot about. The main reason why it is not a good idea is that it will contain a large number of unknown words, most of which will be well outside your current knowledge and which will not be repeated in the book or even in the next book you read.

1. **Number of unknown words:** Depending on your vocabulary size, there will be a very large number of unknown different words. If, for example, the book is an average length novel (around 100,000 words long), there will be well over 1,000 different words beyond the 4,000 word level. This is a very heavy vocabulary burden for any reader.
2. **Usefulness of unknown words:** In addition, the unknown words will be spread across a wide range of frequency levels in the language. That is, some words will be words that you should learn because they are from the frequency level just beyond your present level of proficiency. Others however will be words from way beyond your present proficiency level. It would be much more useful to learn other more generally useful words than these. Some of these words will be topic-related words and will occur often in the text, but these are not a big number (Hwang & Nation, 1989; Schmitt & Carter, 2000). They are worth learning for this text, so if you meet a word two or three times and you feel you will meet it again in the text, you should learn it.
3. **Density of unknown words:** The density of unknown words in the text will be high, meaning that you will probably meet an unknown word in almost every one or two lines of the text. This can make reading very laborious.
4. **Repetitions:** While a relatively small number of words will occur many times in the book, around half of the different words in the book will occur only once, and two thirds of the different words will occur only once or twice. This means that for most of the unknown words, there will not be enough repetitions to support learning. This distribution of words is described by Zipf's law, and is typical of all coherent texts and collections of texts. See Sorell (2012) for a very clear description of Zipf's law, and see also Nation (2013, pages 32–35) for its applications to teaching and learning.

For these reasons, reading a book from beginning to end and learning every word is generally not a great idea. The reading would involve too much misdirected and largely unproductive vocabulary work. There are more effective and efficient ways of going about vocabulary learning.

Let us look at some evidence for these four negative points. Table 1 compares vocabulary statistics for five different novels, each cut to be exactly 100,000 tokens long to allow easy comparison. They were all not too far from 100,000 words long in their original form. The five novels examined are:

1. *Captain Blood* by Rafael Sabatini (1922)
2. *Babbitt* by Sinclair Lewis (1922)
3. *A Necessary Evil* by Alex Kava (2006)
4. *A Thousand Suns* by Alex Scarrow (2006)
5. *The Devil's Company* by David Liss (2009).

The choice of the texts was somewhat random but they were all roughly the same length, written by different authors, and included some recently written novels and a couple of older classics.

The analysis is done assuming a reader who already knows the most useful 3,000 word families of English. A reader knowing fewer words would have a much more difficult task. The 3,000 word level was chosen because this consists of the high frequency words of English (Schmitt & Schmitt, 2014), and is the typical end point of series of graded readers. The mid-frequency words include the 4th 1,000 to the 9th 1,000, and the words beyond that level are the low frequency words of English. Table 1 summarizes the results. Note that the last row in the table deals with the low frequency words.

Table 1. Vocabulary statistics for five different novels each cut to 100,000 tokens long

	1 Blood	2 Babbitt	3 Evil	4 Suns	5 Devil
Total different words	5,210	6,487	4,477	4,920	4,474
Different words beyond 3,000 level	2,690	3,314	1,740	2,086	1,980
One-timers beyond 3,000 level	1,579	1,365	866	970	1,025
Words beyond 3,000 level occurring 5 or more times	431	369	180	284	215
Unknown words per 100 tokens *	8.49%	9.23%	4.96%	6.34%	5.18%
Different words beyond 9,000 level	640	672	265	322	384

* assuming 1st 3,000 word families are known

The unit of counting in this analysis is the word family, so this increases the chance of repetitions because the occurrence of a different family member is counted as a repetition of that word family. Using word families also reduces the number of different words because all the members of a family are counted as totaling one word family. Word families thus provide a somewhat optimistic pic-

ture of the vocabulary load of the texts. Here is an example of a typical high frequency word family from the second 1,000 words. Words beyond the 3rd 1,000 on average contain less than six family members. The less frequent the word, the fewer the family members.

access

accessed

accesses

accessibility

accessible

accessing

inaccessibility

inaccessible

The word lists used in the analysis are the BNC/COCA lists available from Paul Nation's web site. *Total different words* refers to the total number of word families in each novel and include high frequency words, mid-frequency words, low frequency words and other words including proper nouns, transparent compounds and marginal words. In Table 1, in terms of the four reasons described above, *Different words beyond the 3,000 level* looks at the number of unknown words. These range from 1,740 to 3,314 different word families – a very large number of unknown words to deal with. *One-timers beyond 3,000 level* and *Words beyond 3,000 level occurring 5 or more times* look at repetition. The one-timers (words that only occur once on the novel) average over a 1,000. These are words that would need to be guessed or looked up but which would not be met again in that book. Dealing with these unknown words would be largely wasted effort. *Words beyond 3,000 level occurring 5 or more times* counts the words which stand a chance of being learned because they are repeated. Although five repetitions are probably not enough for learning (at least 12 would be better), there are relatively small numbers of words occurring more than five times and very few indeed occurring more than 12 times. If we look over all five books, there are only 126 word families beyond the 3,000 word level occurring in every book, and only 1,215 occurring in three or more of the five books. This is out of a total of over 9,000 word families beyond the 3,000 level in the five books. Clearly, although some unknown words are repeated, most do not get enough repetitions to support learning. *Unknown words per 100 tokens* looks at the density of unknown words. This percentage coverage figure does not include proper nouns, marginal words or transparent compounds. Two of the books have moderately manageable densities of around 5% (1 unknown word in every two ten word lines), while the other three have higher densities of unknown words, requiring dealing with an unknown word or two in at least every sentence. *Different words beyond 9,000 level* involves the usefulness

of the unknown vocabulary. These range from 265 to 672. The words from the 4,000 to 9,000 word levels are called mid-frequency words (Schmitt & Schmitt, 2014), and are seen as being the most useful words to learn after knowing the high-frequency words (the first 3,000). Most native speakers of English beginning secondary school already know these words. The words beyond the 9,000 level are called low-frequency words. For learners with a vocabulary size of 3,000 words these words are generally of very low priority. Each book contains several hundred of these, and learning them at this stage of knowledge is largely misdirected effort.

The data in Table 1 shows that the vocabulary burden of a novel is too heavy in a range of aspects for a reader with a small vocabulary size.

3. How many words do you have to know before reading becomes easier?

Although there are many factors affecting the difficulty of reading, the more words you know, the easier it will be to read. For learners of English, a vocabulary size of around 3,000 word families would still leave at least one word in every 20 as an unknown word. That is around one unknown word in every two lines of text (see Table 2). That is still a high density of unknown words and a lot of unknown words in a typical novel (1,980 word families in Table 2). A vocabulary size of 9,000 word families means that less than one word in every 50 running words will be unknown. This is about one in every five lines or six per 300 word page – an easily manageable density. As Table 2 shows, a reader with a vocabulary size of 2,000 words would not know 2,570 word families occurring in the book which on average would be a density of three unknown words in every two 10 word lines of text. As vocabulary size increases the number of unknown words and the density of unknown words drops, but even with a vocabulary size of 5000 word families, there would be 1,265 unknown word families in the book and just under three unknown words in every 10 lines.

Research (Hu & Nation, 2000; Schmitt, Jiang & Grabe, 2011) and common-sense suggests that at least 98% coverage by known words is a desirable known word density, and thus 2% or fewer unknown words (one unknown token in every 50 tokens) would be helpful for reading while still allowing the opportunity for some useful unknown words to be met.

As Table 2 suggests, a vocabulary size of at least 7,000 words is needed before the vocabulary burden of reading, in terms of the density and number of unknown words, is reduced enough to make dealing with unknown vocabulary not too burdensome.

Table 2. Density and number of unknown words in a typical novel at various vocabulary sizes* (*The Devil’s Company* by David Liss)

Vocabulary size	Unknown word density	Number of unknown words
1,000 word families	14.20%	3,326
2,000	8.12%	2,570
3,000	5.18%	1,980
4,000	3.61%	1,551
5,000	2.90%	1,265
6,000	2.34%	998
7,000	1.93%	826
8,000	1.56%	663
9,000	1.34%	532

* Proper nouns, marginal words, transparent compounds and acronyms are counted in known words

4. Are there any books that make it easier to learn new vocabulary?

We have seen that the factors of a large number of unknown words, the low usefulness of the unknown words, a high density of unknown words, and a lack of repetition of the unknown words make reading a whole book a time-consuming, difficult, and inefficient method of learning vocabulary. If there were books where at least some of these negative factors did not occur, then vocabulary learning through reading would be easier. The factor of lack of repetition cannot be significantly changed within a book because normal text seems to follow Zipf’s law, meaning that there will always be a substantial number of words that occur only once or twice in the book. Whenever a story moves on or new facts are dealt with, new vocabulary is needed to talk about these ideas. The only way to avoid Zipf’s law is say the same sentence made up of different words over and over again. This results in a very repetitive and boring text, but every word in the text will be repeated exactly the same number of times. However, because normal texts need to develop ideas, additional vocabulary is needed.

There are texts, however, that deal very effectively with the other three factors of number of unknown words, usefulness of the words, and unknown word density. These are texts that are written within a strictly controlled vocabulary, that is, any words outside a particular vocabulary level are not used in the book. These books are called graded readers and most major ELT publishers have one or more series of such readers, with several different vocabulary levels within each series, and with several texts at each level. For example, Oxford University Press has several series of graded readers. One of them, the Oxford Bookworms series, has six

levels of 400 words, 700 words, 1,000 words, 1,400 words, 1,800 words and 2,500 words. So, a learner with a vocabulary size of less than 400 words can read a level one book and meet only a few unknown words, and these unknown words will be words that are well worth learning now.

If a text is written within a controlled vocabulary, and the vocabulary lists that are used to guide the control are reasonably well made (as most are), then it becomes easy to deal with the factors of number of unknown words, usefulness of the words and unknown word density. Because each book is written within a controlled vocabulary, the number of unknown words is low. This of course also depends on readers choosing books at the right level for them. Because the number of unknown words is low, the density of unknown words is also low. Learners also know that almost any word that they meet within the book will be a useful word for them and is well worth noting for deliberate learning. Graded readers thus make vocabulary learning much easier largely because they contain only useful vocabulary at the right level for the learner. Graded readers have only a small effect on vocabulary repetition, but this issue can be dealt with by reading several books at the same level before moving on to the next level. Nation and Wang (1999) suggest reading about five at a level for the Oxford Bookworms series.

Graded readers for learners of English focus largely on the high frequency words of English – the most frequent 3,000 words. There are however free mid-frequency readers that are adapted to the 4,000, 6,000, and 8,000 word levels for learners who are more advanced in their reading and vocabulary knowledge (Nation & Anthony, 2013). These readers are freely available from Paul Nation's web site.

Graded readers are well worth reading from cover to cover, not only because their design makes them very suitable for vocabulary growth, but also because many are very well written, gripping stories (see the Extensive Reading Foundation language learner literature awards at <<http://erfoundation.org/>>).

5. Arguments in favour of working your way through a difficult book

We have looked at the disadvantages of reading your way through a difficult book and one of the alternatives, graded reading. Let us now look at the advantages of working your way through certain kinds of books. Such reading is a good idea, if one or more of the following conditions apply.

1. The book is a set text for your course. If you are required to read the book and will be examined on this reading then you have no choice. Where English is taught as a foreign language and this requirement exists, it is not unusual to see an L1

translation of the book quickly become available for sale. Reading an L1 translation first can help bring the L2 reading task closer to the reader's experience (Nation, 2007) and make the reading a bit easier. We look at other ways of making such reading easier in the next section of this article,

2. The book is an important text in the academic field that you are studying. For example, you may be studying economics and need to read an economics text in your L2. This is a very important reason for doing such reading. Reading the book will not only provide you with important information about your content area, but it also will help you learn the technical vocabulary of your field. This technical vocabulary is highly likely to occur in other related texts you may need to read, and some of it will be repeated quite often. That is, the two factors of repetition and the usefulness of the vocabulary will be favourable. Research on technical vocabulary (Chung & Nation, 2003; Quero, 2015) shows that the density of technical vocabulary in a technical text is usually very high, with roughly one word in every three to five running words being a technical word. Similarly, the number of technical words in a technical text also is high (several hundred to several thousand different words depending on the subject area). Balanced against this is the overall smaller vocabulary of cohesive texts compared to the diversity of vocabulary in a collection of unrelated texts (Sutarsyah, Nation & Kennedy, 1994; Quero, 2015). A text in a single topic area typically contains about half the number of different words than a similarly sized collection of texts on a wide variety of topics. Working your way through an important text in your field is a worthwhile thing to do, both for content and language.

3. The subject matter of the book is very familiar to you and the technical terms are similar to those in your L1. The greater the amount of relevant knowledge you bring to a text, the easier it will be to read the text. Substantial knowledge of the content and at least partial familiarity with the technical vocabulary will greatly reduce the difficulty of reading a text in another language. Tests of English as a foreign language show that learners do much better on cognates or loan words than on words that have no form and meaning relationship with their L1 (Elgort, 2013; Cobb, 2000). Such words are more like known words than unknown words even though they may not have been previously met or learned as L2 words. False cognates may be misleading but there are typically many more true cognates than false cognates. Cognates and loan words reduce the number and density of unknown words in a book. They are useful enough items to learn because the effort involved is minimal.

4. The book you have chosen to read is one that you have already read in your first language, so you can more easily guess or skip a lot of the unknown words. This situation affects the density and number of unknown words, but not the use-

fulness or repetitions of the unknown words. It is still not an ideal reading situation, but easier than reading without background knowledge. There are fortunately additional ways of reducing the difficulty of reading a text and we will now look at these.

6. What can you do to make it easier to read an unsimplified text which is really important for you to read?

We have already looked at the strategy of reading an L1 version first.

You can buy an electronic version of the text or scan a hard copy you have bought to make an electronic version. In some countries scanning a text will break copyright. You should not distribute a scanned electronic version to other people as this will definitely break copyright. Fortunately a very large number of texts in electronic form which are out of copyright are available free from Project Gutenberg <www.gutenberg.org>. An electronic version of the text can be read in an electronic reader like Kindle or Kobo that allows very quick dictionary look-up or a translation in one of several languages simply by touching a word. You can also set Kindle to provide a meaning in a small font above the word (the setting is called *word wise*). See Nation (2014b page 20) for instructions on sending an electronic text to your Kindle app to read.

If you are a teacher and have to provide a set text for your learners to read, you might consider adapting the text first if you have an electronic copy of it. The freely available program AntWordProfiler makes the adaptation relatively painless, and a full length novel for example can be adapted in a few hours. The adaptation involves replacing the unrepeatable low frequency words with already known words. Once you set the word level that you want to adapt the text to, the program quickly finds all the words you may need to change and allows you to make the changes there and then. The program also provides you with a vocabulary analysis of the text and tells you how often each selected word is repeated in the text, so that you can decide whether you want to replace a particular word or leave it there because it is repeated enough times to have a chance of being learned. AntWordProfiler was used to make the Mid-frequency Readers (Nation & Anthony, 2013).

7. Is re-reading the same book a good idea?

So far, our focus has been on making a text easier to read so that not only is new content understood, but vocabulary learning has a chance to occur. If vocabulary learning is the major concern, then doing a large quantity of reading can provide

reasonable opportunities for vocabulary learning. This would involve reading the equivalent of between three to twenty novels per year, depending on your vocabulary size. In the early stages of language learning, this involves around a half an hour to an hour per week doing meaningful reading, and this time should increase as proficiency develops (Nation, 2014a). But, is it worthwhile sometimes re-reading a book you have recently read so that you can meet the same words again and thus have a chance of strengthening your knowledge of them?

To answer this question we need to understand some of the conditions that favour the learning of vocabulary. Table 3 lists these conditions (Webb & Nation, 2017).

Table 3. Vocabulary learning conditions and example activities

Initial occurrence/Repetition		
Quality of attention	Incidental attention	Deliberate attention
Noticing	Guessing from context Noticing a gap when speaking or writing	Text highlighting Focusing on a form or meaning using word cards Dictionary or glossary look up Being taught words
Receptive or productive retrieval	Meeting a previously met word while listening or reading and recalling its meaning Recalling and using a recently met word in conversation or writing	Retrieval using word cards Doing vocabulary exercises after reading a text Recalling using a word wall
Varied meetings or use	Meeting a previously met word in a new form or context while listening or reading and recalling its meaning Recalling and using a recently met word in a new way in conversation or writing	Looking at a concordance Filling the blanks True/false sentences
Elaboration	Meeting and using a word in genuine high stakes communication Meeting and using a word in relation to pictures Interactive reading (Shared blown-up book)	Using the keyword technique Semantic mapping Word part analysis Focusing on words in teacher-led intensive reading Find the core meaning

Vocabulary learning depends on two major conditions – repetition and the quality of the meetings with words. The more repetition, the more likely learning is to occur. The richer the quality of the meetings with words, the more likely learning is to occur. In Table 3, the quality of meetings increases as you go down the table. So, successfully retrieving the meaning of a word you have met before will have a stronger effect than just noticing it, and meeting it again in a new con-

text (varied retrieval) will have a stronger effect than meeting it again in exactly the same context (verbatim retrieval).

The advantages of re-reading a book are

1. it will be a lot easier than reading it the first time
2. it will guarantee repetition of the vocabulary
3. it will provide an opportunity for retrieval to occur (verbatim retrieval)
4. it may help develop reading fluency
5. it will take much less time than reading a different book of the same length.

The advantages of reading a different book are

1. some of the words met in the previous book will be met again in the current book
2. the previously met words will be met in different contexts (varied meetings)
3. there will be some new words to learn.

These lists suggest that re-reading is a good idea, and that it would be good to have a mixture of some re-reading and a lot of different reading in a language learning program.

8. Conclusion

In this article we have examined a technique for vocabulary learning, reading your way through a book. To evaluate the technique we looked at information about the nature of language, in particular Zipf's law, and information about the conditions needed for vocabulary learning, namely repetition and the quality of encounters with words. We also looked briefly at research evidence for vocabulary learning from reading. The general conclusion was that learning all the vocabulary in a particular book was not a good idea unless the book was a graded reader, a technical text in a relevant subject area, or a set text that would be examined as a part of assessment. In situations where the text must be read, it is worthwhile for learners to be strategic in their vocabulary learning, concentrating on the higher frequency unknown vocabulary and giving these words some deliberate attention, and quickly passing over the many less useful words. Making use of electronic resources while reading will make the task much easier.

All language learning techniques and activities can be examined in this way (Webb & Nation, 2017). The value of such an examination is that it can lead to the improvement of the activity or to its replacement by a more efficient and effective technique. Sustained reading of a difficult text is best replaced by extensive graded reading for low and intermediate proficiency learners.

References

- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136–163.
- Burling, R. (1968). Some outlandish proposals for learning a foreign language. *Language Learning*, 18(1–2), 61–76. <https://doi.org/10.1111/j.1467-1770.1987.tb00390.x>
- Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(2), 103–116.
- Cobb, T. (2000). One size fits all? Francophone learners and English vocabulary tests. *Canadian Modern Language Review*, 57(2), 295–324. <https://doi.org/10.3138/cmlr.57.2.295>
- Elgort, I. (2013). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing*, 30(2), 253–272. <https://doi.org/10.1177/0265532212459028>
- Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21, 225–241. <https://doi.org/10.1017/S0272263199002041>
- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Hwang, K., & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6(1), 323–335.
- Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research*, 15(4), 391–411.
- Mondria, J. A. (2003). The effects of inferring, verifying and memorising on the retention of L2 word meanings. *Studies in Second Language Acquisition*, 25(4), 473–499. <https://doi.org/10.1017/S0272263103000202>
- Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253. <https://doi.org/10.2307/747758>
- Nation, I. S. P. (2014a). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26(2), 1–16.
- Nation, I. S. P. (2014b). *What do you need to know to learn a foreign language?* Available free from <<http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>>.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd edn). Cambridge: Cambridge University Press.
- Nation, I. S. P. (2007). Vocabulary learning through experience tasks. *Language Forum*, 33(2), 33–43.
- Nation, I. S. P., & Anthony, L. (2013). Mid-frequency readers. *Journal of Extensive Reading*, 1(1), 5–16.
- Nation, P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355–380.
- Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL programme. *TESL Canada Journal*, 11(1), 9–27.
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do *Things Fall Apart*? *Reading in a Foreign Language*, 22(1), 31–55.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1–28.

- Quero, B. (2015). *Estimating the vocabulary size of L1 Spanish ESP learners and the vocabulary load of medical textbooks* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand.
- Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, 9(1), 4–9.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43.
<https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.
<https://doi.org/10.1017/S0261444812000018>
- Sonbul, S., & Schmitt, N. (2009). Direct teaching of vocabulary: Is it worth the effort? *English Language Teaching Journal*, 64(3), 253–260. <https://doi.org/10.1093/elt/ccp059>
- Sorell, C.J. (2012). Zipf's law and vocabulary. In C. A. Chapelle (Ed.), *Encyclopaedia of Applied Linguistics*. Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeah1302>
- Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based study. *RELC Journal*, 25(2), 34–50. <https://doi.org/10.1177/003368829402500203>
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130–163.
- Webb, S. & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

Address for correspondence

Paul Nation
 Victoria University of Wellington
 School of Linguistics and Applied Language Studies
 PO Box 600
 Wellington 6140
 New Zealand
paul.nation@vuw.ac.nz