

How vocabulary is learned

Paul Nation

*School of Linguistics and Applied Language Studies,
Victoria University of Wellington, New Zealand*

Abstract

Vocabulary learning requires two basic conditions – repetition (quantity of meetings with words) and good quality mental processing of the meetings. Other factors also affect vocabulary learning. For example, learners may differ greatly in their motivation to engage in learning, and words may differ greatly in their learning burden. However, without quantity and quality of processing, learning cannot occur. The greater the number of repetitions, the more likely learning is to occur. The deeper and more thoughtful the quality of processing, the more likely learning is to occur. This paper explains quantity and quality, and shows how teachers and learners can increase the quantity and quality of their processing of vocabulary, thus increasing their vocabulary size.

Keywords: vocabulary learning, mental processing, vocabulary size, repetition

Introduction

Vocabulary learning requires two basic conditions – repetition (quantity of meetings with words) and good quality mental processing of the meetings. Other factors also affect vocabulary learning. For example, learners may differ greatly in their motivation to engage in learning, and words may differ greatly in their learning burden. However, without quantity and quality of processing, learning cannot occur. The greater the number of repetitions, the more likely learning is to occur. The deeper and more thoughtful the quality of processing, the more likely learning is to occur. This paper explains quantity and quality, and shows how teachers and learners can increase the quantity and quality of their processing of vocabulary, thus increasing their vocabulary size.

Quantity and quality are closely related to each other. Many of the conditions affecting quality, such as retrieval, productive use, and varied meetings, depend on a word having been met before, and repetition usually involves changes in the quality of the mental processing of a word.

Quantity of meetings (Repetition)

Repetition occurs when a word is met more than once. There are other quantity factors such as the density of unknown words in a spoken or written text, the number of unknown words and the overall frequency of words in the language. However, these factors have a more indirect effect on learning than repetition. The density of unknown words can affect the opportunity to guess from context clues. The number of unknown words can affect motivation and the manageability of the task. The overall frequency of words can affect the usefulness of the words that are learned.

Repetition occurs after an initial meeting with a word. A critical factor in repetition is the spacing of repetitions within a task or learning session and between tasks. The overwhelming finding is that spaced repetition results in better long-term retention than massed repetition (Nakata, 2015). That is, it is better to meet a word now and then meet it again several times later at spaced intervals than to concentrate on learning a word by putting all the repetitions in one concentrated learning session. Early research suggested that the repetitions should be increasingly spaced, but more recent research suggests that evenly spaced repetitions are effective. The important factor is that the repetitions should be spaced.

The importance of repetition for learning suggests that the first meeting with a word should be seen as only one small step in the eventual learning of a word. The first meeting with a word should result in enough learning to last until the next meeting. This means that teachers need not present words as if the presentation will result in full learning. Rather than be concerned about how to present a word for the first time, teachers should be more concerned with providing opportunities for words to be repeated. Nevertheless, a brief effective initial meeting can increase the effectiveness of later meetings because there is some useful knowledge of the word to draw on and build on. Effective initial meetings can come from guessing from context while listening or reading, deliberate learning from word cards or flash card programs, look-up in a bilingual or monolingual dictionary or gloss, brief L1, L2 or pictorial explanations by the teacher (Elley, 1989), the teacher showing the relationship of the L2 word to an L1 cognate or loan word (Daulton, 2008), the teacher doing word part analysis to relate the unknown L2 word to a known L2 word (Wei & Nation, 2013), or using a mnemonic technique like the keyword technique (Pressley, 1977).

Repetitions can be verbatim or varied. A verbatim repetition involves meeting the word again in exactly the same form, context and circumstances as the previous meeting. Verbatim repetition is most likely to occur during flash card learning and when re-reading or listening again to the same text. In general, most repetitions are varied repetitions, but there is a wide range

of degrees of variation. We will look at this more closely in the following section on quality of meetings, but here we will consider the role of word families in repetition.

A word family consists of a word stem and its closely related inflected and derived forms (Bauer & Nation, 1993). ‘Closely related’ means that as well as a regular form relationship between the members of a family, the meaning of the stem remains roughly the same in the inflected and derived family members. Let us look at an example, the word family based on *amaze*, to clarify this.

amaze
amazed
amazes
amazing
amazingly
amazement

Every member contains the stem *amaze*. The family contains the stem form, three inflected forms, and two derived forms. Note that to be a member of the *amaze* family, all the members need to share the stem meaning of “unexpected surprise”. The size of a word family depends on the learners’ knowledge of the affixes of English. It is possible for a very short period of time in their learning, learners may consider *amaze* and *amazed* as different words. *Amazingly* might be a challenge for some learners even when they know *-ly*. Being able to comprehend the members of a word family involves an increase in the quality of knowledge of a word. It also means an increase in the potential repetitions of a word because meeting different family members is effectively a repetition of the same family. Developing learners’ knowledge of the most common affixes of English is a very important means of increasing the opportunities for learning words through repetition.

How many repetitions are needed for learning? The safest answer to this question is the more the better, although there are diminishing returns for each successive repetition. Receptive knowledge of some words is established with as few as three to five repetitions. In an innovative study, Pellicer-Sanchez (2016) used eye tracking technology to measure how many repetitions it took when reading a text before learners spent the same amount of time focusing on a previously unknown word as focusing on already well established words. She found that around 3 to 5 repetitions there was a noted increase in speed of retrieval as evidenced by fixation time. With 8 occurrences (the maximum in her study) retrieval time was close to that of known words. A follow-up set of vocabulary tests showed learners scored

well on form and meaning recognition and moderately well on recall. It seems safe to assume that for many but not all words around ten to twelve repetitions may be a useful goal (Horst, Cobb & Meara, 1998; Webb, 2007; Pellicer-Sanchez & Schmitt, 2010). Deliberate learning, especially flash card learning of vocabulary requires fewer repetitions, though it must be borne in mind that deliberate learning is more likely to involve verbatim repetition which is likely to be less enriching in aspects of word knowledge than the more varied repetition which is typical of meetings in listening and reading input.

Single meeting learning may occur, especially when a word is a cognate or loan word, but teachers need to see repetition as being essential for all words, and vocabulary learning of each word being a cumulative process.

How can we increase repetition in a course?

In many ways planning for vocabulary learning can be seen as a battle between repetition and Zipf's law. Zipf's law (see Sorell, 2012 for a clear explanation) shows that while there is a rather small number of words in a text or collection of texts that are often repeated, around half of the different words in the text occur only once. That is, around half of the words in a text are not repeated. A well-planned vocabulary course deals with this in several ways. The following recommendations are ranked in order of importance.

1. The vocabulary in the course is controlled in carefully designed stages so that words which are way beyond the learners' current level do not occur in the course material. This ensures that low frequency one-timers do not occur in the material, and reduces the density of unknown words. Graded readers are an essential part of this approach.
2. There are large amounts of vocabulary controlled input so that high frequency and mid-frequency vocabulary have plenty of opportunities to be repeated. Nation (2014) estimates that around 300,000 tokens of input are needed to get at least twelve repetitions of words at the 3rd 1000 level, and around one million tokens of input are needed to get at least twelve repetitions of words at the 5th 1000 level. One million tokens of reading requires around 33 minutes of reading per day, five days a week for forty weeks of the year. Vocabulary-controlled texts still demonstrate Zipf's law with large numbers of words occurring only once. However, in a

vocabulary-controlled text, every word is worth learning and is likely to be repeated in other texts.

3. Learners are taught, memorize and have practice in recognising the most frequent affixes of English, beginning with inflections and quickly moving to the most regular, frequent and productive derivational affixes (see Bauer & Nation (1993) for an ordered list).
4. One-quarter of the time in a well-balanced course should be spent on deliberate learning. Some of this should involve individualised independent vocabulary learning using flashcards. Such learning is not affected by Zipf's law because it does not involve texts.
5. Learners are trained in deliberate vocabulary learning strategies so that the quality of their learning reduces the need for substantial repetition. We will look at these strategies in the next section of this paper. Learners who are old enough should know about the importance of repetition and how to add quality to their mental processing of vocabulary.
6. One-quarter of the time in a course should be spent on developing fluency across the four skills of listening, speaking, reading and writing. Increasing reading fluency through speed reading courses will increase the amount of reading input that learners can get. Fluency, quantity of input and vocabulary repetition can also be increased by re-reading texts that have been read before and listening to texts that have been read or listened to before. Sonia Millett's web site contains a range of free vocabulary-controlled speed-reading courses. There are also commercially produced courses (Malarcher & Nation, 2017).
7. There are teaching activities and approaches to lesson and course design that increase the opportunities for repetition. Perhaps the best example is the linked skills activity (Nation, 2013c, Chapter 15 presents a lot of examples and a rationale for this activity). In a linked skills activity, learners focus on the same topic three times across three different skills. For example, they read about the topic, talk about it and then write about it. This allows the same topic vocabulary to occur in all three steps of the activity. Similarly, content-based instruction (also called Language through the Curriculum and Content and Language Integrated Learning) increases the opportunity for repetition and a reduction in vocabulary load through a narrowing of topic focus (Sutarsyah, Nation & Kennedy, 1994). Narrow reading (Gardner, 2008) and narrow listening (Krashen, 1996) may have similar effects. The common feature in all these activities and approaches is the sustained focus on the same topic or topic area.

Quality of meetings

The importance of quality of processing in memory research was given impetus by Craik and Lockhart's (1972) Levels of Processing hypothesis. The Levels of Processing hypothesis says that what really determines what is remembered or not is the level or quality of mental processing at the moment that learning takes place. If the processing is deep and thoughtful, then the learning will last. If it is superficial, then it will soon disappear.

Table 1 lists levels of processing for vocabulary. The major distinction is between incidental attention and deliberate attention. This distinction is not an easy one to make as many instances of incidental attention contain some elements of deliberate attention. Nonetheless, we can distinguish vocabulary learning while engaging in meaning-focused use of the four skills of listening, speaking, reading and writing (incidental learning) from deliberate attention to words as words rather than as part of the message. Typically deliberate attention is more efficient and effective than incidental attention. This is not surprising because incidental attention assumes that most attention is focused on something else.

The conditions in column one of Table 1 are listed in order of quality of processing with noticing being the most superficial, followed by retrieval, varied meetings and use, and the deepest level of elaboration. Each of these four levels can apply to both incidental and deliberate attention, and each level can be divided into receptive and productive with productive attention being deeper than receptive attention. So, retrieval for example can be receptive retrieval as in reading and listening when the learner meets the word form and has to retrieve its meaning, or productive retrieval as in speaking and writing where the learner has a meaning to express and needs to retrieve the appropriate word form.

Noticing involves giving attention to a word. It does not involve recalling anything about the word or analysing it in any way. Noticing occurs if we study a list of words and their meanings or if we meet an unknown word in a text. Often our first encounter with an unknown word involves noticing.

Retrieval includes noticing but also includes an attempt to recall something that we have already noticed about the word. Retrieval can only occur if we have met the word or a related word before. For example, if we have looked up the meaning of a word in a dictionary, the next time we meet the word in listening or reading we can try to retrieve its meaning. Similarly, if we have made word cards with the word on one side and its translation into the L1 on the other side, then when we go through the cards looking at the words, we can try to retrieve its L1 translation before turning over the

card to see if we are correct. Retrieval is a powerful learning condition and each successful retrieval strengthens the form-meaning association we trying to learn. Retrievals work best if they are spaced, and so listening and reading set up good conditions for retrieval because the occurrences a word do not occur immediately after each other but are spaced by other parts of the text. When deliberately studying words on flash cards, it is good to be working on several words because this means that the other words occur between the repetitions of any particular word thus spacing its occurrences.

Table 1: Vocabulary learning conditions and example activities

Quality of attention	Incidental attention	Deliberate attention
Noticing	<i>Guessing from context Noticing a gap when speaking or writing</i>	<i>Text highlighting Focusing on a form or meaning using word cards Dictionary or glossary look up Being taught words</i>
Receptive or productive retrieval	<i>Meeting a previously met word while listening or reading and recalling its meaning Recalling and using a recently met word in conversation or writing</i>	<i>Retrieval using word cards Doing vocabulary exercises after reading a text Recalling using a word wall</i>
Varied receptive meetings or varied productive use	<i>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</i>	<i>Looking at a concordance Filling the blanks True/false sentences</i>
Elaboration	<i>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)</i>	<i>Using the keyword technique Semantic mapping Word part analysis Focusing on words in teacher-led intensive reading Find the core meaning</i>

Receptive retrieval occurs when we see the word form and need to retrieve its meaning. Productive retrieval occurs when we need to express a meaning and thus need to retrieve the word form. Productive retrieval is

more difficult than receptive retrieval (Griffin & Harley, 1996) (see Nation (2013a: 46-58) for a discussion of possible reasons for this).

The deepest quality condition, elaboration, includes a variety of ways of providing elaborative, analytical and enriching processing of vocabulary. In incidental learning involving normal language use, elaboration relates to the memorable nature of the language use and to the combination of visual and language related aspects of the use. Elaboration is likely to occur during genuine communication, especially that related to the here-and-now in the presence of objects such as when learning how operate something, following directions or buying something. In deliberate learning, elaboration occurs when words are analysed for their known word parts and the meaning of the parts is related to the meaning of the word. It also occurs when using the well-researched keyword technique which similarly relates form and meaning, and uses visualisation of a linking image. Semantic mapping makes deliberate connections between the visual and mental relationships between ideas and thus sets up the condition of elaboration. As with varied meetings and varied use, we could create a scale of degrees of elaboration, with greater elaboration resulting in stronger learning.

How can we increase the quality of meetings?

The quality of meetings depends on how much the teacher and learners can include the conditions of deliberate attention, retrieval, varied meetings and use, and elaboration to the opportunities for learning that occur.

- 1 Deliberate learning using word cards or flash card programs quickly provides a basic amount of knowledge of each word that is then available for retrieval or varied meetings and varied use. Such deliberate learning using flash cards makes use of the conditions of receptive and productive retrieval and in some cases elaboration when the keyword technique or some other mnemonic trick is used. This deliberate learning can be planned so that the most useful vocabulary is covered systematically.
- 2 Graded reading and where possible graded listening provide the condition of retrieval through varied meetings. If this learning through meaning-focused input is related to vocabulary learned through word cards and flash card programs, then repetition and quality of processing are nicely combined.
- 3 Relating input and output allows varied meetings to lead to varied use. In other words receptive knowledge can become productive knowledge. It is important that this relating of listening and reading

input to speaking and writing does not restrict the amount of meaning-focused input. Dealing with the same content in a variety of different ways not only provides repetition of vocabulary but also varies the contexts in which words occur. This varied occurrence enriches and strengthens word knowledge,

- 4 Learners should be trained in how to learn vocabulary. This training should include the guessing from context strategy, the word card and flash card programs strategy with an understanding of the importance of spaced retrieval and receptive and productive retrieval, the word part strategy, the keyword technique, and an elaborative dictionary use strategy that involves looking for core meaning, related words and extra examples of use. Learners should also know how to test their own vocabulary size and how to choose the most useful vocabulary to learn based on this knowledge of their vocabulary size.
- 5 We can regard fluency of retrieval as an aspect of quality. Some high frequency vocabulary such as numbers, days of the week, months of the year, greetings, and survival vocabulary (Nation & Crabbe, 1991) should get targeted fluency practice, largely involving deliberate repeated retrieval. For example, the teacher says numbers in a random order and the learners point to them. In addition there should be fluency practice in each of the four skills of listening, speaking, reading and writing so that vocabulary knowledge is readily available for use.
- 6 Genuine communication can involve both receptive and productive varied meetings and use. It can also provide visual and episodic associations make words memorable.

Repetition and quality of processing

Table 2 combines the recommendations given above for increasing repetition and the quality of processing. The recommendations are ranked in order of effectiveness in providing repetition and quality of processing. The justifications typically explain how each recommendation provides opportunities for repetition and quality of processing.

The first recommendation, on vocabulary control, ensures that learning effort is directed towards what is useful. The major effect of vocabulary control is to deal with Zipf's law in ensuring that learners do not have to deal with large numbers of non-repeated words that are well outside the learners' current vocabulary needs. In many ways the recommendations are macro-level recommendations that apply to course design and lesson planning. It is also worth considering the need for repetition and quality of processing at a micro-level which applies to what the teacher and learners

are doing at any particular moment in the classroom or in independent activities outside class. The sixth recommendation of strategy training and awareness of how to learn encourages this micro-level thinking. Teachers should regularly look at what is happening in the classroom and ask themselves “At this moment are the learners applying good learning conditions? If not how could I adjust the activity so that there are opportunities for repetition and deeper quality processing?”. For an example of these questions applied to spoken activities see Nation (2013a, 190-199).

Table 2: Recommendations for repetition and quality

Recommendation	Justification
1 Control the levels of vocabulary in the course to match the learners current level and needs	Vocabulary control ensures that all words met in a course are useful and that time is not wasted on less useful vocabulary.
2 Include vocabulary learning using word cards or flash card programs	Deliberate learning is both efficient and effective, and provides opportunities for elaboration. Cards involve retrieval and allow for plenty of repetition.
3 Run a substantial, well-monitored extensive reading program and an extensive listening program	Although learning from meaning-focused input is not as efficient as deliberate learning, large quantities of input provide plenty of opportunities for repetition and varied meetings.
4 Use theme-based learning and activities that give repeated attention to the same topic across different skills	Limitation of topic areas reduces the number of different words, and revisiting the same topic increases repetitions and opportunities of varied meetings and varied use.
5 Teach the most frequent affixes and give plenty of practice in recognising them	Being able to deal with words as a part of word families greatly increases repetitions and quality of processing through varied meetings.
6 Train learners in strategy use and understanding of how to learn	Knowing why you are doing a particular activity increases motivation and focus. Awareness and strategy use increases elaboration and allows learners to consciously apply the conditions needed for learning.
7 Have a strong fluency development component in the course	Fluency development increases the amount of input and output and thus repetition and varied meetings and use.

The most notable omission from Table 2 is recommendation 6 from the section on quality of processing about genuine communication. One reason for the omission is that that recommendation affects only quality of processing and does not have positive effects on repetition. Another reason is that there is a lack of research evidence to support this recommendation, and my observation of the effect is based on personal anecdotal evidence. Nevertheless, it seems a reasonable goal for a language course to be relevant and engaging, and for language use to involve real-world high-stakes communication.

Final Thoughts: Vocabulary learning and the four strands

While writing this paper I was struck how the recommendations cover all the four strands of meaning-focused input, meaning-focused output, language-focused learning, and fluency development (Nation, 2007, 2013b). While this is at least partially a persistent bias in my thinking about language teaching and learning, it does suggest that a well-balanced course sets up good conditions for learning. Table 3 classifies the recommendations in Table 2 into the four strands.

Table 3: The four strands and the conditions for vocabulary learning

Strand	Recommendation
Meaning-focused input	1 Vocabulary control
	2 Extensive reading and extensive listening
Meaning-focused output	3 Repeated focuses on the same or related material
Language-focused learning	4 Word card learning
	5 Learn the most useful affixes
	6 Training in strategies and learning how to learn
Fluency development	7 A fluency development strand

Recommendation 1, vocabulary control, is placed in meaning-focused input because it most easily applies to the use of graded readers. However it is relevant to all the other strands as well. Similarly, recommendation 4 is placed in meaning-focused output when it really applies across all four strands. Recommendation 4 suggests an important addition to the principle of the four strands. The principle states that a well-balanced language course should provide opportunities to learn across the four strands and roughly equal time should be given to each strand. To better take account of the conditions for learning, the content and language focuses within and between the strands need to be well integrated so that learners get plenty of opportunities to get focus on same material. This integration of

reading with listening, speaking with reading, fluency development with deliberate learning and so on would increase repetition of vocabulary and quality of processing.

Research on vocabulary learning shows that there are clear guidelines to follow when planning vocabulary learning in a course. This paper has outlined these guidelines and suggested how they can be applied. Their application is not difficult and the effects should be beneficial for vocabulary learning.

The author

Paul Nation is Emeritus Professor of Applied Linguistics in the School of Linguistics and Applied Language Studies at Victoria University of Wellington, New Zealand. His books on vocabulary include *Teaching and Learning Vocabulary* (1990) and *Researching and Analysing Vocabulary* (2011) (with Stuart Webb) both from Heinle Cengage Learning., His latest books on vocabulary are *Learning Vocabulary in Another Language* (second edition 2013) published by Cambridge University Press, *Making and Using Word Lists* (2016) from John Benjamins, and *How Vocabulary is Learned* (2017) (with Stuart Webb).. Two books strongly directed towards teachers appeared in 2013 from Compass Media in Seoul –*What should every ESL Teacher Know?* (available free from www.compasspub.com/ESLTK) and *What should every EFL Teacher Know?* He is also co-author of Nation, P. and Malarcher, C. (2017) *Timed Reading for Fluency. Books 1-4*. Seoul: Seed Publishing.

References

- Bauer, L., & Nation, I.S.P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Daulton, F.E. (2008). *Japan's Built-in Lexicon of English-based Loanwords*. Clevedon: Multilingual Matters.
- Elley, W.B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24(2), 174-187.
- Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading. *Reading in a Foreign Language*, 20(1), 92-122.

- Griffin, G.F., & Harley, T.A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics*, 17, 443-460.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357-377.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Krashen, S. (1996). The case for narrow listening. *System*, 24(1), 97-100.
- Malarcher, C. & Nation, P. (2017). *Timed Reading for Fluency*. Seoul: Seed Learning.
- Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning: Does gradually increasing spacing increase vocabulary learning? *SSLA*, 37, 677-711.
- Nation, I.S.P. (2014). 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. (2013a). *Learning Vocabulary in Another Language*. Second edition. Cambridge: Cambridge University Press.
- Nation, P. (2013b). *What should every EFL Teacher Know?* Seoul: Compass Publishing.
- Nation, P. (2013c). *What should every ESL Teacher Know?* Seoul: Compass Publishing.
- Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 1-12.
- Nation, P., & Crabbe, D. (1991). A survival language learning syllabus for foreign travel. *System*, 19(3), 191-201.
- Pellicer-Sánchez, A. (2016). Incidental vocabulary acquisition *from and while* reading: An eye-tracking study. *SSLA*, 38(1), 97-130.
- 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.
- Pressley, M. (1977). Children's use of the keyword method to learn simple Spanish vocabulary words. *Journal of Educational Psychology*, 69(5), 465-472.

- Sorell, C.J. (2012). Zipf's law and vocabulary. In C.A. Chapelle (Ed.), *Encyclopaedia of Applied Linguistics*. Oxford: Wiley-Blackwell.
- Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based study. *RELC Journal*, 25(2), 34-50.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Wei, Z., & Nation, P. (2013). The word part technique: A very useful vocabulary teaching technique. *Modern English Teacher*, 22(1), 12-16.