

# Vocabulary: principles and practice

1

**Norbert Schmitt**  
applies a principled approach  
to teaching words.

All teachers realise that their students need to know vocabulary in order to function in English. But the best way to help our students learn this vocabulary most effectively remains an unresolved issue. Some commentators (such as Stephen Krashen) assert that sufficient vocabulary can be learnt simply from exposure to English, particularly reading. Other commentators (such as Batia Laufer) argue that much vocabulary needs to be *taught*. As is usually the case, some middle ground is probably the best course. This would include maximum exposure to English, both inside and outside the classroom, and also explicit teaching of selected vocabulary.

There is also the question of how to carry out the explicit teaching: what are the most effective teaching techniques and tasks? Luckily, there has been an explosion of research into second language vocabulary learning in the last 25 years, and this series of articles will cover some of the newest findings.

In this initial article, I will introduce some of the key principles for vocabulary teaching that have come out of this research, and many of these themes will be taken up throughout the series.



## Principle 1

**A lot of vocabulary is necessary to function well in English.**

How much vocabulary do your students need in order to function in English? The short, simple answer is 'a lot', but it depends on your students' learning goals. An extremely ambitious goal might be to have a vocabulary size comparable to that of an educated native speaker (15,000–20,000 word families). Most learners will never achieve this, but luckily, students can flourish in English with far smaller vocabulary sizes than this. Paul Nation summarises the current consensus on vocabulary size requirements:

- To listen to everyday conversational English: 6,000–7,000 word families;
- To read a range of novels and newspapers: 8,000–9,000 word families;
- To watch films aimed at children (eg *Shrek*): 7,000 word families.

It is important to understand that these figures don't refer to just any random words, but represent the 6,000–9,000 *most frequent* word families in English. This is because research has shown that the most frequent words in English do the bulk of the work. For example, the most frequent 8,000–9,000 word families will make up around 98 percent of all

the words in English novels and newspapers. It is also worth noting that the figures are counted in *word families*, which include a base word (eg *stimulate*), its inflections (*stimulated*, *stimulating*, *stimulates*) and its derivative forms (*stimulation*, *stimulative*, *stimulatingly*). Many teachers mistakenly interpret Nation's figures as simply the number of words their students need to acquire. In fact, they need to learn multiple word family members, which entails learning many more words than the figures above suggest.

The reality is that most learners do not achieve these targets, and lack of vocabulary remains one of the greatest hurdles in advancing our learners' language proficiency. The implication of this is that we need to do everything we can to help increase the size of our students' vocabularies.

## Principle 2

### Learning vocabulary entails more than just learning a word's form and meaning.

Knowledge of a word starts with making the link between its form and meaning (eg the word *ecstatic* means 'extremely happy or joyful'). While most teachers may think that teaching vocabulary is mainly about teaching meaning, in fact it is mainly about teaching the forms of words (spelling and/or pronunciation), at least initially. Learners typically already know the meanings of most of what they want to say from their experience of the world and their L1. The gap is that they don't know the words in English to express these meanings. This suggests that teaching the 'form-meaning' link will mainly entail teaching the English word forms.

But 'knowing' a word well requires much more than just the form-meaning link. Paul Nation describes comprehensive word knowledge as being inclusive of a word's form and meaning, but he also adds several *more* aspects, including:

- The grammatical characteristics of a word (eg its part of speech);
- The collocations of a word (eg we can say *strong coffee*, but not *\*powerful coffee*);
- The contextual constraints on use (appropriate connotations and tone), eg a *beautiful woman*, a *handsome man*.

The word knowledge aspects your students need to know depend on what

they want to do with their vocabulary. If they only want to read or listen, then the form-meaning link might be enough. They need to recognise the form which is already provided in the reading passage or spoken discourse, and then retrieve the meaning linked to it. But production (writing or speaking) is more complex. Your students will have to recall a word that is appropriate not only for the desired meaning, but also for the particular grammatical, collocational and contextual situation. Thus *productive* use requires a wider array of word knowledge than *receptive* use.

## Principle 3

### Vocabulary learning is incremental in nature.

Word knowledge is not a black-and-white proposition where words are either 'known' or 'unknown'. Rather, words are learnt gradually over time. There are several reasons for this. First, the various word knowledge aspects are not usually learnt at the same rate; some aspects are typically learnt before others. For example, the form-meaning link is generally mastered before collocation knowledge. Thus, it takes time for the later-mastered aspects to catch up with the earlier-mastered aspects. Second, each word knowledge aspect also develops over time, progressing along a continuum from no knowledge at all to full mastery. Third, word knowledge typically develops from receptive mastery (comprehension in listening and reading) to productive mastery (ability to use appropriately in speaking and writing). For all these reasons, we generally should not speak of words as being 'unknown' versus 'learnt'. It is more fruitful to think in terms of how well our students know words, what they can do with them, and which word knowledge aspects need enhancing to achieve greater mastery.

An incremental view of vocabulary learning leads to some pedagogical implications that many teachers might find surprising. For example, many teachers believe that vocabulary should always be taught in context. Context is certainly helpful in enhancing knowledge about many aspects, such as collocation and connotation. But initially, it is most important for students to begin fixing the form-meaning link in their minds. For this, word lists can be highly

effective. Since word learning will happen gradually anyway, it makes sense for students to begin the process in their own time as homework. Using word lists, they can quickly make the first step towards learning the form-meaning link. Then, in class, teachers can build on this budding knowledge by using the word in context, without having to teach the form-meaning link first.

Likewise, as the purpose of word lists is introducing the form-meaning link, one reasonable way of doing this is with L1 translations. Translation is especially useful for lower-level students who don't yet know much defining vocabulary; it can help them more quickly reach the point where the word makes sense in the L2 contexts the teacher provides. For higher-level students, using L2 definitions has the advantage of recycling the L2 vocabulary used in those definitions.

The 'how well' approach to vocabulary knowledge also affects the way we should interpret *test* results. No test format can measure all aspects of word knowledge, and so we should not interpret a correct answer as showing a word is 'known'. Rather, we should think of what each format shows about the student's degree of mastery.

Let's take two typical test formats: multiple-choice and translation:

#### rage

- a type of car
- an angry feeling
- a tool used to build a house
- a bad smell

rage = \_\_\_\_\_ (translation)

In the multiple-choice test, the students only have to recognise the correct meaning (b), but what does this tell us about their ability to comprehend the word when reading? Reading texts don't have options, so it is difficult to know whether they could understand the word or not. Conversely, in the translation test, the students have to recall the word's meaning when presented with its form, a process much closer to what happens when reading. So the translation test actually gives a better indication of the level of mastery required to employ the word when reading.

# Vocabulary: principles and practice 1

## Principle 4

### Recycling is crucial to vocabulary learning.

Because vocabulary is learnt incrementally, it is essential that students are exposed to words often over time. This helps to consolidate the learning which has already taken place, so that it is not forgotten and lost. But the additional exposures also serve to *enhance* words and push the students to higher levels of mastery in all aspects of word knowledge. The timing of the repetitions is important, too. People tend to forget much of what they have learnt relatively quickly after the learning episode, and so it is very beneficial if the first review happens soon after the learning. Some researchers suggest it should occur within an hour, or at least the same day.

## Principle 5

### There is a role for both explicit instruction and incidental learning.

From Principles 2 and 3, we can see that vocabulary knowledge is complex and multidimensional. Different kinds of exposure affect different facets of this knowledge. Some word knowledge aspects, such as the form–meaning link, are relatively responsive to explicit teaching and intentional learning. So it makes sense to focus our students' attention on these aspects, either in class or as homework, because conscious attention can lead to successful learning. Other aspects, however, are less responsive to explicit attention. For example, collocation is largely about intuitions about which words happen to co-occur together regularly in English. While it is possible to teach a limited number of collocations, it is basically impossible to teach the whole range of every word's collocations. This can only come from large amounts of exposure, where collocations come up naturally in a wide variety of contexts. So for collocations, the incidental learning

which comes from a vast amount of exposure is likely to be more effective than explicit teaching in building these intuitions. We thus need both *explicit instruction*, to fast-track the initial learning of vocabulary, and *large amounts of exposure*, to provide the recycling and enhancement which supports that instruction, and also to cover the word knowledge aspects that we cannot teach.

## Principle 6

### Vocabulary is more than just individual words.

When discussing vocabulary, most teachers probably think in terms of individual words. But research has shown that language also consists of large amounts of multi-word phrasal vocabulary, such as idioms, phrasal verbs and collocations. These phrasal units are usually referred to as *formulaic language*, and are common across a range of languages. It is estimated that between one third and a half of all English discourse consists of formulaic language. It is so widespread because it carries out key communicative functions, such as:


- social interaction (*I understand, how nice*);
- functional language use (*how can I...? I am sorry to hear that, I'd be happy to ...*);
- organising discourse (*on the other hand, first of all*).

Formulaic language also plays a key part in facilitating fluency. Because formulaic language is very common, it is both expected and predictable. This makes production easier, as students merely need to remember a formulaic phrase for the function/meaning they are trying to express, rather than having to create new language from scratch. The predictability also eases the cognitive load receptively, helping listeners/readers to recognise chunks of language, rather than having to process discourse in a word-by-word manner.

Despite the importance of formulaic language in discourse, it has been found to present problems for learners. This is particularly true if it carries idiomatic meanings – for example, *look up* = 'search for information, as in a dictionary' or *top drawer* = 'of very good quality'. Given the importance and difficulty of formulaic language, it is a

logical target for some explicit teaching, and research is just beginning to develop principled approaches to teaching it. To date, the most tangible outcomes are a number of lists of various types of formulaic language which merit attention. These include the *PHaVE List* for phrasal verbs, the *PHRASE List* for high-frequency formulaic items and the *Academic Formulas List* for phrases common in academic language.



You may think that some of the principles I have described are common sense, and indeed this is true. Teachers generally have good intuition about how their students learn, and probably have some idea of many of the principles outlined in this article. However, seeing their intuitions validated by research can give teachers confidence. In some cases, research may push teachers to challenge some of their existing assumptions, eg that vocabulary should always be taught in context. Research is most useful if it is applied in the real world, and I hope this article and the series it introduces will help you to think more productively about your own teaching practices, and how vocabulary fits into your teaching context. 

### References

- Krashen, S *The Input Hypothesis: Issues and Implications* Longman 1985
- Laufer, B 'Focus on form in second language vocabulary acquisition' In Foster-Cohen, S H, García-Mayo, M P and Cenoz, J (Eds) *EUROSLA Yearbook 5* John Benjamins 2005

### Further reading

- González-Fernández, B and Schmitt, N 'Vocabulary acquisition' In Loewen, S and Sato, M (Eds) *The Routledge Handbook of Instructed Second Language Acquisition* Routledge 2017

Also see [www.norbertschmitt.co.uk](http://www.norbertschmitt.co.uk) for a large variety of research and discussion of various aspects of vocabulary acquisition, pedagogy and use.



Norbert Schmitt has authored over 100 publications on second language vocabulary issues, but he is most proud of his former and current PhD students who are pushing vocabulary research forward, and who are contributing to this series on vocabulary, which will continue in subsequent issues of ETP

[norbert.schmitt@nottingham.ac.uk](mailto:norbert.schmitt@nottingham.ac.uk)