

DEVELOPING ACADEMIC VOCABULARY OBJECTIVES: CORPUS ANALYSIS AND WORD LISTS

Tanju Deveci and Adam Simpson

Introduction

This chapter is concerned with the development of a vocabulary syllabus to accompany and supplement a coursebook series. To further enhance the use of these books, we decided to initiate a project to develop an academic vocabulary syllabus. This chapter will outline the process of analyzing corpus data, deciding on lexis to cover and how to spread this across the course of study, creating support materials to facilitate learning as well as incorporating this into the wider curriculum .

Deciding on the foundations of our work

The fact that we had to base our vocabulary syllabus on the existing coursebooks taught on our mainstream courses helped us in the sense that our learners' needs had already been analysed and general objectives had been decided on. What we needed to do was to choose the particular lexical items from the coursebooks and organise them according to existing levels.

Taking into consideration the word lists available

In deciding the words that would make up our vocabulary syllabus, we decided to consult the Academic Word Lists (AWL) (Coxhead, 2000: 213-238) and Xue and Nation's University Word List (1984: 215-219). We also consulted the General Service List (West: 1953), although we did not prioritise certain words in this list in order to keep our focus on academic words rather than general vocabulary. This also helped us avoid overloading our lists.

Text analysis

In order to make a decision on what head words should be covered in the lists, we used a vocabulary profiler, which enabled us to decide on the frequency of words in each unit of the coursebooks. Of the variety of available profilers, we decided to use the University of Hong Kong's webpage (http://ec.hku.hk/vocabulary/tools_cp.htm) since it is user-friendly and offers multiple functions such as concordancing.

The data we obtained using the profiler showed that 80 to 90% of the texts in our coursebooks were made up of the most common 2000 words. In addition to this, this figure increased to above 90% when the AWL words were included. This suggested that we could actually base our vocabulary syllabus work on some sound basis as, in a typical academic text, coverage would be as follows:

Vocabulary Coverage in an Academic Text

Level	Text coverage
High-Frequency Words (top 2,000)	87%
Academic Vocabulary	8%
Technical Vocabulary	3%
Low-Frequency Words	2%

Coady & Huckin, 1997: 239

We also noticed that there were some additional topic-specific words which did not appear in any of the lists. We felt that a certain number of these words should be included since they seemed to be important in the context of our learners' future studies.

Deciding on the words to cover

First of all, we carried out a frequency analysis of the words in each unit of all the books. This showed us both the lists of all words used in the texts and how many times particular words appeared in particular units and/or levels. This kind of an approach allowed us to identify different examples of words being used in similar or different contexts, typical sentence patterns in which the words appeared and different word forms they were used in. This also gave us the opportunity to see whether or not any important academic vocabulary items appeared with very low frequency in a certain unit. This in turn helped us decide if we should include them in our vocabulary syllabus, and if so, where to include them.

Analysing the texts in this way helped us identify the kind of words we could aim to cover for different levels. In doing so, we tried to answer certain questions, such as:

- How frequently does a word appear in different texts/units/levels?
- How easy is it to illustrate the meaning of the word?
- Could learners use it in immediate situations as well as in different contexts?
- Is it common in both spoken and written language?
- Does it have false friends?
- Is there a close L1 cognate? If so, is it better not to include it in order to avoid overloading learners?

Production of vocabulary lists

Once we were able to decide which words to include for each of the units of the coursebooks, we decided to have two versions of the vocabulary lists. The teachers' version included the frequency of each word and page numbers where they appeared for ease of reference. It also indicated which words were 'off-list' words i.e. topic-specific words which did not appear in the lists. The learners' version of the lists, on the other hand, were a list of the target words for each unit, without any indication of frequency of occurrence or whether any of them were off-list words. We did this in order to prevent learners from ignoring any words that did not belong to AWL, Nation's University Word List or General Word Lists since they were some important content words they might encounter in their academic studies at faculty level.

Also, we felt that some of the most frequent words should be repeated in different units and levels for further exposure. Any incidence of such words was indicated in the words lists, explicitly stating where these items had been encountered previously. This gives learners the idea that their learning is facilitated more if they are exposed again to the words they have previously studied. Upon feedback from learners and teachers, we decided to categorise words according to which text they appeared in the units. These categorised lists were then added to the Teaching and Learning Programmes (see Chapter on TLP) which make it easier for both learners and teachers to refer to the words in context.

Going beyond word lists

While we wanted the lists as a reference point for teachers and learners, this did not do enough to fulfill our intention to effectively support the learners' studies in terms of

vocabulary. Therefore, we needed to develop a more substantial document to further develop learners' linguistic competence. In order to achieve this, we decided to draw learners' attention to the contexts of the words used in the coursebooks but also add additional contexts to facilitate learners' comprehension. We thought more visuals would foster this and that learner training activities would be useful for guiding learners to develop their awareness of word formation, collocation etc.

What are 'vocabulary companions'?

All of the above factors played a part in the next stage of the project, which involved the production of documents called 'vocabulary companions'. The vocabulary companions contain the same words as the lists but act as a natural progression from the word list, where words are presented in isolation, to a document that would foster a deeper level of understanding. With this in mind the companions were written to contain many pieces of relevant information such as:

- definition(s);
- example(s) in context;
- word formation;
- collocations;
- antonyms/synonyms; and
- reference to examples in other units where the item appeared.

Consequently, the learners could refer to these to gain further information about lexis; basically they serve as a mini-dictionary accompanying the coursebook unit. An example page of a vocabulary companion can be found in Appendix 1.

Exploiting the companion in class

The companion had to be something that could act as a springboard to greater development of vocabulary acquisition. In other words, materials were required to exploit the companion and further justify its place within the overall syllabus. We did not feel it would be enough to simply refer learners to the vocabulary companion for further information on a particular vocabulary item. A series of materials were needed to take learners on the journey from first encountering the lexis to the point where they would be able to use it in written/spoken production. This led to the development of companion activities.

Companion Activities

We decided to build up a bank of activities that would take learners through the process of knowing a word:

- the form (spoken, written and word parts);
- the meaning (the concept of the word, its referents and its associations);
- usage (grammatical functions, collocations and constraints such as register and frequency) (Nation (2001: 27)).

Furthermore, the activities would form structured levels of vocabulary development, taking into consideration factors such as:

- recognising the word in spoken/written form;
- recognising affixes and relating these to meaning;
- knowing the meaning in a particular context, and that meanings are influenced by context;
- knowing that certain words are related;
- recognising when the word has been used correctly;
- recognising common collocations and connotation (Nation (2001: 26-28)).

Here is a description of each of the four stages of the companion activities:

Stage 1: a first encounter with the word

Stage One activities were developed so as to treat the lexical items as if it were the learners' first encounter with the vocabulary. Clearly, with the existence of the companion documents, the learners were receiving a greater level of support than a learner meeting the word in an everyday uncontrolled situation, especially as the examples used in the companions were taken directly from the reading texts and listening scripts. Therefore, the Stage One activities stressed such strategies as guessing meaning from supported/controlled context. An example of a stage 1 activity can be seen in Appendix 3.

Stage 2: getting to know the word

The Stage Two activities addressed such issues as targeting productive vocabulary and addressing collocations, usage restrictions and other elements related to using a word, and necessitated that learners develop a deeper level of understanding of the lexical item. The example in Appendix 3 shows an activity used to develop a greater level of cognition. In this example, the learners are given the opportunity to develop their knowledge of some of the strong collocations and lexical chunks associated with each item. The learners, in other words, are moving past viewing the words in isolation.

Stage 3: using the word

Stage Three activities created a need for the learners to use the vocabulary productively. Such activities required the learners to complete such tasks as making sentences from prompts and writing quizzes for other learners and classes. (See Appendix 4).

Stage Zero: learner training

One of our fears in using only one source for examples of vocabulary was whether it would provide good examples of the grammar of the lexis and indicate typical strong collocations. One way in which this issue was addressed was in the development of Stage Zero activities, for learner training. A typical activity was to assign particular lexical items to groups of learners and have them research that word on the Internet. By searching using the lexical item as the keyword, a list of news headlines were shown, from which collocations could be observed. As this example activity shows, these collocations were then used by learners to prepare materials to share with their friends. Other, more typical learner training resources were devised such as materials discussing methods of recording vocabulary. An example of what we called a stage 0 activity can be seen in Appendix 5.

Feedback

An essential part of the process was the trialling of the materials and collecting feedback on how they were working.

Feedback from teachers

We asked teachers for semi-formal feedback, in the form of feedback sheets. Several issues became evident:

- The number of subject-specific items was considered too high, and it was feared that learners were spending too much time studying non-essential vocabulary. For instance, a teacher on a basic level course asked *“How could my learners learn this many words in just a week?”*. To address such issues, the unit specific vocabulary was revised and the number of items reduced by leaving some words to later units or deleting some of the words which appeared in earlier lists.
- There was insufficient recycling of essential vocabulary. To ensure recycling of such vocabulary items presented at lower levels, were presented again at higher levels.
- Further assessment practice was required for learners to measure their progress. Therefore, quizzes were prepared for each unit.
- The presentation and practice almost exclusively involved learners seeing the words in written form and then using them in activities. One way to counter this was the development of the ‘voiced companion’ which was created for learners to be able to click on an online document and listen to the word in isolation, and listen to the word as it appears in an example sentence.

Feedback from learners

We also asked learners for their feedback on the resources produced. Feedback was taken in the form of group interviews and samples were taken from learners at all levels of proficiency. As with the teachers’ feedback, several interesting points arose as a result.

- The materials had to be motivating. The goals and objectives of the materials were intended to be as relevant to the learners as possible, but should also contain a degree of enjoyment, fun, creativity and develop a sense of achievement. One strategy we used to do this was placing it in contexts that the learners could relate to. For instance, using pictures of known celebrities meant that the vocabulary being presented became more meaningful, and the task more enjoyable. Learners noted that tasks containing visual stimuli helped them to retain the vocabulary and made doing the exercises more pleasurable. When asked what learners thought of such materials, one learner said *“I really like game worksheets. They are fun and motivate me”*.

- Another important issue raised was that of spending time on vocabulary work in class. One learner said *“It is easier to remember if we use the material in class with groups and pairs”*. This indicates that learners wanted to engage with the materials in a collaborative manner. Learners strongly indicated that they felt time spent working with vocabulary materials in class was more valuable than having to work through the materials themselves. For instance, another learner said *“Teacher’s input is invaluable”*.
- Interviewees particularly noted the motivational benefits of being assessed on vocabulary, and that this increased their inclination to study independently. An intermediate level learner said *“I study at home, repeat what (I have) done in class. (This is) essential”*. To support learners in this way, we created vocabulary quizzes. We also got learners to write their own quizzes for their peers, which we observed increased their motivation greatly.

Addressing the future vocabulary needs of learners

The next question to be addressed was the extent to which the words in the syllabus prepared learners for their continuing studies. A separate project was consequently undertaken to analyse the coursebooks that learners study in the compulsory courses in the first year of their degree programmes.

First of all, the coursebooks were scanned and converted into a ‘Notepad’ document. This enabled a corpus analysis of the texts using the *AntConc* software tool (<http://www.antlab.sci.waseda.ac.jp/software.html>) which is a downloadable corpus analysis software which examines the frequency of a word, as well as its’ collocates. For each subject area, lists of key words were generated which indicated the most common words in the first year coursebooks but less frequent in the coursebooks used in the Foundation Development Year English programme. The results informed us about items which could be added to our syllabus to make it a better preparation for the following years of study. For example, the key word lists for the social sciences course indicated that there were only a dozen or so words which were not already covered in our programme. Therefore, these words were incorporated into the vocabulary programme in a suitable

place. Similarly, these key word lists also informed the vocabulary requirements of the English of Maths and Science course (see the chapter on EMS).

GUIDING PRINCIPLES

Other institutions may wish to consider the following guiding principles when developing an academic vocabulary syllabus in their own context:

- **Refer to appropriate word lists**

Depending on the context, one or more of the following lists might be useful:

- The Academic Word Lists (Coxhead, 2000);
- University Word List (Xue and Nation, 1984);
- General Service List (West, 1953).

Such word lists have been developed on scientific data and represent a good starting point when deciding what words need to be included.

- **Use an available vocabulary profiler**

Profiling the texts will inform decisions about the appropriacy of including a particular word in any given unit of study. There is a variety of readily available profilers such as *AntConc*, and some simpler online tools, like the Hong Kong University's vocabulary profiler, that will provide useful data about the composition of a corpus of texts.

- **Identify specific criteria**

Clear criteria are needed to define the vocabulary items to be included in the vocabulary syllabus. Therefore, criteria need to be decided on first. These criteria should include an approach as to what to do with topic-specific vocabulary not included in the lists. Another important issue is how to go about recycling certain words which are deemed useful for the target groups of learners. In addition, the vocabulary load needs to be realistic according to the time available for learners.

- **Develop Support Materials**

Having vocabulary lists might encourage learners to translate words into their mother tongue and go no further in their vocabulary studies. In order to avoid this, it is advisable to create material to build up their knowledge from the point where they first meet the words to the point where they are able to use them productively. The materials need to encourage learners to learn useful phrases, collocations as well as lexical chunks, which will increase their competence substantially. Also, the programme should include multiple encounters the target words, which will enhance the learners' vocabulary acquisition.

ACKNOWLEDGEMENTS

We wish to thank Deniz Özden Rodriguez, who provided invaluable assistance throughout the project and spent a considerable amount of time writing quizzes as a part of the process. We are also grateful to our colleagues and learners who were always ready to give us feedback at different stages of the project. A special thanks also goes to those colleagues who were involved in the process of dividing the words lists into inputs in the TLPS. We are also grateful to Pinar Gündüz for initiating the voiced companion as well as to those who offered to help with the recording of the documents.

Useful publications

- Anthony, L. - AntConc software <http://www.antlab.sci.waseda.ac.jp/software.html>
- Coady, J. & Huckin, T. (1997). *Second language vocabulary acquisition*. USA: Cambridge University Press
- Coxhead, A. (2000) 'A New Academic Word List'. *TESOL Quarterly*, 34, (2), 213 - 238.
- Graves, K. (Ed.), (1996) *Teachers as Course Developers*. Cambridge: CUP
- Miller, D (2011) 'ESL Reading Textbooks vs. University Textbooks: Are we giving our students the input they may need?' *Journal of English for Academic Purposes*, 10, (1), 32-46
- Nation, I.S.P. (2001) *Learning Vocabulary in another Language*. Cambridge: CUP
- University of Hong Kong Centre for Applied English Studies - vocabulary profile tools http://ec.hku.hk/vocabulary/tools_cp.htm
- West, M. (1953) *A General Service List Of English Words*. London: Longman
- Xue Guoyi and Nation, I.S.P. (1984) *A University Word List*. *Language Learning and Communication* (3), 215-229.
- Zwier, L.J. (2002) *Building Academic Vocabulary*. Michigan: University of Michigan Press

Appendices

Appendix 1. Sample page of a vocabulary companion.

Level: Two	Book: One	Unit: One
Inputs: 1) People need help (R) / 2) Flow theory (L) / 3) Effects of colour and scents (L) / 4) Prozac issues (R)		

Word 1	Definition / Examples	Related Words / Collocations
<u>abnormal</u> (ADJ)	<p>different from what is usual or average, especially in a way that is bad:</p> <p><i>I am taking two psychology courses. In one of them we're studying abnormal psychology and in the other we're learning about children's mental development as they grow up. (Grammar Focus 1)</i></p> <p>See also: <u>abnormality</u> (N) and <u>abnormally</u> (ADV) BASIC Unit 9</p>	OPPOSITE: <u>normal</u>

Word 2	Definition / Examples	Related Words / Collocations
to <u>appreciate</u> (v)	<p>to recognize or understand that something is valuable, important or as described:</p> <p><i>First, when resources are unevenly distributed, people compare themselves to richer people instead of appreciating the money they <u>have</u>. (I1)</i></p> <p>See also: <u>appreciation</u> (N)</p>	to appreciate sth (that) you <u>have</u>

Word 3	Definition / Examples	Related Words / Collocations
to <u>assert</u> (v)	<p>to say / believe that something is certainly true:</p> <p><i>Calling for psychologists to promote what delivers happiness only echoes what a great philosopher once asserted. (I1)</i></p> <p>See also: <u>assertion</u> (n)</p>	

Word 4	Definition / Examples	Related Words / Collocations

<p>challenge (N) [C or U]</p>	<p>(the situation of being faced with) something needing great mental or physical effort in order to be done successfully and which therefore tests a person's ability:</p> <p>“One of the most important challenges psychologists face,” said Csikszentmihalyi, “is to help people discover alternatives to material wealth that lead to more rewarding lives.” (I1)</p> <p>See also: to challenge (V) and challenging (ADJ) BASIC Unit 5</p>	<p>to face a challenge</p>
-----------------------------------	--	----------------------------

Appendix 2. Sample of a stage 1 activity

Vocabulary in Context

Look at these examples. In each example, one word is in **bold**. Try to understand this word from the context of the sentence, and answer the questions.

1. To have an **altering** effect on sth

Experts on scent have found that particular scents do seem to have certain mood **altering effects** on people.

Do you think **altering** means **a) causing the characteristics of something to change**, or **b) causing an unusual situation**?

2. To **appreciate** sth (that) you have

When resources are unevenly distributed, people compare themselves to richer people instead of **appreciating** the money they have.

If you appreciate something, do you **a) not realise or understand that something is valuable**, or **b) recognise or understand that something is valuable**?

3. To be **capable** of doing sth

Not all patients should resort to Prozac, because if psychotherapy is **capable of curing** depression among patients, then that is the best way to overcome their problems.

If something is capable, is it **a) able to do things effectively and skilfully, and to achieve results**, or **b) unable to do things effectively and skilfully, and to achieve results**?

4. To do **extensive** research

Kwallek has also done extensive research on her own about office colours. Kwallek looked at all of the available information.

In this sentence, does extensive have a similar meaning to **a) covering a small area; having a small range**, or **b) covering a large area; having a great range**?

5. To **indicate** a relationship between X and Y

Although research **indicates** no clear relationship between material wealth and happiness, many people still equate more money with more happiness.

Does this word mean to **a) to hide or make unclear in some way**, or **b) to show, point or make clear in another way**?

Appendix 3. Sample of a stage 2 activity

Let's Collocate!

You should complete this task after you have studied the companion. Look at the words in the column on the left. These are words from the Unit 1 Companion. Match these words to suitable phrases in the right column to make collocations. One example is given.

appreciate	to lead a _____ life
capable	a good / bad _____
commit	to do _____ research
consistently	to be _____ Of curing a disease
extensive	to how _____ important sth is
indicate	to appreciate the things you have
outcome	to _____ a relationship between two things
pleasant	To _____ an error
realise	a _____ smell
rewarding	to _____ claim that sth is true

Use the Unit 1 Companion to check your answers. Then complete the exercise on the next page.

Fill in the gaps in these sentences using the collocations you have made.

1. When resources are unevenly distributed, people compare themselves to richer people and don't _____ the money they _____, even though they might have good lives.
2. One of the most important challenges psychologists face is to help people discover alternatives to material wealth that allow people to _____ lives that are more _____.
3. Stewart McGinnikus is an academic who has done _____ _____ relating to office colours.

4. Despite the fact that research _____ no clear relationship _____ material wealth and happiness, many people still think more money means more happiness.
5. If psychotherapy is really _____ of _____ depression among patients, then that is the best way to overcome their problems.
6. We don't even _____ sometimes how _____ and influential scent is. It is a necessary part of our lives.
7. White coloured workspaces are the worst because people actually _____ more _____ in all-white offices than in offices of different colours.
8. A casino in Las Vegas found that when they started infusing areas with a _____, the amount of money gambled in those areas rose by 45%, because of the scent!
9. The test subjects _____ that the scented shampoo was easier to rinse out, foamed better and left their hair shinier even though it was exactly the same shampoo!

Appendix 4. Sample of a stage 3 activity

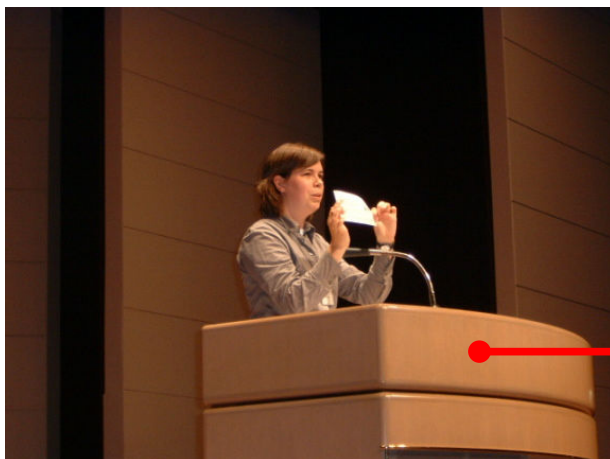
1 picture = 1000 words

7) Imagine you are doing this bungee jump. Would you hesitate before you jump? Why?



They are...

8) Look at these two pictures. One shows people having a debate. The other shows a person making a speech. Which is which?



She is...

Appendix 5. Sample of a stage 0 activity

Vocabulary Learning Training Materials

Task 1: How many words do you need to know?

1. Read the following questions and circle the answer you think is correct.
 - a) How many words are there in the English language?
54,000 200,000 500,000 1,000,000 2,000,000
 - b) How many words does an educated native speaker of English know?
10,000 20,000 50,000 200,000 500,000
2. How many words do you think you know in Turkish? How do you add to your vocabulary knowledge in Turkish?
3. How many words do you think you need to know in order to follow your academic studies in English?

Task 2: What does it mean to know a word?

1. Look at the following points about what it means to 'know' a word:
 - a) to understand it when it is written or spoken
 - b) to recall it when you need it
 - c) to use it with the correct meaning
 - d) to use it in a grammatically correct way
 - e) to pronounce it correctly
 - f) to know which other words you can use with it
 - g) to spell it correctly
 - h) to use it in the right situation
 - i) to know if it has positive or negative associations

You will not need to know all of the above aspects about every word or phrase. It may depend on whether you only need to **recognize** the word/phrase or whether you want to be able to **use it actively** in speaking or writing.