# CREATING A BUSINESS WORD LIST FOR TEACHING BUSINESS ENGLISH

*Nikolaos Konstantakis*
*Swansea University, UK*

*Foreign language students planning to study business at English-speaking universities have little or no previous training in English for Academic Purposes in general, or in Business English in particular. Nonetheless, these students rely heavily on their ability to read and fully comprehend a variety of business course books and other materials in the field. Reading comprehension, however, depends to a great extent on an individual's vocabulary level. Vocabulary may be acquired by learning word lists. One very useful list of this kind is the General Service List (West, 1953) which includes the 2,000 most common words for teaching and learning English. Another potentially useful list, which is more specialised, is the Academic Word List (Coxhead, 2000), which provides the most common words in academic texts. These two lists, however, fail to provide the prospective students of business with the technical vocabulary that is particular to their area of study. A third list, therefore, more specialised and business-oriented is required in order to cover this gap. The present paper describes the attempt to create a Business Word List.*

*Key words: Business English, technical vocabulary, corpora, word lists, lexical coverage*

*Los estudiantes de lenguas extranjeras que planean estudiar finanzas en universidades de países de habla inglesa tienen poco conocimiento del inglés con fines académicos en general, y del inglés para los negocios en particular. Sin embargo, estos estudiantes deben leer y comprender perfectamente gran variedad de libros de texto y materiales relacionados con el mundo de los negocios. Pero la comprensión de textos depende del*

*nivel de vocabulario de cada individuo. El vocabulario puede ser adquirido a través de listas de palabras. Una lista muy útil es la General Service List (West, 1953), que incluye las 2.000 palabras más comunes para enseñar y aprender inglés. Otra lista, más especializada, que puede ser de utilidad es la Academic Word List (Coxhead, 2000) y ofrece las palabras más comunes en textos académicos. Estas dos listas, sin embargo, no proveen a los futuros estudiantes de negocios del vocabulario técnico específico de su área de estudio. Por lo tanto, una tercera lista más especializada y orientada hacia los negocios es necesaria para cubrir esta necesidad. El presente artículo describe el intento de crear una Lista de Vocabulario para los Negocios.*

*Palabras clave: inglés para los negocios, vocabulario técnico, corpus, listas de vocabulario, cobertura léxica*

## 1. Background

Foreign students studying business subjects at English-speaking universities rely heavily on their ability to read and fully comprehend a variety of business course books and other materials in the field. Before starting business courses at universities, these foreign students usually have little or no training in Business English specifically or in English for Academic Purposes (EAP) in general. An example is given by Tschirner (2004) who conducted a study with students in secondary schools in Sachsen, Germany, and found that even though the students had been studying English for more than eight years they did not necessarily meet the vocabulary thresholds required for understanding academic texts. Another case is mentioned by Nurweni and Read (1999) who conducted a study on how many words first-year students of an Indonesian university knew. The results showed that the students knew 986 words out of the General Service List's (GSL) (West, 1953) 2,000 words, and 240 out of the University Word List's (UWL) (Xue and Nation, 1984) 800 words, making a total of 1,226 words. This, however, was much less than the 4,000 words which were prescribed by the national curriculum for high school. Students who have inadequate or no training will, sooner or later, find themselves facing difficulties when they have to

cope with specialist texts. Reading comprehension, consequently, becomes a determining factor of their success.

## 1.1. Reading Comprehension and Vocabulary

Students' reading comprehension abilities are largely affected by their lexical level. Laufer and Sim (1985a), for example, while investigating how students with partial linguistic knowledge interpret a text, concluded that learners need mainly a strong vocabulary base in order to use contextual clues effectively. In another study, Laufer and Sim (1985b) showed that in their effort to understand texts learners relied firstly on word meaning, then on subject knowledge, and last of all on syntax. In their research, Ulijn and Strother (1990) suggest that reading depends chiefly on understanding the meaning of the text's words and its subject matter. Laufer (1992b) conducted a study in order to find out whether the learners' general academic ability or their lexical level is a better predictor of reading in L2, and her results indicated that it is their lexical level. Finally, Haynes and Baker (1993), in an experiment with college students on a text that included both linguistic and non-linguistic lexical familiarisations, concluded that L2 readers need not only better reading strategies but also greater vocabulary knowledge for comprehension. Vocabulary, therefore, appears to be the basic aspect in understanding a text.

Foreign students preparing for university level studies in business can be no exception to the above conclusions; they need adequate vocabulary in order to be able to understand and cope with their business course books. But how much vocabulary is exactly necessary for them to achieve successful comprehension?

## 1.2. Vocabulary Size

Laufer (1989) attempted to determine what the minimum lexical coverage for comprehension of authentic texts is. She found in her study that the students who could understand 95% of the words in the text she used for

analysis scored 55% and above on a reading comprehension test, which was an acceptable score for her experiment. In a further study (Laufer, 1992a), she tried to ascertain this time what vocabulary size is required by learners to achieve an acceptable score in a reading comprehension test. The results showed that a vocabulary size of 3,000 word families – or 4,800 lexical items – predicts a minimum reading comprehension test result of 56%, again an acceptable score. The two studies therefore suggest that a lexical coverage of 95% in a text, or knowledge of a vocabulary of 3,000 word families, is the much desired threshold for learners to be able to achieve satisfactory reading comprehension. However, Hirsh and Nation (1992), in a study conducted using three novels intended for younger readers, have suggested that a coverage of 98% is the desirable level for pleasurable reading. Hu and Nation (2000), using four different levels of coverage for a fiction text which was tested on learners, produced results which showed that at the 95% level some learners achieved adequate comprehension but the majority did not. They also concluded that learners would need to know 98% of the running words of the text in order to be able to read for pleasure, which supports the 98% coverage claim made earlier by Hirsh and Nation.

What kind of words, however, must be taught and learned in order to achieve either of the above percentages? The most frequent words in English is a first answer. These are words that make up the core vocabulary that is used in all kinds of texts. A collection of such words is the GSL. Since the present study has prospective university students of business as its target group, an additional number of words which are found in academic texts is a necessary supplement. These can be found in the Academic Word List (AWL).

**1.3. The GSL and the AWL**

The GSL is a list developed by West (1953) from a corpus of 5 million words and it is based on word counts carried out in the 1930s. It contains about 2,000 headwords mainly chosen on the basis of frequency (Nation and Waring, 1997). These 2,000 headwords are high frequency words and they

contain most of the 176 function words of English (Coxhead and Nation, 2001).

The AWL was developed fairly recently by Coxhead (2000) from the Academic Corpus. This is a corpus of 3.5 million running words which Coxhead compiled and it consists of texts from four subcorpora which are representative of academic subjects, i.e. Arts, Commerce, Law, and Science. The list contains 570 headwords that have been expanded to word families according to the word family definition of Bauer and Nation (1993), up to and including level 6. These are words which are not included in the GSL.

## 1.4. Validity of the GSL and the AWL

A question which readily comes to mind at this point is to what extend the GSL and the AWL are valid. Nation (2004) conducted a very interesting study to this end in which he compared the GSL and the AWL with three frequency lists he compiled from the BNC. Regarding the GSL, numbers, days of the week and months of the year were not included in the original list but were added by Nation for the purposes of his study. The headwords were expanded to word families using the word family definition according to Bauer and Nation (1993), up to and including level 6. As a result, the 2,000 headwords were reduced to 1,986 headwords in the end. Added to these were the 570 word families from the AWL, making a total of 2,556 headwords or word families. As for the BNC lists, he named the first list BNC 1000 because it contains the 1,000 most frequent words from the BNC, the second list BNC 2,000, and the third one BNC 3,000. It must be noted that the BNC lists contained 444 more word families (3,000 in total) than the combined GSL and AWL.

These two list groups were tested on four different corpora in order to find the coverage they provide. The results showed that the BNC lists provide only a slightly better coverage ranging from 0.9% to 2%. These findings are sound indication that the GSL and the AWL are extremely valid, most surprisingly given the date the GSL was developed.

### 1.5. How Much Coverage Do the GSL and the AWL Provide?

Nation and Waring (1997), based on a number of previous studies, report that the GSL provided a coverage of 78% to 92% in various kinds of written texts, giving an average of 82%. The AWL on the other hand gives a coverage of 10% in the Academic Corpus (Coxhead, 2000). But when tested on the four subcorpora of the Academic Corpus individually, i.e. Arts, Commerce, Law, and Science, it gives a surprising 12% for the commerce subcorpus and 9.0% to 9.4% for the other three academic areas (Coxhead, 2002). As Coxhead rightly remarks (2002), the AWL benefits the commerce students more than the others. The real potential of the lists, however, can be seen when these two are used together. When the GSL is combined with the AWL, they provide a coverage of 88.8% for the commerce subcorpus in the Academic Corpus, the highest in all four academic subjects (Coxhead, 2002). When tested on the whole Academic Corpus, i.e. all subcorpora included (3.5 million words), the two lists give 86.1% (Coxhead , 2000). Also, Nation (2004) reports that when he tested the GSL and the AWL on four other corpora, the two lists combined provided a coverage of 85.5% to 91.4%.

The above results therefore strongly indicate that knowledge of these two word lists could provide a very satisfactory basis for reading comprehension. These lists are a short-cut to efficient comprehension because they are the most important words that students need to know to cope with university texts.

### 2. Aims

The aims set for the present study are, therefore, the following:

1. To ascertain whether the GSL and AWL combined can provide a percentage of lexical coverage in business texts that reaches 95%.

2. If the two lists do not provide a 95% coverage, to attempt to create a new list made specifically of business-related words.

As has been pointed out earlier, two percentages have been suggested regarding the threshold for successful reading comprehension: 95% and 98%. However, the 95% was chosen as the aim because it is the minimum lexical level required for comprehension, and also because it was determined using academic texts, and this research addresses prospective university students of business.

## 3. Methodology

### 3.1. The Corpus

For the present study, the Published Material Corpus (PMC) compiled by Nelson (2000) was chosen. It contains about 600,000 words and it is made up of texts coming from 33 Business English course books selected over a period of ten years (1986-1996). These course books were selected according to a popularity criterion since it was considered important that the corpus reflected the actual books that Business English teachers and students use. It was upon this criterion that the particular corpus was chosen because it represents what the students of Business English were actually taught.

### 3.2. The Software Tool

The software used for the purposes of this study is RANGE (Heatley, Nation, & Coxhead, 2002). This software uses the GSL and the AWL as files. The words in both lists are expanded into their word families. The GSL is divided into two files, one containing the first 1,000 most common words and another file containing the second 1,000 most common words. The words of the AWL are all in a third file. The user can load one or more texts and the program will compare the text(s) against these word lists. The outcome of coverage is presented in percentages for each list. The words that are not found in any of these lists are also provided. One of the very useful features of the software is that it can also sort the words in alphabetical order and in order of range or frequency.

### 3.3. Procedure

The first aim of the study was to verify whether the GSL and the AWL would provide a percentage in coverage for the PMC close to the 90% which, as mentioned above, is characteristically the standard percentage that these two lists provide, or 95% to cover the business vocabulary as well. If the coverage proved to be 95%, it would mean that the GSL and the AWL words would be enough for prospective students of business to know. If, however, it proved to be 90%, then a Business Word List (BWL) would have to be developed in order to cover the gap. To that end, the PMC was run in RANGE and the results are presented in Table 1:

| 1k    | 80.26% |
|-------|--------|
| 2k    | 5.46%  |
| AWL   | 4.66%  |
|       |        |
| TOTAL | 90.38% |

Table 1: Percentage of coverage for the PMC

The results showed that the GSL and the AWL words cover a little more than 90% in the PMC, so the combination of the two word lists is not enough to reach 95%. This result provided an answer to the first aim of this research. It also meant that a BWL would have to be created to reach a 95% coverage. However, the percentage the BWL actually needed to cover in order to reach 95% was only about 5%, since more than 90% was already covered by the words of the GSL and the AWL.

The next step was to sort the words of the corpus not belonging to the GSL and the AWL as this was the pool from which the words of the BWL would come from. It was decided to start with the words that have a

high range of occurrence in the corpus. For the PMC, range represents the 33 course books that make up the corpus, and, consequently, a limit had to be set. It was decided that words that did not appear in less than five course books were not eligible.

        The following step was to take the selected words through a process of clearing. During this stage, all proper names, numericals, Latin words, nationalities, acronyms, and interjections were taken out. These are words which are considered known or with very low learning burden. This is a procedure that Coxhead (2000) has also followed for the AWL. Another category that had to be taken out was abbreviations. This was a difficult task since, apart from abbreviations like *Co* (for Company) or *Ltd* (for Limited), there were a number of others that although strictly speaking they are abbreviations, they are part of everyday vocabulary and accepted nowadays as such. Examples are *TV*, *CV*, *memo*, and *ad*. The criterion of abbreviation was strictly imposed and such words were also taken out. The remaining words were then expanded to their word families according to Bauer and Nation (1993), up to and including level 6, and that produced a list of 480 headwords. These comprised the first version of the BWL which had to be tested now for coverage. They were saved and added to RANGE as a fourth list. The PMC was run again in RANGE; this time it checked the corpus's words against the GSL, the AWL, and the newly developed BWL. The results are presented in Table 2.

| | |
|---|---|
| 1k | 80.26% |
| 2k | 5.46% |
| AWL | 4.66% |
| BWL | 2.55% |
| | |
| TOTAL | 92.93% |

Table 2: Percentage of coverage including the BWL

The results were very encouraging. 480 word families provided a coverage of 2.55% which gave a total text coverage of 92.93%, not very far from the original 95% aim. It was decided then that the BWL should be enriched with words coming from the frequency count of the corpus. For this reason, it was determined that only words appearing in the PMC ten times and more were eligible, and, as with the range count, they should be from those not belonging to the GSL and the AWL. After the frequency count was produced by RANGE and the words were selected according to the above criterion, the same process of clearing was applied as before, i.e. all proper names, numericals, Latin words, nationalities, acronyms, and abbreviations were taken out. The remaining words were then checked against the range list for cases of overlapping. Surprisingly, only 80 words were not in the range list. These 80 words were again expanded into their families and added to the range list making a new Business Word List with a total of 560 headwords. This was again added to RANGE as a fourth list file.

The PMC was run again in RANGE with the new version of the BWL and the final results are presented in Table 3:

| 1k | 80.26% |
|-------|--------|
| 2k | 5.46% |
| AWL | 4.66% |
| BWL | 2.79% |
|  |  |
| TOTAL | 93.17% |

Table 3: Percentage of coverage including revised BWL

The Business Word List gave a final 2.79% in coverage. The 80 words from the frequency list added to the range list a 0.24% increase. The final total of all the lists reached a 93.17% in coverage.

## 4. Discussion

The first aim of the present study was to check whether the GSL and the AWL combined could provide approximately 95% of lexical coverage in Nelson's PMC. The results showed that the two lists provide a 90.38% coverage. The second aim was to create a Business Word List which, added to the GSL and the AWL, would cover the gap and provide a total 95% coverage in the PMC. The final percentage achieved was 93.17%. This result is only partially successful. The words that comprise the Business Word List give an increase of 2.79% in the coverage of a business-related corpus, beyond the 90% provided by the GSL and the AWL.

Regarding the methodology followed, Nation (2001) suggests that in order to reach a coverage of 95% in academic texts, a vocabulary of around 4,000 word families would be necessary. These 4,000 word families consist of the 2,000 most frequent words taken from the GSL, the 570 word families from the AWL, and 1,000 or more words of technical nature. Cobb and Horst (2001: 319) suggest the same method, i.e. "extracting GSL and UWL (University Word List) terms from a corpus of domain texts, leaving a residue of terms which characterise the domain". This residue should account for 5%. In the present study, taking out the words of the corpus that belong to the GSL and the AWL left a considerable number of words which are specific to the business field that the corpus represents. Applying a set of selection criteria to this residue resulted in the creation of a new list of words which provides an increase in the coverage of the corpus. However, the number of these words was not 1,000, as Nation suggests above, but 560.

The 2.79% percentage achieved in this study also confirmed the prediction of Coxhead and Nation (2001) and Nation (2001) who claim that the technical vocabulary for any subject provides coverage of up to 5%. The original goal was indeed 5% but this was not finally reached. One reason

why the 5% was not achieved might be the selection criteria of the range and frequency counts. Setting the limit of range to five means that the Business Word List includes words that appear in five course books but also words that appear only five times in the whole corpus (i.e. once in five individual course books). Likewise, setting the frequency limit to ten means that certain words may appear ten times in the whole corpus but only in one course book. These shortcomings may account for a number of words that found their way to the list but which contribute a very low percentage to the total.

Another problem could be the corpus itself. The results showed that the GSL and the AWL cover a little more than 90% in the PMC. In the Academic Corpus, the GSL provided an average of 76.1% and the AWL an average of 10% (Coxhead, 2000) (see Table 4). In the PMC, the coverage provided by the GSL is much higher, reaching 85.72%, while the AWL provided a mere 4.66%, much lower than the respective percentage in the Academic Corpus. It must be stressed, however, that the PMC is not a corpus of academic texts but of texts taken from Business English course books. The above results could clearly be an indication of the vocabulary preferred by authors of these course books; they focus more on the frequent words, and, as a result, words of lower frequency are neglected. This should not be happening, however, in this case, because these books are supposed to be written for a specific purpose and target group in mind. As a result, more business terms should be there. The particular corpus of course books could not provide the required 5% which means that learners are not adequately trained in business terms. An additional disadvantage is the fact that the percentage of words belonging to the AWL is also much lower than usual. Consequently, learners are not well prepared for academic level language either.

|           | Academic Corpus | PMC    |
|-----------|-----------------|--------|
| GSL       | 76.1%           | 85.72% |
| AWL       | 10.0%           | 4.66%  |
|           |                 |        |
| TOTAL     | 86.1%           | 90.38% |

Table 4: Coverage of PMC and Academic Corpus

An interesting question which arises at this point is what kind of words the remaining 6%-7% in the corpus are and whether they are significant enough to be learned or not. To begin with, they include the words which were cleared out during the word list development process, i.e. all proper names, numericals, Latin words, nationalities, acronyms, and abbreviations. These were taken out because they are considered known or with very low learning burden. But exactly because they are not included in any of the file lists, they are regarded by the software as unidentified words, so they are part of the remaining 6%-7%.

Attempting to ascertain exactly what is their coverage percentage in the corpus was a point worth investigating. It was decided to measure in particular the categories of proper names, nationalities, and acronyms and abbreviations. For this purpose, an attempt was made to create respective lists that could be added to the software. The abbreviations and the acronyms that were taken out during the clearing process of the BWL were kept in a separate file. As a result, a list was made of them and was added to the software. When the excluded acronyms and abbreviations are added to the Business Word List, they contribute an increase of 0.30%. We have to bear in mind though that this is a list containing the abbreviations and acronyms within the set range and frequency limits. For the proper names, a list which was developed by Nation was used. This is a list which comprises proper names taken from the BNC. This list is part of Nation's RANGE software (Heatley, Nation, & Coxhead, 2002) which uses files from the BNC. When

added to the original RANGE files, the list of proper names gave a percentage of 1.63%. Regarding nationalities, a new list was developed containing the most common nationalities. When this list was added to the software, it gave a percentage of 0.55%.

The valid question to ask now is whether any of these three lists is worth including in the original Business Word List. It was thought that the acronyms and abbreviations list was a legitimate list to add. The reason is that the words of the GSL, the AWL and the BWL are supposed to be taught to students in order to prepare them for university studies. However, when the word 'company', for example, is taught to students, it does not necessarily mean that the abbreviation *Co* will also be taught. Other examples are the words *Limited – Ltd*, *Brothers – Bros* etc. As far as the acronyms are concerned, they must be regarded as multi-word expressions. *VAT – Value Added Tax*, for example, will be taught as such because it conveys a meaning that needs all three words to be used. But is it going to be taught with the word *value*, *add* or *tax*? The same goes for *AGM – Annual General Meeting*, etc. Personal experience has shown that it is not going to be taught with any of the three words that make up the expression, but simply as a combination, and only when that is encountered in a text. Unless, therefore, specific mentioning is made in the teaching material, acronyms and abbreviations do not stand a good chance of being taught. Since the Business Word List aspires to serve as teaching material to prospective students of business, it was deemed important to include the abbreviations and acronyms in the list in the end. This short list included also the common abbreviations, such as *TV*, *ad*, or *CD*, which were taken out in the first place together with the business abbreviations. This way, the final coverage provided by the BWL is 93.47% (see Table 5). Concerning the other two categories, i.e. proper names and nationalities, they have no or very little learning burden, and for this reason they were not finally included.

| | |
|---|---|
| 1k | 80.26% |
| 2k | 5.46% |
| AWL | 4.66% |
| BWL | 2.79% |
| TOTAL | 93.17% |
| Acronyms/Abbreviations | 0.30% |
| NEW TOTAL | 93.47% |

Table 5: Coverage including acronyms etc

However, when one takes into consideration the negligible learning burden of proper names and nationalities, a very interesting result occurs: with the BWL's lexical coverage of 93.47%, and provided students can understand words such as *London*, *Peter*, and *Mexican*, the comprehension percentage coming from the BWL is actually 95.65%, when the percentages of proper names and nationalities are also added (see Table 6). Although the 95% in lexical coverage was not achieved, it appears that a comprehension coverage of 95% is feasible. What this means is that although students may have been taught all the words of the GSL, the AWL, and the BWL, they can actually understand more owing to the extended coverage that proper names and nationalities provide.

| 1k | 80.26% |
|---|---|
| 2k | 5.46% |
| AWL | 4.66% |
| BWL | 2.80% |
| Acronyms/Abbreviations | 0.30% |
| Proper names and Nationalities | 2.18% |
| NEW TOTAL | 95.65% |

Table 6: Percentage coverage including proper names and nationalities

## 5. Conclusion and Future Research

Students preparing for university business subjects need to be trained with technical vocabulary that will help them cope with the texts of their university course books. A corpus with texts from Business English course books was used in order to find this required technical vocabulary. As a result, a Business Word List was developed using a set of selection criteria. When this list is added to the GSL and the AWL, they all provide a percentage of 93.47%. This list includes the acronyms and abbreviations which are so necessary for effective comprehension of business texts. Assuming also that students have relatively no difficulty in identifying proper names and nationalities, the actual comprehension coverage reaches 95.65%. This percentage goes slightly beyond the 95% set by Laufer's (1989) study for successful comprehension in academic texts, and that was one of the aims of this study. What this means in practice is that students, provided they already know the words belonging to the GSL and the AWL, will be in an advantageous position to achieve greater comprehension regarding business texts, once they also acquire the words coming from the BWL.

A possible way to raise the coverage percentage of the Business Word List even more would be to lower the range and frequency criteria. If the limit goes further down, i.e. less than 5 for range (which means including words which appear in less than five Business English course books) and less than 10 for frequency (include words which occur less than ten times in the whole corpus), then a considerable number of words will be added to the already developed BWL, possibly contributing an even greater percentage of coverage. The question, however, is exactly how useful words of frequency of 3, for example, can be.

The present BWL must be tested with students. A study exploring whether this BWL is enough for students to achieve satisfactory results in actual reading comprehension tests of business texts is essential, and will verify the validity of the list. In the future, a study using another corpus may yield a higher percentage of coverage. This other corpus may be made up of texts coming from actual university business course books, and not from Business English course books, as was the one used for the present study. This way a more representative list will be developed and hopefully serve as a very effective material to be taught to prospective students of business.

**References**

Bauer, L. & Nation, I.S.P. (1993). Word families. *International Journal of Lexicography*, 6/3, 253-279.

Cobb, T. & Horst, M. (2001). Reading academic English: Carrying learners across the lexical threshold. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 315-329). Cambridge: Cambridge University Press.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-239.

Coxhead, A. (2002). The academic word list: a corpus based word list for academic purposes. In B. Ketteman & G. Marko (Eds.), *Teaching and Learning by doing corpus analysis* (pp. 73-89). Amsterdam: Rodopi.

Coxhead, A. & Nation, I.S.P. (2001). The specialised vocabulary of English for Special Purposes. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 252-267). Cambridge: Cambridge University Press.

Haynes, M. & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English text. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second Language Reading and Vocabulary* (pp. 130-152). Norwood, NJ.: Ablex.

Heatley, A., Nation, I.S.P. and Coxhead, A. (2002). RANGE and FREQUENCY programs. Retrieved March, 28, 2006, from http://vuw.ac.nz/lals/staff/Paul_Nation

Hirsch, D. & Nation, I.S.P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a foreign language*, 8,2, 689-696.

Hu, M. & Nation, I.S.P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13, 1, 403-430.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Laurén & M. Nordmann (Eds.), *Special language: from humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.

Laufer, B. (1992a). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 126-132). London: Macmillan.

Laufer, B. (1992b). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability. *Journal of Research on Reading*, 15, 2, 95-103.

Laufer, B. & Sim, D.D. (1985a). Taking the easy way out: non-use and misuse of contextual clues in EFL reading comprehension. *English Teaching Forum*, 23, 2, 7-10.

Laufer, B. & Sim, D.D. (1985b). Measuring and explaining the threshold needed for English for academic purposes texts. *Foreign Language Annals*, 18, 405-413.

Nation, I.S.P. (2001). *Learning Vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, Acquisition and Testing* (pp. 3-14). Amsterdam: Benjamins.

Nation, I.S.P. & Waring R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.

Nelson, M. (2000). *A corpus-based study of Business English and Business English teaching materials*. Unpublished PhD Thesis. Manchester: University of Manchester.

Nurweni, A. & Read, J. (1999). The English vocabulary knowledge of Indonesian University students. *English for Specific Purposes*, 18, 2, 161-175.

Tschirner, E. (2004). Breadth of vocabulary and advanced English study: An empirical investigation. *Electronic Journal of Foreign Language Teaching*, 1, 1, 27-39. Retrieved May 31, 2006, from http://e-flt.nus.edu.sg/v1n12004/tschirner.htm

Ulijn, J.M. & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading*, 13, 38-54.

West, M. (1953). *A General Service List of English Words*, London: Longman, Green & Co.

Xue, G. & Nation, I.S.P. (1984). A university word list. *Language Learning and Communication*, 3, 215-229.

**Appendix**

**Business Word List**

| | | |
|---|---|---|
| above-mentioned | in-depth | taxation |
| accessories | inflation | teamwork |
| acid | ingredients | technicians |
| adverse | inhabitants | technologies |
| aerospace | in-house | telecommunications |
| after-sales | injection | telefax |
| agenda | insist | television |
| aggressive | install | telex |
| aircraft | insurers | tennis |
| airline | interview | terrific |
| airport | irrevocable | territory |
| airways | jacket | textiles |
| alarm | jeans | theft |
| alliance | junction | prawns |
| amazing | junior | preferably |
| announced | keen | premises |
| antispy | keyboard | premium |
| anti-theft | kids | presentations |
| appeal | know-how | prestige |
| appliances | laboratories | pre-tax |
| appraisal | lading | price-list |
| aptitude | lamb | printer |
| assemblers | lane | processor |
| asset | large-scale | professor |

atmosphere
audit
authorised
automotive
autonomy
award
awful
bankrupt
barriers
battery
beach
beer
beforehand
bet
bid
champagne
charter
chat
check-in
chemist
chocolate
cigarette
cinema
circuit
circulated
clearance
clerical
client
climate
coach
cocktail
cocoa
column
commercials
compact
competent
competitor
compromise
concert
concrete
confidential
conglomerate
conjunction
cons
consignment
continental

launch
layout
leaflet
leasing
left-hand
leisure
lemonade
liability
lieu
lifestyle
inventory
invoice
literally
logo
long-term
lounge
lucrative
luggage
lunchtime
luxury
madam
magazine
re-engineering
reference
refund
remittance
remote
remuneration
rental
resort
resume
retail
ridiculous
risen
robots
rocket
managerial
manpower
marine
marital
marketplace
mark-up
massive
maternity
medium-sized
medium-term

profile
profitability
promotional
prone
proposition
protectionist
prototype
provisional
pub
queries
questionnaire
radius
rally
ratings
recall
reception
recession
recipe
reckon
recruit
recycled
redesign
redundancy
resacked
salad
salesman
sandwich
satellite
score
scrap
seasonal
secretarial
self-employed
seminar
senior
session
setback
set-up
shampoo
shareholder
shipper
shortlist
short-term
sirs
site-sharing
ski

| | | |
|---|---|---|
| conveyor | memorandum | slice |
| cope | menu | slump |
| correspondent | merchandise | smart |
| cosmetics | merge | software |
| cost-effective | mess | sophisticated |
| counter | metropolitan | specialising |
| counterparts | misunderstood | spectacular |
| county | mobile | spicy |
| crane | modem | spokesman |
| crates | monetary | sponsorship |
| crazy | morale | squash |
| credibility | mortgage | staffing |
| far-reaching | motorway | stake |
| fax | movie | starters |
| feasible | multinational | start-up |
| feed | museum | state-of-the-art |
| feedback | nationality | static |
| festival | negotiate | stationery |
| fiscal | nervous | stereo |
| fittings | niche | sterling |
| flexitime | notify | stimulating |
| fluent | nutshell | storage |
| fluorescent | objectives | strain |
| follow-up | obliged | strengths |
| foothold | obstacle | tick |
| forecast | old-fashioned | tiles |
| foremost | on-board | timing |
| fortnight | opera | tiny |
| fragile | opposition | token |
| franchise | optimistic | tone |
| freight | organiser | toner |
| fringe | outdoor | tonnes |
| fuel | outlets | tourism |
| fulfil | outset | traffic |
| full-time | overdraft | trainee |
| gamble | overhead | traits |
| garments | overnight | transaction |
| gateau | overtime | traveller |
| gear | overview | tremendous |
| electronic | pace | trolley |
| embarrassing | packaging | truck |
| emotional | palace | tumble |
| engaged | panic | tunnel |
| enquiries | paperwork | turnover |
| enterprise | par | underground |
| enthusiasm | part-time | unexpected |

entitled
entrepreneurial
equity
executive
exhibitors
expenditure
fabric
fairs
fake
fantastic
genuine
geographical
giant
gin
glance
golf
goods
goodwill
graduate
graph
grocers
gross
gulf
guy
halt
ham
hardware
hard-working
headaches
headhunter
headquarters
height
holdings
homemakers
homework
hopefully
hospitality
hostile
household
huge
humour
import
impressed
inconvenience
incurred

passport
pasta
patent
payable
peak
peer
penalty
pence
pension
personality
personnel
petrol
pharmaceutical
photocopy
phrases
pie
pizza
planet
plastic
platform
port
portable
portfolio
postage
potatoes
potteries
rotation
routine
rumours
studio
subsidiaries
suite
super
superior
supermarket
supervision
surplus
swap
switch
tackle
tactics
takeover
talent
tank
tariffs

update
upgrade
up-to-date
upturn
urgent
vacancies
vacation
vast
vegetables
vegetarian
venture
venue
vertical
viable
vice
vice-president
video
viewpoint
vital
walkman
warehouse
waybill
weaker
well-known
whatsoever
whisky
wholesale
withdraw
workforce
workplace
workshop
worldwide
worthwhile
yacht
yen
zone

**Abbreviations List**

| | | |
|---|---|---|
| Ab | Ft-Sef EEC | Ref |
| Ad | Enc | Reg |
| Advert | Encl | Rep |
| Ag | Ffr | Sa |
| AGM | GATT | Tel |
| Asap | Gdp | Telecom |
| Asn | Gmbh | TV |
| Attn | GNP | USD |
| Auto | Hi-Fi | VAT |
| Bros | Inc | |
| Cc | Ltd | |
| CD | MBA | |
| CEO | MD | |
| Cif | Memo | |
| Co | PA | |
| Corp | PC | |
| CV | Plc | |
| Del | Pp | |
| Dept | Pr | |
| DTI | Pty | |
| EC | Recap | |