Creating and Validating an Advanced Practical Word List

BY

Wang Chen

School of Linguistics and Applied Language Studies

A dissertation

submitted to the Victoria University of Wellington

in fulfillment of the requirements for the degree of

Master of Arts in Applied Linguistics

Victoria University of Wellington

2012

Abstract

This study aims to create and validate a list of highly practical words used in the daily lives of people living in an English-speaking environment. These words are known to most native speakers of English, but are unknown to most advanced non-native speakers. The resulting list will be of great value to advanced non-native users of English who may function well in their respective domains of work but encounter many embarrassments for not having the vocabulary for daily communications.

Three contemporary dictionaries were used in the initial development of the Advanced Practical Word List (hereafter APWL). Seven native speakers were then asked to rate the usefulness of the initial word list. Sixty advanced non-native speakers and 40 native speakers of English were asked to identify the words that were known to them. Based on these criteria, 867 word families have been identified by all native speakers as of high practical value but were unknown to most of the advanced non-native speakers in this study. The APWL comprises words from a wide range of frequency levels, and includes both words that should have been known to the advanced learners, and those that are beyond their proficiency level.

The APWL is divided into 25 categories, each representing a type of words useful in daily life scenarios. Additionally, the APWL is divided into two types of sublists, one based on the words' frequency levels and another based on the potential usefulness of the words.

The APWL is an important aid for advanced learners in an English-speaking context. It helps them cope with words that are frequently used by native speakers in daily life. The APWL is also important, as it attempts to look for another type of vocabulary in addition to the four tiers of words identified by Nation (2001).

Acknowledgements

I wish to thank my supervisor, Dr. Peter Gu, for his encouragement, support and kindness during the writing of this dissertation. I am also grateful to the participants who devoted their time in the project. Finally, my thankfulness goes to my parents and my husband.

Table of Contents

List of Tables

1　Introduction

The focus of this study is on words that are known to most native English speakers, but unknown to most non-native speakers. These are words that are used in daily communication. This group of words has not been identified in Nation's four-level description of vocabulary (Nation, 2001), although they are all low frequency words. This type of words is not easily learnt by non-native speakers, even at a high proficiency level. This is partially because these words are not as frequently encountered as high frequency words, and therefore, they are not included in the learning programs for intentional learning. Furthermore, these words are known to native speakers from early on in life through authentic use of the language. However, second language learners in foreign learning context mostly study the language in classrooms through textbooks. The chances for them to use the words for real life communication are very limited. Therefore, daily life words are unlikely to be repeated enough for incidental learning to occur (See Webb, 2007, for a review).

Practical words are important for advanced learners studying, or living, in an English-speaking context, or for teachers looking for words to prepare their learners for such a context. The number of low frequency words is far too large to be learnt one by one, and therefore, advanced learners and teachers need a word list outlining those words that should be learnt primarily.

Research into the practical words required for daily life communication is quite limited. Crabbe and Nation (1991) developed a syllabus for people visiting an English-speaking country for a short period of time. It includes 120 terms that are easily learned by those with very limited prior knowledge of the language. This syllabus is salient for those who wish to visit the English speaking countries. However, it would not be as salient for advanced learners of English, whose vocabulary sizes exceed the most frequency 2000 words.

The aim of the present study is to identify words that are known to most native speakers, but unknown to most advanced non-native English speakers. These words

include those used in a wide range of daily life scenarios, and have not been studied in any prior research on word lists. The determination of these words is based on a set of criteria especially developed for the purpose of the present study. Both native speakers and advanced non-native speakers were invited to use their knowledge to validate the list.

2   Literature review

Nation (2001) classified words into four levels: high-frequency words, academic words, technical words, and low-frequency words. The classic high frequency word list is the General Service list of English Words (West, 1953). It is agreed that high-frequency words are important for all learners of English, because of their wide and consistent coverage of texts of most genres. Curriculum designers and graded readers have used the 2000 high frequency words as a basis for writing textbooks and set up learning goals.

Moving to higher learning needs, Coxhead (2000) developed the Academic Word List, which contains 570 word families from a corpus covering a wide range of academic topics and materials. The AWL covers around 10% of all the words in the academic corpus, and thus, is important for all learners studying English for academic purposes. For words found in specialised contexts, Chung and Nation (2003) used two different texts, one on anatomy and the other on applied linguistics, to estimate the size of technical vocabulary. They found that technical words account for a larger coverage of the language than the 5% suggested by Nation (2001). The low frequency words comprise the vast majority of vocabulary in English that does not fit into any of the previous categories. Nation (2001) suggested that these words are far too numerous for classroom learning, and that learners should use strategies to cope with these words.

This literature review has two aims. The first is to determine the goals and purposes of existing word lists. The goal and purpose directly influence the type of corpus that the words are chosen from, and the criteria that were used to select the target words. The second aim is to look at these criteria, and their implications for the present word list.

2.1 High frequency words

High frequency words refer to the words which are used very often in most uses of the language. Hwang and Nation (1995) used the criteria of frequency, coverage and range to look for high frequency words from LOB (Lancaster-Oslo-Bergen corpus) and Brown corpora, each containing 1,000,000 running words. They, then, compared the words chosen from the two corpora with the GSL, and provided evidence that 2,000 words was the reasonable cutting point for high frequency words. The cumulative coverage of this group of words in the corpora is 83.4%. This suggested that a general service vocabulary of around 2,000 words is appropriate for learners for general language use purposes, but that a more specialised vocabulary is needed for specific learning purposes (Nation & Hwang, 1995).

2.1.1   The General Service List of English Words

The classic word list for high frequency words is the General Service List of English words (West, 1953). West developed the GSL list based on the 1936 Interim Report on Vocabulary Selection, and incorporated the semantic counts. The GSL consists of around 2,000 word families, and covers around 80% of general texts (Nation, 2001). The GSL has considerable overlap with other high-frequency word lists built from other corpora (Nation & Hwang, 1995).

Because the purpose of developing the GSL is to select words for teaching of English to foreigners, the selection of words takes into consideration the learning aspects of language. The words in the list were selected from a corpus as large as 5,000,000 running words, based on written materials from various sources, such as encyclopedias, magazines, textbooks, novels, essays, biographies, etc (West, 1953, pp.xi). The word list consists of semantic counts of each meaning of a given word to estimate the frequency of usage of that meaning. The classification of meaning was based on the thirteen-volume *The Oxford English Dictionary.* One notable point about the GSL is that it outlined both the frequency of occurrence of each word and the percentage of occurrence of each meaning of the word. Through this, the list

attempted to help teachers recognise the necessity of the learning different words and reduce the learning burden of multiple meanings of polysemous words. The semantic counts were based on several criteria in addition to frequency of occurrence. They include:

1. *Ease of learning:* Easy words are selected over the difficult ones.
2. *Necessity:* A low frequency word may be selected if it covers a certain range of necessary ideas.
3. *Cover:* A frequency word may not be selected if it has close resemblance in meaning to another word.
4. *Stylistic level:* Common words are selected over its literary or colloquial counterparts.
5. *Intensive and emotional words:* Words to express emotions were of secondary importance, thus, they were not selected.

(West, 1995, pp. x)

These criteria accentuate the purpose of the word list; that is, to help learners at the early stage of learning English. Also, it is innovative, in the sense that it took into consideration the learning burden of polysemous words. The *necessity* criterion gave preference to words with multiple meanings that cover several necessary ideas. Recent studies on cognitive aspects of vocabulary acquisition suggest that the learning burden of a polysemous word could be largely reduced by introducing the mental link between the core and peripheral senses (see e.g. Verspoor & Lowie, 2003; Tyler & Evans, 2004). When learners encounter new words, it is easier for them to learn these words' multiple meanings, rather than learning several words each with one meaning. This is because, in most cases, there are links between the core or conceptual sense, and the meanings.

However, the criteria also present some problems. Firstly, the *stylistic level* criterion means that learners keep following the "middle path", using words with little or no stylistic preference. However, in real usage of English, people are more

likely, for example, to use "fellow" instead of "person" to refer to someone in their conversations. The exclusion of colloquial words may make learners sound funny in real-use of the language. Secondly, *intensive and emotional words* were not included in the list. West used this criterion based on the idea that learners study the language in order to express ideas, not emotions. This may be true for learners at the survival level of learning; however, as they advanced, it is inevitable that they will need more subtle words to express their emotions and ideas. Despite these issues, the GSL is still a useful guide for learners at the beginners' level. But for those who are more advanced, a word list with more vivid and precise words is required.

2.1.2 The BNC 2000

A more recent effort to develop a word list is the written and spoken word frequency list, based on British National corpus (Leech, Rayson & Wilson, 2001). There are two purposes of this word frequency list. Firstly, it aims to address the criticism of the older word lists, like the GSL. The GSL has been the subject of much criticism since the 1970s (e.g. Richard, 1974；Hwang, 1989). This is partially due to the age of the list, and in particular the age of the materials used to build the list (Leech, Rayson & Wilson, 2001). Most of the materials that the GSL used were collected in early 1990s. There is a need for a more updated word list which can reflect the current usages of the language. Secondly, the word list attempts to make use of the modern computer technology to develop a word frequency list for research on style, register and psychological processing of the language. These purposes lead to two major differences between the two word lists.

The first difference is the size of the corpus and the sources of language samples. The BNC word frequency list is based on the British National Corpus (hereafter the BNC). The BNC is twice the size of the corpus s used to develop the GSL, and uses samples from more varied sources. The corpus is said to be "a finite, balanced, sampled corpus" (Leech, Rayson & Wilson, 2001). The BNC corpus consists of materials from both written and spoken sources. The written component includes

both informative texts and imaginative texts. The informative texts make up 80% of the written texts, and are classified into 8 domains: world affairs, social science, leisure, applied science, arts, commerce, natural science, belief and thought. The sample taken into the corpora attempts to reflect the book published in these fields. On the other hand, the imaginative texts were sorted into different genres of poetry, prose and drama.

The spoken sources of the corpus consist of conversational data and task-oriented data. For the conversational data, over 1,000 speakers were asked to record their daily conversational interactions with other interlocutors. In contrast, for the task-oriented data, it recorded more formal activities of daily life, for example, lectures, consultations, sermons, broadcasting, etc. The majority of texts dated from 1985 to 1994, with spoken data dating no earlier than 1991. Together, the written and spoken sources make up the 100,000,000 running words corpus, which could be regarded as the most comprehensive representation of the present-day usage of English.

Another difference between the BNC word frequency list and the GSL is the criteria used to develop the list. The BNC word frequency list uses frequency, coverage and dispersion to reflect the usage of words in present-day English. Leech *et al.* (2001) sorted the 100 million words corpora into 100 one million word sub-corpora, with similar types of texts in each one. All the words in their list have been lemmatized. A lemma includes the base form of a word and its inflections. For example, the words *run, ran, runs, running*, belong to one lemma, run. Each lemma and its inflections in the list have been accompanied by three figures. These are for frequency, which shows how many times the lemma and its inflections have occurred in the whole corpora; coverage, how many sub-corpora the lemma and its inflections appear in; and, dispersion, how evenly the lemma and its inflections have been used across all the sub-corpora.

Using frequency, coverage and dispersion of the headwords, Nation (2004) identified the 3000 most frequent words in the BNC. One difference between Leech *et al.*'s word frequency list and Nation's word lists is the unit of counting for an entry.

Instead of using lemma, Nation (2004) used word families. A word family is a word, plus related inflections and derivations. For example, *access, accesses, accessible, inaccessible,* and *accessibility* are all included in the word family under the headword *access*. Nation took over 6,500 headwords with a frequency equal to, or higher than, 10,000 in the whole 100 million token corpus. This step ensures that the words chosen are of the highest frequency in the whole corpus. He then used the range criterion to remove those words which appeared in less than 98 of the 100 sub-corpora. This ensures that the words chosen are widely used in all types of texts. Finally, dispersion was used to further remove those words which are lower than the threshold. This step ensures that all the lemmas in the list are evenly distributed in the 100 one million sub-corpora. Nation also developed the second and third 1000 words in the same way with the remaining 6500 headwords.

2.1.3 Conclusion

This literature review has examined previous studies on high frequency words, with a specific focus on the word lists developed for the learning needs of non-native speakers of English. Special attention was also given to the size and sources of the corpus, and the criteria used to select potential words for the word lists. The two word lists examined, the General Service List (West, 1953) and the BNC 2000 (Nation, 2004), have been developed for somewhat different purposes, which has some implications for the present study.

Firstly, both lists use word families as the unit of counting words, while the target audiences of the two lists are all learners of English. Therefore, the unit of counting words needs to take into consideration how learners organise words in their mental lexicon. There is research evidence showing that learners psychologically group members of a word family together (Nagy *et al.*, 1989; Schmitt & Zimmerman, 2002). Knowing a word in the word family can also greatly assist the learning of other words in the family. In light of this, the present study will also use word families (further definitions in the methodology section).

Secondly, the GSL used *stylistic level* as a criterion to select potential words. Stylistic level is important because it helps distinguish the type of words selected for the list. GSL preferentially includes words that could be used in common settings to avoid any clear stylistic preferences, because common words are versatile and would ease the learning burden of beginning L2 learners. The present study attempts to identify words that are used in more casual settings for advanced learners, and therefore, will use stylistic level as a criterion to select potential words.

Finally, both GSL and BNC 2000 are reliable word lists for high frequency words. They are used as a criterion in the present study to exclude any potential words present in these lists, as words in these two lists are used for general purposes of English and should be already known to all advanced learners of English.

2.2 Technical vocabulary

When beginner learners of English have mastered the first 2000 high frequency words, they need vocabulary for more specialised purposes. An example is academic research in an English context where a specific subject area requires a more specialised range of vocabulary. Therefore, learners face the need to know a more specialized range of vocabulary. Nation (2001) introduced two types of specialised vocabulary: technical words and academic words. In this section, the size and types of words in the technical vocabulary are covered. The methods used to create this group of words, and the methodology and purposes of the Academic Word List (Coxhead, 2000) are discussed in the next section.

Technical vocabulary includes subject related words that are important for learners of that particular subject. There is much dispute on what type of words could be defined as technical words. Moreover, categorising technical vocabulary requires subject related knowledge to distinguish words that carry special meanings in that subject. Despite this, there is a wide range of technical vocabulary, and it is indispensible for relevant students (Chung & Nation, 2003).

The coverage of technical words in a text varies by subjects. Chung and Nation

(2003) compared two texts on different subjects, anatomy and applied linguistics. The coverage of technical vocabulary in the anatomy text was twice that in the applied linguistic text (37.6% and 16.3% respectively). This report used a four-level rating scale to identify technical words. These four scales distinguished two types of technical words: those used only in the specific type of text, and those also used in general texts. Step one of the rating scale includes function words that are not related to the subject area, while step two includes words that are minimally related to the subject area. Step three includes words closely related to the subject area, and step four includes words unlikely to be known in general use of the language. This scale was later examined in a studying comparing the reliability of four methods used to identify technical words (Chung & Nation, 2004). These four methods were the rating scale, using a technical dictionary, using clues provided in the text, and using a computer-based approach with term extraction software. Rating scale was found to have the highest reliability in identifying technical terms.

2.3 The Academic Word List

The most widely used academic word list in the past decade is the Academic Word List (AWL) developed by Coxhead (1998, 2000). The AWL consists of 570 word families. The list is not simply made up of past word lists, as University Word List was (Xue and Nation, 1984). Rather, it is built from a large, specially complied academic corpus, with 3,513,330 running words that cover written academic texts from four academic disciplines and 28 subject areas.

The AWL was developed for two reasons. Firstly, Coxhead (1998) reviewed past research on academic word lists, and found that they were developed from small corpora of scientific English. This was not large enough to be representative of academic use of the language. Further, those word lists only used a limited range of criteria to select words. Secondly, the AWL is designed to help learners of English cope with tertiary study in an English context. These dual purposes of the AWL are reflected in the size of the corpus, and the criteria for word selection.

2.3.1 Academic Corpus

The AWL is based on the Academic Corpus of Written English (hereafter Academic corpus), which was specifically compiled for the research. The corpus was based on three criteria: size, range and representation (Coxhead, 1998). With a total of 3,513,330 words, the corpus is almost ten times the size of previous academic corpora (Campian & Elley, 1971; Praniskas, 1972). This size ensures that the corpus includes sufficient samples of academic texts. The Corpus incorporates four major academic disciplines of similar size: the arts, commerce, law, and science. Within each of the four disciplines, there are seven subject areas. Thus, in total, there are 28 subject areas. The corpus also includes a wide range of academic texts of different genres to assimilate authentic academic uses of the language. The corpus collected samples from textbooks, as other academic corpora did, and also from other academic readings that the learners are likely to encounter in their studies (e.g. laboratory manuals, book chapters from set readings, work books, lecture notes, journal articles and technical reports). This ample variety of academic texts ensures that the Academic Corpus achieves a high representation of words encountered in academic study.

2.3.2 Four criteria: specialised occurrence, frequency, range, uniformity

The AWL used four criteria to choose words: specialised occurrence, range, frequency, and uniformity. The first criterion, words included in the GSL are excluded, ensures that the AWL is for the special purposes of academic study, instead of being a general word list which is needed by all learners of English. Additionally, it can be assumed that all learners studying at tertiary level in an English speaking country should already have good receptive knowledge of the GSL words. The second criterion is that words included in the AWL should appear in all four disciplines and also in 15 of the 28 subject areas. This ensures that the words are useful to academic study in most of the subject areas and disciplines.

The third criterion is that the words included in the AWL occur at least 100 times

in the whole corpus, in other words, once in every 35,000 running words. This cut-off point is for word families rather than lemmas. The AWL gives special concern to word families with only one member, like *forthcoming*, because the frequency of these words could be underestimated compared to other word families with many members. As such, the cut-off point for word families with only a single word is set at 80. A notable point is that frequency is secondary to range, allowing for the bias that may be caused by relying too much on frequency. This consideration enhances the validity of the AWL in that frequency is heavily context-dependent. A word that appears frequently in one subject may not be useful to another. For example, *collage* may be frequently used in a text talking about arts; however, it is hard to expect the same word with high frequency in an economic text.

The fourth and final criterion is that the words included in the AWL should appear at least ten times in any one of the sub-corpora. This resembles the dispersion criterion in BNC 2000, which also aims to find the words that are used widely across the corpus. This is to ensure that the words chosen are useful to all learners studying for academic purpose regardless of their specific subject areas.

2.3.3 Conclusion

The Academic Word List is important for learners studying at a tertiary level in most subject areas, and especially those looking for best return of their efforts in learning words for academic purposes. This section examined the corpora and the criteria used to select the potential words for the list. The development of the Academic Word List has several implications for the present study.

The first implication is that the size of the Academic Corpus ensures that the language samples collected could reflect the real life use of the language. The size of the corpus is important, in that it should be large enough to include substantial language samples consisting of ample frequently used words. The Academic Corpus is almost ten times the size of the corpora used to build other previous academic word lists. The present study uses dictionaries compiled from the two largest corpora

to ensure that the words selected will be representative of real life use of English.

Secondly, the AWL used specialised occurrence to remove those words that are part of general language usage. The AWL excludes any words that appear in the GSL to ensure that all the words on the list are used for academic purposes. Hwang and Nation (1995) suggest that learners with special language needs should study the vocabulary beyond the most frequent 2,000 words. The present study removes all the words that appear in the AWL to ensure that the words selected are not used specifically for academic purposes, but for communication in daily life scenarios.

2.4 Low frequency words

This group of words is by far the largest group of English words, and contains many thousands of words. Until now, this category has been understudied. Practically speaking, as English teaching time is limited, it is impossible to go through the low frequency words one by one. It is therefore advisable that learners be equipped with learning strategies to handle this enormous group of words (Nation, 2001). However, given the size of this category, and the needs of advanced learners, there should be studies into this level of vocabulary. McCarthy (2001) calls for study of the "long-tail vocabulary", which refers to all the infrequent words, stating that this group of vocabulary is too massive in size to be left for learners to handle on their own.

2.4.1   The size of low frequency words

The definite size of low frequency words has not been clearly stated, because this size depends on many aspects. One of the aspects is the definition of a word itself (Nation, 2001). Defining a word in lemma or in word family would no doubt affect the size of the vocabulary. For example, using the lemma definition, *enjoyment* would be a different word from enjoy, because they belong to different word classes. However, using the word family definition, *enjoy* and *enjoyment* are the same word.

Further, the number of low frequency words may depend on the group of people it refers to. Low frequency words for native speakers and non-native speakers

are likely to be dramatically different due to the difference in the vocabulary size of the two groups. Nagy and Anderson (1984) estimated that there were 88,500 word families based on the American Heritage Word Frequency Book compiled by Carroll *et al.* (1971). This number is what the native speakers might be exposed to in the school materials. It is estimated that five-year-old children may have a vocabulary size of 4,000 to 5,000 word families at the start of the schooling. The average vocabulary size of a university graduate is around 20,000 word families (Waring & Nation, 1997). However, non-native speakers studying in EFL context would be highly unlikely to gain such an enormous vocabulary. The concept of low frequency words here is for non-native speakers learning English as a foreign language. In SLA, it is now agreed that low frequency words refer to those words which are not in the most high frequency 2,000 words (Nation, 2001).

2.4.2 Type of words at the low frequency level

The category of low frequency words include many different kinds of words(Nation, 2001); for example, middle frequency words (not frequent enough to be listed as the first 2000 most frequent words), technical words for different fields of interests, proper nouns and words that are genuinely low frequency and seldom used.

There are two things to note about the words in this category. First, the compiling of any frequency list is based on the language samples in the given corpus, and therefore, the frequency varies from corpus to corpus. If the language samples collected for the corpus are from spoken sources, words like *er, yeah,* and *got*, would be among the most frequency words used. These words would be much less frequent in written corpus (Leech, *et al.,* 2001). A word not included in the high frequency list of one corpus would not mean that its significance for the learners should be underestimated. On the other hand, as the difference in vocabulary size between native speakers and language learners is huge, words infrequency used by language learners may be frequently used by native speakers, and vice versa. If the purpose of learning the language is to use the language like a native speaker, then, it

is important to be cautious about the boundary between the high and low frequency words.

2.4.3 Existing study into low frequency words

The role of low frequency words have been played down pedagogically; as such, past research produced little findings compared to that on other word categories. Also, the sheer size and diversity of the low frequency words makes studying this group a daunting and demanding task.

Until now, the mostly used low frequency word list in the relevant research is the BNC-20. This list is based on the British National Corpus, and made up of twenty one-thousand words, from the highest to the lowest frequency. The word list was developed in a similar way to BNC 2000 by Nation (2004). As such, frequency, range and dispersion were used as the criteria to choose words for each frequency level. This word list has been used in developing vocabulary size tests to make learning plans for students, and as a research tool for studies.

2.4.4 Conclusion

The number of low frequency words is too big to be learnt in any program. However, this does not mean that these words are not necessary for learners. Advanced learners, who have mastered the most frequent 2,000 words and academic words, need to study low frequency words to expand their vocabulary size and improve their proficiency in language learning. However, the present studies into low frequency words do not yield many findings to help learners learn such words efficiently. Without a valid word list grouping words for special needs together, learners must pick up such words in incidental learning, which is ineffective. Current low frequency word lists, like BNC 20, have their drawbacks. The size of the word list is much bigger than the high frequency words lists or the AWL to be realistically learnt. For a certain group of learners, some words are more useful for them than others. For example, for a learner entering a bakery, the high frequency word *bread*

is clearly not specific enough. Words, like *éclair* and *ciabatta,* are probably more useful in this case. However, these two words are dispersed in the frequency list, and are therefore unlikely to be learned together if the learning materials are developed only based only on frequency level of words.

2.5 Implication for the present study

This literature review evaluates research on word lists developed for three levels of vocabulary. Both basic and academic needs of learners have been analyzed, with word lists developed accordingly. For beginner learners, the GSL and BNC 2000 are great guides to master essential words for general English use. For learners with academic needs, the AWL is an excellent source of words which will be useful in a wide range of academic fields and topics. There is, however, a need for a new word list, catering to the needs of advanced learners, who require additional words for daily use in English-speaking contexts.

2.5.1 Survival Learning Syllabus

Previous research has focused on the lexical items necessary for communicative needs of foreign visitors to the English-speaking countries. The survival learning syllabus (Crabbe & Nation, 1991) was developed for visitors with little or no knowledge of English, who need only the most basic words and expressions to help them accomplish basic conversations during short term visits. The focus of this syllabus was on survival, travel, and social needs (Crabbe & Nation, 1991:192). The 120 expressions are classified into eight sections: greeting, bargaining, reading signs, getting to places, finding accommodation, ordering food, talking about yourself and to children, and controlling and learning language. However useful for survival needs, the items on the list would not be helpful for engaging and maintaining learners in conversations with locals. As such, a new word list is needed to satisfy the higher social and communicative needs of advanced learners.

2.5.2　The need for a practical word list

The difference between the size and composition of L1 and L2 learners' lexicon is largely due to the quality and quantity of input. The major source of input for L2 learners is written materials in textbooks, many of which are abridged to suit the proficiency levels of the learners. This results in a blank area in the learners' lexicon, i.e. the daily life words. Daily life words are used in everyday communication, and cover words used in various social settings, such as conversations between doctors and patients, when grocery shopping, in television programs and movies, at leisure events, and in talks between friends. Without these words, the ease and pleasure of natural language use will be compromised. Learners who do not know these words will have difficulty with both language input and output. They may only understand partial information in messages, or constantly use discourse strategies, like confirmation checks and clarification requests which undermine the fluency and naturalness of communication. For output, learners may have to resort to communicative strategies like paraphrasing which compromise the clarity of messages.

Surprisingly, there is no existing research in the SLA field that documents these words. Advanced learners of English already have a vocabulary greater than the 2000 most frequent words level, and need practical words that can help them achieve ease and pleasure in communication, while avoiding frustration and embarrassment. These words, though highly practical, may not be frequently used (e.g., snot), and may therefore not be easily picked up in natural use situations. These could be words that have been acquired by native speakers in childhood or through schooling, but have become less often used in adulthood. They could also be words from specialised domains of use that have entered into daily use, but have remained low frequency (e.g., sty). Singling out these words would have tremendous practical implications for advanced learners, and users of English living in a native speaker environment.

2.6 Research Question

The main task of this study is to develop and validate a word list for practical use in daily life by advanced learners/users of English. Advanced learners are operationally defined in this study as those learners from various L1 backgrounds, who are currently enrolled in an MA or PhD program in New Zealand, Australian, or British universities. The major research question is:

Which lexical items are perceived as practical in daily life and therefore known to all native speakers and yet unknown to most advanced non-native speakers living in an English-speaking context?

3 Development of the Advanced Practical Word List

3.1 Steps involved in developing the word list

Nation and Webb (2011) suggested a set of six steps for making word lists and demonstrated how to use these steps to create word lists by showing how the word frequency lists were made for British National Corpus. These steps are:

1   Decide on the research question the list will be used to answer, or the reason for making the list.

2   Decide on the unit of counting you will use

3   Choose or create a suitable corpus.

4   Make decisions about what will be counted as words and what will be put into separate lists.

5   Decide on the criteria that will be used to order the words in the list.

6   Cross-check the resulting list on another corpus or against another list

This study followed the steps above in building the Advanced Practical Word List (hereafter APWL), with some changes in the details of implementations. The present study uses the word family as the unit of counting, and the members in each word family include all the derivational forms in the six levels developed by Bauer and Nation (1993). Instead of using corpora to select potential words, this study used three contemporary dictionaries as the primary source. Five criteria were used to narrow down the initial word list developed from dictionaries to a manageable size. These five criteria are: special occurrence, stylistic level, time constraint, usefulness, and dialects. Both native speakers and advanced non-native speakers were invited to participate in the study to validate the list. In the following section, all these steps are discussed in great detail.

3.2 Unit of word for selection

In this study, the words were selected based on word family. Word family has been adopted by many recent word lists as the unit for selection, including the General Service List (West, 1953), the Academic Word List (1998), and BNC 2000 (Nation, 2004).

3.2.1 What is a word family?

The reason behind using word family as the unit of counting is the assumption that, once one word in the family (especially the base word) is known, the acquisition of other words in the family would not present much difficulty to learners (Nagy *et al.*, 1989). In this study, I use the definition proposed by Bauer and Nation (1993) on word family. They pointed out that:

"From the point of view of reading, a word family consists of a base word and all its derived and inflected forms that can be understood by a learner without having to learn each form separately. "

(pp. 253)

In their study into the inflections and affixations of English words, Bauer and Nation (1993) used findings on English productivity, frequency, regularity and predictability to support the levels of affixation as a basis for teaching and learning words.

In this study, the word family includes all related words up to level six of affixation. In the process of looking for potential words in the dictionaries, identifying word family was not generally a problem, since dictionaries usually group members of a word family closely together, often in the same entry. At the same time, receptive morphological knowledge was required to identify the base word and other extended members in a word family.

Table 3.2.1 Different levels of inflections and affixations

(Bauer & Nation, 1993: 254)

Level	Affixation and inflection
2	-s, -ing,-ed,-er,-est, (all inflections)
3	-able, -er, -ish, -less, -ly, -ness, -th, -y, non-, un-,
4	-al, -ation, -ess, -ful, -ism, -ist, -ity, -ize, -ment, -ous, in-
5	-age, -al, -ally, -an, -ance, -ant, -ary, -atory, -dom, -eer, -en, -ence,- ent, -ery,-ese,-esque,-ette,-hood,-I,-ian,-ite,-let,-ling,-ly,-most,-ory,-ship,-ward,-ways,-wise,ante-, anti-, arch-, bi-, circum-, counter-, en-, ex-, fore-, hyper-, inter-, mid-, mis-, neo-, post-, pro-, semi-, sub-, un-
6	-able, -ee, -ic, -ify, -ion, -ist, -ition, -ive, -th, -y, pre-, re-.

3.2.2 Why should word family be used as a unit of counting?

The main reason for using word families in this study is because it is assumed advanced learners of English already have some knowledge of the morphological makeup of words. They should be able to recognise the relationship between words like *ideal* and *idealism,* and their understanding of the latter would not present much difficulty if they have prior knowledge of the base form *ideal* and the suffix *-ism*.

Past research has provided evidence supporting the idea that there is a mental link between the base form, its inflections and derivations. Nagy *et al.* (1989) asked 109 native speakers at college level to perform word identification tasks. The experiment aimed to find out whether the frequency of a single word affects the speed of recognition, or, instead, the frequency of a word and its inflectional and derivational word family as a whole has the effect. The results showed that the speed and accuracy of word recognition is strongly influenced by inflectional and derivational relation between the words. The findings lend support to the presence of inflectional and morphological relations in the mental lexicon, and therefore, support the use of word family as the unit of counting.

There are a few concerns surrounding the use of word families as the unit of counting words. Chief among these is that learners would not have native-like knowledge of morphological relations between words in a family. Schmitt and Zimmerman (2002) carried out a study asking non-native participants at postgraduate and undergraduate levels to write down the derivational forms of stimulus stem words. They found out that only rarely could participants produce all four different derivations of the stimulus words. They suggested that learners only have partial knowledge of derivational forms.

3.3 Sources of words

3.3.1 Why does this study use dictionaries instead of corpora?

Considering the purpose of this study is to determine words used in daily life scenarios that are known to native speakers, but unknown to advanced non-native speakers, it was decided to use dictionaries as the source to select potential words from. There are two main reasons why dictionaries instead of corpora were chosen for this study.

Firstly, the size of the corpora makes it impossible to go through every sample to look for potential words. The present corpora are too large for any one to go through each and every bit of it. The British National Corpus, for example, has a total of 100,000,000 running words. If it is edited as a book, and each page has 400 words, it will be a book of 250,000 pages. If one individual can read 100 pages per day, then, it will take about 2,500 days to finish this book. It is, therefore, quite impossible for the scale of this research to adopt this method.

Secondly, an alternative way of conducting this research would be to compare the words used in a learner corpus to those in corpora with language samples collected from native speakers. However, this approach has two limitations. First, learner corpora mostly collect written samples from exam or essay scripts of second language learners. For example, the Cambridge Learner Corpus collects exam papers

written by students taking ESOL English exams around the world. Collecting samples in this way would not reflect the natural use of language, and, at the same time, would produce samples with mostly formal words. Second, the words used by learners are unlikely to be all the words that they know. Learners would preferentially use the words which they have productive knowledge of, however, generally speaking, the size of the productive vocabulary is usually smaller than those of the receptive vocabulary. Not using a word in an essay does not mean that the word is not known to the learner. As such, there is no way of knowing the exact size of their vocabulary repertoire by studying an essay they have written.

3.3.2 Dictionaries used in the study

After careful and systematic examination of a number of dictionaries, the following three dictionaries were chosen for developing the initial pool of words: Collins COBUILD advanced learner's dictionary (2009), Collins COBUILD advanced dictionary of American English (2008), and Oxford advanced learner's dictionary, 8th edition (2010). These three dictionaries come from three large corpora: the Collins corpus, the Bank of English, and the Oxford Corpus. These three dictionaries cover up-to-date usage of English with a large number of entries, and are tailor-made for advanced learners of English.

Collins COBUILD Advanced Dictionary (2009) is the 6th edition of the Collins advanced dictionary for learners of English. In its 21 year history of building learner dictionaries, Collins has established itself as a reliable learning source for advanced learners of English. The Collins COBUILD advanced dictionary of American English (2008) has over 75,000 entries. These two dictionaries were chosen to balance the British English and American English sources. These two dictionaries come from the Collins corpus and the Bank of English corpus respectively. Collins corpus has in total 2.5 billion running words, while, the bank of English has 650 million running words.

The Oxford advanced learner's dictionary, 8th edition (2010) also has an enormous amount of entries, with 185,000 entries of words, phrases and meanings. The compilation of this dictionary is based on the Oxford corpus. With a total of 2 billion

running words, this corpus is one of the largest of its kind. It collects written materials from many different written and oral sources and diverse geographic locations. With it enormous size and careful selection of word samples, I believe this corpus represents the full picture of real language use at present.

3.4 Defining known words

Another important issue is how to define whether or not a word is known. Knowing a word can mean a lot of different things. Research into depth of vocabulary knowledge has pointed out different aspects of knowledge of a word. Researchers like Richard (1976) and Nation (2001), suggest that there are different aspects of knowledge involved in knowing a word.

In his influential book on teaching and learning vocabulary, Nation (2001) pointed out that there are three aspects of knowing a word: form, meaning, and use, each with three sub-aspects. Each of the nine sub-aspects includes both receptive and productive knowledge of words.

Table 3.4 The nine aspects of knowing a word (Nation, 2001:27)

	Spoken form
Form	Written form
	Word parts
	Form and meaning
Meaning	Concept and referents
	associations
	Grammatical function
Use	Register
	collocations

The definitions of receptive and productive knowledge are mixed, with researchers adopting different definitions in their studies (e.g., Laufer, 1998; Meara, 1990; Wesche, Paribahkt & Paribahkt, 2009). Generally speaking, receptive knowledge involves the comprehension of a word through input, like reading and listening. Productive knowledge involves the use of word through speaking and writing (Nation, 2001). It is now agreed that, rather than being a dichotomy, the concept of receptive and productive knowledge should be seen as a continuum, instead of being completely independent of each other.

In this study, I will use the receptive knowledge of form and meaning as the criterion of defining knowing a word. In other words, if one meaning of a word is known to the participant, then the word is classified as known. Therefore, while searching for potential words in this study, those words with at least one known meaning would not be included. Likewise, when asking participants to go through the list and identify the words that they already know, the same definition would also applied.

One potential problem with this definition is that many words are polysemous. However, cognitive researchers have provided evidence to show that the acquisition of various meanings of polysemous words could be facilitated by introducing the link between the core/conceptual sense and different meanings (e.g., Tyler & Evans, 2004; Verspoor & Lowie, 2003). Verspoor and Lowie (2003) conducted an experiment to test whether providing the core sense of a polysemous word could facilitate guessing of unknown meanings, and short- and long-term memory of the target word. They chose words with at least three senses, with strict control on the link between the three senses. The three senses of a word should have a core sense, a figurative sense, and a more figurative or abstract sense. The findings supported the hypothesis that providing a core sense facilitates the guessing of other senses and long-term memory. Although this experiment was carried out in controlled experimental conditions, it lends support to the facilitating effect of acquisition of meanings of polysemous words.

3.5 Criteria for selecting target words

3.5.1 Criteria for selecting potential words from dictionaries

The first stage of collecting data is reading through the three dictionaries page by page to look for unknown words. This work was done by the author alone, however, three other participants, at a similar level of study, were also asked to go through samples of the dictionaries to validate the results (for information on the three participants, refer to section 3.6.1). The author is currently enrolled in a Master's program, majoring in Applied Linguistics. Also, she has seventeen years of experience of learning English prior to coming to New Zealand. Additionally, she had stayed in Australia for 1.5 years studying for her first Master's degree, majoring in Commerce from the year of 2004 to 2005. At the time of the study, she has stayed in New Zealand for six months, giving a total of two years at the start of the study. All the words that were unknown to the author and the participants were selected and recorded for further analysis.

All proper nouns were excluded from the initial word list. This was because the understanding of proper nouns requires context-specific knowledge, and therefore, may not be useful for all advanced learners living in English contexts. For example, names of places, like *Wellington*, may be used frequently by residents of New Zealand, but it would be rarely used by people in other geographic locations. On the other hand, proper nouns do not pose much of a learning burden for learners, because the concepts could be easily understood through interlocutor cues in conversations.

The initial word list included 7192 word families.

3.5.2 Criteria for narrowing down the initial word list from dictionaries

The initial word list was further analysed through a range of criteria to ensure that it was representative of daily life words in various settings. These criteria were:

Specialty of occurrence. All the words that appears in the GSL, BNC 2000 and AWL lists were excluded. The reasons for using this criterion are two-fold. Firstly, advanced learners' vocabulary size should be beyond the most high frequency words. Secondly, daily life words should not be frequently used in the academic settings.

Stylistic level. The word list is for the purpose of daily life use; therefore, words that are only used for formal settings and literary purposes were excluded. For example, for similar meaning, we might use the word *hate* in common settings; while in formal settings, the word *abhor* might be used. Comparatively, *hate* would be more useful for learners than *abhor.* Likewise, *in the air* might be more useful than *aloft,* which might appear in more formal contexts, poems and literature.

Time constraint. Learners study the word list to cope with everyday contact in English. Therefore, all the words in the list should be currently used, with out-dated words excluded. For example, the French origin word *boudoir* was used in the past to refer to a women's bedroom, but it rarely used now. As such, its usefulness to learners is very low.

Usefulness. Potential words should be those that do not require specific knowledge restricted to certain areas of interests. Therefore, those technical words known and used only by people specialised in certain fields were excluded. Several fields were classified as technical in the selection: agriculture, anatomy, biology, botany, business, chemistry, computer science, journalism, law, mathematics, medical science, military, music, politics and zoology.

Dialects. Potential words should be understood by native speakers from all backgrounds, and therefore, words used only in certain parts of the world were excluded. For example, words which are only used in Britain, America, Canada, Australia, New Zealand, South Africa or India are excluded. For example, the words like *concierge, kibbutz,* and *samurai,* which are mostly used only in some regions, or refer to things in a specific culture, were not included in the list.

Screening the initial word list with these criteria left the word list with 3043 words.

3.6 Cross-checking the word list

There are three steps of cross-checking the word list in this study, with the final version of the Advanced Practical Word List developed out of step 3. There are two purposes of cross-checking the word list: reducing it to a manageable size, and validating it among both native and non-native speakers of English.

3.6.1 Cross-checking the initial word list

To ensure that the selection of unknown words was valid among learners of the same level, three other advanced learners were invited to go through randomly chosen samples of the dictionaries in the same manner. The language learning backgrounds of the three learners is shown in the following table. Each learner was asked to read around 10% of the Collins COBUILD Advanced Learners Dictionary (2009). This was done by selecting the first 10 pages of every 100 pages of the dictionary, for example, page 1 to 10, then, page 101-111, etc. Each learner then read through a copy of the 180 pages of the dictionary selected to look for unknown words. They were asked to underline any words that were unknown to them, including proper nouns and acronyms. This was done so as not to burden the participants with too many details of the research, and allow them to focus only on whether they have any prior knowledge of the words in the dictionary. The results were then compared with what the author had underlined. Both the number and type of the words unknown to the participants were similar to those of the author. The differences were discussed and clarified with each participant, and all of them were added to the initial word list. As such, the initial word list has 7,192 word families (for discussion on the initial word list, refer to section 3.5.1).

Table 3.6.1 Demographic information of participants

Participants	Gender	First Language	Years of learning English	Length of staying in NZ	Level of study
1	Female	Vietnamese	15	11 months	Currently studying Masters
2	Female	Japanese	18	11 months	Currently studying Masters
3	Male	Chinese	18	2 years	Currently studying PhD

3.6.2 Cross-checking the screened word list

The purpose of this step was to cut the word list to a manageable size by using the judgment of the native speakers to determine which words are useful in daily life settings. The screened word list was classified into 28 different categories for two purposes. The first was that many of the words in the list are polysemous, categorising the words into different sections made it easier for participants to identify the meaning referred to in the list. For example, the word *kosher* has two meanings when used as an adjective, referring to different concepts. Categorising the word under the heading *food, drink and cooking* made it explicit that the word is referring to the preparation of food in this case. The second purpose is that categorising words into different sections would make it easier for learners to learn the words and apply them to specific contexts.

This classification was based on Miller and Fellbaum's (1991) framework. In their report on the psychological framework of semantic network of English words, they

point out that it is possible to use word parts to map out the hierarchy, and relationships between a large numbers of English words. Further, they suggested 26 unique beginners for English nouns, and 14 for verbs. It is argued that all the nouns and verbs could start their hierarchies and build up synonyms starting with these labels. The 26 beginners for nouns and 14 for verbs are:

Table 3.6.2.1 Beginners for nouns and verbs (Miller & Fellbaum, 1991: 204-5)

28 beginners for nouns (Miller& Fellbaum, 1991: 215)

act, action, activity	animal, fauna
artifact	attribute, property
body, corpus	cognition, ideation
natural object	natural phenomenon
person, human being	plant, flora
possession, property	process
communication	event, happening
feeling, emotion	food
group, collection	location, place
motive	quantity, amount
relation	shape
society	state, condition
substance	time

14 beginners for verbs (Miller& Fellbaum, 1991: 215)

bodily care	contact
bodily functions	creation
change	motion
cognition	perception
communication	possession
competition	social interaction
consumption	weather

These 40 beginners were revised accordingly with the 3043 word families in the screened word list. The 40 beginners were merged into 28 categories, including nouns, verbs and adjectives. These 28 beginners are:

Table 3.6.2.2 28 beginners of the screened word list

Act, action	Natural object
Animal	Attribute
Natural phenomenon	Artifact, material
Person, profession	Plant
Body parts, dress and accessories	Cognition, mind and thoughts
Process	Communication
Quantity	Event, happening
Feeling, emotion, attitudes	Shape, building
Group, collection	State, condition
Food, drink and cooking	Society, relation
Location, place	Substance, tools
Motive	Time
Bodily experience	Movement, behavior
Religion	Illness, treatment

A questionnaire was made with the 3043 word families in 28 categories, and seven native speakers with diverse backgrounds were invited to go through the word list and underline the words that they thought were useful for daily life communication. The demographic information of these 7 native speakers is as follows:

Table 3.6.2.3 Demographic information of the 7 native speakers

	Gender	Nationality	Age	Education	Profession
1	Female	United States	24	Bachelor	EFL teacher
2	Female	New Zealand	27	Master(studying)	High school teacher
3	Male	New Zealand	29	High school	supermarket worker
4	Male	New Zealand	45	High school	supermarket worker
5	Female	Canada	23	Master(studying)	ESL teacher
6	Female	New Zealand	36	Bachelor	Housewife
7	Female	New Zealand	42	Bachelor	Housewife

The questionnaires were analyzed one by one, and only the words that were chosen by five or more participants were retained. This left 891 words on the list in 26 categories. The two categories, process and time, were removed from the list because no words in these two categories were chosen as useful in daily life scenarios. These 891 words were used in a second questionnaire given to 100 participants for their judgment on the prior knowledge of the words.

3.6.3 Cross-checking with both native and non-native speakers

As the purpose of this study was to identify words that are known to most native speakers, but unknown to non-native speakers, the word list was given to both groups to see if they had prior knowledge of the selected words.

In this step, the word list of 891 words was made into a questionnaire, and given to 60 non-native speakers and 40 native speakers. All 40 native speakers were tertiary students at a New Zealand university, aged from 22 to 30. For non-native speakers, the English proficiency level and length of time staying in English-speaking

countries were considered when looking for participants. Only those who were currently enrolled in Master's or PhD programs in Australia, New Zealand, or UK universities, and had spent at least 12 months in the country were invited. A total of 60 non-native participants, aged from 23 to 39, were selected, and they came from various L1 backgrounds. Fifty participants come from mainland China, four from Vietnam, two from Japan, one from Tonga, one from Pakistan and two from India. The questionnaires collected from the Indian participants were not included in the data analysis, because English is one of the official languages in India, and therefore, Indian learners could not be classified as strictly EFL learners of English. Overall, the study collected questionnaires from 58 non-native speakers. Of them, two were studying at UK universities, twelve at Australian universities and the remaining 44 at New Zealand universities. The participants had from eight to 28 years' experience of learning English before they came to an English speaking country, with an average of twelve years. Before the study, they had spent from one to nine years in New Zealand. Twelve of the participants were enrolled in PhD programs, the rest were in Master's programs.

There are two reasons for inviting only participants fitting these two criteria. The first one is determining which learners are advanced learners of English. Until now, there has been no definite way of determining such learners. Most studies use the number of years spent learning English as the criterion to define advanced level of English proficiency. In a study into lexical richness in advanced learner's written work, Laufer (1991) asked first year English major student in an Israeli university to participate in the study and treated them as advanced learners of English, based on their years spent in learning English. In another study into the lexical units, Arnaud and Savignon (1997) studied three groups of participants, all of whom were defined as advanced learners of English. The first group was first year English major university students, with French as their first language and seven years' experience learning English. The second group was made up of third year university students doing the same course, while the third group included trainee English teachers. Following similar criteria, this study invited participants who are studying Master's or PhD

programs in universities in English speaking countries with diverse L1 backgrounds. They have around 10 years' experience of learning English before enrolling in the program and they have received English-only input in their current study. Therefore, they could be regarded as advanced learners of English.

The second reason relates to the length of time spent living in the English speaking countries. The purpose of this study is to look for words that are used in daily life scenarios. As such, participants need to have some exposure to English-speaking environments. Therefore, the study invited only L2 participants with at least 12 months experience living in the English-speaking countries, to make sure that they had had enough opportunities to use the language authentically.

Paper versions of the questionnaires were given to New Zealand participants, while those in Australia and the United Kingdom received electronic copies. They were asked to complete the questionnaire in their own time, and return it to the author within three days. On average, the questionnaire took 30 minutes to finish. Non-native participants were asked to underline every word which they knew the meaning of, while native participants were asked to underline the words they did not know the meaning of. The questionnaires were collected and analysed for the final version of the Advance Practical Word List.

3.7 The advanced practical word list

The questionnaires collected from native and non-native speaking participants were first analysed separately, and then, compared to see which group of words were known by most native speakers, yet unknown by most non-native speakers.

The 40 questionnaires collected from native speakers showed that all the words listed were known to every participant. Therefore, only words also known to non-native participants were removed from the list.

The questionnaires collected from non-native speakers showed that participants only knew a small amount of listed words. On average, the non-native participants only knew about 15% of words in the questionnaire. There is great variance between

participants' knowledge, with one participant understanding only fourteen words, or 1.5% of the list. In contrast, the participant with the most understanding of the listed words knew 438 words (49%). However, almost 50% of the participants knew less than 10% of the words on the list, that is, less than 88 words out of the 891 words, with a median understanding of only 88 words out of a total of 891.

Table3.7 The percentage of known words for non-native speakers

	Percentage of words known	number of participants
1	1%-10%	26
2	11%-20%	18
3	21%-30%	4
4	31%-40%	8
5	41%-50%	2

Words that were known by non-native speakers were recorded, and those known to more than half of the participants were excluded from the list. 50% was chosen as the benchmark, because the purpose of this study is to identify words that were unknown to most non-native speakers. If a word is known to 50% of the participants, it is likely known by a large number of non-native speakers from various backgrounds, As such, these words would not be necessary in the word list. The result showed that there were only a few words known to most of the participants. In total, only 23 words were removed from the list at this final step. This gave the Advanced Practical Word List a total of 867 word families.

4   Description of the Advanced Practical Word List (APWL)

This part describes the frequency and word class makeup of the Advanced Practical Word List. It then discusses the frequency and type of words in the two sets of three sublists of the APWL.

4.1 Frequencies in the British National Corpus

The APWL has 867 word families, with words coming from a range of different subject areas in daily life settings. The frequency of the word families in the APWL is very interesting. It does not only have words that beyond the current vocabulary size of the advanced learners, but also words that are of moderate frequency, that were expected to be known by advanced learners of English.

The computer program VocabProfile was used to analyse the frequency of word families in the APWL. This program was first developed by Laufer and Nation (1999), and used to analyse the vocabulary makeup of written work of advanced learners of English. It is now available from the website www.lextutor.ca developed by Tom Cobbs. This program was chosen because it could analyse the words, using twenty levels each with 1000 words from the British National Corpus. The development of the BNC frequency list was discussed in section 2.1.2. Although the BNC has its limitations, the validation of the frequency lists with nine other corpus (the LOB, FLOB, Brown, Frown, Kohlapur, Australian, Wellington written, Wellington spoken and LUND corpora) showed that the words are properly ordered in the lists (Nation, 2004).

A total of 827 out of 867 word families appeared in the BNC, with words present in frequency levels three to twenty. Forty word families were not present in the BNC 20, making up 4.61% of APWL. It can be seen from table 4.1.1 that over 50% of the word families appear in the 3rd to the 9th frequency level. The 5th to 11th frequency levels each comprise 5% to 11% of the APWL, with each level taking up $5^{\%}$ to 11% of the whole list. The number of words drops sharply from the 12th level, from over 60

words each level to around 30 words. The most infrequent three levels, the 18th to the 20th, make up merely 3% of the APWL.

Table 4.1.1 Frequency of APWL using BNC-20

Freq. Level	Families	Types	Tokens	Coverage (tokens)%	Cum%
K1 Words :				0.00	0.00%
K2 Words :				0.00	0.00%
K3 Words :	23	23	23	2.65	2.65%
K4 Words :	40	40	40	4.61	7.26%
K5 Words :	63	63	63	7.27	14.53%
K6 Words :	81	81	81	9.34	23.87%
K7 Words :	81	81	81	9.34	33.21%
K8 Words :	99	99	99	11.42	44.63%
K9 Words :	80	80	80	9.23	53.86%
K10 Words :	77	77	77	8.88	62.74%
K11 Words :	67	67	67	7.73	70.47%
K12 Words :	33	33	33	3.81	74.28%
K13 Words :	36	36	36	4.15	78.43%
K14 Words :	23	23	23	2.65	81.08%
K15 Words :	38	38	38	4.38	85.46%
K16 Words :	33	33	33	3.81	89.27%
K17 Words :	30	30	30	3.46	92.73%
K18 Words :	9	9	9	1.04	93.77%
K19 Words :	9	9	9	1.04	94.81%
K20 Words :	5	5	5	0.58	95.39%
Off-List:	?	40	40	4.61	100.00%
Total	827+?	867	867	100%	100%

There are two points worth noting about the frequency level of the APWL as a whole. The first is the words that do not appear in the BNC, the second is the words that are within the vocabulary size of advanced learners at this level of proficiency.

Forty word families in the APWL were not found in the BNC-20 frequency lists. A close look into these words provides an explanation for this (see table 4.1.2). The main reason these words do not appear in the BNC is due to the sources used to compile this corpus. 90% of the language samples in the BNC are formal written materials. Words that are informal in nature are less likely to be found in the BNC-20. For example, the terms *ballsy, dud* and *killjoy* are, in most cases, used in informal interactions, and are not listed in the 20 lists. Alternatively, words like *baguette, éclair* and *pavlova,* all food originating in places other than the UK, were not present, because BNC mainly collected language samples in the UK.

Table 4.1.2 Off-list words

Off-list words

ambidextrous	dexterous	headwind	liposuction	salivate
baguette	ditzy	idyllic	medicore	sandpit
ballsy	dopey	jaywalking	megapixel	starstruck
beanie	dud	jumpstart	pantyhose	sty
blabbermouth	earlobe	killjoy	pavlova	
blindfold	earmuffs	kook	pebble	
breathalyze	éclair	lard	pepperoni	
broth	gumdrop	licorice	pinprick	
cibatta	headrest	lingo	rake	

There is lack of research into the vocabulary size of advanced L2 learners. Their vocabulary size can only be estimated from various studies on the vocabulary gain from research into incidental and intentional learning. Cobb and Horst (2001) found that L2 learners gained 140-180 words over two months using a concordance-based

study program. As such, L2 learners would learn around 800 words per year, or 8,000 words a decade. Nation (2004) estimated that advanced L2 learners need a vocabulary of around 8,000 to 9,000 words to complete postgraduate level study programs in an English-speaking country. As the advanced learners in this study were postgraduate students with at least 10 years' experience learning English, it is interesting in that almost 50% of the APWL words (i.e. words these learners did not know) are from the 9000 most frequency word families in the BNC. These words should have been known to these advanced learners at this level of study. This could be the result of the type of input received by L2 learners in the EFL learning backgrounds.

4.2 Sublists of the Advanced Practical Word List

The reason for making sublists is to ease the burden of learners. Advanced learners in English-speaking countries face pressures, such as from studying, working, and socializing. Singling out words that are most useful to them would allow them to prioritise their learning. Additionally, it would also help them to schedule their pace of learning, and make the learning goal more feasible in a limited amount of time. The 867 word families in the APWL are divided into two sets of sublists. The first is based on the frequency level of the words, and the second on perceived usefulness.

4.2.1 Frequency sublists

The frequency sublists were divided based on the words frequency as shown in the BNC-20. The rationale for dividing the sublists in this manner is the relationship between lexical coverage and comprehension of input. The lexical coverage needed for comprehension varies with the type of input. The APWL contains words that are used for social interaction needs, and therefore, the major discourse type of input should be spoken. Past research suggests that 98% of lexical coverage is ideal for the comprehension of both written and spoken text (Hirst & Nation, 1992; Hu &Nation, 2000; Nation, 2006). In a recent study into the lexical coverage and comprehension

of movies, Webb and Rodgers (2009b) show that a vocabulary size of at least 5,000 word families is needed, while the most difficult genre of movie require a vocabulary size of 10,000 word families. Using similar coverage, McCarthy (2001) studied the vocabulary needed to be included in the learning program for advanced learners, and suggested that 6,000 words would be a good start. As such, four sublists were delineated using the frequency levels of the BNC-20. The first sublist included words from levels 3-5(the 3,000-5,000 most frequent words), the second covered words from levels 6-10, the third levels 11-15, and the fourth comprised levels 16-20. For words that are from 10,000 to 20,000 levels, the cut-off point is the 15,000 level, because it is the middle point of the frequency range and the words in these ten levels are evenly distributed around it. All the words not present in the BNC 20 were incorporated into sublist four.

Sublist one has 126 word families, comprising 14.53% of the APWL. The frequency range of this sublist is from the 3rd to the 5th level in BNC 20. This frequency range could be considered as medium frequency instead of being strictly low frequency. Sublist two has 418 word families (48.21% of the APWL), Sublist three has 197 word families (22.7%), and sublist four has 126 word families (14.53%).

4.2.2 Perceived usefulness sublists

Three sublists based on the native speakers' judgments on its usefulness in the daily life scenarios were also devised. The three sublists are of descending importance, and are made up of different numbers of words. Sublist one contains words judged useful in daily life by all seven native speakers, sublist two has words deemed useful by six speakers, and sublist three has words selected by just five speakers. Sublist one has 108 words, sublist two has 336 words, and sublist three has 424 words.

4.2.2.1 Sublists one

Sublist one contains 108 words, comprising 12% of the APWL. These words have

been identified by native speaker participants as the most useful for daily life communications. These words cover 21 of the 28 categories, with words in the categories of action, food and illness, making up 42%. These words are relatively frequently used, with over 60% appearing in the BNC-20 frequency levels three to nine, as shown in Table 6.2.2.1.

Table 4.2.2.1 Frequency of words in sublist one

Freq. Level	Families	Types	Tokens	Coverage (tokens)%	Cum%
K1 Words :				0.00	0.00%
K2 Words :				0.00	0.00%
K3 Words :	2	2	2	1.85	1.85%
K4 Words :	2	2	2	1.85	3.70%
K5 Words :	10	10	10	9.26	12.96%
K6 Words :	16	16	16	14.81	27.77%
K7 Words :	16	16	16	14.81	42.58%
K8 Words :	9	9	9	8.33	50.91%
K9 Words :	13	13	13	12.04	62.95%
K10 Words :	8	8	8	7.41	70.36%
K11 Words :	5	5	5	4.63	74.99%
K12 Words :	5	5	5	4.63	79.62%
K13 Words :	4	4	4	3.70	83.32%
K14 Words :	2	2	2	1.85	85.17%
K15 Words :	2	2	2	1.85	87.02%
K16 Words :	3	3	3	2.78	89.80%
K17 Words :	5	5	5	4.63	94.43%
K18 Words :	1	1	1	0.93	95.36%
K19 Words :	1	1	1	0.93	96.29%
K20 Words :	1	1	1	0.93	97.22%
Off-List:	?	3	3	2.78	100.00%
Total	105+?	108	108	100%	100%

In sublist one, we could find many words that are considered frequent and common in daily life scenarios. The lack of knowledge of these common words could be due to the living style and cultural discrepancies between native and non-native speakers. For example, the words in the "food" category make up more than 10% of the words in sublist one. Words like *scone, crust* and *marmalade* are terms describing part of Western diets. While these words are from the 5th thousand most frequently used words in the BNC-20, and are within the vocabulary size of advanced learners. However, due to the differences in living styles, these words are unknown to most of them. Learners from different living style backgrounds have to learn the concepts of these types of food before acquiring the word labels for them.

Another example is words in the "religion" category. Words related to religious belief, like *gospel* and *hymn,* are known to many native speakers from very early on in life, but rarely known by non-native speakers. These two words are from the 3rd and 5th frequency levels of the BNC-20. Only 14 of the non-native speaking participants knew the word *gospel*, and a mere eight knew the word *hymn.* The lack of knowledge of these highly frequent religious words may be explained by the backgrounds of the non-native participants of the study. Most of the participants were from China, a country where religious belief has been discouraged by the government publicly in the recent past. It is possible that the L2 learners from this background would not have the world knowledge of such concept, and would therefore not know words of this kind.

4.2.2.2 Sublist two

Sublist two has 336 word families, and comprises 39% of the APWL. This sublist covers words from 26 of the 28 categories, with the most words coming from the "attribute" category, followed by "action" and "food". These three categories make up 47% of words in the sublist. Words from the 3rd to the 9th frequency level account for a little over 50% of all the words in the sublist, as shown in the table 4.2.2.2.

Table 4.2.2.2 Frequency profile of sublist two

Freq. Level	Families	Types	Tokens	Coverage (tokens)%	Cum%
K1 Words :				0.00	0.00%
K2 Words :				0.00	0.00%
K3 Words :	11	11	11	3.27	3.27%
K4 Words :	19	19	19	5.65	8.92%
K5 Words :	24	24	24	7.14	16.06%
K6 Words :	34	34	34	10.12	26.18%
K7 Words :	31	31	31	9.23	35.41%
K8 Words :	33	33	33	9.82	45.23%
K9 Words :	29	29	29	8.63	53.86%
K10 Words :	26	26	26	7.74	61.60%
K11 Words :	27	27	27	8.04	69.64%
K12 Words :	16	16	16	4.76	74.40%
K13 Words :	11	11	11	3.27	77.67%
K14 Words :	11	11	11	3.27	80.94%
K15 Words :	16	16	16	4.76	85.70%
K16 Words :	12	12	12	3.57	89.27%
K17 Words :	10	10	10	2.98	92.25%
K18 Words :	2	2	2	0.60	92.85%
K19 Words :	3	3	3	0.89	93.74%
K20 Words :				0.00	93.74%
Off-List:	?	21	21	6.25	100.00%
Total	315+?	336	336	100%	100%

4.2.2.3 Sublist three

Sublist three contains 424 words and makes up almost 50% of the APWL. These

words cover all the 26 categories, with the most words in the "attribute" category, followed by "action" and "movement". These three categories make up almost 40% of the words in this sublist, as shown in the table 4.2.2.3.

Table 4.2.2.3 Frequency profile of sublist three

Freq. Level	Families	Types	Tokens	Coverage (tokens)%	Cum%
K1 Words :				0.00	0.00%
K2 Words :				0.00	0.00%
K3 Words :	9	9	9	2.12	2.12%
K4 Words :	19	19	19	4.48	6.60%
K5 Words :	31	31	31	7.31	13.91%
K6 Words :	30	30	30	7.08	20.99%
K7 Words :	35	35	35	8.25	29.24%
K8 Words :	57	57	57	13.44	42.68%
K9 Words :	37	37	37	8.73	51.41%
K10 Words :	45	45	45	10.61	62.02%
K11 Words :	35	35	35	8.25	70.27%
K12 Words :	12	12	12	2.83	73.10%
K13 Words :	20	20	20	4.72	77.82%
K14 Words :	10	10	10	2.36	80.18%
K15 Words :	19	19	19	4.48	84.66%
K16 Words :	18	18	18	4.25	88.91%
K17 Words :	15	15	15	3.54	92.45%
K18 Words :	6	6	6	1.42	93.87%
K19 Words :	5	5	5	1.18	95.05%
K20 Words :	4	4	4	0.94	95.99%
Off-List:	?	17	17	4.01	100.00%
Total	407+?	424	424	100%	100%

4.3 Categories in the Advanced Practical Word List

The 867 words in the APWL were organized into 25 categories. These 25 categories and the number of words in each category are as follow:

Table 4.3 25 categories of the APWL

Act, action	117	Natural object	18
Animal	30	Attribute	138
Natural phenomenon	5	Artifact, material	30
Person, profession	41	Plant	7
Body parts, dress and accessories	43	Cognition, mind and thoughts	14
Motive	12	Communication	10
Quantity	12	Event, happening	9
Feeling, emotion, attitudes	33	State, condition	19
Group, collection	4	Society, study	12
Food, drink and cooking	81	Substance, tools	51
Location, place	16	Movement, behavior	56
Bodily experience	30	Illness, treatment	48
Religion	18		

The category with the most words is the "attribute", which includes adjectives to describe both humans and the world around them, for example, *fastidious, grumpy* and *exquisite.* This is followed by "action" and "food". These three categories combined account for 39% of words on the APWL. The category with the least number of words is "shape and building", which contains words related to the architectural features of buildings and houses. There is only one word in this category that has been identified as useful by native speakers. The only word in this category,

pantry, is reorganized into the category of "food and drink". There are two other small categories, "group and collection" and "natural phenomenon", both of which have less than 5 words each.

The 25 categories can be grouped together into two main assemblages, one closely related to human experiences and the other related to the world around us. The first assemblage includes categories that cover the human mind, body, and food, which are central to living. This includes the categories: "person", "body parts", "cognition", "feeling", "food", "motive", "bodily experience", "movement", "religion and Illness". These ten categories contain 374 words from APWL, or 43%. This assemblage includes words which are used to describe human feelings, like *ecstasy, nausea* and *bliss*. It also contains words that are used to describe the way a person moves, like *saunter* and *dodge*. The second assemblage comprises sixteen categories, with 493 words, which are used to describe the world around us and the interaction between the humans and other members in our and world.

4.4 Part of speech of words in APWL

The 867 word families in the APWL belong to four classes, all of which are open classes. None of the words in the APWL are function words, because most of such words are of high frequency in the word list, and have been learnt at the earlier stage of learning the language. At the same time, function words of lower frequency, like *nonetheless, whereas* and *notwithstanding,* have been included in the Academic Word List (Coxhead, 1998) for the purpose of academic learning.

The classification of part of speech in this study is based on the first word class in each word's dictionary. The Collins COBUILD Advanced Dictionary (2009) was used primarily to outline the part of speech of words. If words could not be found in this dictionary, then, the part of speech is based on the definitions in the Oxford Advanced Learner's Dictionary (2010). The majority of the words, a total of 510, or 59%, in the APWL are nouns. The second largest word class is verbs, with 202 words, or 23%. This is followed by adjectives, with 160 words, or 18%. There is only one

adverb in the whole list, *pronto*. The figures above are based on the first part of speech in the definitions of the words in the three dictionaries.

One important issue when classifying words in APWL into parts of speech is that words can have more than one part of speech. A total of 126 words in the APWL could be used as both a noun and a verb, or sometimes, noun and adjective. Of the nouns, 68 had additional uses: 55 as verbs, 11 as adjectives, and two as both verbs and adjectives. Likewise, 45 words in the verb category had additional uses: 43 as nouns and two as adjective. Finally, thirteen words in the adjective category also had additional uses: seven as nouns and six as verbs.

4.5 Summary

The Advanced Practical Word List (APWL) comprises 867 words, compiled to help advanced learners in an English-speaking environment cope with unknown words that appear in daily life scenarios. The words in the list come from a wide range of frequency levels, including some that should have been known to advanced learners already, considering their vocabulary size, and words that are beyond their vocabulary repertoire. All the words in the APWL were orgnaised into 25 different categories to help learners determine the words necessary for various demands. For example, words describing human feelings and attitudes were grouped under "feelings, emotions and attitudes". Learners who are not confident in their knowledge of words in this area can refer to this section of the list. The majority of words in the APWL are nouns, verbs and adjectives, with 15% having multiple usages.

5 Potential uses of the Advanced Practical Word List and implication for further research

This chapter introduces the potential uses of the APWL for teachers, learners, test designers and dictionary compilers. It also includes a number of suggestions for possible future research.

5.1 Potential uses of the APWL

The Advanced Practical Word List is primarily meant for practical use by teachers and learners, especially advanced learners of English, who would otherwise have to learn these words in real use situations, which can be frustrating and embarrassing. The APWL could be incorporated into language courses that are designed to be taken by advanced learners prior to enrolling in university level courses. For this group of learners, not only the academic needs are important for them, but also the social needs. Their need to interact with the world around them should not be underestimated. Learning the words necessary for social interaction in daily life scenarios will help learners venture out of their small social circles of other fellow L2 learners, and step into the authentic English-speaking environment. At the same time, the list could also be useful for test developers, who design tests for immigration purposes, and for other advanced users of English. Finally, the list may also be of use for dictionary compilation purposes, to encourage autonomous learning on the part of the learners.

5.2 Direct learning from the APWL

Direct learning is encouraged for learning of the APWL for a number of reasons. Research into incidental learning of L2 words through reading have found that learning new words requires a number of repetitions (for a review on the research, see Webb, 2007). However, L2 learners may not have enough exposure to the

language to make incidental learning as effective and efficient as it is for native speakers. Further, past research shows that the new words gained from reading is quite limited (Waring & Takaki, 2003). These conservative results lead some researchers to doubt the encouragement of incidental learning as the primary source of learning new words for L2 learners (Laufer, 2001, 2003).

In direct learning, the primary focus is on learning lexical items; therefore, various techniques could be employed to increase the word gain. Learners could employ strategies, such as vocabulary notebooks, to increase the depth of learning. Also, multiple repetitions could be achieved by various tasks designed for learning the words. Nation (2007) suggests a four-strand framework in planning courses for vocabulary learning. The four strands are meaningful input, meaningful output, form-focused activities and fluency development. He suggests that these strands should be balanced in the course to maximise the opportunities for learning new lexical items. Learning words in the APWL would be mainly through listening and speaking, as the list caters primarily for oral communication. The following sections propose some suggestions for the course design and autonomous learning of the lexical items in the APWL.

5.2.1 Meaningful input through listening

This strand mainly aims to help learners to notice, be aware of, and understand new lexical items through meaningful input. The input should be enjoyable, with the primary focus of the learners directed towards understanding and gaining knowledge (Nation, 2007). It is important that the inputs be from the topics that the learners have background knowledge of, and are interested in learning more about. Words in the APWL are mostly come from spoken sources, so, in this strand, learners could receive input through watching movies, television programs, or purpose-made videos reflecting different daily life scenarios.

One important aspect of meaningful input is the coverage of known words in the input. The current vocabulary size of the advanced learners is estimated to be around 8,000 to 9,000 words (Nation, 2006). Recent studies of vocabulary coverage of British

and American television programs show that, a vocabulary size of 9,000 words would be large enough to reach coverage of 98% of words in programs of all genres and episodes (Webb & Rodgers, 2009a). A further study showed that a vocabulary size of 9,000 words would also be sufficient to reach 98% coverage of movies of all genres except animated (Webb & Rodgers, 2009b). At the coverage of 98%, learners would be able to understand the content of the input with relative ease. As such, movies and television programs would be a valuable source of input for advanced learners.

Another more strictly controlled source of input would be purpose-made videos to help learners familiarise themselves with life in English-speaking environment. Different episodes of these videos could be made for different real life situations. For example, a video filmed in supermarket could be used to familiarise learners with the scenario of shopping, and words related to food, drink and other household related items. Words in the categories of "food", "substance" could be depicted in such a video. In the same way, there could be a video filmed in a clinic, showing conversations between a doctor and patients. Specific discourse, such as a patient describing symptoms, or doctor diagnosing an illness could be shown. Words in the categories of "illness and treatment", and "feeling, emotion and attitudes" could be repetitively used in this video.

5.2.2 Meaningful output through speaking

The major principle of meaningful output is that learners consciously use the words that they are intended to learn. This may be achieved through the design of the task in a classroom setting or by steering real-time conversations in daily lives. Again, the topics should be familiar to the learners. Words in the APWL could be used in a wide range of topics that occur frequently in daily life, for example exchanging ideas and opinions on music, movies, and television programs, or describing a past experience which might involve vividly picturing the movements and actions. The task or conversation should occur in a non-threatening atmosphere, where the main purpose is to convey the messages. This will allow learners to feel comfortable enough to use the target words without feeling threatened or embarrassed.

It is important that there is a large amount of output, so that the learners can consciously use the words, and see if they have used them in the appropriate register. Learners should be encouraged to use the words in the APWL whenever possible, and to test their hypotheses on how the words should be used. If the words are not used properly, learners could explain the idea through negotiation of meaning or clarification. Productive use of the words would likely leave deeper marks on memory than receptive use, because more mental activities are involved (Craik & Lockhart, 1975).

Nation(2007) stated that the boundary between meaningful input and output is hard to define in that one learner's input could be another's output, and vice versa. This means that, while performing the task of peer talk or group talk in class, learners could all benefit from the actively use of the words, or passively hearing the words.

5.2.3 Language-focused learning

Learners could benefit from language-focused activities when their primary focus is on learning the form of the words through tasks. Learners could benefit through learning the pronunciation, spelling and grammatical features of the words in the APWL. These aspects of knowledge are complementary to understanding and remembering new words. Schmitt (2001) provided ample evidence that showed learning some formal features of the words, such as pronunciation, can aid recall of the written form of the words.

It is important to note that all the words in the APWL are mostly used in spoken English; therefore, all the language-focused activities should also include a meaning-related element. Learners could be encouraged to consciously use strategies to learn the words in the APWL. They could use word cards, with L2 words on one side, and L1 meaning on the reverse. Word cards have shown to be helpful in both receptive and productive learning. Learners could see the L1 meaning and recalling its L2 form for productive learning, or see the L2 word and recall its L1 meaning for receptive learning. The memory traces of both the form and meaning can be enhanced in word card learning.

Many words in the APWL are nouns that are used to refer to food, drink and equipment that are frequently encountered, like *tongs, cider* and *marmalade.* Learners could attach a little note to the subjects, and memory of form and meaning mapping could be reinforced by every encounter.

Learners should be aware that there are many words that have similar meanings, spellings and pronunciation, such as strut/stomp, swagger/snigger, and squeak/squeal. These types of words can cause confusion in understanding or remembering both the form and the meaning of the new words (Nation, 2001). These words should be learnt separately to reduce the burden of learning.

5.2.4 Fluency development

Faster retrieval of the target words is another important aspect of vocabulary knowledge, and this can best be achieved by repetitively using the target words to develop fluency. In this strand, the learners come under some pressure to speak or respond faster.

To develop speaking fluency, a learner could tell the same story about an everyday event, such as visiting the supermarkets or clinic using words in APWL to different people in a limited period of time each round, like the 4/3/2, which is reduced for each person spoken to (Nation, 2001). For example, they could tell the first person the experience in four minutes, the second in three minutes, and the last in two minutes. Both the story teller and the listener would be exposed to the same words over and over again in this activity.

To develop listening ability, learners could listen to the same video several times, until the video is quite familiar, and all words in it are known. After hearing the video, learners could be asked to retell the main points in the story to check whether they have understood the storyline. When the story is fully understood, they could start to extend what they have learnt in the story to real life (e.g., similar events in their lives), and tell their own stories to learning partners or group members.

5.3 Implications for further research

There is still much to be learnt about the words that occur often in daily life communication. In this section, a number of suggestions for future research will be proposed. Firstly, there is a need for a new learner corpus with a focus on authentic use of the language in English-speaking settings. Secondly, there is more to be learnt about the polysemous words in the APWL. Thirdly, it is important to see which words are usually used together with APWL words, in other words, the collocations of the words. Additionally, there are also a number of methodological possibilities that could be used to explore words of this kind. One such example is surveys which could be carried out to gauge learners' opinions of their communication problems. Also, when a new learner corpus with samples of authentic language use is established, a corpus approach could be used to develop future versions of the APWL.

5.2.1 Build a new learner corpus

The existing learner corpora, like the Cambridge Learner Corpus, collect spoken data from test sources, such as the responses to test questions during IELTS, etc. A corpus with a more natural use of language by learners is needed. Large native speakers' corpora like CANCODE (Cambridge and Nottingham Corpus of Discourse in English), collects spoken data from various sources, including casual conversation, people socializing, telephone conversation, shopping and discussions, etc., all of which involve authentic use of English. Such an approach could be used to build a corpus consisting of authentic use of language by learners from various settings in English-speaking contexts. Learners could carry recording devices and record whenever they feel ready. Valuable information could be drawn from such a corpus. Word frequency lists could be made, and compared with those from native speakers' corpora to see the differences in word composition. The corpus could also be used to estimate the productive vocabulary size of learners.

5.2.2 Understanding polysemous words in the APWL

Many English words are polysemous. In the APWL, 328 words (37.78%) are polysemous. Understanding the frequency of each meaning in the list has pedagogical implication. The GSL improved on previous high frequency list by adding a frequency index to each meaning in the words (West, 1953). By looking at the index, it is evident which meaning of the word has been most frequency used, and should therefore be learnt first. Teachers and learners could save a lot of learning effort by referring to the index. In light of this, it would be helpful to determine the frequency index for words in the APWL, and see the meaning that is most often used in spoken language.

5.2.3 Multi-word units and collocations

Collocation is one of the nine aspects of the vocabulary knowledge in Nation's (2001) framework. Learning collocations of the words in the APWL is another important extension of the present research. Learners would be advantaged by knowing which words could be used with other words in the APWL.

5.2.4 Use need analysis to develop the categories

The present list developed 28 categories based on Miller and Fellbaum's 40 beginners for nouns and verbs (1991), and the meanings of the words in the screened word list. There is another way of developing the categories. A future study could carry out need analysis among learners to determine the categories. This approach was adopted in the development of the surviving language learning syllabus of English (Nation & Crabbe, 1991). They carried out a need analysis by interviewing the learners, studying guide books and resorting to personal experience. For further study into word lists for advanced learners, the research could use surveys to ask advanced learners about the scenarios that they find to be relevant to their daily lives and those scenarios that they find difficult. In this way, the list of categories may more reliably reflect the needs of advanced learners in English-speaking contexts and be more helpful to them.

5.2.5 Use corpus to select potential words

The present word list selected words from dictionaries, because of the lack of appropriate learner corpora. Future study could use a corpus approach, if a learner corpus was built with samples from authentic spoken language use by learners in English-speaking contexts. When such a corpus is established, a word frequency list of learners' spoken language could be developed, based on frequency, coverage and dispersion of the words in the corpus. A comparison between frequency lists developed from native speakers' corpora and the learners' corpus could then be done to look for discrepancies. Those words found in the frequency list of native speakers, but not that of learners could be used as potential words for such a list. Learners could then be tested on whether they have any prior knowledge of the potential words. Those words which are not known to most learners could be retained, to make a word list that helps learners cope with daily communication.

6 References

Arnaud, P.J.L., & Savignon, S.J. (1997). Rare words, complex lexical units and the advanced learner. In Coady & Huckin (Eds.), *Second language vocabulary acquisition: a rationale for pedagogy* (pp.157-173). Cambridge: Cambridge University Press.

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, *6(4)*, 253-279.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American word frequency book.* New York: American Heritage Publishing Company Limited.

Chung, T., & Nation, I.S.P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language, 15(2),* 103-111.

Chung, T.M., & Nation, P. (2004). Identifying technical vocabulary. *System, 32(2*), 251-263.

Collins COBUILD advanced dictionary of English. (2009). Boston: Heinle Cengage Learning.

Collins COBUILD advanced dictionary of American English. (2008). Boston: Thomas Heinle.

Coxhead, A. (1998). T*he development and validation of an academic word list.* Unpublished Master's thesis, Victoria University of Wellington, Wellington, New Zealand.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34(2),* 213-238.

Crabbe, D., & Nation, I.S.P. (1991). A survival language learning syllabus for foreign travel. *System, 19(3),* 191-201.

Craik, F., & Lockhart, R. (1975). Levels of processing: a framework for memory research. *Journal of verbal learning and verbal behavior, 11 (6),* 671-684.

Hirsh, 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.

Hwang, K. (1989). Reading newspapers for the improvement of vocabulary and reading skills. Unpublished Master's Thesis, Victoria University of Wellington,

Wellington, New Zealand.

Hwang, K., & Nation, I.S.P. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System, 23(1),* 35-41.

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

Laufer, B. (1991). The development of L2 lexis in the expression of the advanced language learner. *Modern Language Journal, 75(4),* 440-448.

Laufer, B. (1998). The development of passive and active vocabulary: Same or different? *Applied Linguistics, 19(2),* 255-271.

Leech,G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English.* Harlow: Longman.

McCarthy, M. (2001). What is advanced level vocabulary? In Davidson, P., Coombe, C., Lloyd, D., & Palfreyman, D. (Eds.) *Teaching and learning vocabulary in another language (pp. 21-34).* Dubai: TESOL Arabia.

Meara, P. (1990). A note on passive vocabulary. *Second language research, 6(2),* 150-154

Miller, G.A., & Fellbaum, C. (1991). *Semantic networks in English. Cognition, 41,* 197-229.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19(3),* 304-330.

Nagy, W., Anderson, R., Schommer, M., Scott, J. A., & Stallman, A. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly, 24 (3),* 262- 281.

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 Bogaards,P. & Laufer, B. (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing (pp.3-13).* Amsterdam: John Benjamins.

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and

listening? *Canadian Modern Language Review, 63(1),* 59-82.

Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1(1),* 1-12.

Nation, I.S.P., & Webb, S. (2011*). Researching and analyzing Vocabulary.* Boston, MA: Heinle.

Oxford advanced learner's dictionary 8th edition. (2010). Oxford: Oxford University Press.

Praninskas, J. (1972). *American university word list*. London: Longman.

Richards, J. (1974). Word lists: problems and prospects. *RELC Journal, 5(2),* 69-84.

Schmitt, N. (2000). *Vocabulary in Language Teaching.* Cambridge: Cambridge University Press.

Schmitt, N., & Zimmerman, C.B. (2002). Derivative word forms: what do learners know? *TESOL Quarterly, 36(2),* 145-171.

Tyler,A., & Evans, V. (2004). Applying cognitive linguistics to pedagogical grammar: The case of Over. In Achard, M.,& Niemeier, S. (Eds.), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*(pp.257-280). Berlin/New York: Mouton de Gruyter.

Vespoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning, 53(3),* 547-586.

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

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15(2),* 130-163.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28,* 46-65.

Webb, S. (2008). Receptive and productive vocabulary size. *SSLA, 30(1),* 79-95.

Webb, S., & Rodgers, M.P.H. (2009)a. The vocabulary demands of television programs. *Language Learning, 59(2),* 335-366.

Webb, S. & Rodgers, M.P.H. (2009)b. The lexical coverage of movies. *Applied Linguistics, 30(3),* 407-427.

Wesche, M., Paribakht, T., & Paribakht, T. (2009). *Lexical inferencing in a first and second language: Cross-linguistic dimensions.* Britain: Short Run Press Ltd.

West, M. (1953). A general service list of English words. London: Longman, Green.

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication, 3(2),* 215-229.

7 Appendices

Appendix One: Headwords in Advanced Practical Word List

This list includes all the headwords in the APWL. The abbreviation accompanying each word is the part of speech of that word.

abacus	n.	amputate	n.
abduct	v.	anatomy	n.
abrasive	adj.	animosity	n.
abysmal	adj.	anorexia	n.
accomplice	n.	antics	n.
acne	n.	antidote	n.
acorn	n.	antipasto	n.
acquit	v.	antiseptic	n.
adamant	adj.	apathy	n.
adrenalin	n.	apostle	n.
aerobics	n.	appease	v.
aerosol	n.	arithmetic	n.
agile	adj.	aromatherapy	n.
ale	n.	artisan	n.
algebra	n.	asbestos	n.
alibi	n.	asphalt	n.
allspice	n.	assent	n.
ambidextrous	adj.	asteroid	n.
ambivalent	adj.	astrology	n.
amicable	adj.	astute	adj.
ammonia	n.	atheism	n.
amnesia	n.	atrocious	adj.
amphetamine	n.	aubergine	n.
amphitheatre	n.	auburn	adj.

audacious	adj.	blasphemy	n.
aura	adj.	blatant	adj.
autism	n.	blazer	n.
autograph	n.	bleak	adj.
avalanche	n.	bleary	adj.
avid	adj.	blemish	n.
backhand	n.	blindfold	n.
backlash	n.	bliss	n.
backlog	v.	blister	n.
backtrack	v.	blizzard	n.
baffle	v.	bloated	adj.
baguette	n.	blubber	n.
ballsy	adj.	bluff	n.
bandanna	n.	blunt	adj.
bandwagon	n.	bodice	n.
bandwidth	n.	boggle	v.
banter	n.	bogus	adj.
barley	n.	boisterous	adj.
basmati	n.	boomerang	n.
batter	v.	boutique	n.
beanie	n.	braid	n.
bedridden	adj.	brash	adj.
benign	adj.	brat	n.
bereaved	adj.	brawl	n.
berserk	adj.	brazen	adj.
biceps	n.	breaststroke	n.
bimbo	n.	breathalyze	v.
blabbermouth	n.	brew	v.
bland	adj.	brittle	adj.

bronchitis	n.	chauffeur	n.
broth	n.	chickpea	n.
brunette	n.	chime	v.
bulimia	n.	chisel	n.
burlesque	n.	chlamydia	n.
burp	v.	cholesterol	n.
bust	v.	chubby	adj.
busybody	n.	chutney	n.
butch	adj.	ciabatta	n.
buttock	n.	cider	n.
bystander	n.	cinnamon	n.
cabaret	n.	citrus	n.
cagey	adj.	clamp	n.
callous	adj.	clan	n.
camisole	n.	clatter	v.
camouflage	n.	clique	n.
canister	n.	clog	v.
cannabis	n.	clot	n.
canopy	n.	clove	n.
carcass	n.	clunk	n.
cartilage	n.	coax	v.
cashmere	n.	cobweb	n.
casserole	n.	coddle	v.
catapult	v.	codeine	n.
catcall	n.	coleslaw	n.
catnap	n.	collate	v.
chaplain	n.	comfy	adj.
charade	n.	commemorate	v.
charisma	adj.	commotion	n.

compilation	n.	cupid	n.
concoction	n.	custodian	n.
concussion	n.	cyst	n.
condone	v.	dabble	v.
connive	v.	daffodil	n.
connoisseur	n.	dainty	adj.
consecrate	n.	daredevil	n.
constellation	n.	deadpan	adj.
constipation	n.	debauchery	n.
contagious	adj.	decadent	adj.
contraception	n.	delicatessen	n.
convertible	n.	delirious	adj.
cord	n.	delirium	n.
corduroy	n.	deluxe	adj.
corsage	n.	denim	n.
counterfeit	adj.	dent	v.
cove	n.	dentures	n.
coy	adj.	deodorant	n.
cram	v.	deranged	adj.
crass	adj.	despicable	adj.
cremate	v.	detergent	n.
cringe	v.	dexterous	adj.
crinkle	v.	diabetes	n.
crouton	n.	dialysis	n.
crowbar	n.	diarrhea	n.
crust	n.	diehard	n.
cufflink	n.	dilate	v.
cull	v.	dingy	adj.
cult	n.	dire	adj.

diss	v.	edible	adj.
dissect	v.	effeminate	adj.
ditto	n.	effervescent	adj.
ditzy	adj.	electrocute	v.
dodge	v.	elope	v.
dollop	n.	embellish	v.
doodle	n.	ember	n.
dope	n.	embossed	adj.
dopey	adj.	emerald	n.
dosage	n.	encore	n.
dough	n.	enigma	n.
douse	v.	envisage	v.
dowdy	adj.	epilepsy	n.
dreadlocks	n.	epilogue	n.
dribble	v.	epiphany	n.
drool	v.	epitomize	v.
droop	v.	etiquette	n.
dud	adj.	euthanasia	n.
dunk	v.	evangelical	adj.
duo	n.	evaporate	v.
dupe	v.	exasperate	n.
duvet	n.	excruciating	adj.
eardrum	n.	exorbitant	adj.
earlobe	n.	expatriate	n.
earmark	v.	exquisite	adj.
earmuffs	n.	extravaganza	n.
éclair	n.	exuberant	adj.
ecstasy	n.	facelift	n.
eczema	n.	fallacy	n.

falsetto	n.	fluorescent	adj.
fanfare	n.	fluoride	n.
fang	n.	flurry	n.
farsighted	adj.	foetus	n.
fastidious	adj.	foliage	n.
feisty	adj.	fondue	n.
feline	adj.	forage	v.
feud	n.	forensic	adj.
fiddle	v.	forerunner	n.
fidget	v.	foreshore	n.
figurine	n.	fowl	n.
finicky	adj.	foyer	n.
fizz	v.	fraternity	n.
fizzle	v.	fray	v.
flamboyant	adj.	freehand	adj.
flannel	n.	frenetic	adj.
flap	v.	fret	v.
flask	n.	frill	n.
flaunt	v.	fringe	n.
fleece	n.	frivolous	adj.
flimsy	adj.	frizzy	n.
flinch	v.	frumpy	adj.
fling	v.	fudge	n.
flint	n.	fugitive	n.
floppy	adj.	fumigate	v.
floss	n.	funnel	n.
flounder	v.	furnace	n.
fluff	n.	furry	adj.
fluke	n.	gab	adj.

gadget	n.	graze	v.
gaga	adj.	groggy	n.
gala	n.	groin	n.
gallop	v.	grouch	n.
gangrene	n.	growl	v.
garnish	n.	gruff	adj.
gaudy	adj.	grumpy	adj.
gauze	n.	guacamole	n.
gawk	v.	guava	n.
gazebo	n.	gullible	adj.
geek	n.	gumdrop	n.
gelatine	n.	gurgle	v.
genealogy	n.	guru	n.
gin	n.	gush	v.
gland	n.	gusto	n.
glean	v.	guzzle	v.
gloss	n.	haemorrhage	n.
glucose	n.	hallucinate	v.
gluten	n.	hammock	n.
glutinous	adj.	hamper	v.
gnocchi	n.	hamstring	n.
gondola	n.	handlebar	n.
goof	v.	haphazard	adj.
gore	v.	haughty	adj.
gospel	n.	headrest	n.
graffiti	n.	headwind	n.
granite	n.	heirloom	n.
grater	n.	hem	n.
gravitate	v.	henchman	n.

hernia	n.	innuendo	n.
hinge	n.	insomnia	n.
hippie	n.	instigate	v.
hitch	v.	insulate	v.
hockey	n.	intermittent	adj.
holler	v.	intertwine	v.
holocaust	n.	jab	v.
hologram	n.	janitor	n.
homophobia	n.	jaywalking	n.
horoscope	n.	jersey	n.
horrid	adj.	jingle	v.
huff	v.	jinx	n.
hunch	n.	jolt	v.
husky	adj.	jostle	v.
hustle	v.	jot	v.
hymn	n.	judaism	n.
hype	n.	juggle	v.
hyperventilate	v.	jumpstart	v.
hypnosis	n.	kaleidoscope	n.
hypothermia	n.	karma	n.
icicle	n.	keg	n.
icing	n.	kennel	n.
idyllic	adj.	killjoy	n.
implore	n.	kindling	n.
impromptu	adj.	kinky	adj.
incognito	adj.	knack	n.
indent	v.	knead	v.
infatuated	adj.	kneecap	n.
inhale	v.	kook	n.

kosher	adj.	ludicrous	adj.
kudos	n.	lumber	n.
lacerate	v.	lunge	v.
lactose	n.	lurk	v.
ladle	n.	lustre	n.
lanolin	n.	macho	adj.
lard	n.	mahogany	n.
lash	n.	maim	v.
layman	n.	malaria	n.
leash	n.	malice	v.
ledge	n.	mangle	v.
leer	v.	manure	n.
leeway	n.	marble	n.
legible	adj.	margarine	n.
lethargy	n.	marinade	n.
leukaemia	n.	marmalade	n.
liaison	n.	maroon	adj.
licorice	n.	marshmallow	n.
limousine	n.	martyr	n.
lingerie	n.	mash	v.
lingo	n.	masquerade	n.
lint	n.	matinee	n.
liposuction	n.	maul	n.
liqueur	n.	mauve	adj.
liquorice	n.	mayhem	n.
lobotomy	n.	mayonnaise	n.
loophole	n.	meager	adj.
lozenge	n.	meander	v.
lucid	adj.	measles	n.

mediocre	adj.	muzzle	n.
mega	adj.	nausea	n.
megapixel	n.	navel	n.
melanoma	n.	neon	adj.
mellow	adj.	neuter	v.
melodrama	n.	nibble	v.
memorabilia	n.	niggle	v.
menace	n.	nightmarish	adj.
menopause	n.	nimble	adj.
meringue	n.	nirvana	n.
mesh	n.	nocturnal	adj.
mesmerize	v.	nonchalant	adj.
metabolism	n.	nosedive	v.
meticulous	adj.	nostalgia	n.
migraine	n.	oar	n.
molar	n.	obliterate	v.
mole	n.	oblivion	n.
montage	n.	obnoxious	adj.
mope	v.	ointment	n.
mortify	v.	oodles	n.
mosque	n.	ooze	v.
motif	n.	orchestrate	n.
mound	n.	origami	n.
mozzarella	n.	ostracize	v.
muesli	n.	ovary	n.
mumps	n.	overhaul	v.
mundane	adj.	oversight	n.
mural	n.	paddle	n.
mutter	v.	padlock	n.

palpable	adj.	perpetuate	v.
palpitate	v.	perverse	adj.
pamper	v.	philanthropy	n.
pamphlet	n.	phlegm	n.
pantomime	n.	phoney	adj.
pantry	n.	physique	n.
pantyhose	n.	pier	n.
papaya	n.	pinkie	n.
paprika	n.	pinprick	n.
paracetamol	n.	placebo	n.
paralyse	v.	placenta	n.
paramedic	n.	plaster	n.
paranoia	n.	platter	n.
parody	n.	pliers	n.
pastry	n.	ply	v.
pasture	n.	pneumonia	n.
pathology	n.	podium	n.
pavlova	n.	poise	n.
pawn	v.	polystyrene	n.
pebble	n.	poncho	n.
pecan	n.	ponderous	adj.
pediatrics	n.	posh	adj.
pedicure	n.	potpourri	n.
pedigree	n.	pouch	n.
peg	n.	pox	n.
pelvis	n.	prairie	n.
pendant	n.	precocious	adj.
penicillin	n.	premiere	n.
pepperoni	n.	premonition	n.

preposterous	adj.	rapids	n.
preppy	n.	raucous	adj.
pretzel	n.	raunchy	adj.
prim	adj.	ravioli	n.
prissy	adj.	recede	v.
prod	v.	recline	v.
pronto	adv	reconvene	v.
prop	v.	recuperate	v.
pry	v.	reef	n.
psalm	n.	reek	n.
psychopath	n.	regime	n.
puddle	n.	reinstate	v.
puff	v.	resurrect	v.
pulp	n.	resuscitate	v.
pulpit	n.	retaliate	v.
pumice	n.	retentive	adj.
pungent	adj.	revamp	v.
puree	n.	rhinestone	n.
purge	v.	ricochet	v.
purr	v.	rind	n.
quack	n.	rowdy	adj.
quaint	adj.	rupture	n.
quash	v.	sacrilege	n.
queasy	adj.	saga	n.
quicksand	n.	salami	n.
quilt	n.	salivate	v.
rake	n.	salmonella	n.
ram	v.	sanctify	n.
ranch	n.	sandpit	n.

sapphire	n.	simmer	v.
sash	n.	singe	v.
satchel	n.	sinus	n.
saunter	v.	sirloin	n.
savour	v.	skewer	n.
savoury	adj.	skit	n.
savvy	n.	skittish	adj.
scald	v.	sleaze	n.
scam	n.	slime	n.
scavenge	v.	slink	v.
schmooze	n.	slither	v.
schnapps	n.	slob	n.
scone	n.	slop	v.
scrape	v.	slouch	v.
scrawl	v.	slurp	v.
scribble	v.	slush	n.
scrumptious	adj.	smear	v.
scuffle	n.	smog	n.
sedate	adj.	smooch	v.
sedentary	adj.	smudge	n.
serviette	n.	smug	adj.
sesame	n.	snag	n.
sewage	n.	sneak	v.
sewer	n.	snicker	v.
shaft	n.	sniffle	v.
sham	n.	snigger	v.
sheen	n.	snip	v.
sherbet	n.	snot	n.
siesta	n.	snout	n.

snug	adj.	starch	n.
snuggle	v.	starstruck	adj.
socket	n.	staunch	adj.
sodden	adj.	stench	n.
soggy	adj.	stiletto	n.
solitaire	n.	stillbirth	n.
somersault	n.	stirrup	n.
sorbet	n.	stomp	v.
spank	v.	stoop	v.
spasm	n.	stringy	adj.
spatula	n.	strut	v.
spew	v.	stunt	v.
splat	n.	sty	n.
splurge	v.	suave	adj.
splutter	v.	suede	n.
spout	v.	superlative	adj.
sprain	v.	surge	n.
spree	n.	surreal	adj.
sprinkling	n.	swagger	v.
sprint	n.	swat	v.
spunk	n.	swindle	v.
squander	v.	swish	v.
squeak	v.	swivel	v.
squeal	v.	synopsis	n.
squelch	v.	tabloid	n.
squirm	v.	tantrum	n.
squirt	v.	tarpaulin	n.
stale	adj.	tenacious	adj.
stamina	n.	theology	n.

therapeutic	adj.	vaccinate	v.
thwart	v.	valet	n.
tinfoil	n.	vengeance	n.
tinge	n.	ventriloquist	n.
tingle	v.	versatile	adj.
tinker	v.	veterinarian	n.
tomboy	n.	vibe	n.
tongs	n.	villain	n.
tonsils	n.	vinaigrette	n.
toupee	n.	vinyl	n.
tousled	adj.	void	n.
treadmill	n.	volatile	adj.
treble	v.	voluptuous	adj.
trek	v.	voodoo	n.
trickle	v.	wacky	adj.
trinket	n.	waddle	v.
tripod	n.	waggle	v.
trot	v.	weave	v.
tubby	adj.	wheeze	v.
turmeric	n.	whim	n.
tuxedo	n.	whine	v.
tweak	v.	whisk	v.
twerp	n.	wholesome	adj.
twinge	n.	whoosh	v.
umpire	n.	whopping	adj.
umpteen	n.	wield	v.
underdog	n.	wig	n.
upbeat	adj.	wiggle	v.
uproot	v.	wiry	adj.

wishbone	n.	wriggle	v.	
wisp	n.	yam	n.	
wistful	adj.	yelp	n.	
wobble	v.			

Appendix two: Frequency sublists

This appendix includes the four sublists divided based on the frequency levels of the words in the British National Corpus.

Sublist one

adamant	compilation	foyer	hockey	mash
aerobics	contraception	fray	horrid	mega
ale	convertible	frill	hymn	menace
amputate	cord	fringe	icing	mole
arithmetic	cram	fudge	inhale	nibble
baffle	crust	furry	insulate	nightmarish
barley	daffodil	gaga	jersey	ooze
batter	dent	gala	jot	paddle
berserk	diabetes	gallop	juggle	pamphlet
blatant	dire	gin	knack	pantomime
bluff	dosage	gland	ledge	paracetamol
blunt	dough	gloss	leeway	paralyse
brew	dribble	gospel	ludicrous	pastry
bust	envisage	grater	lurk	peg
chauffeur	evaporate	graze	macho	pier
cider	fiddle	grumpy	mahogany	plaster
clamp	flap	hamper	marble	pneumonia
clog	fling	hernia	margarine	posh
commemorate	floppy	hinge	marmalade	pox

prop	scrape	snip	tinker	weave
puff	scribble	socket	treble	wig
ram	sewage	spout	trek	wiggle
rapids	sewer	sprinkling	trot	wobble
regime	shaft	squeal	versatile	
saga	slop	stringy	veterinarian	
scone	sneak	surge	void	

Sublist two

abduct	antiseptic	bereaved	buttock
abrasive	apathy	bimbo	bystander
abysmal	apostle	bland	cabaret
accomplice	aromatherapy	blazer	cagey
acne	asbestos	bleak	callous
acorn	asphalt	bliss	camouflage
acquit	assent	blister	canister
adrenalin	astrology	bloated	cannabis
aerosol	atrocious	bodice	canopy
agile	aubergine	boggle	carcass
algebra	aura	bogus	cartilage
alibi	autograph	boisterous	casserole
ambivalent	avalanche	braid	catapult
amicable	avid	brat	chaplain
ammonia	backlash	brazen	charisma
anatomy	backlog	brittle	chime
animosity	backtrack	bronchitis	chisel
anorexia	bandwagon	bulimia	cholesterol
antics	bedridden	burp	chubby
antidote	benign	butch	chutney

cinnamon	delicatessen	epitomize	forensic
clan	delirious	etiquette	forerunner
clatter	denim	evangelical	fowl
clot	deodorant	exasperate	fraternity
coax	despicable	excruciating	fret
cobweb	detergent	exorbitant	frivolous
collate	dialysis	exuberant	funnel
comfy	dilate	fallacy	furnace
commotion	dingy	fanfare	gadget
concoction	dissect	fang	gangrene
concussion	ditto	feud	gauze
condone	dodge	fidget	glean
connive	dollop	fizz	glucose
consecrate	doodle	flannel	gluten
constipation	dope	flask	graffiti
contagious	dowdy	flaunt	granite
cove	drool	fleece	groggy
coy	droop	flimsy	groin
crass	duo	flinch	growl
cremate	dupe	flint	gruff
cringe	duvet	floss	gullible
crinkle	eardrum	flounder	haemorrhage
crowbar	ecstasy	fluff	hallucinate
cull	eczema	fluorescent	hammock
cult	edible	fluoride	hamstring
custodian	embossed	flurry	handlebar
dabble	emerald	foetus	haphazard
dainty	encore	foliage	hem
decadent	epilepsy	forage	hippie

holler	leukaemia	motif	pedigree
holocaust	liaison	mound	pelvis
hologram	limousine	muesli	pendant
homophobia	liqueur	mundane	penicillin
horoscope	loophole	mural	perpetuate
huff	lucid	mutter	perverse
hunch	lumber	muzzle	philanthropy
hustle	maim	nausea	phlegm
hype	matinee	navel	phoney
hypothermia	malaria	neuter	physique
impromptu	malice	niggle	pliers
indent	mangle	nimble	ply
infatuated	manure	nirvana	podium
instigate	maroon	nostalgia	poise
intermittent	martyr	oar	polystyrene
intertwine	masquerade	obliterate	pouch
jab	mauve	obnoxious	prairie
jingle	mayhem	ointment	premiere
jolt	mayonnaise	orchestrate	prim
jostle	meager	ovary	prod
judaism	measles	overhaul	pry
kaleidoscope	mellow	oversight	psalm
kennel	menopause	padlock	psychopath
kinky	meringue	pamper	puddle
ladle	mesh	pantry	pulp
lash	metabolism	paranoia	pulpit
layman	meticulous	pasture	purge
leash	migraine	pathology	quack
lethargy	mope	pawn	quaint

quilt

ranch

recede

recline

reconvene

recuperate

reef

reek

reinstate

resurrect

resuscitate

retaliate

retentive

revamp

rind

rowdy

rupture

sacrilege

salami

salmonella

sanctify

sapphire

sash

savour

savoury

scald

scrawl

scuffle

sedate

sesame

sham

sheen

siesta

simmer

sinus

sleaze

slime

slob

slouch

slurp

slush

smear

smog

smug

snag

snigger

snout

snug

snuggle

soggy

somersault

spank

spasm

spew

splutter

sprint

squander

squeak

squelch

squirt

stale

stamina

starch

staunch

stiletto

stirrup

stoop

strut

stunt

suede

swindle

swish

swivel

tabloid

tantrum

theology

therapeutic

tinge

tingle

tongs

tonsils

trickle

tripod

tweak

twinge

umpire

upbeat

uproot

vaccinate

vengeance

vibe

villain

vinyl

volatile

voluptuous

wacky

waddle

waggle

wheeze

whim

whine

whisk

whoosh

wield

wisp

wriggle

Sublist three

abacus	citrus	euthanasia	gush
amnesia	clique	expatriate	gusto
amphetamine	clove	exquisite	guzzle
amphitheatre	clunk	extravaganza	haughty
appease	connoisseur	facelift	heirloom
artisan	constellation	falsetto	henchman
asteroid	corduroy	fastidious	husky
astute	counterfeit	feline	hypnosis
atheism	cupid	figurine	icicle
auburn	cyst	finicky	implore
audacious	daredevil	fizzle	incognito
autism	deadpan	flamboyant	innuendo
backhand	debauchery	fluke	insomnia
bandwidth	delirium	foreshore	karma
banter	deluxe	frenetic	keg
biceps	dentures	fugitive	kindling
blasphemy	deranged	fumigate	knead
bleary	douse	garnish	kosher
blemish	dreadlocks	gaudy	kudos
blizzard	dunk	gazebo	lacerate
boutique	effeminate	gelatine	lactose
brash	effervescent	genealogy	leer
brunette	electrocute	glutinous	legible
busybody	elope	gondola	lingerie
camisole	embellish	gore	lint
cashmere	ember	gravitate	liquorice
catcall	enigma	gurgle	lozenge
charade	epilogue	guru	lunge

lustre	poncho	sirloin	tenacious
marinade	ponderous	skewer	thwart
maul	precocious	skit	tomboy
meander	premonition	slink	tousled
melanoma	preposterous	slither	treadmill
melodrama	pronto	smudge	trinket
memorabilia	pumice	sodden	tubby
mesmerize	pungent	sorbet	tuxedo
molar	purr	spatula	umpteen
montage	quash	sprain	underdog
mortify	queasy	spree	valet
mosque	quicksand	squirm	ventriloquist
neon	raucous	stench	vinaigrette
nocturnal	raunchy	stillbirth	voodoo
nonchalant	ricochet	stomp	wholesome
oblivion	satchel	suave	wiry
ostracize	saunter	superlative	wistful
palpable	scam	surreal	yam
parody	scavenge	swagger	yelp
placebo	scrumptious	swat	
placenta	sedentary	synopsis	
platter	singe	tarpaulin	

Sublist four

allspice	bandanna	blubber	broth
ambidextrous	basmati	boomerang	burlesque
antipasto	beanie	brawl	catnap
baguette	blabbermouth	breaststroke	chickpea
ballsy	blindfold	breathalyze	chlamydia

ciabatta	guava	medicore	rhinestone
coddle	gawk	megapixel	salivate
codeine	geek	mozzarella	sandpit
coleslaw	gnocchi	mumps	savvy
corsage	goof	nosedive	schmooze
crouton	grouch	oodles	schnapps
cufflink	guacamole	origami	serviette
dexterous	gumdrop	palpitate	sherbet
diarrhea	headrest	pantyhose	skittish
diehard	headwind	papaya	smooch
diss	hitch	paprika	snicker
ditzy	hyperventilate	paramedic	sniffle
dopey	idyllic	pavlova	snot
dud	janitor	pebble	solitaire
earlobe	jaywalking	pecan	splat
earmark	jinx	pediatrics	splurge
earmuffs	jumpstart	pedicure	spunk
éclair	killjoy	pepperoni	starstruck
epiphany	kneecap	pinkie	sty
farsighted	kook	pinprick	tinfoil
feisty	lanolin	potpourri	toupee
fondue	lard	preppy	turmeric
freehand	licorice	pretzel	twerp
frizzy	lingo	prissy	whopping
frumpy	liposuction	puree	wishbone
gab	lobotomy	rake	
gaga	marshmallow	ravioli	

Appendix three: Usefulness sublists

This appendix includes the three sublists divided based on native speakers' judgment on their usefulness in daily life.

Sublist one

acne	cringe	fluorescent	nausea
adrenalin	crouton	fringe	ointment
aerosol	crust	gospel	pamper
amnesia	custodian	grumpy	paramedic
anorexia	denim	guacamole	pedicure
aromatherapy	dent	gullible	penicillin
atrocious	deodorant	hammock	premiere
autism	diabetes	homophobia	puddle
avalanche	diarrhea	horoscope	pungent
biceps	ditzy	hymn	queasy
bluff	doodle	infatuated	raunchy
boutique	dough	insomnia	recuperate
brawl	ecstasy	kennel	retaliate
bronchitis	edible	licorice	sapphire
bulimia	emerald	limousine	savour
canopy	epilepsy	lozenge	scone
casserole	etiquette	mahogany	slurp
chaplain	evangelical	marble	smear
charade	excruciating	marinade	smog
chisel	farsighted	marmalade	smooch
citrus	figurine	megapixel	smudge
coax	flannel	metabolism	snip
cobweb	fleece	mozzarella	spasm
constellation	flimsy	muesli	spew

stamina	tonsils	vaccinate	wishbone
tarpaulin	tuxedo	vengeance	
tinfoil	umpire	whine	
tomboy	uproot	wholesome	

Sublist two

abrasive	backhand	breaststroke	coleslaw
accomplice	backlash	broth	comfy
acorn	backtrack	burp	compilation
aerobics	baguette	bust	concussion
agile	bandanna	butch	constipation
ale	banter	camisole	contagious
algebra	batter	camouflage	contraception
alibi	bedridden	carcass	corduroy
allspice	benign	cartilage	coy
ambidextrous	bereaved	cashmere	cram
amputate	berserk	charisma	cremate
anatomy	bland	chauffeur	crinkle
animosity	blasphemy	chickpea	cufflink
antics	blatant	chime	daffodil
apathy	blazer	chubby	dainty
appease	blemish	chutney	deadpan
arithmetic	blindfold	ciabatta	decadent
asphalt	bliss	cider	delicatessen
astrology	blizzard	cinnamon	delirious
astute	bloated	clique	deluxe
atheism	blubber	clog	dentures
aubergine	brash	clove	detergent
audacious	brat	codeine	diehard

dingy	fiddle	goof	knead
dire	fidget	gore	kudos
dissect	fizz	granite	lactose
ditto	flamboyant	gravitate	lard
dodge	flask	growl	leer
dollop	flaunt	gruff	legible
dopey	flinch	gurgle	lethargy
dosage	fling	hallucinate	leukaemia
dreadlocks	floppy	headrest	lingerie
drool	floss	headwind	liposuction
dunk	flounder	heirloom	lucid
duvet	fluff	hippie	ludicrous
earlobe	foetus	hockey	lunge
earmark	forerunner	holler	lustre
earmuffs	foreshore	hunch	macho
éclair	frivolous	hype	maim
electrocute	frizzy	hypnosis	malaria
elope	frumpy	icing	manure
ember	fudge	impromptu	maroon
envisage	fugitive	inhale	marshmallow
epitomize	fumigate	innuendo	mash
euthanasia	funnel	insulate	maul
evaporate	gadget	jab	mayhem
exquisite	gallop	jaywalking	mayonnaise
extravaganza	gazebo	jersey	meander
fallacy	gelatine	judaism	measles
fastidious	genealogy	keg	mediocre
feisty	gloss	kindling	mega
feline	glucose	kinky	mellow

memorabilia	pastry	recede	somersault
meringue	pasture	recline	sorbet
mesmerize	pathology	reconvene	spatula
meticulous	pavlova	resurrect	splurge
mole	pebble	retentive	sprain
motif	pecan	saga	sprint
mundane	pedigree	salmonella	squeak
mutter	pendant	sandpit	squeal
muzzle	pepperoni	savoury	squirm
nibble	perpetuate	scald	squirt
niggle	platter	scam	stale
nirvana	pneumonia	scrape	starch
nocturnal	podium	scrawl	stench
nonchalant	polystyrene	scribble	suave
nosedive	poncho	scrumptious	suede
nostalgia	posh	serviette	superlative
ooze	preppy	sesame	swindle
orchestrate	prim	sewage	synopsis
origami	prissy	siesta	tabloid
ostracize	prod	skewer	tenacious
ovary	pronto	slither	therapeutic
overhaul	psychopath	slush	tinge
padlock	pumice	smug	tingle
pantomime	purge	snag	trickle
pantry	quicksand	sneak	trinket
pantyhose	rake	snicker	trot
paprika	ranch	snuggle	tweak
paralyse	rapids	soggy	twerp
paranoia	ravioli	solitaire	upbeat

versatile	vinaigrette	weave	whoosh
veterinarian	voluptuous	wheeze	wiry
villain	voodoo	whisk	yam

Sublist three

abacus	ballsy	busybody	connive
abduct	bandwagon	buttock	connoisseur
abysmal	bandwidth	bystander	consecrate
acquit	barley	cabaret	convertible
adamant	basmati	cagey	cord
ambivalent	beanie	callous	corsage
amicable	bimbo	canister	counterfeit
ammonia	blabbermouth	cannabis	cove
amphetamine	bleak	catapult	crass
amphitheatre	bleary	catcall	crowbar
antidote	blister	catnap	cull
antipasto	blunt	chlamydia	cult
antiseptic	bodice	cholesterol	cupid
apostle	boggle	clamp	cyst
artisan	bogus	clan	dabble
asbestos	boisterous	clatter	daredevil
assent	boomerang	clot	debauchery
asteroid	braid	clunk	delirium
auburn	brazen	coddle	deranged
aura	breathalyze	collate	despicable
autograph	brew	commemorate	dexterous
avid	brittle	commotion	dialysis
backlog	brunette	concoction	dilate
baffle	burlesque	condone	diss

dope	flap	gland	holocaust
douse	flint	glean	hologram
dowdy	fluke	gluten	horrid
dribble	fluoride	glutinous	huff
droop	flurry	gnocchi	husky
dud	foliage	gondola	hustle
duo	fondue	graffiti	hyperventilate
dupe	forage	grater	hypothermia
eardrum	forensic	graze	icicle
eczema	fowl	groggy	idyllic
effeminate	foyer	groin	implore
effervescent	fraternity	grouch	incognito
embellish	fray	guava	indent
embossed	freehand	gumdrop	instigate
encore	frenetic	guru	intermittent
enigma	fret	gush	intertwine
epilogue	frill	gusto	janitor
epiphany	furnace	guzzle	jingle
exasperate	furry	haemorrhage	jinx
exorbitant	gab	hamper	jolt
expatriate	gaga	hamstring	jostle
exuberant	gala	handlebar	jot
facelift	gangrene	haphazard	juggle
falsetto	garnish	haughty	jumpstart
fanfare	gaudy	hem	kaleidoscope
fang	gauze	henchman	karma
feud	gawk	hernia	killjoy
finicky	geek	hinge	knack
fizzle	gin	hitch	kneecap

kook	menace	parody	pry
kosher	menopause	pawn	psalm
lacerate	mesh	pediatrics	puff
ladle	migraine	peg	pulp
lanolin	molar	pelvis	pulpit
lash	montage	perverse	puree
layman	mope	philanthropy	purr
leash	mortify	phlegm	quack
ledge	mosque	phoney	quaint
leeway	mound	physique	quash
liaison	mumps	pier	quilt
lingo	mural	pinkie	ram
lint	navel	pinprick	raucous
liqueur	neon	placebo	reef
liquorice	neuter	placenta	reek
lobotomy	nightmarish	plaster	regime
loophole	nimble	pliers	reinstate
lumber	oar	ply	resuscitate
lurk	obliterate	poise	revamp
malice	oblivion	ponderous	rhinestone
mangle	obnoxious	potpourri	ricochet
margarine	oodles	pouch	rind
martyr	oversight	pox	rowdy
masquerade	paddle	prairie	rupture
matinee	palpable	precocious	sacrilege
mauve	palpitate	premonition	salami
meager	pamphlet	preposterous	salivate
melanoma	papaya	pretzel	sanctify
melodrama	paracetamol	prop	sash

satchel	slouch	strut	valet
saunter	sniffle	stunt	ventriloquist
savvy	snigger	sty	vibe
scavenge	snot	surge	vinyl
schmooze	snout	surreal	void
schnapps	snug	swagger	volatile
scuffle	socket	swat	wacky
sedate	sodden	swish	waddle
sedentary	spank	swivel	waggle
sewer	splat	tantrum	whim
shaft	splutter	theology	whopping
sham	spout	thwart	wield
sheen	spree	tinker	wig
sherbet	sprinkling	tongs	wiggle
simmer	spunk	toupee	wisp
singe	squander	tousled	wistful
sinus	squelch	treadmill	wobble
sirloin	starstruck	treble	wriggle
skit	staunch	trek	yelp
skittish	stiletto	tripod	
sleaze	stillbirth	tubby	
slime	stirrup	turmeric	
slink	stomp	twinge	
slob	stoop	umpteen	
slop	stringy	underdog	

Appendix four：Headwords in Advanced Practical Word List in Categories

Act, Action

abduct	crinkle	hype	paralyse	slurp
acquit	cull	indent	pawn	smear
assent	dabble	inhale	perpetuate	snip
backhand	dent	insulate	ply	spank
backlash	dilate	intertwine	prod	splurge
backlog	dissect	jaywalking	prop	spout
backtrack	douse	jingle	pry	squander
batter	dribble	jot	puff	squeak
bluff	droop	juggle	purge	squeal
brawl	dunk	jumpstart	quash	squelch
burp	dupe	knead	ram	squirt
bust	earmark	lacerate	recede	stunt
catnap	envisage	leeway	reconvene	trek
charade	facelift	maim	recuperate	trickle
chime	fizz	mangle	reinstate	tweak
clog	flap	meander	revamp	uproot
clunk	fray	mortify	ricochet	vengeance
coax	fumigate	niggle	savour	weave
coddle	glean	nosedive	scam	whisk
collate	goof	obliterate	scrape	whoosh
commemor	gravitate	ooze	scrawl	wield
ate	gurgle	overhaul	scribble	wiggle
connive	gush	palpitate	singe	
cram	guzzle	pamper	slop	

Natural object

asteroid	emerald	ledge	marble	rapids
avalanche	flint	lumber	pebble	sapphire
boomerang	granite	lustre	pumice	
ember	kindling	mahogany	quicksand	

Animal

blubber	forage	growl	maul	sirloin
carcass	fowl	kennel	muzzle	slime
cobweb	furry	lanolin	neuter	snout
fang	gallop	lard	pedigree	stirrup
feline	gore	leash	purr	wishbone
foetus	graze	manure	ranch	yelp

Natural phenomenon

blizzard	constellation	evaporate	headwind	puddle

Artifact

asbestos	duvet	frill	mesh	poncho
braid	figurine	gauze	montage	rhinestone
canopy	flannel	hammock	motif	sash
cashmere	fleece	heirloom	mural	suede
denim	floss	hologram	origami	tarpaulin
doodle	fluff	kaleidoscope	pendant	trinket

Person and profession

accomplice   chauffeur   geek   layman   underdog

artisan   connoisseur   grouch   martyr   valet

bimbo   custodian   guru   paramedic   ventriloquist

blabbermouth   daredevil   henchman   preppy   veterinarian

bleary   diehard   hippie   psychopath villain

brat   expatriate   incognito   slob

brunette   forerunner   janitor   tomboy

busybody   fraternity   killjoy   twerp

bystander   fugitive   kook   umpire

Attribute

abrasive   bleak   decadent   finicky

adamant   blunt   deluxe   flamboyant

agile   bogus   deranged   flimsy

ambidextrous   boisterous   dexterous   floppy

amicable   brash   dingy   fluke

astute   brazen   dire   fluorescent

atrocious   brittle   ditzy   freehand

auburn   butch   dopey   frenetic

audacious   cagey   dowdy   frivolous

aura   callous   dud   gab

avid   camouflage   effeminate   gaga

ballsy   charisma   effervescent   garnish

bandwagon   chubby   epitomize   gaudy

bedridden   comfy   exquisite   gloss

benign   coy   falsetto   glutinous

bereaved   crass   farsighted   gruff

berserk   dainty   fastidious   grumpy

bland   deadpan   feisty   gullible

gusto	mundane	queasy	suave
haughty	neon	raucous	superlative
husky	nimble	raunchy	surreal
impromptu	nocturnal	retentive	tantrum
intermittent	palpable	savvy	tenacious
jinx	perverse	sedate	tousled
kinky	phoney	sedentary	treadmill
knack	poise	sheen	tubby
legible	ponderous	skittish	versatile
loophole	posh	snag	void
macho	precocious	snug	volatile
maroon	preposterous	soggy	voluptuous
mauve	prim	splat	wacky
meager	prissy	spunk	wiry
mediocre	pungent	starstruck	wistful
mellow	quack	staunch	
menace	quaint	stench	

Plant

acorn	corsage	foliage	yam
clove	daffodil	liquorice	

Body parts, dress and accessories

adrenalin	bodice	corduroy	earlobe	hamstring
bandanna	buttock	cufflink	fringe	hem
beanie	camisole	cyst	frizzy	jersey
biceps	cartilage	dreadlocks	gland	kneecap
blazer	clot	eardrum	groin	lash

lingerie	navel	physique	sinus	tuxedo
menopause	ovary	pinkie	snot	wig
molar	pantyhose	placenta	stiletto	
mole	pelvis	satchel	toupee	

Cognition, mind and thoughts

ambivalent	enigma	homophobia	nostalgia	synopsis
baffle	fallacy	horoscope	oversight	whim
boggle	haphazard	hunch	stamina	

Communication

appease	epilogue	innuendo	pamphlet	skit
banter	extravaganza	lingo	parody	solitaire
catcall	fanfare	lucid	saga	tabloid
ditto	feud	matinee	schmooze	
encore	graffiti	melodrama	scuffle	

Quantity

dosage	oodles	surge	umpteen
mega	pinprick	tinge	whopping
megapixel	sprinkling	treble	wisp

Event and happening

| cabaret | gala | holocaust | premiere | stillbirth |
| flurry | hockey | masquerade | sprint | |

Feeling, emotion and attitudes

abysmal	ecstasy	ludicrous	premonition
animosity	excruciating	mesmerize	sleaze
antics	exorbitant	meticulous	smug
apathy	fret	mope	twinge
blatant	frumpy	nausea	upbeat
bliss	horrid	nightmarish	vibe
debauchery	idyllic	nonchalant	
delirious	infatuated	obnoxious	
despicable	kudos	pantomime	

Food, drink and cooking

ale	coleslaw	guava	mozzarella	savoury
allspice	crouton	gumdrop	muesli	schnapps
antipasto	crust	icing	papaya	scone
aubergine	dollop	keg	paprika	scrumptious
baguette	dough	kosher	pantry	serviette
barley	éclair	lactose	pastry	sesame
basmati	edible	ladle	pavlova	sherbet
brew	fondue	licorice	pecan	simmer
broth	fudge	liqueur	pepperoni	skewer
casserole	gelatine	margarine	platter	sorbet
chickpea	gin	marinade	pretzel	spatula
chutney	glucose	marmalade	pulp	stale
ciabatta	gluten	marshmallow	puree	starch
cider	gnocchi	mash	ravioli	stringy
cinnamon	grater	mayonnaise	rind	turmeric
citrus	guacamole	meringue	salami	vinaigrette

wholesome

Society and study

algebra	arithmetic	forensic	ostracize
alibi	astrology	genealogy	philanthropy
anatomy	etiquette	liaison	regime

Group and collection

clan	compilation
clique	duo

State and condition

burlesque	exuberant	hamper	pronto	rowdy
commotion	fizzle	lethargy	reek	sham
embellish	flounder	mayhem	resurrect	sodden
embossed	groggy	oblivion	resuscitate	

Location and place

amphitheatre	foreshore	mound	prairie
boutique	foyer	pasture	pulpit
cove	gazebo	pier	reef
delicatessen	gondola	podium	sandpit

Substance

abacus	ammonia	asphalt	bandwidth	blindfold
aerosol	antiseptic	autograph	blemish	canister

chisel	fluoride	memorabilia	pouch	socket
clamp	funnel	oar	quilt	tinfoil
concoction	furnace	paddle	rake	tongs
convertible	gadget	padlock	sewage	tripod
cord	handlebar	peg	sewer	vinyl
crowbar	headrest	plaster	shaft	
detergent	hinge	pliers	slush	
earmuffs	icicle	polystyrene	smog	
flask	limousine	potpourri	smudge	

Motive

counterfeit	flaunt	instigate	swindle
elope	huff	orchestrate	thwart
exasperate	implore	retaliate	tinker

Bodily experience

aerobics	concussion	electrocute	ointment	spasm
amputate	contraception	gangrene	paranoia	sprain
aromatherapy	cremate	hallucinate	pedicure	therapeutic
bloated	dentures	lint	salivate	tingle
breaststroke	deodorant	liposuction	scald	
breathalyze	dope	metabolism	siesta	

Movement and behavior

catapult	cringe	drool	flinch	hitch
clatter	diss	fiddle	fling	holler
condone	dodge	fidget	gawk	hustle

hyperventilate	nibble	snicker	stomp	waggle
jab	recline	sniffle	stoop	wheeze
jolt	saunter	snigger	strut	whine
jostle	scavenge	snuggle	swagger	wobble
leer	slink	somersault	swat	wriggle
lunge	slither	spew	swish	
lurk	slouch	splutter	swivel	
malice	smooch	spree	trot	
mutter	sneak	squirm	waddle	

Religion

apostle	consecrate	evangelical	karma	sacrilege
atheism	cult	gospel	mosque	sanctify
blasphemy	cupid	hymn	nirvana	theology
chaplain	epiphany	judaism	psalm	voodoo

Illness and treatment

acne	cholesterol	haemorrhage	migraine
amnesia	codeine	hernia	mumps
amphetamine	constipation	hypnosis	paracetamol
anorexia	contagious	hypothermia	pathology
antidote	delirium	insomnia	pediatrics
autism	diabetes	leukaemia	penicillin
blister	dialysis	lobotomy	phlegm
bronchitis	diarrhea	lozenge	placebo
bulimia	eczema	malaria	pneumonia
cannabis	epilepsy	measles	pox
chlamydia	euthanasia	melanoma	rupture

salmonella

stye

tonsils

vaccinate