SELECTING AND CREATING A WORD LIST FOR ENGLISH LANGUAGE TEACHING

by **Deny A. Kwary** and **Jurianto**

Universitas Airlangga

Dharmawangsa Dalam Selatan, Surabaya 60286, Indonesia

d.a.kwary @ unair.ac.id / juri.jurianto@gmail.com

Abstract

Since the introduction of the General Service List (GSL) in 1953, a number of studies have confirmed the significant role of a word list, particularly GSL, in helping ESL students learn English. Given the recent development in technology, several researchers have created word lists, each of them claims to provide a better coverage of a text and a significant role in helping students learn English. This article aims at analyzing the claims made by the existing word lists and proposing a method for selecting words and a creating a word list. The result of this study shows that there are differences in the coverage of the word lists due to the difference in the corpora and the source text analysed. This article also suggests that we should create our own word list, which is both personalized and comprehensive. This means that the word list is not just a list of words. The word list needs to be accompanied with the senses and the patterns of the words, in order to really help ESL students learn English.

Keywords: English; GSL; NGSL; vocabulary; word list

1. Introduction

A word list has been noted as an essential resource for language teaching, especially for second language teaching. The main purpose for the creation of a word list is to determine the words that the learners need to know, in order to provide more focused learning materials. In the English language, the word list which has been widely used for creating learning materials is the General Service List (GSL) created by West (1953). GSL contains a list of 2,000 words which has the highest frequency in a general English text. A computer analysis shows that about 80% of the individual words in most written English texts are members of these 2,000 words (Nation, 2001). This means that if an English language learner knows these 2,000 words, he/she can have a fair comprehension of a general English text. The 2,000 English words will look very small if we compare this number with the more than 600,000 entries listed in the *Oxford English Dictionary* (2009).

Realizing the significance of a word list, lexicographers of current English monolingual learner's dictionaries have created a word list, called 'defining vocabulary', to define the headwords in the dictionaries. The first dictionary that uses defining vocabulary is the *Longman Dictionary of Contemporary English* (1978), where the words put in the defining vocabulary are based on GSL. The *Oxford Advanced Learner's Dictionary* started using the so-called 'Oxford 3000' from its seventh edition (2007), as defining vocabulary. This means that all of the definitions of the more than 180,000 headwords in an English learner's dictionary are only defined using a list of 2,000-3,000 words.

Since language develops over time, the frequency of use of some words may also change. As for GSL, the critiques on the fact that it is outdated have been put forward for several decades. For example, Richards (1974: 71) pointed out the words *fear*, *loyal*, and *mannerism* in GSL to be of limited utility, and suggested the more common words, such as *astronaut*, *helicopter* and *pilot*, which are missing from GSL. Nevertheless, it can also be argued that the words suggested by Richards (1974) were only common in the 1970s, and are of limited utility at the moment. The data from the *Corpus of Historical American English* (Davies 2010-) and the *Corpus of Contemporary American English* (Davies 2008), show that the word *astronaut* occurred 16.67 times per million words in the 1970, but it decreases to about a half of it, i.e. 7.99 times per million words in 2010.

Another study (Kwary 2011) makes a comparison of the word *potato*, which is not in GSL, and the word *virtue*, which is included in the list. Based on the *Corpus of Historical American English* (Davies 2010), in 1950s the word *potato* occurred 22.49 times per million, while the word *virtue* occurred 33.20 times per million. For 2010-2012, the *Corpus of Contemporary American English* (Davies 2008) shows that the frequency of the word *potato* is 45.45 times per million, while the frequency of the word *virtue* is only 14.51 times per million. In other words, the word *virtue* has a very limited utility and should be excluded from a new frequency word list, while the word *potato* has the potential to be included in the new word list as it has a higher frequency per million words.

Sixty years after the publication of GSL, there are at least two sets of word lists that call themselves as the New GSL. In this paper, these New GSLs are called NGSL1 and NGSL2. NGSL1, created by Browne, Culligan and Phillips (2013), contains approximately 2,800 words, selected from the 273 million-word subsection of the Cambridge English Corpus (CEC). NGSL1 has been available online since February 2013 (http://newgeneralservicelist.org). NGSL2 was released in August 2013 with the online/advance access publication of the article written by Brezina and Gablasova (2015). It was created based on four English language corpora (the Lancaster-Oslo-Bergen Corpus, the British National Corpus, the BE06 Corpus of British English, and the EnTenTen12) of a total size of over 12 billion running words. NGSL2 is also available online at http://corpora.lancs.ac.uk/vocab/index.php.

In addition to NGSL1 and NGSL2, there are still several other word lists that are created based on a corpus or corpora. The examples are the word list created from the BNC (*British National Corpus*), which is available at http://ucrel.lancs.ac.uk/bncfreq, and the word list created from the COCA (*Corpus of Contemporary American English*), which is available at http://www.wordfrequency.info. A combination of these lists (BNC and COCA) has also been in use since 2012, and can be downloaded from the personal website of Paul Nation at http://www.victoria.ac.nz/lals/about/staff/paul-nation.

Each of the word lists claims to have covered the most important words that English language learners need to know. Consequently, teachers may be baffled as to which word list they should rely on. In the next section, a comparison of these word lists is made, so that teachers can decide which word list they should (or should not) consider when creating teaching materials for ESL or EFL students.

2. Comparing the word lists

When NGSL1 was published, a comparison between the coverage of GSL and that of NGSL1 in a text has been made by Browne (2013). The comparison shows that the coverage of GSL is only 84.24%, while the coverage of NGSL1 is 90.34% (Browne 2013: 16). However, the different percentages can be due to two main factors. The first is the difference in the number of word families used in the computer analysis. GSL only has 1,964 word families (the 2,000 words are regrouped by Browne into 1,964 word families), while NGSL1 contains 2,368 word families. The bigger number of word families in NGSL1 can be the cause of the higher percentage in the coverage of NGSL1 than GSL. The second factor is the text used in the computer analysis. The text analysed is from the CEC corpus, which is the basis to make NGSL1. Consequently, the higher coverage of NGSL1 for the CEC text can be due to the fact that NGSL1 was created using the text from CEC.

In another article, Browne (2014) made a comparison between GSL, NGSL1, and NGSL2. The result shows that GSL offers slightly better coverage for texts of classic literature (about 0.8% better than NGSL1 and 4.5% more than NGSL2), while NGSL1 offers 5-6% more coverage than either word list for two more modern corpora, i.e. *Scientific American* and *The Economist* (both are the names of magazines). However, the difference may also be due to the difference in the number of headwords, i.e. GSL has 1,986 headwords (the 2,000 words are regrouped by Browne into 1,986 headwords), NGSL1 has 2,801 headwords, and NGSL2 has 2,228 headwords. Again, the bigger coverage of NGSL1 could be due to the higher number of headwords in NGSL1 than the other word lists.

In order to compare the coverage of GSL, NGSL1, and NGSL2, as well as the other new word list called BNC-COCA, in a general English text, a small corpus compiled from five news articles published at MTV Asia website (http://www.mtvasia.com/news) on 1 April 2015 was created. The calculation of the coverage of the word lists for the MTV news articles is shown in Table 1. The calculations for GSL, NGSL1, and BNC-COCA are done by using the Vocabulary Profilers available at http://www.lextutor.ca/vp/comp/ (retrieved on 8 April 2015). The results shown in the Vocabulary Profilers start from the level of 1,000 words, then 2,000 words, and so on. NGSL2 is not available in that Vocabulary Profilers web page, so it results in different levels (see Table 1). The calculation for NGSL2 is done using http://corpora.lancs.ac.uk/vocab/analyse.php (retrieved on 8 April 2015). The three levels available are the 500 words, 1,000 words and 2,500 words.

Table 1. The coverage of the word lists in the MTV news articles

GSL		**NGSL1**		**NGSL2**		**BNC-COCA**	
Level	**Cumul.**	**Level**	**Cumul.**	**Level**	**Cumul.**	**Level**	**Cumul.**
1,000w	81.70%	1,000w	79.96%	500w	62.3%	1,000w	82.98%
2,000w	85.17%	2,000w	85.61%	1,000w	67.6%	2,000w	88.25%
2,570w	87.87%	2,801w	89.61%	2,500w	75.6%	3,000w	91.08%

For a similar comparison between the word lists, we shall focus on the results for the 1,000 words. As we can see in Table 1, the coverage for GSL is 81.7%, NGSL1 is 79.96%, NGSL2 is 67.6%, and BNC-COCA is 82.98%. The highest percentage, thus the greatest coverage, is obtained by BNC-COCA. This could be due to the fact that the MTV news articles are closely related to the American English, so a word list compiled from a bigger proportion of American English text will obtain the highest coverage. As the name suggests, BNC-COCA is made from BNC (British National Corpus) and COCA (Corpus of Contemporary American English). BNC contains approximately 100 million words, while COCA contains about 450 million words, which is more than four times bigger than that of BNC.

Looking at the results shown in Table 1, we can question whether NGSL1 and NGSL2 are really significant updates of GSL. At the 1,000 word level, GSL has a better coverage than NGSL1 and NGSL2. If we make the comparison at the 2,000 word level for GSL and NGSL1, and the 2,500 word level for NGSL2, the highest coverage is achieved by NGSL1. However, the difference is not significant. NGSL1 is only less than 1 percent higher than GSL (i.e. 85.61% and 85.17%). This small difference may reflect two possible aspects. The first is that the high frequency words have not changed dramatically after 60 years. The second is that the 2,000 word level is quite a stable level to obtain an approximately 80% coverage of the words in a general English text.

If we relate the results shown in Table 1 and those attested by Browne (2013), we can infer that the coverage of a word list largely depends on the source of text analysed. In the research done by Browne (2013), the coverage of GSL is only 84.24%, while the coverage of NGSL1 is 90.34%, because the text analysed is from CEC, which is the same as the corpus used to create NGSL1. In the results shown in Table 1 the highest coverage is the BNC-COCA word list, because the text analysed is from American English news articles, which is similar to most of the source texts used to create the BNC-COCA word list.

3. Towards a personalized and comprehensive word list

Realizing the differences in the corpus data used to create the word lists and the differences in the coverage of the word lists, English language teachers may face difficulties in deciding which word lists to use. To determine which word list to use, we need to return to the fundamental purpose of creating a word list, i.e. to determine the words that the learners need to know, in order to provide more focused learning materials. This means that the word list should be created from the text that the students will use. For example, if we teach students who want to take the IELTS (International English Language Testing System) exams, the word list created should be based on the exam papers which have been used in the IELTS exams. In a similar case in some countries, when we teach school students who want to take the national exam, the word list created should be based on the national compulsory textbooks and the exam papers.

Creating merely a word list, however, will not give much benefit for the students. It is hard to know the meaning of a word when it occurs in isolation. Hanks (2000: 214) states that “words have meaning potentials, rather than just meaning. The meaning potential of each word is made up of a number of components, which may be activated cognitively by other words in the context in which it is used.” This means that a word will be meaningful when it is used with other words. Therefore, it is necessary to create a personally functional word list to aid learners learn rather than an exclusive source of learning vocabulary. The following are the suggestions that teachers can follow if they want to create a personalized and comprehensive word list.

3.1. From a reading text to a word list

A textbook usually contains reading texts which become the basis of each chapter in the book. Creating a word list from a reading text could be troublesome if we do not know the right tools. Textbooks which have been available in a .pdf format can be directly converted to a .txt file using several computer programs that can be obtained for free. If the textbooks are not available in a .pdf format, we need to scan the pages using a scanner with an OCR (Optical Character Recognition) function, so that all the characters can be read and the file can be saved in a .pdf format.

To create a word list, we need to convert the .pdf file into a .txt file. One of the ways to convert a PDF file into a .txt file is by using the computer program or software called AntFileConverter (Anthony 2013). This software can be downloaded for free from www.laurenceanthony.net/software/antfileconverter/ and can be run directly on any computer, without installing it. Figure 1 shows the screenshot of AntFileConverter.

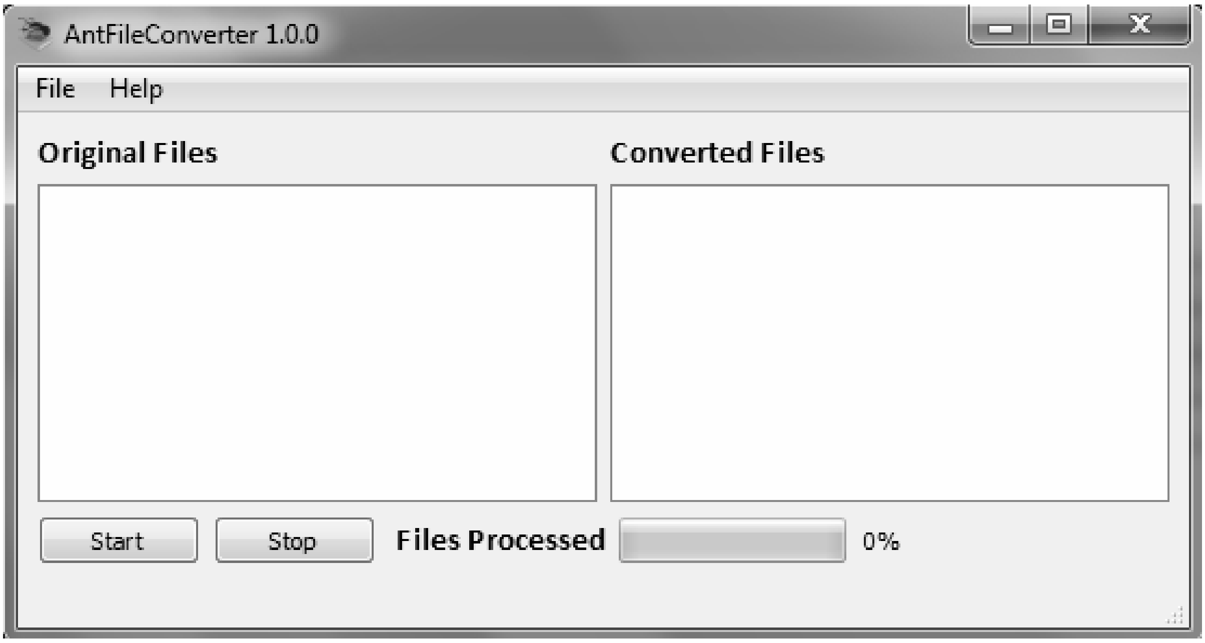
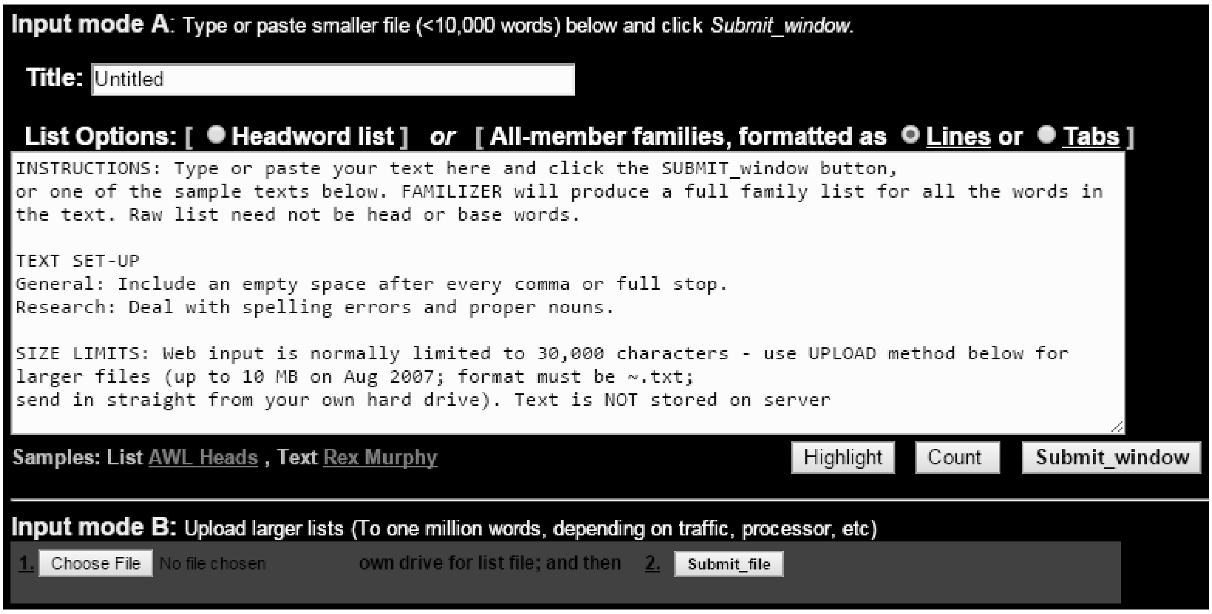


Figure 1. The screenshot of AntFileConverter

The process for converting a .pdf file into a .txt file can be done fast and easily by using the AntFileConverter. We only need to open the file (Under the menu 'File' in Figure 1), and click the 'Start' button. We can also convert several files directly by choosing the files using the 'Ctrl' button on the keyboard or by opening a directory that contain several files. After opening the files, we just need to click 'Start', and the .pdf file will be converted into a .txt file within seconds. The result file will be shown on the right side of the window (the 'Converted Files' in Figure 1). The converted files are automatically saved in the same folder as the original files.

With the file in a .txt format, we have several options or tools to create a word list. The one recommended in this paper is the tool called 'Familizer', which is available online at http://www.lextutor.ca/familizer. It can handle a corpus up to the size of 10MB in a .txt format (i.e. a corpus of 1 million words). The interface of Familizer is shown in Figure 2.



Input mode A: Type or paste smaller file (<10,000 words) below and click *Submit_window*.

Title: Untitled

List Options: [● Headword list] *or* **[All-member families, formatted as ○ Lines or ● Tabs]**

```
INSTRUCTIONS: Type or paste your text here and click the SUBMIT_window button,
or one of the sample texts below. FAMILIZER will produce a full family list for all the words in
the text. Raw list need not be head or base words.

TEXT SET-UP
General: Include an empty space after every comma or full stop.
Research: Deal with spelling errors and proper nouns.

SIZE LIMITS: Web input is normally limited to 30,000 characters - use UPLOAD method below for
larger files (up to 10 MB on Aug 2007; format must be ~.txt;
send in straight from your own hard drive). Text is NOT stored on server
```

Samples: List AWL Heads **, Text** Rex Murphy

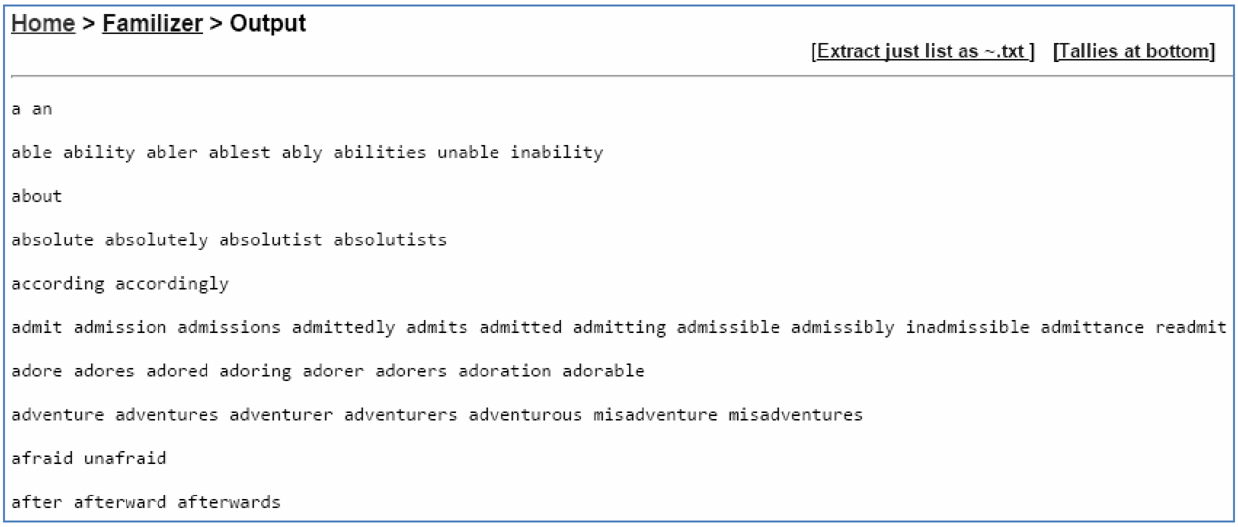
Input mode B: Upload larger lists (To one million words, depending on traffic, processor, etc)

1. Choose File No file chosen own drive for list file; and then 2. Submit_file

Figure 2. The screenshot of the webpage of Familizer

As we can see in Figure 2, there are two ways to create a word list, i.e. using the Input mode A or the Input mode B. The first one, i.e. the Input mode A, is suitable for a short text because the upper limit is 10,000 words. This can be done if we want to create a word list of a particular reading text or a corpus which contains no more than 10,000 words (roughly 20 pages of single spacing text). We only need to copy paste the text into the box provided, and then press 'Enter' on the computer keyboard.

The second one, i.e. the Input mode B, is useful when we want to create a word list from a bigger corpus (more than 10,000 words). As we can see in Figure 2, the Input mode B can handle a corpus up to one million words. The word list can simply be created by clicking the 'Choose File' button, and then the 'Submit file' button (see Figure 2). By using the Familizer, we will get a word list containing the words found in our text, together with all the family members. The result of the word list of the MTV News article is shown in Figure 3.



Home > Familizer > **Output**

[Extract just list as ~.txt] [Tallies at bottom]

```
a an
able ability abler ablest ably abilities unable inability
about
absolute absolutely absolutist absolutists
according accordingly
admit admission admissions admittedly admits admitted admitting admissible admissibly inadmissible admittance readmit
adore adores adored adoring adorer adorers adoration adorable
adventure adventures adventurer adventurers adventurous misadventure misadventures
afraid unafraid
after afterward afterwards
```

Figure 3. The word list created by the Familizer

As we can see in Figure 3, the words are already listed alphabetically with the family members. In the corpus analysed in Figure 3, actually the corpus only has the word *adorable*, but the result from the Familizer shows not only the word *adorable*, but also the other word family members, i.e. *adores*, *adoring*, *adorer*, etc. in the same line as its headword (i.e. *adore*). Consequently, we do not only have the list of the words that are in the text, but the words have been alphabetically rearranged based on their headwords, and added with the family members. This will make it easier for us to make vocabulary exercises or a personal dictionary.

3. 2. From a word list to word senses and patterns

A word list usually only consists of a list of words. This is hardly useful, especially if we want to use the word list to assist ESL students in reading and writing a text. In this case, the word list needs to be accompanied with the word senses and patterns. The current word sense information is needed to assist in reading a text. The pattern information is necessary to assist in writing a text.

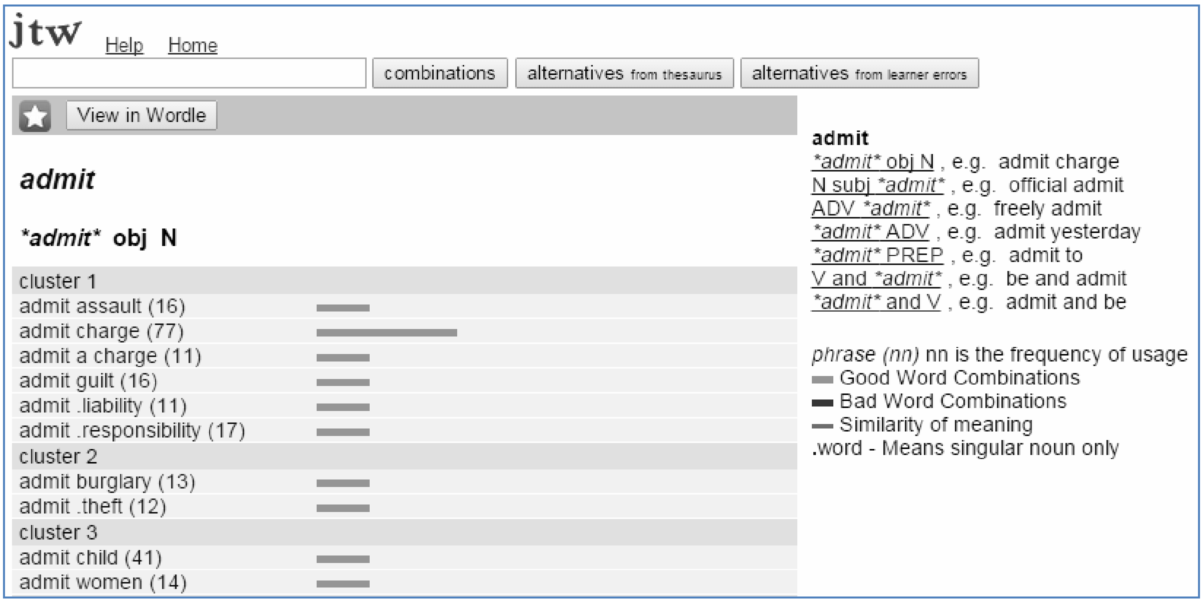
One example of the variety in the word sense is the word *gay*. GSL contains the word *gay*, but the sense of that word in the 1950s (when GSL was published) is different from the sense in the 2010s. By looking at the concordance lines of the text in 1950s (e.g. from the *Corpus of Historical American English*), as shown in Table 2, we can assume that the word *gay* in GSL would most likely to refer to the sense 'cheerful'. The following are the examples of the concordance lines.

Table 2. The concordance lines of the word *gay*

Year	**Concordance Lines**
1951	had jumped from the terrace outside Laura's pretty, **gay** apartment?
1952	her way of saying it, and how he'd been so charmed by her **gay**, girlish voice
1953	Why can't you be **gay** and lighthearted and cheerful, like Leo? " Norman
1954	Joan smiled, the **gay** confiding smile which had won many a heart, and said,

Nowadays, or in 2010s, the word *gay* is more likely to be connected with the sense 'homosexual' or 'sexually attracted to people of the same sex' (cf. *Oxford Advanced Learner's Dictionary*, 2010). Consequently, having a mere word list will not be really useful. It is necessary to provide the sense, the definition, or the explanation of the words in order to assist the students in understanding the right meaning of the word, especially when reading a text. If we want to provide the students with a more thorough meaning of a word, we can list all the senses for each word. However, the senses must be ordered based on frequency, i.e. the more frequent sense is listed first. For example, *Oxford Advanced Learner's Dictionary* (2010) lists five senses of the adjective *gay*, in which the sense 'sexually attracted to people of the same sex' is listed first. We can also choose to list only a few senses that we think appropriate and will likely be encountered by the students during their studies, e.g. *Oxford Learner's Dictionary of Academic English* (2014) only lists two senses of the adjective *gay*.

In addition to the word senses, we also need to add the patterns to the word list. When a student wants to write a text, he/she will need to know what words usually occur with a particular word. Take for example, the word *admit*. Knowing its meaning may not be very helpful if the student wants to write an idiomatic sentence using that word in a particular genre. The student may want to know the patterns and the words that usually go with the word *admit* (or the collocations). One way to obtain the pattern and collocation data is by using the website http://www.just-the-word.com. The result of the search for the word *admit* is shown in Figure 4.



admit

admit **obj N**

cluster 1	
admit assault (16)	▬
admit charge (77)	▬▬▬▬
admit a charge (11)	▬
admit guilt (16)	▬
admit .liability (11)	▬
admit .responsibility (17)	▬
cluster 2	
admit burglary (13)	▬
admit .theft (12)	▬
cluster 3	
admit child (41)	▬
admit women (14)	▬

admit
admit obj N , e.g. admit charge
N subj **admit** , e.g. official admit
ADV **admit** , e.g. freely admit
admit ADV , e.g. admit yesterday
admit PREP , e.g. admit to
V and **admit** , e.g. be and admit
admit and V , e.g. admit and be

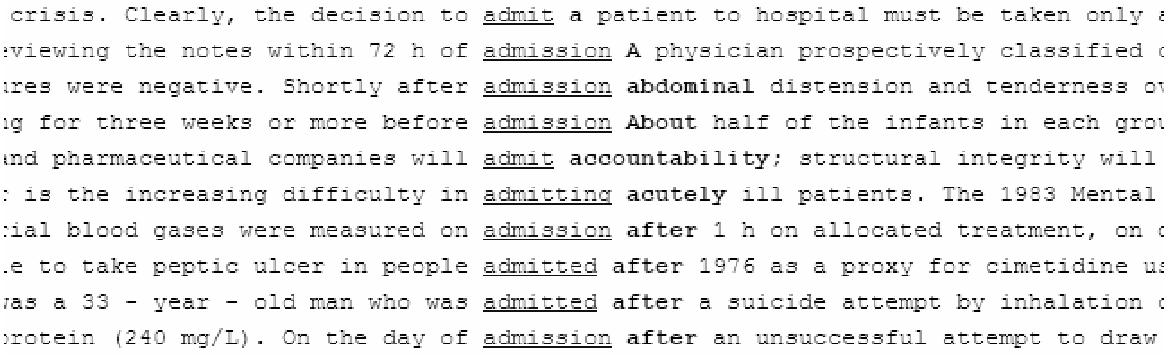
phrase (nn) nn is the frequency of usage
▬ Good Word Combinations
▬ Bad Word Combinations
▬ Similarity of meaning
.word - Means singular noun only

Figure 4. The result of the search for the word *admit*

On the right side of the screen (see Figure 4) we can see the patterns of the word *admit*. We can figure out that it can be used with an object noun (**admit** obj N), it can be used with an adverb (ADV **admit** and **admit** ADV), etc. In the middle of the screen, we can see the examples of the patterns together with the most common words that are usually used with the word *admit* (e.g. there are 16 example sentences for the word *admit* followed by *assault*; there are 77 example sentences for the word *admit* followed by *charge*; etc.). If we click one of the word pairs, the concordance lines (example sentences) will be shown. The concordance lines are taken from the British National Corpus, the 100 million word corpus, which present the actual use of the words.

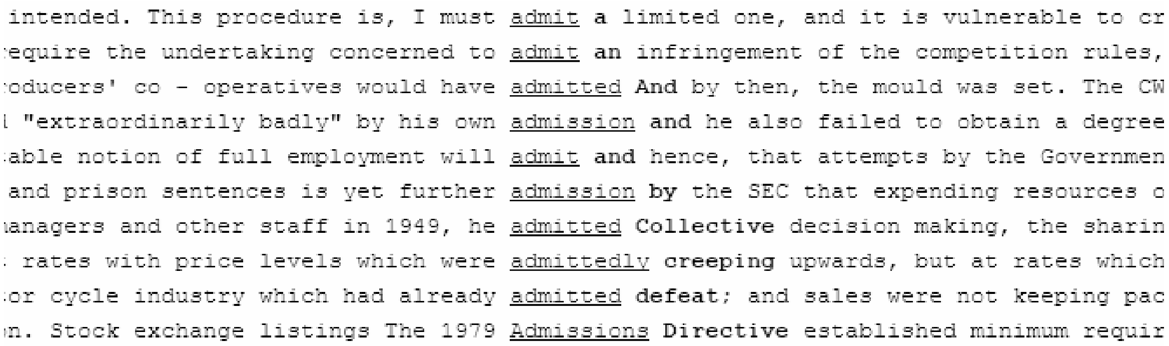
If we are interested in knowing the use of a particular word in a particular genre, we can use http://www.lextutor.ca/conc/eng, which provides access to several corpora. This website has divided the British National Corpus into five genres, i.e. medical, commerce, humanities, law, and social sciences. Therefore, if we want to see how the word *admit* is used in medical text, we can type in the word *admit* and choose the BNC Med corpus. The result is shown in the first part of Figure 5. The sentences there are quite different from those shown in the second part of Figure 5, which are taken from commerce texts (BNC Commerce).

A Sample from the BNC Med



```
crisis. Clearly, the decision to admit a patient to hospital must be taken only a
eviewing the notes within 72 h of admission A physician prospectively classified c
ures were negative. Shortly after admission abdominal distension and tenderness ov
ng for three weeks or more before admission About half of the infants in each grou
and pharmaceutical companies will admit accountability; structural integrity will
: is the increasing difficulty in admitting acutely ill patients. The 1983 Mental
ial blood gases were measured on admission after 1 h on allocated treatment, on c
.e to take peptic ulcer in people admitted after 1976 as a proxy for cimetidine us
ras a 33 - year - old man who was admitted after a suicide attempt by inhalation c
rotein (240 mg/L). On the day of admission after an unsuccessful attempt to draw
```

A Sample from the in the BNC Commerce



```
intended. This procedure is, I must admit a limited one, and it is vulnerable to cr
equire the undertaking concerned to admit an infringement of the competition rules,
oducers' co - operatives would have admitted And by then, the mould was set. The CW
l "extraordinarily badly" by his own admission and he also failed to obtain a degree
able notion of full employment will admit and hence, that attempts by the Governmen
and prison sentences is yet further admission by the SEC that expending resources o
anagers and other staff in 1949, he admitted Collective decision making, the sharin
: rates with price levels which were admittedly creeping upwards, but at rates which
or cycle industry which had already admitted defeat; and sales were not keeping pac
n. Stock exchange listings The 1979 Admissions Directive established minimum requir
```

Figure 5. The concordance lines of the word *Admit*

From the concordance lines shown in Figure 5, we can see that there are differences between the usage of the word *admit* in medical and in commerce discourse. Consequently, by using the concordance lines from BNC, we can write more appropriate sentences for a particular genre.

4. Conclusion

Currently, there are several word lists which can be used as a basis for creating language teaching materials to determine the vocabulary profile of a text. The analysis in this paper shows that the coverage of the word lists varies from one to another, most likely due to the difference in the corpus data used to create the word lists and the text analysed. The new word lists, e.g. NGSL1 and NGSL2, do not necessarily contain more significant coverage when compared with GSL. In an example of a text from the MTV news articles, the coverage of the BNC-COCA word list surpasses GSL, NGSL1, and NGSL2. However, this can be due to the similarity between the text used to create the BNC-COCA word list and the news published in the MTV news articles, i.e. from the United States of America.

Instead of discussing further which word list to be selected from the current word lists, this paper suggests that teachers should create their own word list based on the needs of their students. Given the current technology, the process of creating a word list from the main textbooks or learning materials used by the teachers can be done quite easily. However, teachers should not create a mere word list, i.e. a list of words. The meaning of a word may change over time and is likely to differ from one context to another. A word can only be meaningful when it is used with other words. This calls for a comprehensive word list, which means that it is necessary to add the senses and the patterns to the words, so that the word list created can be more useful to assist the learning process.

References

Anthony, L. (2013). *AntFileConverter* (Version 1.0.0) [Computer Software]. Tokyo: Waseda University. Retrieved August 1, 2016, from www.laurenceanthony.net/software/antfileconverter/.

Brezina, V., & Gablasova, D. (2015). *English Vocabulary Tool*. Retrieved August 1, 2016, from http://corpora.lancs.ac.uk/vocab.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, 36(1): 1-22.

Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *The Language Teacher*, 37(4): 13-16.

Browne, C. (2014). A New General Service List: The better mousetrap we’ve been looking for? *Vocabulary Learning and Instruction*, 3(2): 1-10.

Browne, C., Culligan, B., & Phillips, J. (2013). *A New General Service List*. Retrieved August 1, 2016, from http://www.newgeneralservicelist.org.

Corpus Concordance English. Retrieved August 1, 2016, from http://www.lextutor.ca/conc/eng.

Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved August 1, 2016, from http://corpus.byu.edu/coca/

Davies, M. (2010). The Corpus of Historical American English: 400 million words, 1810-2009. Retrieved August 1, 2016, from http://corpus.byu.edu/coha/

Familizer. Retrieved August 1, 2016, from http://www.lextutor.ca/familizer.

Frequency Lists. Retrieved August 1, 2016, from http://ucrel.lancs.ac.uk/bncfreq/flists.html.

Hanks, P. (2000). Do word meanings exist? *Computers and the Humanities*, 34: 205-215.

Just the Word. Retrieved August 1, 2016, from http://www.just-the-word.com.

Kwary, D. A. (2011). From defining vocabulary lists to English equivalents for L3-L2 mini-dictionaries. In: The *Proceedings of the ASIALEX*. Kyoto: ASIALEX.

Longman Dictionary of Contemporary English. 1st ed. (1978). London: Longman.

Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

News. Retrieved August 1, 2016, from http://www.mtvasia.com/news.

Oxford Advanced Learner’s Dictionary. 7th ed. (2007). Oxford: Oxford University Press.

Oxford Advanced Learner's Dictionary. 8th ed. (2010). Oxford: Oxford University Press.

Oxford English Dictionary. 2nd ed. (2009). Oxford: Oxford University Press.

Oxford Learner's Dictionary of Academic English. (2014). Oxford: Oxford University Press.

Richards, J. C. (1974). Word lists: Problems and prospects. *RELC Journal,* 5(2): 69–84.

The BNC/COCA word family lists. Retrieved August 1, 2016, from http://www.victoria.ac.nz/lals/about/staff/paul-nation.

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

Word Frequency Data. Retrieved August 1, 2016, from http://www.wordfrequency.info.