



An Investigation into the Use of Frequency Vocabulary Lists in University Intensive English Programs

Theodore Burkett

¹ Academic Bridge Program, Petroleum Institute, Abu Dhabi, United Arab Emirates

Received 13th Oct. 2015, Revised 14th Oct. 2015, Accepted 19th Oct. 2015, Published 1st Dec. 2015

Abstract: The use of vocabulary frequency lists in English language teaching and learning has been an area of continued research for roughly the past 170 years. There has been an increased focus on these lists in the last 25 years with a number of new and revised lists being developed, with three in 2013 alone. However, there has been comparatively little focus on how these lists are utilized. In order to cast light on this area, this exploratory study focuses on the extent of the use of these lists in university intensive English programs and shows that there is a substantial belief in the use of these lists, despite a number of limitations, and that they were used by a majority of the programs surveyed. Coxhead's Academic Word List was the most used list by far, but a range of lists was used depending upon the individual programs. Summaries of four individual programs and areas for future research are discussed.

Keywords: Vocabulary, Lexis, Frequency Lists, AWL, GSL, New General Service List, EAP, Intensive English Programs (IEP), Academic Bridge Programs, corpus, CEFR.

1. INTRODUCTION

Word lists are one way to help direct vocabulary teaching and learning. Indeed, as Nation highlights (Bogaards & Laufer, 2006), "Making word lists in the field of L2 learning and teaching is usually done for the purpose of designing syllabuses and in particular it is an attempt to find one way of determining necessities (what needs to be learned) as a part of needs analysis." These lists are often utilized by publishers as well, as can be seen by the use of lists like the Oxford 3000 and the Academic Word List, which are used to inform choice of lexis used in a variety of texts.

Word frequency seems an obvious candidate for prioritizing the acquisition of lexis. As Nation and Waring state concisely in Schmitt and McCarthy (1997, p.17), "Frequency information provides a rational basis for making sure that learners get the best return for their vocabulary learning effort by ensuring that the words studied will be met often. Vocabulary frequency lists which take account of range have an important role to play in curriculum design and setting learning goals." The question of course, is exactly what role they do play, especially in English for

Academic Purposes (EAP) contexts. As such, this study provides an overview of key word-frequency lists, looks at some of the important issues regarding list development, highlights some of the most recent developments in this area and discusses some of the ways that have been proposed to use them. After that, it outlines the rationale behind the study into the use of frequency lists in university foundation and intensive English programs, presents the methodology and goes through the findings of this exploratory study.

2. FREQUENCY-BASED VOCABULARY LISTS- AN OVERVIEW

A. History and Current Developments

In order to understand the current day status of these lists, it is necessary to give a brief overview of the history and some of the key points about how some of the most important lists were developed. According to McArthur (1998), work on frequency-based vocabulary lists in English of some form or another has been going on for around 170 years. However, the word frequency list that laid the foundations for much modern day work was Michael West's General Service List of English Words, or GSL, which was published in 1953 after almost 20 years in construction. This list



was so named, because it was designed to be of “general service” to learners of English as a foreign language (Browne, Culligan and Phillips, 2013). It was noted quite early on that this roughly 2,000 headword, word family based list was focused on written English and did not represent spoken English in the same fashion. Furthermore, Howatt (1984) stresses that it is not strictly a word-frequency list as it has other components including the elimination of specialist items, potentially offensive and slang words among others. However, as Espinosa (2003) points out, the GSL list remained the predominate word-frequency list for almost the next 50 years and was used as the basis for computer programs such as VocabProfile (Laufer & Nation, 1999) that reference it to carry out lexical analysis including frequency studies.

The next most influential development was most likely Coxhead’s Academic Word List (AWL). This list was based on the foundation of the GSL but acknowledged that students of EAP had far different needs than those studying general English as academic corpora vary noticeably from non-academic ones. A total of 570 word families were identified for inclusion in the AWL, which covers about 10% of the 3.5 million word academic corpus she constructed, and together with the GSL, they account for more than 86% of this corpus (Coxhead, 2000). When compared to another non-academic corpus of 3.7 million words based on fiction texts, these words represented only 1.4% of the coverage.

Moving forward, one of the newer lists of import is the Oxford 3000. With this list, Oxford University Press created a unique list specifically for the needs of English language learners. It is not strictly a frequency list, but one that also takes into account range and familiarity and one that employed at least 70 experts to provide input. Two important notes are that frequency ratings are not provided for words on this list, and it seems to include both lemmas and some word families – combining adjectives and adverbs together in many cases under a single headword.

In 2013, at least 3 new lists of note were released, two of them in advance publication in Oxford University Press’s *Applied Linguistics*. The first of these was ‘*A New Academic Vocabulary List*’ by Gardner and Davies, which involved the creation of the new 500 lemma list based on the 120 million-word COCA academic corpus. They recognized that a new Academic Word list was needed for a number of reasons, with the most pronounced being Coxhead’s AWL’s relationship with the GSL list and its use of

word families to determine word frequencies (p.3). They highlight the facts that since the GSL is actually based on corpus work from the early 1900’s, the AWL includes many words in in the highest –frequency lists of the British National Corpus (BNC; Nation 2004; Hancioglu et al 2008; Nation 2008; Cobb 2010; Neufield et al. 2011; Schmitt & Schmitt 2012)(p. 4).”

The second list of note that was made available on the *Applied Linguistics* website on August 26th 2013 in advance of its publication was the ‘*New General Service List*’ (*new-GSL*) created by Brezina and Gablasova. They also comment on the limitations of the GSL, and unlike West, they chose to construct their new list on transparent, replicable and quantitative criteria. Like the AVL, this new-GSL relies on lemmas, and reports a common lexical core between the 4 wordlists of 2,122 words with almost a 71% overlap between these texts. In addition, as two of the corpora were based on more modern sources, they identified another 378 lemmas that were not included within the first 3000 words on the older lists. Therefore, in total, the new GSL includes 2,494 lemmas.

One apparent limitation of this lemma-based approach is the limited breadth of these lists. When looking at the combination of the new-GSL and the new AVL, is that there is a great deal of overlap between the two lists with only 54 words on the AVL that are not included in the new-GSL, and of the 500 words on the AVL, 201 are included in the first 1000 lemmas of the new-GSL. Therefore, these lists are much more limited in range than the combination of West’s GSL and Coxhead’s AWL.

The third important frequency word list developed in 2013 is another new General Service List (NGSL) (Browne, Culligan & Phillips, 2013). This list was based on a carefully selected 273 million word subsection of the 2 billion word Cambridge English Corpus (CEC), and in its creation, followed many of the same steps as West. In this approximately 2800-word list, they aimed to update the size of the corpus, and to provide a more modern list providing the highest amount of coverage possible.

It is clear that these new lists, if embraced by teachers and publishers, may provide a valuable contribution in the future. However, this seems unlikely to happen in a fashion similar the original GSL as the construction of lists has become much more manageable.



B. Uses in English language teaching and learning

Throughout the development of these lists, different ideas on how to best use them for English language learning and teaching have been expressed, and several key areas have emerged. Even with Thorndike's first lists prior to the GSL, (Lorge & Thorndike, 1963), the use of these lists by publishers was quickly acknowledged, especially for helping develop reading skills for elementary school students. Additionally, while students can utilize these lists to target vocabulary development, Nation noted that these lists are not intended to be given directly to students, but rather to serve as a guideline for teachers and book makers (Nation, 1997). This has included dictionaries and course books that are tied directly to these lists. They can also be used to help guide the development of curriculum and assessment materials. Along with a tool such as the Compleat Lexical Tutor (Cobb, 2013), they can be used to check the appropriateness of the lexis in texts to match a certain level of learner, and where they are found to be too difficult, they can be simplified to be more suitable. Corpora and frequency lists are certainly utilized by English language teaching and learning publishers and have been at least since Thorndike's book of 20,000 words (McKee, 1937). They are also used in the creation of high stakes exams, for example, for the Pearson Test of English. On Pearson's website, they note that they use their own corpus of international English as well as an academic collocation list that they provide.

C. Issues regarding the development of word frequency lists

Before any program or organization decides to utilize a list, there are a number of key questions regarding the lists that should be considered, such as:

- Are they worth using?
- What principles should users of the lists look for?
- What problems need to be overcome?
- Are there other options besides frequency?

The value of using these lists depends on a number of factors including desired goals of the institution or organization in conjunction with the lists, the size of the institution, the profile of the learners, available resources, and needs for standardization and transparency. For publishing companies, larger exam boards, and larger academic programs that require a certain level of standardization, it is certainly arguable that by clarifying the expected lexis with an explicit frequency focus, it will certainly help all involved stakeholders to be fully aware of what is expected of them. If this list is aligned with assessment and

curriculum for different levels inside an academic program, it becomes more transparent to fairly assess what the students should know at any given point. Of course, the use of these lists can also help guide learners to acquire a solid base of the most frequent vocabulary.

However, in some cases, the argument does not seem to be quite as strong. If a program is smaller or if individual instructors are given full autonomy for assessment and materials, it might be better to focus more on the specific perceived needs of the students. This could include a wide range of possibilities including specific educational or professional needs (i.e. training for the hospitality industry, or medical English), constraints of the course itself (e.g. being required to use specific materials upon which the learners might be assessed), or the need to focus on specific lexical issues possibly based on widening awareness of lexemes, delving into L1 related issues or building on previous English language education. Additionally, if the course is composed of learners who have mastered most of this vocabulary, these lists could be avoided and instead, teachers could teach vocabulary learning strategies so they can learn less frequent words on their own (Schmitt and Schmitt, 2012).

Secondly, there are a number of principles that need to be considered when looking at the construction of a frequency list before adopting or working with it. Nation and Waring mention the following (1997, p.18):

- 1) representativeness – including both written and spoken corpuses as well as a sample of representative text types,
- 2) frequency and range - including not just overall frequency but also range across a variety of text types and genres,
- 3) word families – dealing with the distinction between lemmas versus word families, thinking about issues like regularity of form and learners seeing the relationship between words in the same word family,
- 4) idioms and set expressions – phrases like “good morning” and “set out” might need to be included as separate entries,
- 5) range of information - just how much information is provided along with the word on the list, are collocations, alternate word forms, variations in meanings, etc. all included with the list, and
- 6) other criteria – like ease or difficulty of learning, necessity, cover, stylistic level and emotional words.



Representativeness is relevant because spoken and written corpora can vary greatly and because a more limited sample of text types will not provide a representative sample – most of the existing lists cover this with a broad range of samples. Frequency and range highlight the reality that in some types of texts, such as business English, some terms might be overrepresented compared to a more general sample and thus not be suitable for a more general list. The previous focus on word families was further elaborated in 2013 as differences between lemma based lists, which provide a much more restrictive selection were developed. This can cause dramatic differences in what might be expected to be “known” to learn a word. Idioms and set expressions can certainly be expanded on with multi-word expressions (Martinez and Murphy, 2011) and academic collocations (Durrant, 2011), which none of the major lists include. West (2012) also helps shine a light on the range of information that might be included with a list and how this information goes far beyond a discrete list (i.e. parts of speech, collocations, etc.). The section of “other criteria” is perhaps the one that has been least dealt with as there are a wide range of considerations that might be only of consideration in specific contexts (e.g. alcohol related terms in Islamic contexts, the focus on formal language in academic writing).

Another area that is vital as it may cause significant differences in the frequency list is the source sub-corpus, corpus or corpora. There may be significant differences based on the source and genre of the texts, the age of the texts, and the country of origin, among others. While there are many potential examples of this, the most obvious can probably be seen in the GSL, which contains no modern technological terms like computer, laptop, or Internet.

D. Difficulties to overcome

Certainly, there are a number of challenges that frequency based lists have to overcome if they are to be used in an English language teaching and learning context. Some of these include the fact that in general, these lists were not designed as lists for language learners, they may include a great number of words related to more advanced concepts or grammatical structures, there is no agreement in the lists about covering items like days of the week, numbers, basic grammatical words like pronouns, and as mentioned previously, they generally do not include set expressions, phrasal verbs, collocations and formulaic language. Not least of all is the very pertinent question of how English language learners will interact with the list.

Looking away from frequency lists, other models exist that might provide other options to direct student vocabulary learning. One is to use the topic-based approach that is commonly found in many English language course books. This has the advantage of introducing vocabulary around a related subject, which allows an appropriate context, but may limit topics that lack a range of appropriate vocabulary. Another notable model that has recently become possible is using a learner’s corpus like the Cambridge Learners Corpus and the related English Vocabulary Profile (2013) which allows users to search by word, level, exam, nationality, and type of error. These could be utilized to help construct an alternative list to help guide learning that might more accurately reflect how “typical” learners acquire lexis. In addition, this website allows the mapping of vocabulary onto the Common European Framework (CEFR) with its established leveling, which could be used to help plan appropriate vocabulary development. Another valuable resource in this regard is the Word Family Framework (West, 2012), which allows the mapping of individual word forms and different lexemes onto the CEFR, which could certainly allow a very clear picture of expected learner acquisition of specific word forms.

3. RATIONALE FOR THE STUDY/RESEARCH QUESTIONS

While there has been extensive work in the development of frequency-based vocabulary lists for English language teaching and learning purposes, there appears to have been little research on which lists are being used and how they are being used. One area that seems to be completely lacking in research is how these lists are used in university level English language teaching programs. As such, this study aims to examine how these lists are used in independent, university level, foundation/ intensive English programs (IEPs) around the world and which lists are used. Independent programs were selected to avoid repeated similar responses from franchise programs, and the focus is on intensive university programs because these programs are usually extended programs and may focus on both EFL and EAP; therefore, a wider range of frequency lists for both general and academic programs might potentially be utilized.

4. METHOD AND METHODOLOGY

To answer these questions an exploratory research study was conducted. A survey was the main instrument in the study, and it was distributed to as many independent, university based IEPs as possible. The aim of the survey was to construct a general view of what kind of IEP programs exist (size, age, etc),



what lists are used, how these lists are used, and general perceptions of the potential value of these lists and how they might be improved for use in this general context. To achieve this, a primarily open-ended 17-question survey (Appendix A.) was set up using the online survey application Survey Monkey was constructed with four sections corresponding to these areas. A follow-up interview was also used to probe into the specifics of how the lists are used and to investigate any unsuspected responses.

Then an extensive, but probably not exhaustive list of more than 600 independent IEPs located at universities was compiled. To do this, the following steps were taken:

- 1) A list of all the countries in the world was printed.
- 2) Using the Google search engine, "list of universities" and the country name were entered, providing a list of the universities in each country, often with direct links available to each university's webpage.
- 3) The webpages for individual universities from each country were examined. For countries with readily evident programs (e.g. Australia, Canada), each university was examined for an IEP; typically with approximately 20 hours of class a week. For countries with no obvious programs, some of the larger universities were examined and then the search was ceased, so the search was not exhaustive for these countries. The one exception to this step was for the United States, where the website intensiveenglishusa.org, which provides information about intensive programs in every state, was utilized. Only independent programs affiliated with universities were used.
- 4) After the most suitable email address was located for each program, an email was sent with a link to the online survey. The order of preference was: curriculum/ assessment supervisor, academic coordinator, program director, head teacher, and other teaching faculty member.

The majority of programs identified were located in English speaking countries, with more than 350 programs at US universities, more than 100 at UK universities, and relatively large numbers in Australia and Canada. For the majority of university-based, independent, intensive English/ foundation programs in the U.S., U.K., Australia and Canada that were directed towards international students, direct contact information was typically readily available on the website, often for the English Language Center (ELC) or English Language Institute (ELI).

For programs in other countries that were aimed more at a domestic market, the contact information was often not immediately available. Very few programs were identified in Asia, Africa, Europe (outside of the U.K. and Ireland) and South America probably because there are either few English medium universities in the country or because there is a generally high level of English instruction in primary and secondary schools. In regards to countries where English was not the native language, the highest number of programs was located in Turkey, with more than 40 institutions with English preparation or 'hazirlik' programs existing at English medium universities. Additionally a number of GCC countries, such as the United Arab Emirates have more than just a few programs.

The final question in the survey asked if the respondents would be willing to be contacted for a follow up interview via Skype. These interviews were conducted and transcribed for a summary of the key points.

A total of 95 responses were received to the survey, representing a return rate of slightly less than 14.5 %. As can be seen in Table 1 below, the majority of the responses were from the U.S., which is not surprising considering the significantly larger number of programs identified there.

Table 1 = Country of Program, # of Programs Contacted, and Response Rate

Country	Programs Contacted	Responses
United States	358	57 (15.9%)
Canada	35	10 (28.6%)
UK	104	10 (9.6%)
United Arab Emirates	18	5 (27.8%)
Turkey	41	4 (9.8%)
Australia	39	3(7.7%)
Kuwait	3	3(100%)
Bahrain	9	1(11%)
China	1	1 (100%)
The Czech Republic	1	1 (100%)
Other Countries	48	0
Total	657	95 (14.4%)

The size of the programs ranged from having no full time teachers (two programs in the U.S.) to up to 200 (one program in the UAE), with an average of 21.7 full time teachers. For student numbers, there was a considerable range from a minimum of 12 students per year, up to a maximum of 2500, with an average of 492 students on an annual basis. There was also a



considerable range in the age of the programs ranging from a program in the US in its first semester to another one in the US, which was 102 years old, with an average age of 21.6 years.

Table 2 gives the position(s) of the individual responding to the survey. Indeed, a number of the respondents seemed to wear several hats at the institution.

Table 2 - Position(s) of individual completing the survey

Classroom teacher/ lecturer	60 (63.82%)
Administration	49 (52.13%)
Curriculum/Assessment Development	45 (47.87%)
Other (Director, Coordinator, Researcher)	10 (10.64%)

5. RESULTS

The awareness of word lists and attitudes towards the lists that are currently available is the first area that will be addressed, as seen in Tables 3-5. Table 3 indicates the self- assessed familiarity with the lists that are currently available, showing that more than 80% of the respondents feel they are somewhat or very familiar with the available lists.

Table 3: Familiarity with available frequency-based word lists

Very familiar	23 (25.8%)
Somewhat familiar	50 (56.2%)
Not very familiar	13 (14.6%)
Not familiar at all	3 (3.4%)
No response	5

Table 4 seems to reflect this awareness with an acknowledgement that these lists are either somewhat or very important with more than 90% of the responses in these categories.

Table 4: Beliefs about the importance of frequency-based word lists

Very important	35 (39.3%)
Somewhat important	46 (51.7%)
Not very important	6 (6.7%)
Not important at all	2 (3.4%)
No response	5

Following on this, Table 5 reflects the belief that a similar percentage of more than 90% believe that the existing lists are either somewhat or very suitable.

Table 5: Belief about suitability of existing frequency-based word lists

Very suitable	16 (18.2%)
Somewhat suitable	64 (72.7%)
Not very suitable	6 (6.8%)
Not suitable at all	2 (2.3%)
No response	6

Moving to the crux of the research, we find that 60% of the programs surveyed that responded to the relevant question reported formally using a vocabulary frequency list and 40% did not. Looking in more detail, we see that out of the 94 respondents, 48 indicated that they used a list, 32 indicated that they did not use a list and 14 did not leave a response.

Some of the 18 write-in comments included the following general themes:

- 1) Lists were used or suggested in some part of the program but were not mandatory across the program
 - We recommend AWL but don't insist. (5 similar responses)
 - Some courses use lists, but others don't. (3 similar responses)
- 2) "Lists" were developed from the course work or related materials.
 - We use lists based on course work. (2 similar responses)
 - ...we utilize vocabulary from "found" words in communicative contexts and academic articles and texts...
- 3) Lists were utilized in assessment.

In terms of the actual lists used, in Table 6 we see that of the programs surveyed, slightly more than half use the AWL, which far surpasses the use of any other list. It is utilized more than four times as much as the next individual list, the GSL. The third largest response category was 'Another list,' which included combinations of previously mentioned lists, lists based on Nation's 1,000 and 2,000 word lists, lists based on class needs, and other lists including Webster's 3000, Longman's 3000 and the Medical Word List.

Clearly, there is a good deal of variety in between institutions, but the use of the AWL predominates, perhaps due to the university and academic affiliation of all these programs.

Table 6: List Used

The Academic Word List	46 (51.7%)
No list used	30 (33.7%)
Another List	16 (18%)
The General Service List	11 (12.4%)
The Oxford 3000	6 (6.7%)
An institutionally developed list	6 (6.7%)
A list based on the British National Corpus	2 (2.3%)
A list based on the Brown Corpus	2 (2.3%)
The Word Frequency List of American English	2 (2.3%)
No response	5

In Table 7, we can see that these lists are primarily used by classroom teachers and by students, contrary to Nation's suggestions. However, this might simply be a commercial book focusing on the AWL. Only roughly a quarter of the respondents use these lists for curriculum and assessment development, and these largely overlap, so this seems to be a relatively uncommon use for these lists. Interestingly, out of the 20 that use it for assessment development, 8 are in the Middle East or Turkey out of the 12 programs in the region that responded.

Table 7: Who uses the list

Classroom teachers	51 (59.3%)
Students	44 (51.2%)
Material Writers	24 (27.9%)
Testing/ Assessment Developers	20 (23.3%)
We don't use a list	29 (33.7%)
No response	8

Table 8 shows how the lists are used with students, and there seems to be quite a range of responses with none that predominate. Some of the categories like "utilized in the classroom" in retrospect seem to lack specifics as they could be used with a list discussed with students or as part of a commercial course book.

Table 8: How the list is used with students

Utilized in the classroom	32 (37.2%)
Provided as a list with practice materials for self-study	26 (30.2%)
Provided as a discrete list for self-study	25 (29.1%)
Not explicitly provided but incorporated into materials	17 (19.8%)
We don't use a list	29 (33.7%)
No response	8

The first open-ended question asked if any one list was more suitable than others. Out of the 54 who responded to this question, 22 replied with no difference, or that they were not familiar enough to distinguish. 13 responded that the AWL or a combination of the AWL and GSL were preferred.

Some of the additional responses included the following themes:

- 1) Using different lists for different purposes
 - mostly indicating the AWL was used for higher level students and another list for lower level students
 - the need for a variety of lists based on level and context
- 2) One list was better for a specific reason.
 - Oxford 3000 is more useful than most as it is based on a learner's corpus as well as the BNC.
 - Webster's 3000 words incorporate words from both the GSL and the AWL
 - The AWL ... seems to hold more validity because it is more immediately practical in the higher levels
 - The new GSL, as it is a basis to start, and provides wide coverage.

The responses generally seem to work around a more pragmatic and flexible approach instead of adhering to a single list. Clearly, although the AWL was the most frequently used list, it is not a one-size-fits-all approach with some different lists being used at different levels.

When asked if there was anything lacking in the current lists, out of the 51 respondents to this question, 26 respondents, or slightly more than half, indicated that there wasn't or that they didn't know. For those who indicated that something was lacking, the responses seemed to fit in several categories:

- 1) something was lacking in the actual "words" included in the list
 - basic vocab like the days, months, colors, etc. / common phrases, idiomatic expressions / lexical chunks / phrasal verbs, which are high frequency but not specifically included
- 2) more specific academic lists were needed
 - more discipline specific academic word list
 - subject-specific references for Higher Education (3 similar responses)
 - Even some "academic" lists are not specialized in university vocabulary.
- 3) additional information was required along with the lists
 - Not knowing which definition of the word /or the use of multiple definitions of the word affecting its position in the list
- 4) a comment on the list not being appropriate for use with learners for some reason.
 - for learners, the list lacks any sort of "usefulness" discrimination



- an understanding of the learner's first language/ Learners with Latin based vocab share more English words than say Chinese learners
- The words in word lists are not necessarily the words our learners need.

When asked what needs were not being served by the current lists, only 44 participants responded, and of these, 17 had a response of N/A, don't know or something similar. Of the remaining 27 responses, two general categories emerged:

- 1) Lacking a context to introduce the vocabulary (4 similar responses)
- 2) A mismatch between the lists and the specific needs of the students (e.g. everyday language, current class related vocab, BASIC or new beginner vocabulary needs, a narrow use of the AWL will leave students at a loss on campus, World Englishes)

A few other interesting answers included the following:

- differences in academic genres across different faculties
- manageability (the AWL is very long); well-rounded coverage
- collocations/ words with different meanings appearing on the list

When asked if it would benefit their institution to develop an institutionally specific list, out of the 58 responses, exactly half responded that it would not, giving responses that generally fit in the following categories:

- 1) The available lists are suitable. (e.g. the existing lists have been developed with a great deal of research; we are preparing our students for academic study anywhere in the world; I don't think our university's culture is drastically different than that of other institutions).
- 2) They don't need an institutional list. (e.g. they need to learn words in context; it would be too top down, without relevance to students; our teachers are experienced enough to bring the necessary resources; our learner demographics and needs vary too much to make this practical).
- 3) A general frequency list wouldn't be of use, but discipline specific ones for advanced students would. (e.g. the academic lists may be helpful to be institution specific if they were also department/field specific; it would be redundant, unless it were for our very advanced students in the pre-university academy who need to master words specific to their own field of study; our students would benefit from more ESP-related

lists having to do specifically with various fields of engineering, biology, and business).

Out of the other 29 who indicated they would be interested in such a list, six responses indicated that they already had some form of program list(s) at their institution, showing that in slightly more than 5% of the programs that responded, work has been done starting with one of these frequency lists. Again, some of the general categories for these responses included:

- 1) It would help provide structure for lexis teaching and learning in our program. (e.g. it would help our new and temporary hires, of which we have many; to help build student vocabulary and better prepare them for the rigors of academic study; it helps focus vocabulary learning; if specific majors each had a core list, this could be helpful).
- 2) It would aid curriculum/ assessment development (e.g. could be used as a placement/exit criterion; it would systematically identify words, which could then be recycled in reading, writing, and listening materials, based on core texts being used at each level which also provide lots of repetition and practice material).
- 3) A specific list would be more appropriate for the students (e.g. because our learners lives (what they do, what they need or what they expect) are different from those of other people in English speaking countries).

The penultimate question inquired what resources would be necessary to transform a vocabulary list into something appropriate to direct student vocabulary learning. The 56 responses to this mostly fell into the following two broad categories:

- 1) Sample texts/ materials using and revising the vocabulary (e.g. reading passages and sentences that use and re-use - something that is not done a lot in vocabulary texts; extensive readers, exercises, games, conversations, listening scripts etc.; a access to online programs (AWL exercises are very useful; many students have used them for projects); a web site with lots of fun, interesting, effective exercises to maximize exposure; "placement" tests that show students their level of mastery of a given list).
- 2) Materials to simplify/ automatize its use and/or help provide a rationale. (e.g. a short, clear convincing rationale for its use; links from words to levels-based corpus data, and a schedule of reminders taking students back repeatedly to words learned; auto-instructional



materials that do not require teacher oversight; close collaboration with faculty lecturers).

The final question invited the respondents to participate in a follow up interview via Skype, and although 24 of the respondents indicated that they would be interested in participating, only 4 completed the roughly 15- minute, semi-structured interview.

Of these four, two were located in Canada, one in the USA and 1 in the Middle East. All four of these programs used a list, and each used the AWL in some fashion. Below is a summary of the key information from each interview.

Interview 1 – from Canada

A mid-sized EAP program preparing primarily Chinese, Brazilian and Japanese students to study at the institution. They use primarily the GSL and AWL, which are incorporated into their textbooks, with one commercial text focused on the Academic Word List. They indicated that the inclusion of research driven data like the AWL helped drive the purchase of commercial textbooks. They liked the fact that the use of these lists helped expose the learners to key vocabulary, but did raise concerns about the amount of vocabulary the students were exposed to, as well as raising the question of what was really expected for students to “know” a word. They were aware of the new-GSL and were looking forward to seeing it used in the future.

Interview 2- from the Middle East

A large English language foundation program almost entirely for female national students (Arabic L1), in which 80% of the students entering the university need to go through at least part of the 6-level program to prepare to study almost entirely in English at the university. They modified several lists for their own use which include one based on the BNC, one based on the first 3 sub-lists from the AWL and a list derived from the course books (with cross-over between the 3), but are currently working to revise their approach and list largely because the BNC did not end up being suitable because of the range of some of the lexis (politics, local government, etc.). They are debating using a list that includes material from a learner’s corpus. They believe very strongly in the use of a list with a frequency component and use it to help direct and standardize vocabulary outcomes and to help grade the lexis appropriately in assessments and supplementary materials. The main challenge is the sheer volume of vocabulary necessary for students to acquire and the difficulty of having a frequency based list that was separate from what the course books

taught. They indicated that a pure frequency approach had not been successful in their experience and that other sources, such as Cambridge’s Learner’s Corpus should be consulted to help create a more learner appropriate approach.

Interview 3- from Canada

This 3-level EAP program prepares the majority of their students for undergraduate study at the institution. They use the GSL for lower level students and the AWL for higher-level students and chose some commercial materials focused on the AWL. They feel that the use of these lists makes students and teachers aware of key, frequent vocabulary and that this provides “the most bang for their buck.” They highlighted the difficulties in going through vocabulary outside of a set context and that it was more difficult for students to acquire the vocabulary by working through a list or a book focused on a list. They also use Cobb’s Compleat Lexical Tutor to help analyze some of their supplementary texts and also use assessments with these lists to make sure lexis is appropriate.

Interview 4 – from the United States

This intensive, outcomes-driven, English program with individual skill classes (Reading, Writing, etc.) prepares primarily Chinese and Saudi students to study at their institution. Their 7-level program ranges from complete beginners to upper intermediate level students. They use the AWL extensively at their upper levels with a commercial book and in-house materials to support this. For their lower levels they use a sub-list they developed from COCA’s academic corpus. They expressed concern about the length of their word lists, especially at the lower levels. In the past they had used the GSL, but moved away from this, as it seemed to have a good deal of lexis that was less academically oriented. One difficulty was the need to develop support materials for their in-house developed lists. Another was the potentially list-driven approach that the lack of support materials might cause. However, they felt that the use of these lists benefited the program by providing a standard across each level, that it helped in the creation of appropriate departmental final exams, that it established a list of words that learners were expected to spell correctly and that it helped in the selection of lexis to focus on in class. When they created their list, to be used prior to the AWL, they felt a focus on the academic genre was important, and they were flexible in reorganizing the frequency list so that it matched their level grammar outcomes or created semantic groups to help organize the lexis in more logical groups based on lemmas for the lower levels.



6. DISCUSSION

One of the first noticeable points is that there does not appear to be a homogenous or standard foundation/intensive English university program. Based on the wide swath of responses, it is clear that these programs range from ones in which students have difficulty spelling basic words or are unaware of some of the most frequent words to those who have high-level, subject-specific lexical needs. As such, the range of lists used seems expected, and as the purpose of most of these programs is to prepare students to study at the university level, the dominance of the AWL is not surprising. Another reason for the popularity of this list might be the ready access of a range of commercial materials that are based on it. However, this is despite the fact that the AWL is based on top of the GSL, which is based on corpus work up to 100 years old, which seems to be the key reason why these new lists have emerged.

Additionally, while more than 90% of the respondents believe these lists are either somewhat or very important, the actual formal use in programs does not seem to support this as only 60% acknowledge their use. There may be a number of underlying reasons for this including teacher experience, inappropriate lists, a lack of resources or lack of time in the program.

While this study shines an initial light on which lists are currently used by programs, a detailed examination as to specifically how they are utilized is needed. While it is evident in the sample that classroom teachers and students use the lists more than curriculum and assessment, the scope of this research did not cover specific details about how they are used in class. Are they used on a regular basis or as a one-off supplement? Are they assessed formally or informally in class to measure student vocabulary growth? Are there commercial products that appear to push the use of some lists more than others? All of these questions are certainly of interest as well as a deeper understanding of how the lists are being at a program wide level – for assessment or materials development, for example, and whether there are practices that might be adopted by a wider range of programs.

The small percentage of programs that have compiled their own lists seem to echo similar concerns about the use of these lists in their programs and dealing with the sheer number of items that need to be covered along with the need to create quality materials to help teach the lexis. A dissatisfaction emerges in

some cases with using a purely frequency based approach in a decontextualized list. Despite this, the participants do seem to feel that the value of these lists outweighs these difficulties.

Another finding that seems to be of note is the apparent lack of awareness of higher level, discipline-specific frequency lists for science, engineering and business among others. A growing number of these lists exist, but many institutions did not seem to be aware of this or found the existing lists lacking.

An obvious limitation of this study is the possibility that the respondents might be more interested in the use of these lists and thus, more likely to complete the survey. This might mean that the sample might have a higher percentage of programs using these lists that in the broader population; however, as an exploratory study, this study is not expected to be predictive of the use in the population at large.

Ideas for further research on the subject include: repeating the survey with different EFL/ESL teaching groups like language schools and primary/secondary schools, doing more in-depth inquiries into the small number of programs that use the lists for curriculum and assessment development, focusing specifically on how the lists are used in class with students, surveying commercially available ELT materials to see which lists are most prevalent, and repeating the survey in a few years to see if there is a shift towards the use of these new lists.

7. CONCLUSIONS

This study showed that there is no clear standard in regards to the use of high frequency list in university IEP programs. In terms of lists used, Coxhead's Academic Word List has a slight majority share, which may correspond directly with the academic nature of these intensive programs. There was a strong indication that these lists are important for use in ELT, but there was a wide range of opinion on how they should be used in programs, perhaps reflecting the variety of programs surveyed. Interestingly, there seems to be a call for frequency driven discipline specific lists to help prepare students to study for specific subjects, which does not seem to appear in the literature. It is clear that these lists have some sort of a role in the majority of the programs surveyed, and that this may be even larger when publications that explicitly use these lists are taken into account.



ACKNOWLEDGMENT

I would like to thank Dr. Philip Durrant for his insightful comments and feedback throughout the research process.

I would also like to thank my wife Tuba for her support and encouragement throughout the extended period of my research.

REFERENCES

- Billuroğlu, A. & Neufeld, S. (2007) Billuroğlu & Neufeld List 2709: The Essence of English (Fourth edition). Lefkosa, Cyprus. Rustem Publishing.
- Bogaards, P. & Laufer, B. (2004) Vocabulary in a Second Language: Selection, acquisition and testing. Philadelphia, John Benjamins, B.V.
- Brezina, V. & Gablasova, D. (2013). Is there a Core General Vocabulary? Introducing the New General Service List. *Applied Linguistics*, 2013, 1-23. Retrieved from <http://aplii.oxfordjournals.org/content/early/2013/08/25/applin.amt018.full.pdf+html>
- Browne, C., Culligan, B & Phillips, J. (2013) A New General Service List, Retrieved from <http://www.newgeneralservicelist.org>
- Cambridge University Press, English Vocabulary Profile, Available at <http://www.englishprofile.org> . Accessed on December 28, 2013.
- Cobb, T. (2013). Compleat Lexical Tutor, available at <http://www.lextutor.ca/>. Accessed 1 October 2013.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34, 213-238.
- Coxhead, A. (2011). The Academic Word List 10 Years On: Research and Teaching Implications. *TESOL Quarterly*, 45, 355-62.
- Davies, M. (2012). Corpus of Contemporary American English (1990-2012), Retrieved from <http://corpus.byu.edu/coca/>.
- Durrant, P. (2011). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, 28, 157-169.
- English Profile: CEFR for English. Available at <http://www.englishprofile.org>.
- Espinosa, S. (2003). Vocabulary: Reviewing Trends in EFL/ESL Instruction and Testing. *Odissea*, 4, 91-112.
- Gardner, D. & Davies, M. (2013). A New Academic Vocabulary List. *Applied Linguistics Advanced Access*:1-24.:
- Good, H. (1928). Three Lists of English Words. *Educational Research Bulletin*, 7, 273-275.
- Goulden, R., Nation, P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-363.
- Green, T. (2011). The Password Test – Design, Development and Reliability. Retrieved from: http://www.englishlanguagetesting.co.uk/uploads/The-Password-Test-Design-Development-and-Reliability_5.pdf
- Howatt, A. (1984). *A History of English Language Teaching*. Hong Kong, Oxford University Press.
- Institute of International Education (2011) Intensive English USA, Accessed from : <http://www.intensiveenglishusa.org/index.asp>
- Laufer, B. & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual and “bilingualised” dictionaries in the comprehension of new words. *The Modern Language Journal*. 81, 189-196.
- Lorge, I., & Thorndike, E. L. (1963). *The Teacher's Word Book of 30 000 Words*. Columbia University Press.
- Martinez, R. & Murphy, V. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, 45, 267-290.
- McArthur, T. (1998). *Living words: language, lexicography and the knowledge revolution*. Exeter: University of Exeter Press.
- McKee, P. (1937). Word Lists and Vocabulary Difficulty in Reading Matter. *The Elementary English Review*, 14, 241-245
- Milton, J. & Treffers- Daller, J. (2013). Vocabulary size revisited: the link between vocabulary size and academic achievement *Applied Linguistics Review*, 4, 151-172.
- Nation, P. (1997). Vocabulary size, text coverage, and word lists. in Schmitt, McCarthy. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge, Cambridge University Press, 6-19.
- Pearson, Research, <http://pearsonpte.com/research/Pages/home.aspx>. Accessed December 29, 2013
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge, Cambridge University Press
- West, M. (1953). *A General Service List of English Words: with Semantic Frequencies and a Supplementary Word-List for the Writing of Popular Science and Technology*. Longman
- West, R. (2012). *Work Family Framework*. <http://www.teachingenglish.org.uk/article/word-family-framework>. Accessed February 26, 2014.



APPENDIX A

Questionnaire on Use of Frequency-based Vocabulary Lists in IEP/ Foundation Programs**Demographics**

- 1) What country are you located in?
- 2) Approximately how many teachers work full-time in your program on an annual basis?
- 3) Roughly how many students are enrolled in your program on an annual basis?
- 4) How old is the IEP/foundation program at your institution?
- 5) What is your position at your institution?
 - Classroom teacher/ lecturer
 - Curriculum/Assessment development
 - Administration
 - Other _____

Questions on use of Word Lists at Your Institution

- 6) Does your institution utilize a frequency based word list to help learners prioritize suitable lexis?
 - Yes, we use a specific list
 - No, we don't use a list.
 - Other _____
- 7) (If yes) Which list do you use?
 - A list based on the British National Corpus (BNC)
 - A list based on the Brown Corpus
 - The Oxford 3000
 - The JACET list
 - The Academic Word List (AWL)
 - Word Frequency List of American English
 - The General Service List (GSL)
 - An institutionally developed list _____
 - Another list _____
- 8) Who uses the list in your program? (check all that apply)
 - Classroom teachers
 - Students
 - Materials Writers/ Curriculum Developers
 - Testing/ Assessment Office
 - Other _____



- 9) (If used by students) How is the list used with students?
- Provided to students for self-study as a discrete list
 - Provided as a list with practice activities.
 - Used in the classroom
 - Not explicitly provided but incorporated into class materials
 - Other _____

Attitudinal Questions

- 10) How familiar are you with the frequency based vocabulary lists that are currently available?
- Very familiar
 - Somewhat familiar
 - Not very familiar
 - Not familiar at all
- 10) How important do you feel that frequency based vocabulary lists are directing student vocabulary learning?
- Very important
 - Somewhat Important
 - Not very important
 - Not important at all
- 11) How suitable do you feel that the existing lists are for directing student vocabulary learning?
- Very suitable
 - Somewhat suitable
 - Not very suitable
 - Not suitable at all

Open Ended Questions

- 12) Do you feel that any one list is noticeably better than another? If so, which list and why?
- 13) Is there anything specific you feel is lacking from currently available lists? What is it?
- 14) What needs are not being served by the current lists?
- 15) Do you think it would benefit the students at your institution to develop an institutionally specific list (if your institution hasn't done so yet)? Why?
- 16) What resources are optimally needed to transform a vocabulary list into something that is appropriate to direct student vocabulary learning?
- 17) Would you be willing to be contacted for a more in-depth discussion on the use of word lists in your program?

