All Works

2-25-2013

The use and development of academic vocabulary in L2 writing: a longitudinal investigation

John Turlik
Zayed University

Follow this and additional works at: https://zuscholars.zu.ac.ae/works

Part of the Education Commons

The use and development of academic vocabulary in second language writing: a longitudinal investigation

John Turlik

Zayed University, UAE

Introduction

Students using English as a second language (L2) entering foundation programmes in Gulf universities tend to have one common problem: how to improve their language proficiency. During the foundation courses, their study is dedicated to this end but as they move into the post-foundation stages, there is an increase in pressure on them as they must not only consolidate what they have learnt but also take on the additional burdens of academic study in their chosen disciplines. One aspect of the pursuit of increased proficiency is vocabulary acquisition, which is made all the more difficult, post-foundation, by the challenges of the reading texts prescribed and the level of writing required. Students therefore face substantial challenges: to continue to increase proficiency in general English, to expand their general and academic vocabulary and, at the same time, to apply their developing knowledge of English in the pursuit of their chosen discipline.

Laufer & Nation (1995), on the basis of wordlists developed by Xue and Nation (1984), developed their Lexical Frequency Profile (LFP) on the premise that an L2 user's richness of vocabulary is reflected in the number of words used from four frequency bands (derived from typical texts such as novels or academic textbooks). The least rich vocabulary uses mainly words from the first 1,000 words ('K1': high frequency, general service vocabulary such as *the*, *in*, *number*, *words* or *used*), while a richer vocabulary uses more from the next most frequent band ('K2', examples being *reflected, frequently* or *upper*), until the richest vocabulary uses an increasing number of words from the University Word List (UWL), such as *interval, implemented* or *assessment*). The UWL comprises words which do not appear in the other lists and are more common in university/upper secondary contexts than in everyday contexts. 70-80% of a typical text in an academic textbook or newspaper consists of general words from the K1 list, with perhaps 5-12% from the K2 list and a similar range from the UWL (the relative proportions depending on how 'academic' the text is).

Coxhead (2000) refined the University Word List (UWL) by condensing it somewhat and dividing it into ten sub-lists to form the Academic Word List (AWL). Each sub-list comprises 60 'word families', each of which is listed by its headword (or baseword). Thus, *analysis* is the first (base)word in Sub-list 1; its word family includes not only *analysis* but also *analyze, analytically* etc., and this family of words are among the most frequent non-general words in academic texts. Sub-list 10 (which, incidentally, contains only 30 headwords) contains the lowest frequency set of academic words. As a guide, vocabulary from sub-list 1 (e.g. *analysis, approach, area*, *assessment*) account for approximately one third of the academic words in a typical text; sub-lists 2 (e.g. *achieve, administration*, *affect*) and 3 (e.g. *alternative, circumstances, compensation*) account for somewhat less than this, while Sub-lists 4 to10 (respective examples for sub-lists 4 to 9 are *access, academic, abstract, adaptation, abandon* and *accommodation*) include words that are infrequent to the extent that they would be unlikely to be learned simply through incidental encounters in reading. In other words, they would, in all probability, need to be taught. Examples of words from (the shorter) sub-list 10 are *assembly, collapse,*

forthcoming, inclination and *integrity.* Given the coverage provided by lists 1, 2 and 3, therefore, sub-lists 5 or 6 cannot be considered the 'half-way mark' in acquisition of the words from the total AWL.

Literature review

In general, development of vocabulary over time might mean (a) the acquisition of more high frequency words (b) the acquisition of more low frequency words or (c) both. Referring to the AWL, Nation (2001) advocates giving most attention to sub-list 1 and "where possible to deal with these words before moving on to other academic words" (p. 386). Second language English students preparing for studies at undergraduate level require an extensive, sound, working knowledge of general (very high frequency) and academic (lower frequency) vocabulary in order to deal with reading texts and required essays. As Nation (2001:197) points out, "learners need the opportunity to use [academic vocabulary] in meaning-focused output activities; that is in speaking and writing in academic contexts". During a foundation course, the expectation is that this output will reflect lexical development stemming from increased lexical knowledge, which clearly puts an onus on both the teacher, to pursue the most effective methods of teaching, and the students, to consciously strive to learn and make use of newly acquired academic vocabulary.

Paquot (2010), however, asks whether there really is an 'academic vocabulary'. Referring to Coxhead's (2000) AWL, he cites Hyland & Tse (2007) who point out that academic disciplines (from which the AWL was derived) may use a word in a particular context, thereby giving that word different specificities. The word *process*, for example, may more likely be taken as a noun by engineering and science students than by students of social science; another example given is *strategy* and how this might have different applications in the business world (*marketing strategy*), in education (*learning strategy*) and in sociology (*coping strategy*). Read (2003) provides a similar example of what he terms conceptualizing a word, citing the differences in the various meanings of *fresh* as in *fresh bread, fresh ideas, fresh supplies* and a *fresh breeze*.

Further evidence of the issue of specificity in vocabulary use in another discipline is provided by Diaz-Gilbert (2004). She conducted an investigation into what she termed 'word knowledge' of pharmacy students whose first language was not English, with the intention of identifying problematic vocabulary, both general and health/pharmacy-related. She also wanted to determine the linguistic obstacles that hinder word knowledge and defined this knowledge by the respondents' capacity to provide "acceptable"[1] definitions or synonyms (sometimes in a context) for words in five categories: patient state, patient complaint, treatment/behaviour, general and pharmacy-related.

From the pharmacy department of her university, 105 words were selected but these were not exclusively related to pharmacy as they also included general service and non-science related items. The respondents were twenty-five students from eight language groups who had lived in the United States from 2.9 to 19 years with an average time of 7 years. It was found that twenty-five respondents provided acceptable definitions or synonyms for 52% of the words. Broken down into the groups, the five pre-professional/first year students provided acceptable definitions or synonyms for 37% of the words; the nine second-professional-year students, an average of 53.8%; the seven third-professional-year,[3] 76%; and the four fourth-professional-year, an average of 56.3%. Surprisingly, perhaps, the researcher does not comment on the relatively poorer performance of the final year group compared to the preceding group but the use of synonyms as a criterion for knowing a word in a language is criticised generally by Waring (2002) on the grounds that some languages (e.g. German) do not have a 'rich store' of them, rendering the synonym requirement language specific. It is not clear if he considers this might create difficulties in, for example, a German speaker being asked to provide synonyms in English.

One conclusion reached by Diaz-Gilbert (2004) is that students may have a false sense of word knowledge. While some words that related to patient's state such as *perkier*, *groggy* and *lethargic* (which, incidentally, do not appear in K1, K2 or the AWL) were either unknown or unfamiliar to "the majority of the respondents"[2], while others demonstrated weak knowledge of words that they believed they knew, such as *drowsy*, *crampy* and *anxious*. Some also responded with unacceptable or problematic definitions or synonyms, even though they had indicated they knew the words, and some also showed a weak knowledge of words or phrases related to patient complaints, treatment and behaviour. As she points out, a false sense of word knowledge has implications for reading, listening and note-taking and if words are unfamiliar or only *thought* to be known, both anxiety and errors in production are likely, especially if the production takes the form of an examination. One important point to note from Diaz-Gilbert's work is that acquiring a high level of knowledge in a *particular, specialised* sphere does not necessarily mean high *general* language proficiency.

Read (2003) makes reference to "familiar words that occur with high frequency" (p. 230) in a particular discipline and quotes Flowerdew's (1993) conclusion that words deemed specific to a discipline (technical vocabulary) need to be learnt in association with words which may be used together with them (sub-technical or general vocabulary), a reminder of Diaz-Gilbert's (2004) conclusions. The example of a *fresh* breeze for a student of meteorology and a f*resh* sample of blood for a medical student certainly appear to confirm the importance of "acquiring the particular meanings of sub-technical terms in their fields of study" (Flowerdew, 1993, p. 236). Paquot (2010) includes terms such as *fresh* in his Academic Keyword List, on the premise that they are *potential* academic words, while Stubbs (1986) points to the inadequacy of frequency lists on their own in distinguishing words for basic (or, as he terms it, 'nuclear') vocabulary. Perhaps academic vocabulary development could and should include increased knowledge of how a single word within a word family is used or, in other words, which items of nuclear vocabulary support the academic.

With reference to teaching vocabulary, Bachman & Palmer (1996) distinguish between language knowledge (knowledge of meaning and grammar) and *communicative competence*: being able to use a word in an accepted context – or, perhaps as Read (2003) succinctly phrases it "*precision* of knowledge".

Nation (2001) points out that there is a large number of words that occur only once in any given text, compared with "a very high frequency of the few most frequent words" (p. 9): in other words, frequency of use suggests that a word in an earlier sub-list (1 or 2) should appear *more often* in a sample of output than a word in a later sub-list (8, 9 or 10). Nation advocates that after having "mastered the 2000 - 3000 words of general usefulness in English" (p. 187), vocabulary learning should be directed to more specialized areas, one of which is academic vocabulary. This is defined as the vocabulary most frequently found "in academic learning such as in school or university" (Daller et al., 2007) and is distinct from technical vocabulary and low-frequency words (Nation, 2001).

The preparation of the AWL was based upon a corpus of 3,500,000 words; and in order to identify academic vocabulary used across a range of subject areas this corpus was grouped under four headings (Arts, Science, Law and Commerce), with each group consisting of seven academic disciplines, such as psychology, mathematics and history (Nation, 2001). The frequency of the words in these sub-groups was compared with their frequency in a corpus of novels (also 3,500,000 words), thereby confirming which were truly academic. West (1953) on the other hand questioned whether range and frequency alone were adequate criteria for allocating a word to a particular word list, suggesting that ease (or difficulty) of learning ought to be considered. Nation (personal communication) endorsed this view, adding the need to also consider ease or difficulty of teaching a particular item.

As teachers, we constantly seek resources that will advance our students' learning and it seems that word lists have a place in vocabulary acquisition and development; however, in the final analysis, how we use frequency lists and how we explain the vocabulary of our subject needs to be considered very carefully since the lists on their own can only be a tool. It would be our wish that over time our students increase the number of words at their disposal *and* increase proficiency in their use but, as can be seen from the foregoing, there are potential areas of difficulty. Learning is a complex activity and learning vocabulary and learning how to use vocabulary in another language are very much part of that complexity. We hope, perhaps, that our students take incremental steps in acquiring proficiency in language and that learning is a process that results in linear advancement….. but does it? Do our students begin their majors with sufficient, perhaps adequate general and academic vocabulary? Do they continue to acquire academic vocabulary by building on a preponderance of words from sub-list 1, then with appropriate numbers from the lower-frequency sub-lists? Does any increase in lexical knowledge include the ability to call upon suitable synonyms or collocations, as well as on other words in the same word family?

Methods

The students

The students were selected by convenience from the foundation programme at Zayed University – one of a number of institutions in the United Arab Emirates providing higher education at the time of this study for ladies only, so all subjects are female. To be included in the study, all participants were required to give their consent for their writing to be copied and used and all had to have ultimately completed the examinations for the final level of the foundation programme (level 8).

The essays

340 formal examination essays were collected, by 42 students and written at regular intervals between March 2003 and June 2005, yielding a corpus of 94,623 words. The mean length of the 340 essays was 270 tokens (words on the page). Of the complete set of 340 scripts, the first 41 were essays from CEPA (Common English Proficiency Examination, a mandatory test for those wishing to enter tertiary education) and 299 were end-of-term formal assessments given at the end of ten-week terms in the Foundation Programme. There were four of these essays per year, all written under examination conditions and, during the course, students would write in response to a prescribed genre, starting with 'descriptive', then 'cause and effect', followed by 'problem/solution' and, in the final two levels, 'position'.

In total, there are *complete* sets for 31 students as eleven essays are unaccounted for and hence not part of the corpus. Even though complete sets were available until the final level of the foundation programme, essays by students 6 and 9 eventually had to be removed, as their final course examination essays were found to have been destroyed. As a result, although there are forty-two participants, student numbers run from 1 to 44. One student completed high school a year before CEPA was implemented and started at the university after a 'gap' year. Her first essay is therefore the first essay of her FP and, where relevant. is referred to as Essay 2, since she has no (CEPA) Essay 1.

The difference in number of essays from students requires further explanation, since on entry into the university students are placed into different levels of the foundation programme. It would be unrealistic to find 42 students who started at the same level, progressed at the same rate and ended the course in Level 8 at exactly the same time, and as a result some of the students reached Level 8 within the two

years. As mentioned, the requirement for inclusion in the study was that a student must have completed the final level (level 8); whether she passed or failed was not relevant. For writing development, the view was taken that if a student started at a relatively high or low level but took two years to complete and exit at Level 8, having repeated each level twice, or if a student started at a low level and worked her way through to Level 8, level by level, writing was still being practised. Furthermore, a student might start at Level 5 and complete the whole programme within a year (a term in each level), providing information on how her writing changed in the shorter period of time. For the present study, however, it is important to remember that at any given stage of the two-year programme, *the students would all have received the same number of teaching hours in that programme*.

Procedures and key concepts

Since the essays were hand-written, they had to be transcribed, verbatim, into computer-readable text and then edited according to a set of procedures to ensure consistency. These procedures included, for example, the correcting of spelling on the basis that spelling is part of the learning process and a word used in the correct context but spelt incorrectly should be acceptable. The essays were analyzed using Nation's RANGE software, which identified, *inter alia,* the number of types (different words) and the number of tokens (total words in the essay) from the different word (sub-)lists.

I report here on *AWL types* as evidence of any use of an individual word from the sub-lists, and *AWL tokens*, as evidence of frequency of use. In other words, 'types' are words which may appear once or more than once, whereas 'tokens' refers to the total number of words on the page. Thus for example, the sentence *A cat sat on the mat* comprises six tokens and six different types, but *one cat is like another cat* consists of six tokens but only five types, since *cat* has been repeated. Note also that the present study interprets different members of a word family as different types: *analyze* and *analysis* are two different types (although the AWL groups them under the same family) but *analyze* and *analyzed* are considered the same type.

According to the assumptions of many writers, in a piece of writing, and hence across all essays, words used from the 'earlier' sub-lists (sub-lists 1, 2, 3 etc.) should be greater in number and appear more frequently than words in the higher sub-lists (Sub-lists 8, 9 and 10). There is of course, also the issue of an individual student's use of a word compared to its use by the students as a whole: if an *essay group*[3] shows a statistical preference for words from one sub-list, it does not necessarily follow that every student's essay in that group shows the same preference.

Analysis of data

Repeated measures ANOVA

Once the number of types and tokens was known for each participant, a repeated measures ANOVA was conducted to determine whether there had been an overall increase across the group in the use of academic vocabulary, compared with previous occasions. It should be noted that the number of participants dropped from forty-two to thirty-nine to twenty-nine between the sixth and eighth essays as students exited the programme, so, to ensure statistical integrity, the ANOVA uses only the range from the first essay to the seventh – March 2003 to March 2005. All other computations make use of the full corpus.

Measurement of academic vocabulary (Students)

Measurements of academic vocabulary in each essay were taken to assess acquisition and use of vocabulary by sub-lists and to determine whether that use was incremental.

Measurement of academic vocabulary (Native speaker)

Since the Word Lists are predicated on the academic vocabulary used in the writings of university/professional native speakers, a sample of such writing was also measured and its sub-list vocabulary determined.

Results

Change in use of AWL types, by essay group

A repeated measures ANOVA shows that the percentage of AWL types increased significantly across the essays, over two years (Greenhouse-Geisser correction applied, $\varepsilon = .68$):

Advanced Types (AWL, no outlier): $F(4.052, 153.961) = 16.783$ (p < .001)

A high value was obtained by one student, in Essay 4 (16.2%) and Figure 1 illustrates the resulting curve if this value is regarded as an outlier and excluded. The rationale for removing the value (16.2%) was based on the absence of any other similarly high value in any essay among the other 339. From her questionnaire, it was found that the student concerned attended a private, English-medium school and her family speak English, with parents educated to tertiary level and employed by the government. Her other essays, however, while of similar fluency, did not capture such high AWL scores (maximum 4.8%) suggesting that 16.2% was a chance occurrence. In addition, both lexical and holistic ratings were high for this essay (8.0), and the raters expressed concern that it might have been written, in part at least, from memory.

Given the significance shown in the repeated measures ANOVA, Figure 1 shows a statistically significant trend (R^2/R^2 Adjusted: Log: .176, .173. Linear: .179, .176]), with the linear regression line indicating a little more of the variance than the logarithmic (curved) line.
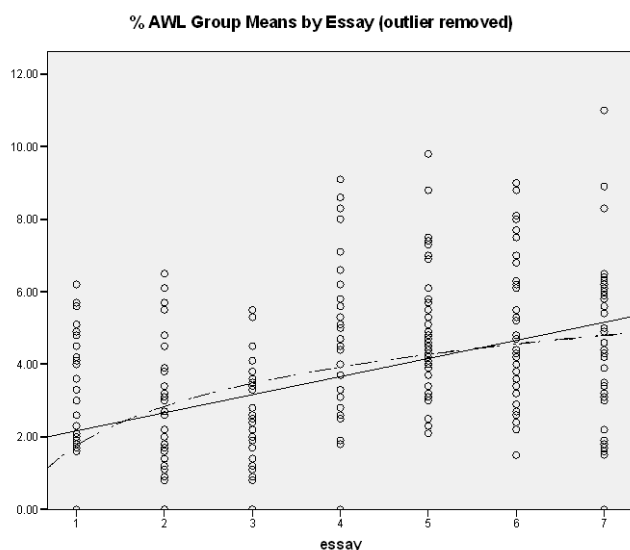

Figure 1: % AWL: group means by essay (outlier removed).

From the group perspective, therefore, it can be seen that there is progress over a period of two years in the use of academic vocabulary from the CEPA essays in March 2003 (Essay 1) to Essay 7 in March 2005, a conclusion supported by ratings of the essays (not included in the present paper) by two IELTS trained examiners.

Use of AWL types by sub-list (all essays)

Figure 2 shows the distribution of AWL types, by sub-list, across the 342 essays. To clarify, the largest group of academic words (505 different types) are drawn from sub-list 2, followed by sub-list 1 (309 types), with the lowest contribution coming from sub-lists 8 and 10.
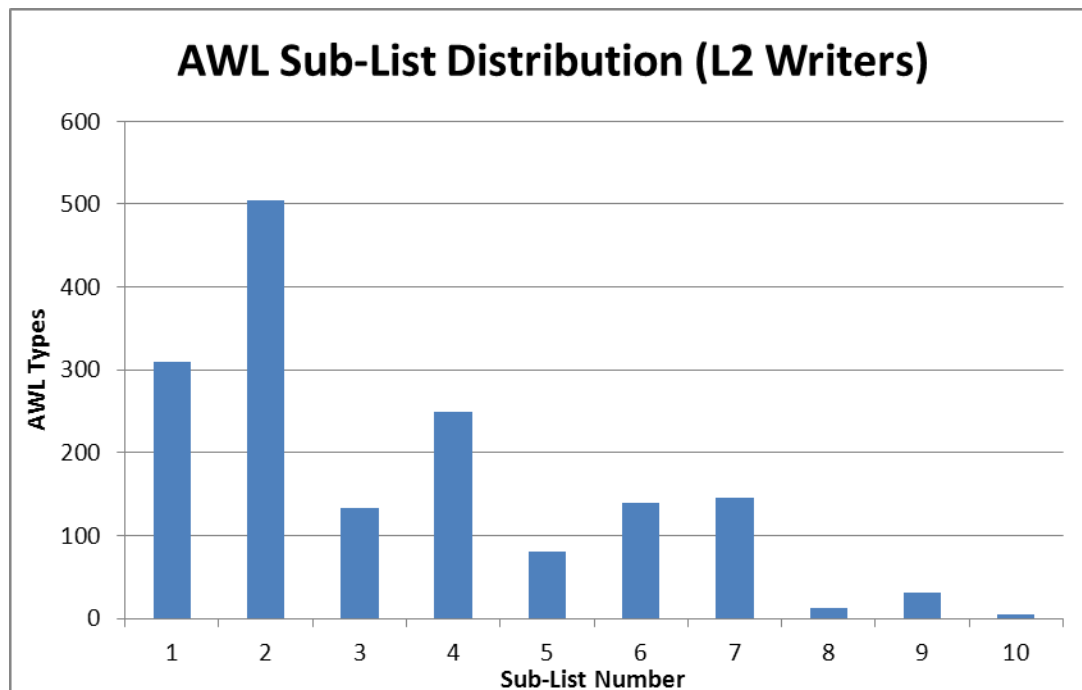

Figure 2: AWL sub-list distribution of types.

If there is statistical preference for words from one sub-list in one essay group, it does not necessarily follow that every student has shown the same preference. These are (it is acknowledged) pooled data and, while the implication seems to be that the most frequent types come from sub-list 2, it is not suggested that every student drew from that sub-list. There are two significant points, however. First, and contrary to the sub-list rationale, most types do come from Sub-list 2 (as opposed to Sub-list 1) and second, there is considerable variation in the number of types taken from the remaining sub-lists, although as expected sub-list 10 contributes the fewest.

Use of AWL tokens (all essays)

Figure 3 shows the distribution of AWL tokens, by sub-list, across the 342 essays. Students overall showed a preference for types from sub-list 2 (636) followed by sub-list 1 (562) and so on, with the lowest numbers coming from sub-lists 8 and 10 (22 and 5 respectively).
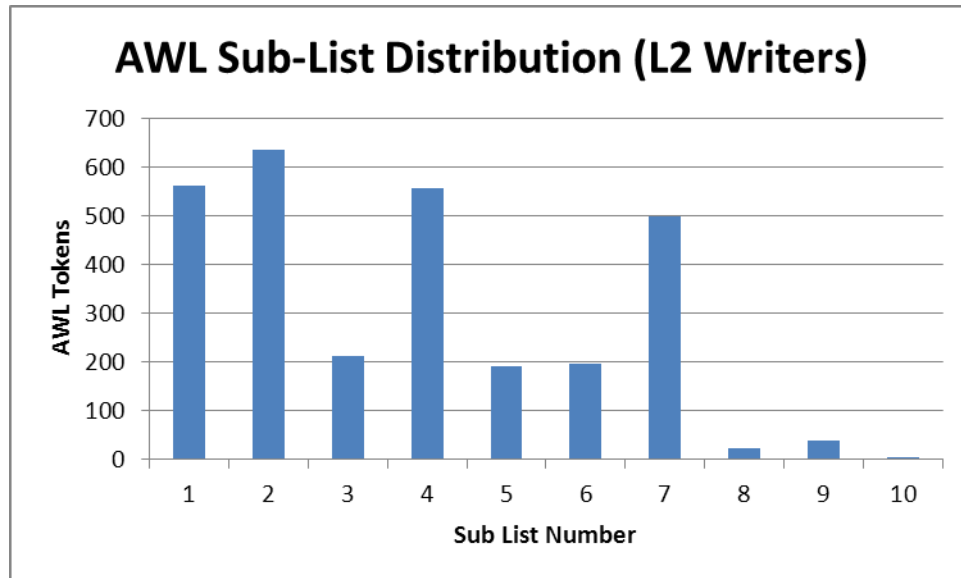
Figure 3: AWL sub-list distribution of tokens.

As with types, there is variation among the sub-lists but still Sub-list 2 is the prevalent source of academic tokens: not only do students overall use more different words from this list, they also make more intensive use of these words in their writing. It must be remembered, however, that these are pooled data.

Use of AWL types (by sub-list and essay group)

In contrast to the overall data above, Table 1 compares the use of AWL types from the sub-lists across separate essays, from Essay 1 (CEPA) to Essay 9.

Table 1: Use of AWL types in each essay, by sub-list.

Types	Essay										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total** (1611 types)	**% of types** in total sample
Sub-list											
1	2	19	20	54	31	56	61	41	25	**309**	**19.2**
2	27	32	44	72	91	66	56	60	57	**505**	**31.4**
3	5	1	6	28	14	31	27	15	7	**134**	**8.3**
4	11	13	23	35	69	34	31	18	15	**249**	**15.5**
5	0	9	2	8	15	23	8	11	4	**80**	**5.0**
6	6	3	3	17	35	31	18	19	8	**140**	**8.7**
7	4	4	11	22	11	10	32	25	27	**146**	**9.1**
8	0	1	0	4	3	1	1	1	1	**12**	**0.75**
9	3	5	2	1	1	7	8	1	3	**31**	**1.9**
10	0	0	0	1	1	2	1	0	0	**5**	**0.3**

The trend across the nine essays again suggests that sub-list 2 makes the largest contribution to the vocabulary evidenced in these essays (31.4% of the total types) followed by sub-list 1 (19.2%), then sub-lists 4 (15.5%), 7 (9.1%), 6 (8.7%), 3 (8.3%), 5 (5.0%), 9 (1.9%), 8 (0.75%) and 10 (0.3%). Note that Essay 7 is the only essay in which sub-list 1 use is predominant.

It is clear that the L2 writers in the present study do not draw on the sub-lists in the way assumed by some writers (making most use of sub-list 1 – the words most frequently used by native speaker/writers – and progressively less use of sub-list 2, then 3 then 4 etc.), although as expected the fewest words come from Sub-list 10.

Use of AWL tokens (by sub-list and essay group)

Table 2 compares the frequency of use of AWL tokens from the sub-lists across separate Essays, from Essay 1 (CEPA) to Essay 9.

Table 2: Use of AWL tokens in each essay, by sub-list.

Tokens	Essay										
Sub-list	1	2	3	4	5	6	7	8	9	Total (1611 types)	% of tokens in total sample
1	2	68	46	133	48	87	89	56	33	562	19.2
2	29	36	56	97	104	87	72	70	85	636	21.8
3	6	1	9	46	26	51	41	25	7	212	7.3
4	17	26	50	93	180	72	64	27	28	557	19.1
5	0	11	2	8	43	93	12	13	9	191	7.
6	6	4	3	18	53	52	21	31	9	197	6.8
7	4	11	12	78	15	17	192	118	52	499	17.1
8	1	1	1	4	3	1	9	1	1	22	.75
9	4	5	2	1	1	12	10	1	3	39	13.4
10	0	0	0	1	1	2	1	0	0	5	.2

The trend suggests a descending order of frequent use of tokens from sub-list 2 followed by sub-list 1, then sub-lists 4, 7, 3, 6, 5, 9, 8 and 10. Note that Essays 4 and 7 are the only essays in which Sub-list 1 use is predominant.

As with types, Sub-list 2 rather than Sub-list 1 is the dominant source of academic vocabulary, Sub-list 10 is the least accessed and the remaining sub-lists contribute on a seemingly *ad hoc* basis.

Genre

In any study of writing, essay genre (e.g. descriptive, or position paper) may be examined as a factor in the proficiency level of the output. In a previous study (Turlik, 2008) it was shown that genre (in the sense it is used in the present study) was not a factor that contributed to increased use of academic lexis, in that in a particular genre (let us say Genre X), some writers could record increased use of academic types compared to a previous essay (and different genre – Genre Y) while others might record *decreased* use in Genre X compared to a previous essay in Genre Y.

If the essays are classed under four general genre headings, (Descriptive, Cause-Effect, Problem-Solution and Position), then any given student answered at least one from each genre during their time in the foundation programme. Figure 4 shows the distribution of the genres across the two years.
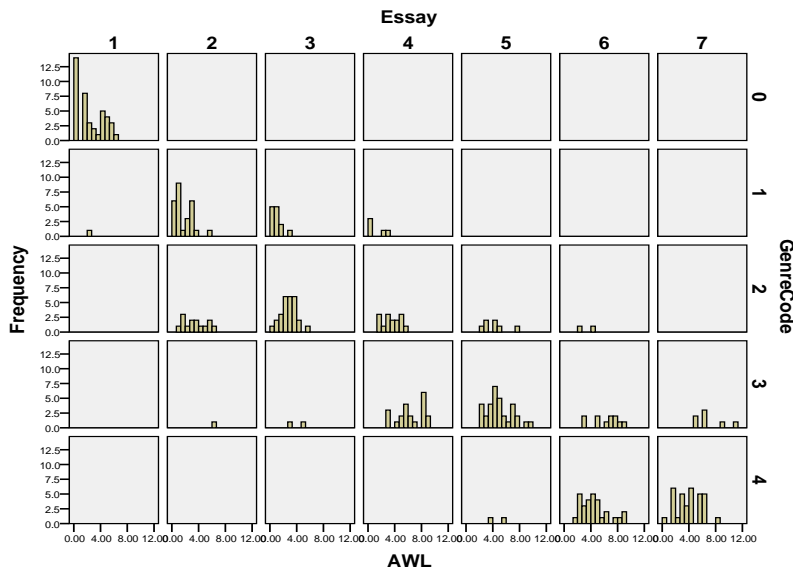
Figure 4: Genre of essays during the course of the study (Genre codes : 0 = Descriptive (CEPA); 1 = Descriptive; 2 = Cause-effect; 3 = Problem-solution; 4 = Position)**.**

At times of increase in mean AWL use (See Figure 1), it can be seen from Figure 4 that students were engaged in up to three different genres, leading to the conclusion that even though there is an increase in the mean use of academic vocabulary over a two-year period, individuals contribute to that rising mean at different times *despite different genres*. Thus, from Figure 4, it can be seen that (for example) for their Essay 4, a number of students wrote Descriptive (Genre 1), more students wrote a Cause and Effect (Genre 2) while the greatest number wrote Problem/Solution (Genre 3). Similarly, for Essay 6, a small number wrote Cause and Effect (Genre 2), rather more wrote Problem/Solution (Genre 3) while the smallest number of students wrote a Position essay (Genre 4).

Native speaker/writer use of academic vocabulary

Since the sub-lists of the AWL are based on texts taken from various academic sources written by native speakers of English (Xue & Nation, 1984; Coxhead, 2000), it was considered appropriate to include a parallel text from such a source.

Figure 5 and Figure 6 show the distribution of types by sub-list from a substantial text written by a native speaker, not to compare first language with second language output but to provide a reference point whereby the sub-lists might be appraised. It is clear that the distributions follow curves that would be expected, based on the rationale for such lists (Xue & Nation, 1984; Coxhead, 2000): overall, the frequency of use decreases from sub-list 1 through to sub-list 10. This is especially true of types (Figure 5). Pearson correlation between the L1 sub-list distribution of types and the present study students' sub-list distribution of types was $p = .77$ (S.D. 8.2 significant at .01 level 2-tailed).
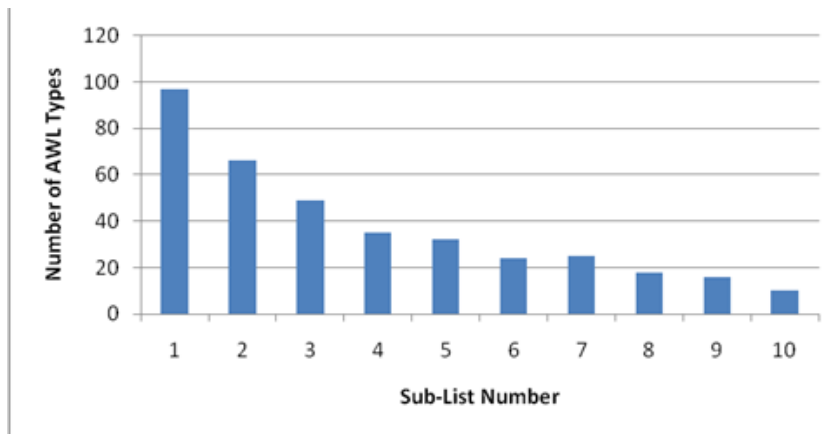
Figure 5: AWL types sub-list distribution (native writer).

Although the distribution of tokens (Figure 6) appears slightly different from that of types, there is nonetheless, an underlying curve from Sub-list 1 to Sub-list 10. Pearson correlation between the L1 sub-list distribution of tokens and these participating students' sub-list distribution of tokens was p = .78 (S.D. 9.1 significant at .01 level 2-tailed).
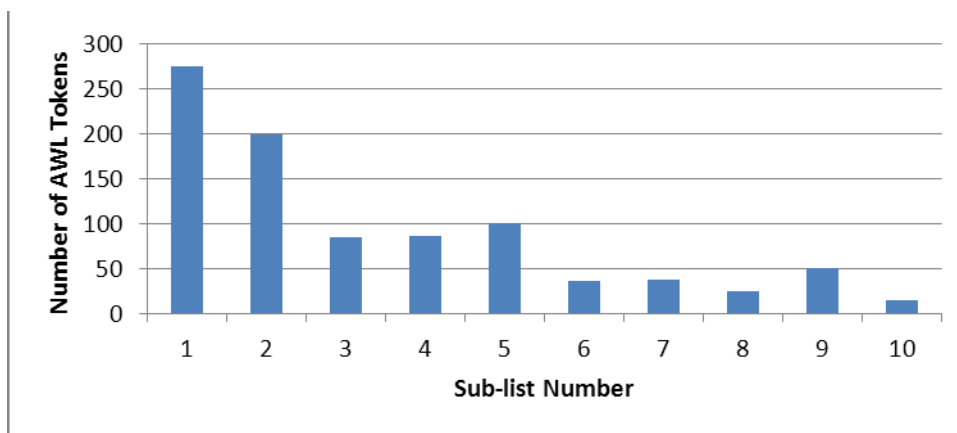

Figure 6: AWL tokens sub-list distribution (native writer).

Discussion and conclusions

As shown above, correlation between the native speaker/writer's use of academic vocabulary and the students' use of academic vocabulary for both types and tokens, is quite high (p= .77 and .78 respectively) which would suggest that the use of academic vocabulary by the second language students does, to an extent, follow the rationale of the word lists. Of particular significance, however, is the predominance of words from sub-list 2 *in every essay group* (except Essay 7) and, it must be stressed, this is real-life data – *actual* output and *actual* lexical selection over a period of two and a half years. At the outset, genre would not appear to be a factor since, in Essay 7, in which sub-list 1 predominated (61 types as opposed to 56 from sub-list 2), there were six problem/ solution and 28 'position' essays. This, of course, is a conclusion based on essays written as part of an English course but it is open to speculation as to whether the same conclusion might be reached if the essays were written in other disciplines.

Having accepted the conclusions in respect of the essays in the present study, a particular feature that occurred in determining the count by essay group was that, on two occasions, there were substantial 'clusters' of comparatively large numbers of tokens. These were in Essay 5 (tokens from sub-list 4) and Essay 7 (tokens from sub-list 7). In the former, 74% of students used tokens from sub-list 4, of which six had a token count of ten or more. On examination, it was found that all essays were from the same writing prompt[4], and most of the total tokens are accounted for by the repetition of one topic-specific word: *job*. Essay 7 also recorded a similar 'cluster': in this case, 60% of students used tokens from sub-list 7; again, the same prompt provided the word *aid*, repeated 83 times across the various pieces of writing.

In view of the foregoing, however, a possible explanation for the predominance of sub-list 2 and the clusters might have something to do with the way a task is presented. Students enter the programme at one of six of the eight 'levels' depending on CEPA score and, at each level, the genre of essay was prescribed, using a format known as 'rhetorical patterns'. Thus, in Levels 2, 3 and 4, 'description' (initially from pictures) was the pattern in their examination, in Level 5, 'cause-effect', Level 6 wrote 'problem-solution' and Levels 7 and 8, a 'position' essay. The names and prescriptive nature of these essays are reflected in the way they are organised: an essay must have an introduction, two or three paragraphs indicating aspects of the issue (e.g. causes or effects for a cause-effect essay) and a conclusion. It can be seen that Essays 2, 3 and 4 and Essays 5 and 6 comprised pieces of writing from three different genres – 1, 2, 3 and 2, 3, 4 – respectively, with Essay Group 7 comprising Genres 3 and 4. This suggests that genre, in its present form, is not a particular influence on the selection of lexical items; however, it is arguable that a particular prompt might result in the use of certain words, giving rise to the clusters mentioned previously.

Frequency of use of tokens implies repetition but the reason for these clusters and their possible implications for teaching and learning probably need further investigation. A point that might be considered is the use of synonyms. Repeating *job* and *aid* could be because the students did not have at their disposal suitable synonyms or perhaps they were unsure whether a particular synonym suited the context. Possible synonyms for *job* might be *employment*, *post*, *position* or *work*. *Employ*, *post* and *work* are all K1 words, theoretically easier to learn and use than those from the AWL, yet the students did not attempt to use them. *Job*, it should be noted, is in AWL sub-list 4, a*id* is in sub-list 7 while *help* is in K1 list and *assistance* is in sub-list 2 of the AWL; but neither *help* nor *assistance* was used. It would appear, therefore, that from a teaching perspective there may be implications not only for word selection but also for the variety of contexts in which those words might be used, which suggests the need to further consider the points mentioned earlier about lexical *precision* (Read, 2003) and some general words as *potential* academic vocabulary (Paquot, 2010). From the learner's point of view, familiarity of use in a particular context provides security, especially in an examination; but this might lead to overuse at the expense of known synonyms, hence the clusters. Another point worth noting is that the word *aid* was used in the essay prompt so, from a student's point of view, it must be contextually safe. Had alternative words been included in the prompts, then it is possible the students would have used them, suggesting that any essay written in response to a prompt is likely to be sensitive to the lexical selection in that prompt. A similar argument may be made in respect of recommended reading texts on which an essay may be based, although the scope for synonyms is far greater in a text than in a prompt. Of course, a counter-argument may be that the more proficient the writer, the more likely will be the use of lexis outside the basic K1 list, the use of lexis in the 'earlier' sub-lists of the AWL *and* the appearance of synonyms, prompts and texts notwithstanding. There would, however, seem to be compelling evidence from the present study to support Paquot's (2010) claim that:

> a number of general service words take on prominent rhetorical and organizational functions in academic discourse and their absence from lists such as Coxhead's (2000) Academic Word List may be highly problematic in academic writing courses. (p. 63)

Read (2003) points to the mindset that links vocabulary to the notion of a single word and suggests that this restricts our view. Although referring to assessment of vocabulary, he does state that there is a need to take account of multi-word lexical items (e.g. *take account of*) and it may be acceptable to conclude that this should be a part of the planning of courses. It is at this point that perhaps we need to consider the strategies we adopt for developing vocabulary in various disciplines. The nuances that a word can take on in a specific discipline would appear to rely so heavily on a fairly intimate knowledge of what we regard as high-frequency vocabulary that this vocabulary is too important to consider it in some way subordinate to academic.

Given the results of the present study, the current rethinking on the content and structure of vocabulary development in courses might be timely. In light of the foregoing, there would appear to be some key points worth considering in respect of vocabulary development. The first is the inclusion of lexical combinations, recognizing that it is not only single words that need to be learned. Part of knowing a word is knowing its correct context, its affixations and correct collocations, and these are issues that might need to be actively addressed through multiple-entry vocabulary items and entries at sentence level. Thus, for business students, the meanings of *marketing* and *strategy* should be presented not in isolation but at the same time as *marketing strategy, lethargic* and *market* with *lethargic market* and *yawning* and *gap* with *yawning gap* (perhaps *in prices* or *between exports and imports*), not to mention the stock market and its relationship to bulls (*bullish*) and bears (*bearish)*! A single word can accommodate considerable variation, with implications for the academic world – *post*-traumatic, *heavy* smoker (health) and *burning* issue (politics) and, for many in the present-day world of the computer, *shift, tab, delete, scroll* and *enter*. Some words and uses will be idiomatic and many will be what we know as phrasal verbs – *look after, look up, look into, take over*, *sort out* – but the meanings of the component words, the combinations *and* their contexts should all be addressed at the learning stage.

In respect of collocations and context, and from a teaching perspective, it would almost certainly be helpful if students were made aware of not only correct collocations but also collocations that are inappropriate and why they are inappropriate, and these courses and dictionaries do just that. An example is the well-known student letter to a friend which began "*Expensive* Charles"! The inclusion of synonyms which has hitherto perhaps been the main means of increasing vocabulary, remains a cornerstone of the learning process, of course, although I would suggest that in using them we, as teachers, should still be mindful of the points discussed above. We should, perhaps, remember that vocabulary lists may not necessarily be mutually exclusive, that a word or words in one may rely on a word or words in another for correct context, and that words deemed 'academic' may require words deemed 'general service' for their correct academic use. It is also worth remembering that perhaps our students, being second language users, may not necessarily access the word lists in the same way as a native speaker and that this access may be affected by the vocabulary we use in the tasks we present to our students, as exams or assignments.

Finally, there remains the matter of how teachers and students can determine whether there has been development in vocabulary acquisition and use. Most essays are written on a computer these days, which makes this easier to check. An online vocabulary profiler such as Cobb (2012) enables the teacher to monitor lexical content in reading materials, as well as the development of the group but more importantly, of the individual, and the student can monitor his or her progress through use of the same resources.

References

Bachman, L. & Palmer, A. (1996) *Language testing in practice*. Oxford University Press.

Cobb,T. (2012). *Web VocabProfile* [accessed 1 February 2012 at http://www.lextutor.ca/vp/eng/], an adaptation of Heatley & Nation's (1994) *Range*. Victoria University of Wellington, NZ. [Computer program, available at http://www.vuw.ac.nz/lals/].

Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly, 34*(2), 213-238.

Daller, H., Milton, J. & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.

Diaz-Gilbert, M. (2004). Vocabulary knowledge of pharmacy students whose first or best language is not English. *American Journal of Pharmaceutical Education 68* (4), Article 91.

Flowerdew, J. (1993). Concordancing as a tool in course design. *System, 21*(2), 231-244

Hyland, K. & Tse, P. (2007). Is there an 'academic vocabulary'? *TESOL Quarterly, 41*(2), 235-53.

Laufer, B. & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307-322.

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

Paquot, M. (2010). *Academic vocabulary in learner writing: from extraction to analysis*.   London. Continuum.

Read, J (2003). *Assessing vocabulary.*   Cambridge: Cambridge University Press.

Stubbs, M. (1986). *Educational Linguistics*. London: Blackwell

Turlik, J. (2008). *A longitudinal study of vocabulary in L2 academic English writing of Arabic first-language students: development and measurement*. Unpublished Ph.D. dissertation. University of the West of England. Bristol, UK.

Waring, R. (2002). *Scales of vocabulary knowledge in second language vocabulary assessment*. Retrieved January 2013 from http://wenku.baidu.com/view/43e085c289eb172ded63b7d5.html

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

Xue, G & Nation, P. (1984). A University Word List. *Language Learning and Communication 3*, 215-229.

---

[1] What is 'acceptable' is not explained.

[2] A precise number is not given.

[3] The expression 'essay group' refers to a set of 42 essays written at a certain time. For example, Essay (Group) 1 refers to the 42 CEPA essays, written in March 2003; Essay (Group) 2 would be the first exam essay written as part of the foundation programme in November 2003.

[4] It must be remembered that, as explained earlier, an essay group does not necessarily include essays of the same genre or prompt due to students' different rates of progress through the course.