Lexical properties in the writing of foreign language learners over eight years of study: single words and collocations

Tami Levitzky-Aviad and Batia Laufer
University of Haifa, Israel

Lexical proficiency has been defined and researched in terms of lexical knowledge, use and fluency. Different studies have shown that use of vocabulary in a foreign language (or L2) develops more slowly than vocabulary knowledge, either passive or active. However, many studies of free production compared learners of two or three proficiency levels and examined single words, not multi-word units, even though the latter are characteristic of idiomatic language, and should be considered a component of lexical use.

The data for the present study was collected as part of the on-going compilation of an Israeli learner corpus of written English. The data was analyzed to examine progress in vocabulary use over 8 years of learning, starting with students at the end of elementary school (grade 6) and ending with English majors at the university. The passages were compared on lexical richness – the proportion of frequent to non-frequent vocabulary, on lexical variation – type token ratio, and on the number of collocations. A total of 290 essays (200 words each) were analyzed using the VocabProfile, a software program that calculates the percentage of a text's words at different frequency levels and provides the text's type-token ratio. Significant increases in the use of infrequent vocabulary and collocations were found only with the university students. A significant increase in lexical variation was found at the end of high school. The lack of substantial progress during school years, on the one hand, and the significant progress during the one year at university, on the other hand, corroborate previous research. In light of this limited progress, recommendations are made for further investigations into the effect of different pedagogical approaches to the teaching of foreign language vocabulary.

1. Vocabulary and writing in a foreign language

The goal of the present study is to examine the development of several 'active' lexical dimensions across eight years of learning English. More specifically, the study aims at investigating developments in active vocabulary knowledge and in three dimensions of vocabulary use: variation, richness and the use of colloca-

tions. Vocabulary is a clear indicator of how well foreign language (FL) learners can communicate (Lewis, 1997; Widdowson, 1989). Effective vocabulary use in writing has been found to have a positive influence on measures of the quality of writing and on one's general language level (e.g. Lee, 2003; Llach & Gallego, 2009; Morris & Cobb, 2004). Also, language learners themselves mention vocabulary as a crucial aspect in writing (Leki & Carson, 1994; Polio & Glew, 1996). It is therefore not surprising that research interest in the importance of vocabulary for writing in a foreign language is growing.

To understand the relationship between vocabulary and writing, we will first explain several key terms in lexical research: lexical knowledge vs. lexical use; depth, breadth and strength of knowledge; passive and active vocabulary knowledge; recall and recognition; lexical variation and lexical richness; and collocations. We will then refer to available research on vocabulary and writing, first for single words, then for collocations.

Vocabulary acquisition can be discussed in terms of both 'lexical knowledge' and 'lexical use'. Lexical knowledge is the information about the word that learners have stored in their mental lexicons, while lexical use is the manifestation of this knowledge in real-time production (Laufer, 2005; Laufer & Goldstein, 2004). This distinction implies that lexical knowledge in a foreign language is typically more advanced than lexical use, because not all words stored in learners' mental lexicons are necessarily activated and used in free writing (Laufer, 1991).

Vocabulary knowledge can be assessed qualitatively, in terms of 'depth' of knowledge, and quantitatively in terms of 'breadth' of knowledge and 'strength' of knowledge. Depth of knowledge refers to the degree of acquaintance with the various form and meaning components of a given lexical entry (e.g. its morphological structure, its grammatical or lexical patterns, and its relations with other lexical items) (Richards, 1976). Breadth of knowledge refers to vocabulary size, i.e. the quantity of lexical entries stored in one's mental lexicon. In measuring vocabulary size, a word is considered 'known' when the correct meaning is associated with the correct word form. However, form-meaning associations can take different forms, reflecting different parameters according to which strength of knowledge is assessed (Laufer, Elder, Hill, & Congdon, 2004; Laufer & Goldstein, 2004). These parameters have been defined along the active-passive and recall-recognition distinctions of meaning-form relationships. More details on how the distinctions were operationalized are provided in the 'Measurement tools' section). The first distinction implies that there is a difference in knowledge between people who can retrieve the FL word form in order to convey a certain meaning ('active' knowledge) and those who cannot do this, but can retrieve the meaning once the FL word is presented to them ('passive' knowledge). The second

distinction implies that there is a difference between those who can recall the form or the meaning of a word and those who cannot do this, but can recognize the form or meaning in a set of options. Four modalities of strength of knowledge thus emerge from these distinctions: active recall, passive recall, active recognition and passive recognition. Of these, active recall is the hardest to achieve, and therefore represents the strongest degree of knowledge, followed by passive recall, active recognition and passive recognition, respectively (Laufer & Goldstein, 2004). In sum, strength of knowledge is a combination of four aspects of knowledge of meaning that constitute a hierarchy of difficulty: passive recognition (easiest), active recognition, passive recall, and active recall (hardest).

Lexical 'variation' and lexical 'richness' are two quantitative measures of vocabulary use. Variation, (or 'diversity'), is a measure of the number of different words (types) used, or, more specifically, the type-token ratio (TTR). 'Richness', on the other hand, is the proportion of low-frequency words in a piece of writing (Laufer, 1994; Laufer & Nation, 1995).

Phraseological analyses suggest that at least one-third to one-half of language is composed of multi-word units (MWU) (Erman & Warren, 2000; Hill, 2000). They are retrieved faster than individual lexical items, indicating perhaps that certain phrases are stored and retrieved as a whole (Erman, 2007; Schmitt, Grandage, & Adolphs, 2004; Wray, 2002). There also seems to be a processing advantage for formulaic sequences, at least in reading (Underwood, Schmitt & Galphin, 2004). Therefore, a good knowledge of formulaic language is advantageous for language learners and users.

Though there are several kinds of MWUs, we focused on the knowledge and use of lexical collocations (henceforth, 'collocations') as it was shown to be one possible indicator of native-like competence (Howarth, 1998; Hill, 2000). We have adopted Nesselhauf's (2003) definition of collocations as word combinations in which one of the words (the 'base' or headword) retains its independent meaning, while the meaning of the other word, (the 'collocate') is restricted to the specific context and can only be used with some semantically related headwords (though not even with all of them). The combinations chosen for investigation in the present research were thus only MWUs which were found compatible with this definition. These included examples such as 'make a decision' or 'heavy rain', but not combinations such as 'eat breakfast' or 'play ball'.

Active vocabulary has been found to be (i) smaller in size, (ii) develop more slowly (Laufer, 1998; Laufer & Goldstein, 2004; Nemati, 2010) and (iii) decay faster (Schneider, Healy, & Bourne, 2002) than passive vocabulary. Accordingly, as mentioned earlier, the most advanced degree of knowledge has been found to be active recall, followed by passive recall, active recognition and passive recognition, respectively (Laufer & Goldstein, 2004). Test results on progress in for-

eign language vocabulary use in writing have shown statistically significant improvements in richness in groups of university English majors (Laufer, 1991; Leńko-Szymańska, 2002), but not so much in school students or in university students who are not English majors. Results of correlation tests between active vocabulary size and use are not as consistent. Some studies have found significant correlations between active vocabulary size and richness and/or variation (e.g. Laufer & Nation, 1995) while others have not (e.g. Laufer, 1998; Lemmouh, 2010).

Knowledge of collocations by FL learners has been found to increase as learning progressed (e.g. Gitsaki, 1999), but not always to the same extent as knowledge of general vocabulary (e.g. Bahns & Eldaw, 1993). In fact, research on the use of collocations by FL learners has demonstrated that even advanced learners have considerable difficulties in producing collocations (Nesselhauf, 2003) and tend to use free combinations where collocations could be used (Nesselhauf, 2005).

2. The study

2.1. Research questions and hypothesis

Our research questions were the following:

1) What developments occur in the following dimensions of lexical proficiency during the years of formal English learning?
   a. the size and strength of active vocabulary knowledge of English words
   b. the lexical richness of learners' written samples
   c. the lexical variation of learners' written samples
   d. the use of collocations in learners' written samples
2) Is there a correlation between the improvements in each of the lexical dimensions over the years?

The basic hypothesis underlying the present research was that improvement would occur in all investigated parameters of vocabulary knowledge and use, throughout the eight years of EFL learning. With regard to collocations, this assumption was made despite the limited use of collocations by foreign language learners found in previous research (e.g. Nesselhauf, 2003, 2005). While acknowledging these findings, it was nevertheless assumed that, at least in some of our data that included participants majoring in English, some improvement would occur in this respect too.

2.2. Sources of data

The primary source of data for the present study consists of 290 passages written by learners of English in Israel during the beginning phases of the compilation of the Israeli Learner Corpus of Written English (ILcoWE, Waldman & Levitzky-Aviad, in preparation). This part of the corpus includes 215 passages written by school-aged students in seven consecutive grades (6-12) and 75 passages written by first year university English majors. The topics of essays varied as learners at very different proficiency levels cannot be expected to write on identical topics. The younger students (grades 5-7/8) were mainly asked to write essays of a narrative or descriptive nature in response to prompts such as 'Describe a family event that you attended' or 'Tell the story of what is going on in the drawing…'. The older students (grades 9-12/university majors) were asked to write descriptive and argumentative essays such as 'What would you do if you got a huge sum of money for your birthday? Explain your choices', 'Which is the most important meal of the day, and why?' or 'Argue for and against the use of computers in the classroom'.

Due to time limitations and the need to keep students' personal information confidential, a longitudinal corpus collection, and accordingly, a longitudinal study were impossible. The study is thus cross-sectional, examining the writing of different students at different years of learning.

The second source of data consists of the results of a bilingual test of active vocabulary knowledge, including both active recall and active recognition. Of the students who provided the corpus samples mentioned above, 101 were also administered a test of active vocabulary knowledge (see section 2.3.1). This sample included students at the end of elementary school (in grade 6), at the end of junior high (in grade 9), at the end of high school (in grade 12) and at the beginning of the first year in the English department.

2.3. Measurement tools

2.3.1. Measuring active knowledge

To measure active knowledge in the present research, we designed an active vocabulary test, modelling it upon two well established and validated vocabulary tests. The items were selected from the monolingual versions of the Vocabulary Size Test (VST, Nation & Beglar, 2007). The number of items was also the same as in VST. The methodology of testing, on the other hand, was modelled on the Computer Adaptive Test of Size and Strength (CATSS, Laufer, 2007).

The monolingual version of the VST tests words sampled from the 7,000 most frequent word families in English, based on the British National Corpus. The list can be divided into seven frequency levels (k1-k7), each comprising

1,000 words (Nation, 2006). In the VST, each of these levels is represented by a sample of 20 words. Hence, VST tests peoples' knowledge of a total of 140 items which represent the above mentioned 7,000 word families. As part of the VST, test-takers show their understanding of each English word tested by choosing the correct option from four options of synonyms and definitions of the word.

Though based on the VST, the test used for the current study was a bilingual test. Since the groups which were compared included beginners and low level learners, a bilingual test was considered more appropriate than a monolingual test. Additionally, while the VST tests passive knowledge, or, more specifically, passive recognition (since learners choose the correct paraphrase of the target item), the test designed for the purpose of the present research tested *active* knowledge.

The other test upon which our test was modelled is the CATSS. The specific feature of CATSS, in addition to testing words at different frequency levels, is that it tests the four modalities of strength of knowledge from strongest to weakest (see section 1): active recall, passive recall, active recognition and passive recognition. The test proceeds as follows: In the first modality (active recall), a prompt appears on screen, which is the L1 translation of the target word. The first letter of the target English word is also provided and the test-taker needs to use this letter and type the English equivalent. Words known in this modality are not tested again in subsequent modalities. Representing the hardest, hence strongest degree of knowledge, each correct answer accounts for 1 point of the final CATSS score. In the second modality (passive recall), the English target word appears on screen for the test-taker to translate into the L1. Words known in this modality are not tested again. Each correct answer accounts for 0.75 points of the final CATSS score. In the third modality (active recognition), the test-taker needs to choose the correct English equivalent for the L1 word out of four English options. Words known at this modality are not tested again. Each correct answer accounts for 0.5 points of the final CATSS score. In the last modality (passive recognition) the test-taker needs to choose the correct L1 equivalent for the English target word out of four L1 options. Representing the 'weakest' degree of knowledge, a correct answer at this modality receives 0.25 points of the final CATSS score. Words not known in any of the four modalities receive zero points in the final score. The items tested proceed from frequent to less frequent. Hence, the final CATSS score has been claimed to represent both size and strength of knowledge as it takes into account not only the number of words test-takers know, but also the 'way' in which these words are known (Laufer et al., 2004; Laufer & Goldstein, 2004).

Modelled upon CATSS, the test designed for the present study also takes into account different strength modalities, yet with several modifications. While

CATSS tests both passive and active knowledge, the test in this study tests only active knowledge (hereafter referred to as ACATSS). Another feature distinguishing ACATSS from the original CATSS is that the Hebrew (L1) prompt words in the ACATSS do not appear in isolation, but rather in between two asterisks *within* a Hebrew sentence. The decision to present the word within a sentence was made so as to avoid ambiguity in cases of polysemy of the Hebrew words. Such an approach also follows the model used in the VST.

In the ACATSS, the learners' task is to provide the English equivalent of the word in asterisks. To do so, the test includes three cycles: two for testing active recall and one for testing active recognition.

First, the target item is tested for active recall without any cues, to mirror a real life situation of independent writing. This is demonstrated in the following example, where the target word is 'lake' and the Hebrew sentence means: *This *lake* is nice.* The instructions for the test were given in both English and Hebrew so that young learners could also clearly understand what they were expected to do.

Example: cycle 1
Translate the words in *asterisks* into English:
ה*אגם* הזה נחמד.

A word known in this cycle is not tested again. If it is not known, it is tested again in the second cycle. Here too active recall is tested, but now with the first letter of the English word provided. Whereas in cycle 1 learners may provide a non-target word which nevertheless fits the context, the first letter in cycle 2 limits word choice, trying to direct the learners to elicit the target word.

Example: cycle 2
Translate the words in *asterisks* into English
(use the first letter of the English word as provided for you):
ה*אגם* הזה נחמד.   l

Based on the assumption that words known in active recall would also be known in active recognition (Laufer et al., 2004; Laufer & Goldstein, 2004), only words which were not known in either one of the active recall stages are tested again for active recognition. In this third cycle, learners are presented with four English words of which they are asked to choose the correct equivalent for the Hebrew word in asterisks. The distracters in the recognition stage were sampled from the same frequency level as the English target word to eliminate the effect that word frequency might have on the choice of the response.

Example: cycle 3

Circle the correct translation for each of the words in *asterisks*:

ה.*אגמ* הזה נחמד.   a. tale   b. rhythm   c. lake   d. lawn

Once all 20 words at one frequency level are tested, the test moves on to the next frequency level. A word scores 1 point if known in the first cycle (active recall with no cue), 2/3 if known in the second cycle (active recall with a cue), 1/3 in the third cycle (active recognition) and 0 for lack of any knowledge. The total score for each frequency level is calculated by adding up the scores learners receive for the 20 words. The total scores of all seven frequency levels are then summed up to provide one total ACATSS score. As in the VST, since the 140 words tested in the ACATSS represent a vocabulary size of 7,000 word families, the total ACATSS score can be multiplied by 50 to provide an indication of active vocabulary size as affected by the strength modalities tested.

2.3.2. Measuring use - VocabProfile

The sampled written passages were analyzed with the experimental BNC-20 version of the Web-VocabProfile (WebVP) program on the Lextutor website (http://www.lextutor.ca Cobb, n.d.). The WebVP is an adapted version of Heatley and Nation's Range program (1994). Both the Range and the WebVP programs match a text with frequency lists and show the relative proportion of words used from different frequency levels. The relative proportion is called LFP (Lexical Frequency Profile). The program also calculates the type-token ratio (TTR) of an essay. The profiles created with these programs present the proportions of k1, k2 and Academic Word List (AWL Coxhead, 2000) words in a text. The experimental BNC-20 version, on the other hand, presents the proportion of words in a text which are taken from the revised 20 frequency levels based on the British National Corpus (Nation, 2006; Cobb, 2007). In this sense, it seems to provide a more detailed and fine-grained profile of the learners' writing. Additionally, as with the use of the ACATSS for active vocabulary knowledge, the experimental BNC-20 WebVP might be more sensitive than earlier versions to developments in vocabulary profiles between different learning stages.

    To use the VocabProfile, various steps had to be taken regarding the corpus data that were used. The profile has been shown to be less stable with essays shorter than 200 words. Such essays were therefore excluded from the present research. Furthermore, in light of the sensitivity of the TTR to composition length (e.g. Kucera & Francis, 1967; Linnarud, 1986), only 200 words of each passage were sampled, even if more words were written.

Three scores were obtained with the VocabProfile. Following the distinction between the first 2000 words (k1-k2) as the most frequent words and the beyond-2000 levels (k3-k20) as the low frequency words (Nation & Kyongho, 1995), we first added up the percentages of k3-k20 to obtain the general percentage of the low frequency vocabulary in the passages. The score obtained was thus considered an indication of how 'rich' the piece of writing was. However, since some of the learners whose essays were sampled for the research were at the very early stages of EFL learning, we also separated the percentages of the 1st and the 2nd 1000 words. Additionally, the TTR obtained with the VocabProfile program was taken to be an indication of variation.

2.3.3. Testing the use of collocations

No measurement tool was employed for testing the use of collocations in the written samples. These were manually counted. Once a word combination was identified as a possible collocation, a further step was taken to check whether these specific combinations were used in native-speakers' language. To this end, three sources based on native-speakers corpora were consulted: the Longman Exams Coach (Summers, Mayor, & Elston, 2006), the Oxford Collocations Dictionary (McIntosh, Francis, & Poole, 2009) and the word frequency list of American English (Davis & Gardner, 2010). If the expression appeared in at least one of these sources, it was considered a collocation.

As we performed the manual check, three things became apparent. First, as in Hsu (2007), the collocations were mostly verb-noun, or adjective-noun collocations. Therefore, only the use of these grammatical combinations was examined. Secondly, the number of collocations in each of the 200-word samples seemed quite small (see table 4.1), and, in many cases, they were the same ones used more than once (in accordance with Nesselhauf, 2005). Counting the total number of such collocations, then, (with many of them repeatedly used), did not seem to be of great value in checking for lexical growth over the years. Hence, following a similar procedure to that used by Zhang (1993) and Hsu (2007), each specific collocation was counted only once even if it was used repeatedly (in much the same way as the counting of 'word types' with single words).

2.4. Data Analysis

When we applied the three procedures outlined in section 2.3, five scores were obtained, each representing one dimension of active lexical proficiency. The total ACATSS scores were used as a measure of active knowledge size and strength. The proportions of k2 words and k3-k20 words in the written samples as calculated by the VocabProfile were used as two measures of lexical richness in writing. The type-token ratio as calculated by the VocabProfile was used as a meas-

ure of lexical variation in writing. Finally, the total number of different verb-noun and adjective-noun collocations was used to examine their prevalence in the written samples.

Four sets of one-way ANOVAs and post-hoc tests were used to compare learners at different points of learning on each of the four dimensions of lexical proficiency: size and strength of active vocabulary knowledge, richness, variation and the use of collocations.

Pearson correlations were then used to test whether the improvements in each of the lexical dimensions over the years correlate with each other.

2.5. Results

Our first research question addressed the developments in each of the dimensions of lexical proficiency. Tables 1.1 – 4.2 show the results for each dimension. As noted in section 2.2, the written data analyzed in the present study consisted of the 290 passages written by school-aged students in grades six to twelve and by first year university English majors. However, the ACATSS results were only obtained for 101 of these students. Thus, tables 1.1 and 1.2, showing the results for active knowledge, refer only to students in grades 6, 9 and 12 and the university students at the beginning of their first year in university. Tables 2-4 then show the results for the different measures of vocabulary use in the written passages for all the school grades tested and for the university students at the beginning and at the end of their first year.

2.5.1. RQ 1a: What developments occur in the size and strength of active vocabulary knowledge of English words during the years of formal English learning?

Table 1 presents the means of the raw scores for each of the English learning stages tested by the ACATSS. Table 2 shows the significance of differences between the different pairs of learning stages. As noted in section 2.2, only 101 of the 290 students were tested with the ACATSS. Accordingly, the results in tables 1 and 2 only refer to these students. Table 1 shows that the mean ACATSS scores increase at each learning stage; table 2 shows that the differences between all pairs of stages are statistically significant.

Table 1. Raw ACATSS scores (out of a maximum of 140) (n=101 learners)

Learning Stage	N	Min	Max	Mean	SD
Grade 6 (end of element. school)	15	9	21	15	4
Grade 9 (end of junior-high)	27	27	50	37	5
Grade 12 (end of high school)	29	30	62	46	9
Eng. Majors- beginning	30	39	74	57	10

Table 2. Differences in mean ACATSS scores between learning stages

Learning Stage	Grade 6	Grade 9	Grade 12
Grade 9 (end of junior-high)	22**		
Grade 12 (end of high school)	32**	10**	
Eng. Majors- beginning	42**	20**	10**

**p<0.01

2.5.2. RQ 1b: What developments occur in the lexical richness of learners' written samples during the years of formal English learning?

Table 3 presents the mean proportions of k3-k20 words in the written samples. Table 4 shows the significance of differences in these proportions between all of the different pairs of learning stages. Table 5 presents the mean proportions of k2 words in the written samples. Table 6 shows the significance of differences in these proportions between all of the different pairs of learning stages.

Table 3 shows a general increase across the learning stages represented by school/university years in the mean proportion of k3-k20 words in the written samples, despite some slight decreases between some of the learning stages (e.g., grade 9 - 3.84%, grade 10 – 3.65%). However, as shown in table 4, in school years all these changes appear to be statistically insignificant. In other words, in the six years between the end of elementary school (grade 6) and the end of high-school there are no statistically significant increases in the use of low frequency words of k3-k20. Statistically significant improvements occur between each of the school grades 6-12 and the English majors at the end of their 1st year in the English department and between each of the school grades 6-10 and the English majors at the beginning of their first year. Another significant improvement occurs in the one year of English studies at the English departments in the college or university.

Table 3. Mean proportions (in %) of k3-k20 words in the written samples (n=290 learners)

Learning Stage	N	Min (%)	Max (%)	Mean (%)	SD
Grade 6	15	1.5	5.45	3.24	1.20
Grade 7	21	.99	5.37	2.85	1.11
Grade 8	35	1	6.40	3.28	1.54
Grade 9	30	.98	6.86	3.84	1.62
Grade 10	39	0	8.16	3.65	1.78
Grade 11	36	.51	7.92	4.04	1.80
Grade 12	39	.50	8.54	4.17	1.78
Eng. Majors- beginning	36	1.49	12.75	5.48	2.74
Eng. Majors-end of 1st year	39	.50	16.58	7.75	3.37

Table 4. Differences in k3-k20 proportions between stages of learning

Learning Stage	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Eng. Majors-beginning
Grade 7	.39							
Grade 8	.04	.43						
Grade 9	.60	.99	.56					
Grade 10	.41	.80	.37	.19				
Grade 11	.80	1.19	.76	.20	.39			
Grade 12	.93	1.32	.90	.33	.52	.13		
Eng. Majors-beginning	2.24*	2.63**	2.21**	1.64*	1.83**	1.44	1.31	
Eng. Majors-end of 1st year	4.51**	4.91**	4.48**	3.91**	4.10**	3.71**	3.58**	2.27**

*$p<0.05$ **$p<0.01$

Table 5 shows a general increase in the use of k2 words. Table 6 shows that significant increases in the use of these words occur already during school years between each of the grades 6-10 and grade 12. Statistically significant improvements also occur between each of the school grades 6-10 and the two university stages.

Table 5. Mean proportions (in %) of k2 words in the written samples (n=290 learners)

Learning Stage	N	Min (%)	Max (%)	Mean (%)	SD
Grade 6	15	2.5	7.35	4.55	1.40
Grade 7	21	1.46	8.37	4.63	2.06
Grade 8	35	1.95	8.29	5.13	1.83
Grade 9	30	0	10.26	4.82	2.64
Grade 10	39	1.46	9.80	5.34	2.88
Grade 11	36	.50	11.50	5.79	2.99
Grade 12	39	1.99	12.56	7.25	2.58
Eng. Majors- beginning	36	2.49	13.93	7.27	3.18
Eng. Majors-end of 1st year	39	2.42	18.65	7.37	3.22

Table 6. Differences in k2 proportions

Learning Stage	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Eng. Majors-beginning
Grade 7	.08							
Grade 8	.58	.50						
Grade 9	.28	.19	.30					
Grade 10	.79	.71	.21	.51				
Grade 11	1.24	1.16	.66	.96	.45			
Grade 12	2.70*	2.62**	2.12*	2.43**	1.91*	1.46		
Eng. Majors-beginning	2.72*	2.64**	2.14*	2.44**	1.93*	1.48	.02	
Eng. Majors-end of 1st year	2.83*	2.74**	2.25**	2.55**	2.03*	1.59	.12	.11

*$p<0.05$ **$p<0.01$

2.5.3. RQ 1c: What developments occur in the lexical variation in learners' written samples during the years of formal English learning?

Table 7 presents the mean type-token ratio reflecting lexical variation, i.e., the percentage of different words in the text. Table 3.2 shows the significance of differences between all the different pairs of EFL learning stages in regard to the type-token ratios.

Table 7. Type-Token ratios (in %) of the written samples (n=290 learners)

Learning Stage	N	Min (%)	Max (%)	Mean (%)	SD
Grade 6	15	41	57.29	50.98	4.79
Grade 7	21	43.41	58.25	49.77	3.83
Grade 8	35	43.37	60.50	53.09	3.97
Grade 9	30	41.09	60.10	53.05	4.63
Grade 10	39	42.36	60.50	52.95	3.78
Grade 11	36	43.07	59.41	52.56	4.19
Grade 12	39	46.83	64.71	56.78	4.06
Eng. Majors- beginning	36	46.77	63.96	56.83	3.93
Eng. Majors-end of 1st year	39	48.74	66.50	56.77	4.32

Table 7 shows a general increase in the type-token ratios in the writing samples, despite some slight decreases which occasionally occur (e.g., grade 6 – 50.98%, grade 7 – 49.78%). The only statistically significant differences, however (table 8) are between each of the grades 6-11 and grade 12 and between each of the grades 6-11 and each of the university stages.

Table 8. Differences in the Type-Token ratios

Learning Stage	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Eng. Majors-beginning
Grade 7	-1.21							
Grade 8	2.11	3.32						
Grade 9	2.07	3.27	.04					
Grade 10	1.97	3.18	.13	.09				
Grade 11	1.58	2.78	.53	.49	.40			
Grade 12	5.80**	7.01**	3.69**	3.73**	3.82**	4.22**		
Eng. Majors-beginning	5.85**	7.06**	3.75**	3.79**	3.88**	4.28**	.05	
Eng. Majors-end of 1st year	5.79**	7**	3.69**	3.73**	3.82**	4.22**	0	.06

*p<0.05 **p<0.01

2.5.4. RQ 1d: What developments occur in the use of collocations in the learners' written samples during the years of formal English learning?

Table 9 presents the raw means of different (not repeated) verb-noun and adjective-noun collocations found in the learners' written samples of 200 tokens each. Table 10 shows the significance of differences between all the different pairs of EFL learning stages in regard to the use of these collocations.

Table 9 shows a general increase in the use of collocations, despite some decreases which occur occasionally (e.g., grade 10 – 0.72, grade 11 – 0.42). However, table 10 demonstrates that the only statistically significant differences are between each of the school grades (6-12) and the English majors at the end of their first year and between each of the grades 6-9 and 11 and the English majors at the beginning of the first year.

Table 9. Raw means of different collocations in the 200-word samples (n=290 learners)

Learning Stage	N	Min (raw)	Max (raw)	Mean (raw)	SD
Grade 6	15	0	1	0.13	0.35
Grade 7	21	0	2	0.38	0.59
Grade 8	35	0	2	0.23	0.55
Grade 9	30	0	2	0.37	0.61
Grade 10	39	0	5	0.72	1.15
Grade 11	36	0	2	0.42	0.60
Grade 12	39	0	4	0.72	0.94
Eng. Majors- beginning	36	0	7	1.31	1.65
Eng. Majors-end of 1st year	39	0	5	1.56	1.57

Table 10. Significance of differences between the raw means of collocations

Learning Stage	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Eng. Majors-beginning
Grade 7	.25	Grade 8	.10	.15				
Grade 9	.23	.01	.14					
Grade 10	.58	.34	.49	.35				
Grade 11	.28	.04	.19	.05	.30			
Grade 12	.58	.34	.49	.35	.00	.30		
Eng. Majors-beginning	1.17*	.92*	1.08**	.94*	.59	.89*	.59	
Eng. Majors-end of 1st year	1.43**	1.18**	1.34**	1.20**	.85*	1.15**	.85*	.26

*p<0.05 **p<0.01

Table 11 shows the results of Pearson product moment correlations between the developments, that is, the mean differences of the various lexical dimensions over the years. Correlations with the ACATSS were conducted only for the 101 students who took this test. All other correlations were conducted for all 290 students.

The table 11 shows that the improvements in almost all lexical dimensions over the years correlate significantly with each other. Lack of significant correlation was found only between the results of the progress on the ACATSS and the progress in the use of collocations.

Table 11. Correlations between the mean differences of the various lexical dimensions

	Active knowledge size & Strength (ACATSS) (N=101)	Variation (TTR) (N=290)	Richness #1 (k3-k20) (N=290)	Richness #2 (k2) (N=290)
Variation (TTR) (N=101)	.380**			
Richness #1 (k3-k20) (N=290)	.207**	.297**		
Richness #2 (k2) (N=290)	.298**	.348**	.316**	
Use of collocations (N=290)	.149	.326**	.222**	.201**

**p<0.01

3. Discussion

The main focus of this study was the similarities and differences in the developmental patterns of several dimensions of L2 lexical proficiency over eight years of study. We will therefore discuss the progress found for each dimension and compare the development of vocabulary knowledge with that of vocabulary use.

Continuous statistically significant improvements were found in active knowledge as reflected in the ACATSS scores across all stages of English learning (see tables 1 and 2). And yet, these significant improvements should also be considered vis-à-vis what they mean in terms of active vocabulary size and its growth, and, even more so, in terms of the manifestation of this knowledge in vocabulary *use*.

An increase in the size of knowledge suggests that there is an increase in the amount of low-frequency words learners know. We can therefore expect that at least those learners who have demonstrated a relatively high command of the language and are accepted to the English department would also possess knowledge of more lower-frequency words than would the general population of school-aged students for whom English is not the major area of study. When multiplying the mean ACATSS score of the first year English majors (see table 1) by 50 to reach the more general estimate of their active vocabulary size (see section 2.3.1), the figure reached is 2850 (57x50). Hence, despite the statistically significant increase in active vocabulary size from the 12th grade to the beginning of the 1st year in the English department (see table 1.2), even the advanced students in the latter group know fewer than 1000 words beyond the 2000 most frequent words in English.

Furthermore, although these figures represent the development in active *knowledge*, they do not necessarily reflect a similar vocabulary growth in free writing. With regards to free writing, the results show a gradual, and sometimes statistically significant, progress in the three dimensions of vocabulary use we tested: richness, variation and the use of collocations. However, while active knowledge demonstrated a *continuous* significant increase throughout the years, our findings, similar to previous ones (Laufer, 1991; Laufer & Nation, 1995; Laufer & Paribakht, 1998; Lemmouh, 2010; Leńko-Szymańska, 2002; Muncie, 2002) indicate that six or more years must pass before students' ability to put this knowledge into use also significantly improves. More specifically, a statistically significant improvement in lexical variation was evident only at the end of high-school (see table 8), whereas statistically significant improvements in the use of the k3-k20 low-frequency words were completely lacking during school years and occur only during the one year of university (see table 4). Lack of significant progress is also evident in the use of collocations, not only during school years, but also during the one

year of university (see table 10). These results corroborate previous findings (Laufer & Waldman, 2011; Nesselhauf, 2003; Pawley & Syder, 1983) and provide a clear indication of the specific difficulty involved in incorporating collocations into the writing of even advanced learners. Laufer and Waldman (2011) explained this difficulty in terms of semantic transparency of collocations and their difference from L1. As many collocations are easily understood, they go unnoticed in the input, and as a collocate in an L2 collocation is often different from L1, learners cannot rely on their L1 and on the knowledge of the individual words in L2.

The lack of statistically significant improvements in students during the six earlier school years, as well as the lack of significant progress in the use of collocations even during the one advanced year at university, are even more puzzling given that richness and variation in vocabulary use can improve even over the course of a single year at university. Since not all school students eventually become English majors, some of them may never again study English in a formal setting. It is hard to accept, then, that what school students end up with is only an active vocabulary size of just over 2000 word families (46X50=2300), and, perhaps, a higher ability to vary the vocabulary they are able to use, without similar increases in the numbers of lower-frequency words or collocations they use.

A few possible explanations can be provided to account for the discrepancies between vocabulary knowledge and use and for the lack of significant progress in vocabulary use during earlier school years. One possible assumption which could have been made is that the nature of vocabulary learning may be such that active knowledge and use are separate traits of lexical proficiency, which develop in totally different ways. However, the moderate correlations we found between vocabulary knowledge and use (see table 11), similar to previous studies (Laufer & Nation, 1995; Leńko-Szymańska, 2002), point to a different interpretation of the results. These correlations indicate that, despite the discrepancies between vocabulary knowledge and use, an increase in learners' active vocabulary knowledge may be moderately reflected in their *use* of richer vocabulary. Also, the statistically significant increase in the use of k3-k20 words during the one year at university suggests that rapid progress in vocabulary use is possible. Hence, taken together, the significant correlations found between active vocabulary knowledge and use and the progress in the use of low-frequency words over the one year of university suggest that the lack of statistically significant growth we found in lexical use could be changed.

Therefore, another explanation for the lack of significant progress in vocabulary use during earlier school years could be the lack of sufficient language training and practice during these years, which could result from learners' writing strategies, the teaching methods applied and/or the time of expo-

sure to English during school years. Coming up with a word to express a certain idea in writing requires learners to know more features of that word than they need when they are asked to provide the word in some controlled setting. However, due to factors such as the rarity of low frequency words, the arbitrary nature of collocations or various incongruencies between L1 and L2 collocations, learners may experience uncertainties regarding the use of such lexical items and may thus simply refrain from using them (Fan, 2009; Hill, 2000; Laufer, 1998; Laufer & Waldman, 2011; Nesselhauf, 2003). Instead, they may resort to using high frequency single words which convey the same, or at least similar, ideas. This strategy is reinforced by teachers who believe that for communication to be effective, foreign language learners' ability to express their ideas using any appropriate vocabulary is satisfactory in many cases. Unfortunately, such a claim, especially when made by teachers, downplays the need for sufficient practice of non-basic vocabulary (Laufer, 2005; Nemati, 2010; Milton, this volume) and, consequently, perpetuates stagnation of vocabulary in free expression. This lack of progress is not something that any education system should welcome.

To achieve progress, specific and realistic goals need to be set, and effective teaching methods need to be implemented. Such teaching methods should involve acknowledging the importance of encouraging FL learners' use of low-frequency vocabulary and collocations in their writing. Previous studies have shown the effectiveness of Form-Focused Instruction (FFI) in activating learners' lexical knowledge and putting some of it to use (Laufer, 2005; Laufer, 2010; Laufer & Girsai, 2008; Lee, 2003; Nesselhauf, 2003; Webb, 2005; Xiao & McEnery, 2006). Such an approach advocates explicit vocabulary instruction, either as part of more general communication tasks (Focus on Form-FonF) or as a goal in itself (Focus on Forms – FonFs). A longitudinal systematic syllabus of FFI which gradually introduces low-frequency words and collocations and encourages their use could be a possible solution for enhancing the knowledge and use of such items at all stages of L2 learning.

Future research could compare the development of EFL vocabulary use in writing in different educational systems, in different classes or in different controlled experimental conditions. Such comparisons might be useful to show the effectiveness of different pedagogical approaches for the development of L2 vocabulary use over the years.

References

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*(1), 101-114.

Cobb, T. (n.d.). *Web Vocabprofile: An adaptation of Heatley & Nation's (1994) Range.* Computer program. Available on-line at http://www.lextutor.ca/vp/

Cobb, T. (2007).*The revised frequency lists of k8-k14.* Available on-line at http://www.lextutor.ca/vp/bnc/cobb_6

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

Davis, M. & Gardner, D. (2010). *Word Frequency List of American English.* Available on-line at www.wordfrequency.com

Erman, B. (2007), Cognitive processes as evidence of the idiom principle. *International Journal of Corpus Linguistics 12*(1), 25-53.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse, 20*(1), 29-62.

Fan, M. (2009). An exploratory study of collocational use by ESL students-A task based approach. *System, 37*(1), 110-123.

Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of collocational knowledge*. San Francisco, CA: International Scholars Publications.

Heatley, A. & Nation, P. (1994). *Range*. Victoria University of Wellington, NZ. Computer program. Available on-line at http://www.vuw.ac.nz/lals/

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching Collocation: Further Development in the Lexical Approach* (pp. 47-70). Hove: Language Teaching Publications.

Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161-186). Oxford: Clarendon Press.

Hsu, J. (2007). Lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. *Electronic Journal of Foreign Language Teaching, 4*(2), 192-209.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press Providence, RI.

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

Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal, 25* (2), 21-33.

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

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook, 5*(1), 223–250.

Laufer, B. (2007). *CATSS: The Computer Adaptive Test of Size and Strength.* Computer program. Available on-line at http://hcc.haifa.ac.il/~blaufer/

Laufer, B. (2010). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, *24*(1), 29-49.

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

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48* (3)*,* 365-391.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing, 21*(2), 202-226.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Journal of Learning Language, 54*(3), 399-436.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29,* 694-716.

Laufer, B., & Waldman, T. (2011). Verb?noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning,* 61(2), 647–672.

Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System, 31*(4), 537-561.

Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *Tesol Quarterly, 28*(1), 81-101.

Lemmouh, Z. (2010). *The Relationship among Vocabulary Knowledge, Academic Achievement and the Lexical Richness in Writing in Swedish University Students of English.* Ph.D. Dissertation, Department of English, Stockholm University.

Leńko-Szymańska, A. (2002). How to trace the growth in learners' active vocabulary? A corpus based study. *Teaching and Learning by Doing Corpus Analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora.* Graz (pp. 19-24).

Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 255-270). Cambridge: Cambridge University Press.

Linnarud, M. (1986). Lexis in composition: A performance analysis of Swedish learners' written English. *Dissertation Abstracts International. C: European Abstracts, 47* (4), 812.

Llach, M. P. A., & Gallego, M. T. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *ATLANTIS, 31*, 129-147.

McIntosh, C., Francis, B., & Poole, R. (Eds.) (2009). *The Oxford Collocations Dictionary.* Oxford: Oxford University Press.

Morris, L., & Cobb, T. (2004). Vocabulary profiles as predictors of the academic performance of teaching English as a second language trainees. *System, 32*(1), 75-87.

Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical frequency profiles across drafts. *System, 30*(2), 225-235.

Nation, I.S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 63(1), 59-82.*

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

Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7): 9-13.

Nemati, A. (2010). Active and passive vocabulary knowledge: The effect of years of instruction. *The Asian EFL Journal Quarterly 12*(1), 30-46.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: Benjamins.

Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics, 7*(5), 551-579.

Polio, C., & Glew, M. (1996). ESL writing assessment prompts: How students choose* 1. *Journal of Second Language Writing, 5*(1), 35-49.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In Schmitt, N. (ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 127-151). Amsterdam: Benjamins.

Schneider, V. I., Healy, A. F., & Bourne L. E. Jr. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language, 46*(2), 419-440.

Summers, D., Mayor, M., & Elston, J. (Eds.), (2006). *The Longman Exams Coach*. Essex: Pearson-Longman.

Underwood, G., Schmitt, N., & Galphin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In Schmitt, N. (ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp 153-172). Amsterdam: Benjamins.

Waldman, T. & Levitzky-Aviad, T. (in preparation). The Israeli Learner Corpus of Written English (ILcoWE).

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*(01), 33 52.

Widdowson, H. G. (1989). Knowledge of language and ability for use. *Applied Linguistics, 10*(2), 128-137.

Wray, A. (2002). *Formulaic Language and the Lexicon.* Cambridge: Cambridge University Press.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103.

Zhang, X. (1993). *English collocations and their effect on the writing of native and non-native college freshmen*. PhD. Dissertation, Indiana University of Pennsylvania.