

First-year university students' receptive and productive use of academic vocabulary

Déogratias Nizonkiza

School of Languages, North-West University, Potchefstroom Campus, South Africa | McGill Community for Lifelong Learning, McGill University, Canada
E-mail: deo.nizonkiza@nwu.ac.za

Abstract

The present study explores academic vocabulary knowledge, operationalised through the Academic Word List, among first-year higher education students. Both receptive and productive knowledge and the proportion between the two are examined. Results show that while receptive knowledge is readily acquired by first-year students, productive knowledge lags behind and remains problematic. This entails that receptive knowledge is much larger than productive knowledge, which confirms earlier indications that receptive vocabulary knowledge is larger than productive knowledge for both academic vocabulary (Zhou 2010) and general vocabulary (cf. Laufer 1998, Webb 2008, among others). Furthermore, results reveal that the ratio between receptive and productive knowledge is slightly above 50%, which lends empirical support to previous findings that the ratio between the two aspects of vocabulary knowledge can be anywhere between 50% and 80% (Milton 2009). This finding is extended here to academic vocabulary; complementing Zhou's (2010) study that investigated the relationship between the two aspects of vocabulary knowledge without examining the ratio between them. On the basis of these results, approaches that could potentially contribute to fostering productive knowledge growth are discussed. Avenues worth exploring to gain further insight into the relationship between receptive and productive knowledge are also suggested.

Keywords: academic vocabulary, receptive knowledge, productive knowledge, collocations, vocabulary dimensions

1. Introduction

The topic of vocabulary has gained popularity since the 1990s (Read 2000), and the topic has been investigated from different angles. Among others, scholars were interested in the conceptualisation of vocabulary knowledge and in measuring it. Vocabulary knowledge was predominantly conceptualised in terms of dimensions. In this regard, Meara (1996) and Henriksen (1999) proposed dimensions of vocabulary knowledge, i.e. vocabulary size, referring to the words one understands; vocabulary depth, referring to how well the words are known; and receptive-productive dimension, referring to the relationship between the words understood and those that can be used. The dimensional approach was supported by many researchers

(cf. Zareva, Schwanenflugel and Nikolova 2005) and since then vocabulary has been measured in terms of dimensions.

The period in which the dimensional approach was adopted was characterised – among others – by testing students' comprehension (receptive knowledge) and use (productive knowledge), as well as the relationship between the two aspects. The immediate consequence of this focus on vocabulary was that it moved from the periphery to being considered as an important component of language (cf. e.g. Daller, Milton and Treffers-Daller 2007, Meara 2002, Zareva et al. 2005). Other major consequences arising from this focus, as observed by Nizonkiza and Van den Berg (2014), are determining the vocabulary to teach at different proficiency levels and using vocabulary tests as placement indicators. This resulted from a predictive relationship established between receptive vocabulary size and linguistic proficiency (Beglar 2010, Meara 1996, Meara and Buxton 1987, Meara and Jones 1988, Nation 1990). However, Nizonkiza and Van den Berg (2014) remarked that this progress was made on the basis of results from just one vocabulary dimension, which is vocabulary size measured receptively and by far the most widely investigated dimension (cf. Daller et al. 2007, Ishii and Schmitt 2009, Milton 2009, Read 2007, Schmitt, Ng and Garras 2011).

The productive component of vocabulary knowledge was also tested and attempts were made to establish a relationship between receptive and productive knowledge. An important finding from studies testing productive knowledge is that it may also predict overall linguistic proficiency (Laufer and Nation 1995, 1999, Meara and Fitzpatrick 2000). Regarding studies that attempt to explore the relationship between receptive and productive knowledge, findings show that learners' receptive vocabulary is larger than their productive vocabulary for both general (Laufer 1998, Laufer and Goldstein 2004, Laufer and Paribakht 1998, Melka 1997, Waring 1997, Webb 2005, 2008, Zhong and Hirsh 2009) and academic vocabulary (Zhou 2010). More importantly, there seems to be parallel growth between receptive and productive knowledge. In other words, the larger the receptive size of learners' vocabulary, the larger their productive vocabulary size (Laufer 1998, Laufer and Goldstein 2004, Laufer and Paribakht 1998, Melka 1997, Webb 2005, 2008, Zhong and Hirsh 2009, Zhou 2010).

Webb (2008: 79) claims that “receptive vocabulary size might give some indication of productive vocabulary size”. However, this may be called into question as the ratio between receptive and productive knowledge varies a great deal. While Waring's (1997) study points to a proportion between receptive and productive knowledge of about 50%, Milton (2009) maintains that the proportion between receptive and productive knowledge ranges between 50% and 80%. This clearly indicates that the relationship between receptive and productive knowledge has yet to be conclusively established (Nizonkiza and Van den Berg 2014).

Furthermore, the available literature shows that the ratio between receptive and productive knowledge may vary according to students' linguistic proficiency level or word frequency (cf. section 2.2). While these findings have great potential with regard to determining the exact relationship between receptive and productive knowledge, they have not been explored any further. I believe that this is a gap that should be bridged and the present study attempts to do so. It involves words of more or less the same frequency (academic vocabulary from Coxhead's Academic Word List (AWL)) and students from the same class, who are thus of comparable linguistic proficiency level. The present study takes up Zhou's (2010) idea to compare the state of receptive and productive knowledge of words from the AWL among English as Foreign

Language (EFL) students, which it complements by making estimates of the ratio between receptive and productive knowledge, and also extends it to English as Second Language (ESL) students. The following research questions are examined:

- What is the state of first-year ESL higher education students' receptive and productive knowledge of academic vocabulary?
- Does productive knowledge of academic vocabulary grow alongside receptive knowledge among ESL students and what is the proportion between them?

2. Related literature

Scholars consider the 1990s as a period in which vocabulary was tested in its own right (Nizonkiza and Van den Berg 2014). Vocabulary knowledge was conceptualised in terms of dimensions that still guide the testing of vocabulary today. Ideally, the tests discussed in this section should be linked to the vocabulary dimensions, but given the focus of the study – measuring receptive and productive knowledge of vocabulary, which is actually one dimension – this section only briefly reviews the testing of receptive and productive knowledge. For a recent overview of the testing of vocabulary under the dimensional approach to vocabulary and the related challenges, as well as new directions, readers are referred to Nizonkiza and Van den Berg (2014) and Nizonkiza and Ngwenya (2015). Moreover, given that the focus of this study is on academic vocabulary, the AWL will also be described.

2.1 Measuring receptive and productive knowledge of vocabulary

One of the important aspects of vocabulary knowledge is the link between receptive and productive knowledge. ESL/EFL researchers define receptive vocabulary as the vocabulary used for comprehension, while productive vocabulary is used for production (cf. Henriksen 1999, Zareva et al. 2005). This definition supports Gairns and Redman's (1986: 64) view that "receptive vocabulary" refers to "language items which can only be recognized and comprehended in the context of reading and listening material" while "productive vocabulary" refers to "language items which the learner can recall and use appropriately in speech and writing". Nation (2001) endorses this view in saying that receptive knowledge is associated with listening and reading tasks that require perception of the form of the word and its meaning, while productive knowledge is associated with speaking and writing. Webb (2008: 79), according to whom, "knowing students' receptive vocabulary size provides teachers with a gauge as to whether those students will be able to comprehend a text or a listening task, whereas knowing their productive vocabulary size provides some indication as to the degree to which students will be able to speak or write", holds the same view.

The receptive knowledge of vocabulary has been measured mainly through matching, multiple choice, and yes/no formats. The commonly used tests adopting the above formats include the Vocabulary Levels Test (VLT), the Vocabulary Size Test (VST), and the Eurocentres Vocabulary Size Test (EVST). The VLT (Nation 1983, 1990) is a matching test requiring test-takers to match words with their definitions. The test has been subject to revisions (Beglar and Hunt 1999, Schmitt, Schmitt and Clapham 2001) especially for validation purposes. The VST is a test designed by Nation and Beglar (2007) and validated by Beglar (2010). It adopts a multiple choice format where learners are presented with target words embedded in a non-defining context. They are instructed to choose the correct meaning from four options – the

correct meaning and three distractors. In both the VLT and the VST, the selection of target words is guided by frequency, which is considered as a key factor in knowing words. The difference between the two is that the VLT retains words from five word frequency bands and the academic vocabulary, while the VST considers 14 word bands from Nation's (2006) word list. With the advent of computers, words are now classified in terms of bands, based on the frequency with which they occur. Nation's (2006) classification, for instance, groups words in frequency bands of 1 000 words each. The initial frequency list consisted of 14 bands; but it has been updated and consists of 25 bands now. The VLT tests words from the 2 000-word, 3 000-word, 5 000-word, 10 000-word bands and the University Word List (Xue and Nation 1984) or the Academic Word List (Coxhead 2000).

The EVST adopts a yes/no format with frequency also being the selection criterion for target words. It is referred to as a checklist test (Read 2000, Schmitt 1994) and a variation of the multiple choice test as the same procedure is applied in selecting test items (Schmitt 1994). However, unlike the multiple choice test formats the test is administered electronically where test-takers are required to demonstrate if they know the word or not by pressing the 'yes' or 'no' button upon reading the word on a computer screen. In order to adjust the scores of test-takers who might overrate their knowledge, the list of target words contains non-words, i.e. one non-word item for every two real words. The EVST exists in different versions, i.e. Meara and Buxton (1987), Meara and Jones (1988), and Meara (1996) – all of which have been validated.

All the above tests have been standardised and can now serve as placement indicators. In other words, with empirical evidence establishing a relationship between the number of words known and overall linguistic proficiency, the above tests can be used to assign students to groups of different linguistic proficiency levels. They have been widely used for both research and pedagogical purposes with the VLT being the most widely used (Ishii and Schmitt 2009, Read 2007, Schmitt et al. 2001). In addition, the tests have allowed researchers to make estimates of the receptive knowledge needed for different activities. Laufer (1992), for instance, suggests that receptive knowledge of the most frequent 3 000 word families may allow students to understand unsimplified texts, which could account for 95% of a running text. The concept of 'word families' refers to head words and their family members, which include their inflections and derivations. For instance, the word family *access* has the following family members: *accessed, accesses, accessibility, accessible, accessing, inaccessible*.

Hirsh and Nation (1992) suggest a threshold of the 5 000 word families for unassisted reading for coverage of 98% of a running text. Hu and Nation (2000) propose the same coverage of 98% for unassisted reading. However, they suggest a much larger vocabulary size, which should be 8 000-9 000 words, echoing Nation's (2006) and Schmitt's (2008) suggestions. On the basis of what precedes, there appear to be minimal and optimal thresholds as highlighted by Laufer and Ravenhorst-Kalovski (2010). The latter should be 4 000-5 000 and about 8 000 word families, respectively for minimal threshold and optimal threshold. The minimal threshold may allow understanding of 95% of a running text, while the optimal threshold may allow understanding of 98% of a running text (cf. Laufer and Ravenhorst-Kalovski 2010, Nation 2006, Schmitt 2008, 2010, Schmitt and Schmitt 2012).

Productive knowledge consists of controlled productive knowledge on the one hand and free productive knowledge on the other. Controlled productive knowledge is measured by means of the Productive Vocabulary Levels Test (PVLVT). It is a productive variant of the VLT and was designed

by Laufer and Nation (1999) to test controlled productive ability. The latter is a cued recall where a sentential context embedding the target word is presented to test-takers. The target word is deleted, but with the first two letters provided and test-takers have to supply it. Free productive knowledge has been measured through lexical richness and association tasks mainly.

According to Schmitt (2010), Laufer and Nation's (1995) Lexical Frequency Profile (LFP) is one of the most widely used frequency-based free productive knowledge tests of vocabulary, while Meara and Fitzpatrick's (2000) Lex30 is the most frequently used association tasks test. The LFP measures lexical richness, which aims to evaluate the relative proportion of words a learner can use in free production. Test-takers are required to write an essay, the LFP of which is calculated in a computerised system. The latter weighs the number of words in each frequency level against the total word families in the piece of writing, and the more words from infrequent bands are used, the more proficient the learners are.

Lex30 consists of 30 words, all from Nation's (1984) first 1 000 most frequent words. They are presented to test-takers one at a time and test-takers are required to give at least three associates. Associates refer to the words test-takers think of upon seeing the target words. The words produced by test-takers are lemmatised through a computerised system, which reports the frequency of each word. Like the LFP, frequency is an important scoring criterion and only words from the 2 000-word band and beyond are given credit. Both the LFP and Lex30 have been validated and proved to discriminate between linguistic proficiency levels (cf. Laufer 2005 for the LFP and Fitzpatrick and Clenton 2010, and Walters 2012 for Lex30). They have been critiqued because they do not test collocations, which are an important aspect of production knowledge. None of the tests have been standardised, which limits the breadth of the generalisability of results from them (Nizonkiza and Van den Berg 2014).

2.2 The relationship between receptive and productive knowledge of vocabulary

As discussed in the previous section, receptive and productive knowledge constitute two important aspects of vocabulary knowledge. Examining the relationship between the two aspects has been the concern of researchers. For some scholars, receptive-productive knowledge should be viewed as a continuum between word comprehension and word use (cf. Henriksen 1999, Melka 1997). The underlying reason is that words are understood first and then pass from receptive to being productively used (Meara 1997, Melka 1997, Schmitt 2010). However, Laufer and Goldstein (2004: 405) remark that, "it is not clear how much knowledge is necessary for a word to move from passive to active status". For other scholars such as Meara (1990), receptive and productive knowledge represent totally different types of associations, where receptive knowledge of a word is activated by external stimuli while productive knowledge of a word is activated by other words.

Studies comparing receptive and productive knowledge (Laufer 1998, Laufer and Goldstein 2004, Laufer and Paribakht 1998, Melka 1997, Webb 2005, 2008, Zhong and Hirsh 2009, Zhou 2010) point to useful observations even though they remain inconclusive with regard to the exact relationship between receptive and productive knowledge (Nizonkiza and Van den Berg 2014). What is particularly striking is that they yielded results that are consistent even though they used different testing instruments. All of these researchers indeed pointed to the observation that receptive knowledge is larger than productive knowledge and that the larger the receptive knowledge, the larger the productive knowledge.

As far as the ratio between receptive and productive knowledge is concerned, different figures emerge. Waring (1997) pointed to a ratio between receptive and productive knowledge of about 50%, while Milton (2009) suggested a much wider gap between the two, which could range from 50% to 80%. This variation between receptive and productive knowledge echoes Laufer's (1998: 257) observation that "passive vocabulary size is considered to be larger than the active size even though no substantiated specification is provided as to how much larger it is". In her study, she found that the ratio between free productive knowledge and controlled productive knowledge may vary depending on linguistic proficiency level. In Laufer's (1998) terms, *passive* means receptive knowledge, while *active* refers to productive knowledge. Laufer (1998) found a ratio of 89% among 10th graders, which decreased to 73% among 11th graders, implying that the ratio widens at a higher level of proficiency. This means that more proficient students understand a lot more than they can actually use, if compared to less proficient students.

The ratio between the two may also depend on the frequency of the word bands from which the test items are selected. Webb (2008), for instance, found a ratio of 88%, 73% and 65% for word band one, two, and three respectively – word band one consisting of the most frequent words and word band three consisting of the least frequent words. This implies that the ratio between receptive and productive knowledge is wider at less frequent word bands. In other words, at more infrequent word bands, fewer of the words learners understand are readily available for productive use as opposed to words from frequent bands where most of the words understood can also be productively used. Webb (2008) further observed that given the parallel growth between receptive and productive knowledge, the latter can be predicted on the basis of the former. However, Nizonkiza and Van den Berg (2014) caution against taking this for granted on the grounds that the relationship between the two is not straightforward and does not seem to have been fully explored.

2.3 The Academic Word List

The Academic Word List (AWL) is a compilation of words frequently used in academic environments and which occur outside the 2 000-word band. It was developed by Coxhead (2000) from a corpus of 3.5 million running words from a wide range of written academic texts. The selection criteria were range and frequency, which refer to the number of disciplines in which a word appears and to the number of occurrences of the word, respectively.

The AWL consists of 570 word families, which account for about 10.0% of the words occurring in academic texts (Coxhead 2000, 2011). The AWL is the most widely academic vocabulary list used or referred to today (cf. Coxhead 2011, Durrant 2009, Nation 2001, Paquot 2007, Schmitt and Schmitt 2005). The AWL constitutes words students need to master in order to cope with the academic demands posed in higher education (Nation 1990). Researchers such as Coxhead (2000) and Nation (2001) suggest teaching these words explicitly to higher education students because they cannot be learned implicitly since they are neither part of the common vocabulary nor part of subject-specific discourse. Despite criticisms that the AWL may not constitute core academic vocabulary as initially intended (Gardner and Davies 2013, Hyland and Tse 2007) and the calls for developing subject-specific lists (Coxhead and Hirsh 2007, Durant 2009, 2014, Wang, Liang and Ge 2008), the AWL remains influential in testing, teaching and designing materials to teach for academic purposes (Coxhead 2011, Durrant 2014).

3. Methodology

3.1 The test battery

3.1.1 The collocation test

In order to measure productive knowledge, the present study administered a previously used collocation test (Nizonkiza 2014). The test was developed with target words selected from the AWL (Coxhead 2000) and collocations selected from the *Oxford collocations dictionary for students of English* (Lea, Crowther and Dignen 2002). The collocation dictionary defines collocations as “the way words combine in a language to produce natural-sounding speech and writing” (Lea et al. 2002: vii), which is the definition adopted in this study. The type of collocation measured is verb-noun and the aspect tested is known in the literature as controlled productive knowledge, which is defined as:

the ability to use a word when compelled to do so by a teacher or researcher, whether in an unconstrained context such as a sentence writing task, or in a constrained context such as a fill in task where a sentence context is provided and the missing target word has to be supplied (Laufer and Nation 1999: 37).

The nouns were selected through systematic random sampling (Babbie 1990). As the AWL consists of 10 sublists with 60 words each, except the 10th which consists of 30 words, the technique was to select six words per sublist for a more balanced distribution of the words covered in the list in terms of their frequency. The sampling ratio was 10 ($60 \div 6 = 10$) and the procedure was thus to select every 10th word from a random starting point. Whenever, the 10th word was not a noun, the next one was selected. As presented in section 2.3, the AWL was chosen for target word selection because it is a vocabulary list that is influential in designing course materials in academic and specific purposes contexts.

The collocation dictionary was used for selecting both the collocations and the sentential contexts in which words were embedded. It presents the collocations of target words following their (target words) meanings in cases of polysemous words, as well as the syntactic categories of the constituents of the collocations. For example, the word *conclusion* is a noun meaning (1) ‘opinion reached after considering the facts’ or; (2) ‘ending of something’. The collocations are presented following the above meanings. For the first meaning, for instance, the collocates are listed in the dictionary, following syntactic categories, such as adjective-noun (e.g. *valid conclusion*), verb-noun (e.g. *reach a conclusion*), etc.

Each noun was looked up in the collocation dictionary and the procedure was to retain the first verb-noun combination where an example sentence was provided. The frequency band of the verbs was cross-checked against Nation’s (2006) frequency list. Most of the verbs were found to belong to either the 1 000-word or the 2 000-word band. Whenever, a verb (collocate) was found to belong to an infrequent band, it was replaced. This echoes Gyllstad’s (2007) suggestion that for testing purposes, a collocate should be more frequent than, or at least as frequent as, the target word. It should be noted that the notion of frequency, as defined in Schmitt (2008), that the cut-off point for frequency should be placed at the 3 000-word band was adopted. This means that any verb (collocate) from the 4 000-word band and beyond was excluded from the sample. This

is the case, for instance, for the verb *enhance* in the sentence: ‘Getting the right qualifications will *enhance* your employment prospects’, where *enhance* (from the 4 000-word band) was replaced with *improve* that belongs to the 1 000-word band.

As Laufer and Nation (1999) suggest the verbs were deleted with only the first two letters provided. As appears in the example below, the instruction makes it clear that participants have to fill in the missing letters of the collocates. The marking pattern was to award 0 points for incorrect or no answer and 1 point for a correct answer; grammatical and spelling mistakes were not considered.

Instruction: Complete the underlined words in the sentences below.

Example: They ma..... a beautiful couple.

They make a beautiful couple.

3.1.2 The Vocabulary Levels Test

For receptive knowledge, the Vocabulary Levels Test (VLT) was adopted. It measures receptive knowledge and adopts a matching format where participants are instructed to match words to their definitions. Schmitt et al.’s (2001) version is the one adopted in this study. The test was chosen because it is the only one that selects academic vocabulary from the AWL. It consists of 30 items at each of the word bands. The words are presented in a column along with a column of definitions. The word column contains six words – three target words and three distractors – while the definition column contains three definitions. Test-takers are instructed to match a word with its definition. The items from the AWL were considered for this study.

3.2 Participants

Participants in this study were selected from the North-West University (NWU), Potchefstroom Campus (N = 204). They were a sample from a diverse population comprising different faculties and institutes at the NWU. All were first-year students taking the course Academic Literacy AGLE111 – a course taught in English. While these students were speakers of many native languages, English was their second language. Their average age ranged between 18 and 20. Students sat the two tests on two different days and were invited through their lecturer. Participants were specifically told that sitting the test was meant for research purposes and would not in any way affect their final mark in the course they were taking. It is worth noting that some of the students who did not write their names and those who did not sit both tests were excluded from the analysis.

4. Results

4.1 Description of items

Language-testing specialists suggest that the ideal test should be reliable, which means that it should measure what it was designed to measure. It should also discriminate between test-takers with different proficiency levels (cf. Alderson, Clapham and Wall 1995, Bachman 1990, Green 2013). In this study, the reliability – mainly measured by means of the Cronbach’s Alpha – was performed for both the receptive and productive tests. Reliability coefficients of .903 and .860, for receptive and productive tests respectively, were found,

which are high and fall within the acceptable range suggested by Pallant (2007), among others. According to Pallant (2007), a well-designed test should have a Cronbach's Alpha of at least .7.

In order to test the discriminating power of the test items, a Corrected Item Total Correlation (CITC) was performed. The CITC ranges from -1 to +1, where the higher the item's coefficient, the better the item discriminates between test-takers. Item coefficients were weighed against Ebel's (1979) scale. The latter has the following four levels: .40 and higher indicates definitely good items; .30 to .39 indicates reasonably good items; .20 to .29 indicates items in need of improvement; and below .19 indicates items to be revised or eliminated. The CITC results are summarised in Table 1 for the receptive test and Table 2 for the productive test. As Table 1 shows, the CITC reveals that 90% of the items function well; only 10% of the items are categorised as items which should be revised or eliminated. This high level of reliability can be accounted for by the fact that Schmitt et al.'s (2001) VLT of which the AWL items were retained for this study is a standardised test which is also used as a placement indicator.

Table 1: Receptive test Corrected Item Total Correlation on Ebel's scale

CITC	.40 and higher	.30 to .39	.20 to .29	Below .19
Item number	1, 6, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	9, 11, 15	4, 5, 8, 10	2, 3, 7
Total items	20 (66.66%)	3 (10%)	4 (13.33%)	3 (10%)

The CITC results for the productive component presented in Table 2 show that about 70% of the items as judged on Ebel's scale function well. The remaining items are categorised as items that should be revised or eliminated. But overall, as 70% of the items can be considered as discriminating fairly well between students with different proficiency levels, the test can be used for the purpose for which it was designed.

Table 2: Productive test Corrected Item Total Correlation on Ebel's scale

CITC	.40 and higher	.30 to .39	.20 to .29	Below .19
Item number	4, 9, 28, 33, 45, 46	3, 7, 18, 19, 20, 26, 37, 38, 39, 41, 43, 43, 44, 47, 48, 49, 50, 51, 54, 57, 59	5, 6, 11, 16, 21, 22, 23, 25, 27, 29, 30, 31, 32, 34, 36, 42, 56	1, 2, 8, 10, 12, 13, 14, 15, 17, 24, 35, 40, 52, 53, 55, 58, 60
Total items	6 (10%)	19 (31.66%)	17 (28.33%)	17 (28.33%)

4.2 First-year students' receptive and productive academic vocabulary knowledge

The first research question examined in this study is the state of receptive and productive knowledge of academic vocabulary among first-year students at the Potchefstroom campus of the NWU. The receptive knowledge was estimated by totalling the scores from the VLT out of 30. They were then weighed against Schmitt et al.'s (2001) threshold that a vocabulary band is mastered if the score for that band is at least 24 out of 30, i.e. 80%. The average score was found to be 27.45, with only 11.38% of the participants, i.e. 14 participants out of 123, scoring below 24. This finding clearly shows that an overwhelming majority of the participants achieved the threshold score, which suggests that the receptive knowledge of academic vocabulary is readily acquired by first-year students. The productive knowledge was explored by analysing scores from the collocation test, which was marked out of 60. The test scores were averaged and show a mean score of 30.44. This score is much lower than the expected one which should be 80%, i.e. 48 out of 60 in this case if we consider Schmitt's cut-off point above. Only about 4.06%, i.e. 5 students out of the total of 123 participants achieved this threshold score. These findings answer the first research question about the state of the receptive and productive knowledge of academic vocabulary of first-year students.

4.3 The relationship between receptive and productive knowledge of academic vocabulary

The second research question examined in this study is the relationship between receptive and productive knowledge of academic vocabulary among first-year students in higher education. This question was answered by comparing the mean scores achieved on both tests – receptive and productive. For comparative purposes, the productive test scores were first averaged out of 30 and then a paired-samples t-test was conducted. As illustrated in Table 3, results indicate that the mean scores of receptive ($M = 27.45$, $SD = 4.26$) and productive ($M = 14.41$, $SD = 3.58$) academic vocabulary knowledge differ. As can also be seen from Table 3, this difference in mean scores was found to be statistically significant at the 0.05 of significance level ($t = 33.432$, $df = 122$, $p = 0.000$). These results imply that first-year students' academic vocabulary receptive knowledge is significantly larger than their productive knowledge.

Table 3: Academic vocabulary receptive and productive knowledge paired

		Mean	N	SD	Difference	t	df	Sig. (2-tailed)
Pair 1	Productive	14.41	123	3.58	13.04	33.43	122	0.000
	Receptive	27.45	123	4.26				

These results were analysed further by determining the proportion between receptive and productive knowledge by calculating the ratio between the two, which amounts to about 51.85%. The relationship between receptive and productive knowledge was also examined by performing a Pearson correlation between scores from the two tests. As Figure 1 indicates, the correlation is positive and linear with a correlation coefficient of 0.404**, statistically significant with a p-value of 0.000, significant at the 0.01 level, 2-tailed. According to Cohen (1988), the strength of correlation might be interpreted as (i) small (.10 to .29); (ii) medium (.30 to .49); and (iii) large (.50 to 1) – the higher the better. In this case, the correlation is medium and significant, which brings another argument in favour of a strong relationship between receptive and productive knowledge of academic vocabulary.

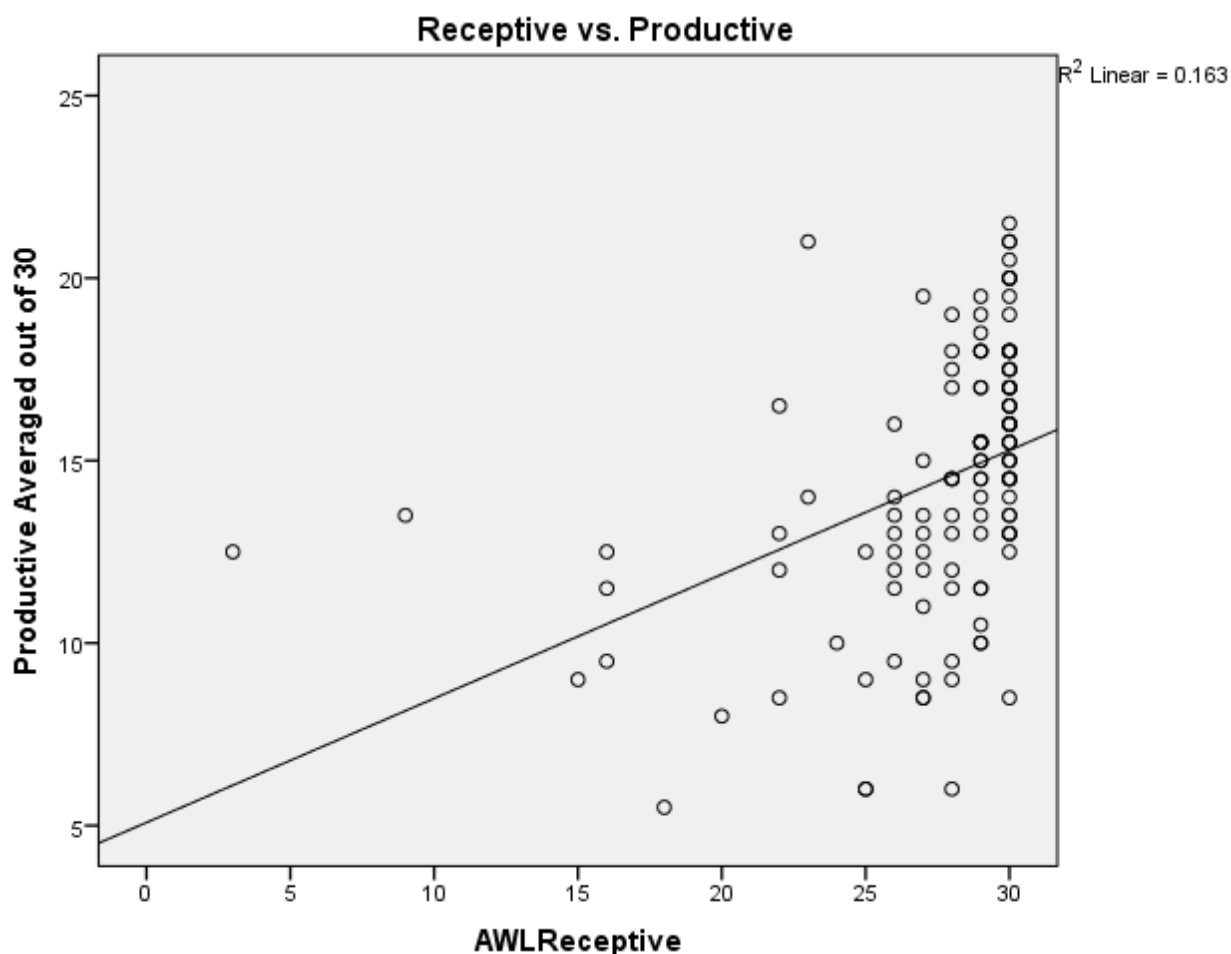


Figure 1: Receptive and productive knowledge correlated

5. Discussion

The present study examined the state of academic vocabulary, operationalised through the AWL, among first-year students at NWU. It set out to investigate both receptive and productive vocabulary knowledge, as well as the relationship between the two aspects. Results show that, while students' productive knowledge is rather low, their receptive knowledge is readily acquired with an overwhelming majority of the students achieving a score of 80% and higher. This entails that receptive knowledge of academic vocabulary is larger than productive knowledge, which confirms previous findings on academic vocabulary (Zhou 2010) and general vocabulary (Laufer 1998, Laufer and Goldstein 2004, Laufer and Paribakht 1998, Melka 1997, Waring 1997, Webb 2005, 2008, Zhong and Hirsh 2009). Results from the productive component also support previous studies, which demonstrated that collocations cause problems for both ESL and EFL students, even at an advanced level – holding for both general (Laufer and Waldman 2011, Nesselhauf 2005) and academic vocabulary (Nizonkiza 2014). Consequently, the call to adopt a more production-oriented approach to teaching academic vocabulary in general (Paquot 2007, Zhou 2010), and collocations in particular (Lewis 2000, Nation and Chung 2009, Nizonkiza and Van de Poel 2014), especially collocations of academic vocabulary to higher education students (Nizonkiza 2014), seems to be warranted.

Furthermore, the ratio between receptive and productive knowledge was found to be slightly above 50%; falling within Milton's (2009) range of 50% and 80%. However, it should be noted that this ratio could be much lower as the productive component measured in this study is controlled rather than free productive knowledge, which generally lags behind (Laufer 1998, Zhong and Hirsh 2009). This lends support to Laufer's (1998: 265) explanation of the absence of a significant correlation between free productive and controlled productive on the one hand and between free productive and receptive knowledge on the other, which she phrases as follows: "learners who could recognise more words than other learners and produce them if forced to, were not necessarily those who would use more infrequent vocabulary in free expression".

Having used words from the same frequency band and having involved students from the same class and thus of comparable proficiency level, this study indicates that the ratio between receptive and productive knowledge could be around 50%. Even though this needs further research for clarification, it can be posited that the variation towards 80% could be attributed to proficiency of students, or word frequency band, or both.

6. Conclusion

The present study attempted to gauge the state of receptive and productive knowledge of academic vocabulary of first-year students in a higher education institution and the proportion between the two aspects of vocabulary knowledge. Participants were administered a receptive and a productive test the scores of which were subsequently compared. Results indicate that the particular cohort of higher education first-year students seem to have mastery of the academic vocabulary at the receptive level. However, they have very little control of the productive academic vocabulary, and the relationship between the two indicates that the larger the size of receptive vocabulary, the larger the productive knowledge, while the ratio between them is slightly over 50%.

These results answer the initially raised questions, but give rise to more questions, which should be explored further for our understanding of the relationship between receptive and productive knowledge. First of all, this study compared receptive and controlled productive knowledge of academic vocabulary. As discussed above, research evidence indicates that free productive knowledge lags behind controlled productive knowledge, which in turn lags behind receptive knowledge. Yet this relationship still has to be confirmed at the level of academic vocabulary. Secondly, a follow-up study is needed to determine the exact role of linguistic proficiency level on the one hand and the frequency of words on the other with regard to the ratio between receptive and productive knowledge. This may contribute towards testing the assumption made in the discussion above that these two aspects may influence the ratio between receptive and productive knowledge of vocabulary. Thirdly, it should be acknowledged that this study involved a small sample of participants and the conclusions reached can by no means be generalised. Thus extending it to a larger sample may shed light on the link between receptive and productive knowledge. Fourthly, adopting a production oriented approach to teaching academic vocabulary and undertaking a study in a pre- and post-experimental design may shed some light on the slow growth of productive knowledge. The latter can indicate, among others, whether the slow growth is due to the nature of collocation or the teaching approaches adopted today, which seem to be predominantly receptive.

References

- Alderson, J.C., C. Clapham and D. Wall. 1995. *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Babbie, E.R. 1990. *Survey research methods*. California and Belmont: Wordsworth Company.
- Bachman, L.F. 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Beglar, D. 2010. A Rasch-based validation of the Vocabulary Size Test. *Language Testing* 27(1): 101–118.
- Beglar, D. and A. Hunt. 1999. Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing* 16(2): 131–162.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.
- Coxhead, A. 2000. A new Academic Word List. *TESOL Quarterly* 34(2): 213–239.
- Coxhead, A. 2011. The Academic Word List 10 years on: Research and teaching implications. *TESOL Quarterly* 45(2): 355–362.
- Coxhead, A. and D. Hirsh. 2007. A pilot science word list for EAP. *Revue Française de Linguistique Appliquée* xii(2): 65–78.
- Daller, H., J. Milton and J. Treffers-Daller (Eds.) 2007. *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.
- Durrant, P. 2009. Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes* 28(3): 157–169.
- Durrant, P. 2014. Discipline and level specificity in university students' written vocabulary. *Applied Linguistics* 35(3): 328–356.
- Ebel, R.L. 1979. *Essentials of educational measurement*. New Jersey: Prentice Hall.
- Fitzpatrick, T. and J. Clenton. 2010. The challenge of validation: Assessing the performance of a test of productive ability. *Language Testing* 27(4): 538–555.
- Gairns, R. and R. Redman. 1986. *Working with words: A guide to teaching and learning vocabulary*. Cambridge: Cambridge University Press.
- Gardner, D. and M. Davies. 2013. A new Academic Vocabulary List. *Applied Linguistics*: 1–24.
- Green, R. 2013. *Statistical analyses for language testers*. Basingstoke: Palgrave Mcmillan.

- Gyllstad, H. 2007. Testing English collocations. Unpublished PhD dissertation. Lund University. Available online: <https://lup.lub.lu.se/luur/download?func=downloadFile&recordOid=599011&fileOid=2172422> (Accessed 25 November 2014).
- Henriksen, B. 1999. Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21(2): 303–317.
- Hirsh, D. and I.S.P Nation. 1992. What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language* 8(2): 689–696.
- Hu, M.H. and I.S.P Nation. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1): 403–430.
- Hyland, K. and P. Tse. 2007. Is there an “academic vocabulary”? *TESOL Quarterly* 41(2): 235–253.
- Ishii, T. and N. Schmitt. 2009. Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal* 40(1): 5–22.
- Laufer, B. 1992. Reading in a foreign language: How does L2 lexical knowledge interact with the reader’s general academic ability? *Journal of Research in Reading* 15(2): 95–103.
- 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. Lexical frequency profiles: From a Monte Carlo analysis to the real world, a response to Meara 2005. *Applied Linguistics* 26(4): 582–588.
- Laufer, B. and Z. Goldstein. 2004. Testing vocabulary knowledge: Size, strength, and Computer adaptiveness. *Language Learning* 54(3): 399–436.
- Laufer, B. and I.S.P Nation. 1995. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16(3): 307–322.
- Laufer, B. and I.S.P Nation. 1999. A vocabulary size test of controlled productive ability. *Language Testing* 16(1): 33–51.
- Laufer, B. and S. Paribakht. 1998. The relationship between passive and active vocabularies: Effects of language learning contexts. *Language Learning* 48(3): 365–391.
- Laufer, B. and G.C. Ravenhorst-Kalovski. 2010. Lexical threshold revisited: Lexical text coverage, learners’ vocabulary size and reading comprehension. *Reading in a Foreign Language* 22(1): 15–30.
- Laufer, B. and T. Waldman. 2011. Verb-noun collocations in second language writing: A corpus analysis of learners’ English. *Language Learning* 61(2): 647–672.

- Lewis, M. 2000. *Teaching collocations: Further development in the Lexical Approach*. Hove: Language Teaching Publications.
- Meara, P. 1990. A note on passive vocabulary. *Second Language Research* 6(2): 150–154.
- Meara, P. 1996. The dimensions of lexical competence. In G. Brown, K. Malmkjaer and J. Williams (Eds.) *Competence and performance in language learning*. Cambridge: Cambridge University Press. pp. 35–53.
- Meara, P. 1997. Towards a new approach to modelling vocabulary acquisition. In N. Schmitt and M. McCarthy (Eds.) *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press. pp. 109–121.
- Meara, P. 2002. The rediscovery of vocabulary. *Second Language Research* 18(4): 393–407.
- Meara, P. and B. Buxton. 1987. An alternative to multiple choice vocabulary tests. *Language Testing* 4(2): 142–154.
- Meara, P. and T. Fitzpatrick. 2000. Lex30: An improved method of assessing productive vocabulary in an L2. *System* 28(1): 19–30.
- Meara, P. and G. Jones. 1988. Vocabulary size as a placement indicator. In P. Grunwell (Ed.) *Applied linguistics in society*. London: Centre for Information and Language Teaching and Research. pp. 80–87.
- Melka, F. 1997. Receptive versus productive vocabulary. In N. Schmitt and M. McCarthy (Eds.) *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press. pp. 84–102.
- Milton, J. 2009. *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Nation, I.S.P. 1983. Testing and teaching vocabulary. *Guidelines* 4(1): 12–25.
- Nation, I.S.P. 1984. *Vocabulary lists*. Victoria University of Wellington, English Language Institute, Wellington, New Zealand.
- Nation, I.S.P. 1990. *Teaching and learning vocabulary*. Rowley, MA: Newbury House.
- Nation, I.S.P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. 2006. How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63(1): 59–82.
- Nation, I.S.P. and D. Beglar. 2007. A vocabulary size test. *The Language Teacher* 31(7): 9–13.

- Nation, I.S.P. and T. Chung. 2009. Teaching and testing vocabulary. In M.H. Long and C.J. Doughty (Eds.) *The handbook of language teaching*. Malden: Wiley-Blackwell. pp. 543–559.
- Nesselhauf, N. 2005. *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.
- Nizonkiza, D. 2014. The relationship between productive knowledge of collocations and academic literacy among tertiary level learners. *Journal for Language Teaching* 48(1): 149–171.
- Nizonkiza, D. and T. Ngwenya. 2015. Challenges of testing deep word knowledge of vocabulary: Which path to follow? *Journal for Language Teaching* 49(1): 223–253.
- Nizonkiza, D. and K. Van de Poel. 2014. Teachability of collocations: The role of word frequency counts. *Southern African Linguistics and Applied Language Studies* 32(1): 301–316.
- Nizonkiza, D. and K. Van den Berg. 2014. Dimensional approach to vocabulary testing: What can we learn from past and present practices? *Stellenbosch Papers in Linguistics Plus* 43: 1–14.
- Pallant, J. 2007. *SPSS survival manual*. Buckingham and Philadelphia: Open University Press.
- Paquot, M. 2007. Towards a productively-oriented Academic Word List. In J. Walinski, K. Kredens, and S. Gozdz-Roszkowski (Eds.) *Corpora and ICT in language studies. PALC 2005*. Frankfurt am main: Peter Lang. pp. 127–140.
- Read, J. 2000. *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. 2007. Second language vocabulary assessment: Current issues and new directions. *International Journal of English Studies* 7(2): 105–125.
- Schmitt, N. 1994. Vocabulary testing: Questions for test development with six examples of tests of vocabulary size and depth. *Thai TESOL Bulletin* 6(2): 9–16.
- Schmitt, N. 2008. Review article: Instructed second language vocabulary learning. *Language Teaching Research* 12(3): 329–363.
- Schmitt, N. 2010. *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.
- Schmitt, N. and D. Schmitt. 2005. *Focus on vocabulary: Mastering the Academic Word List*. London: Longman.
- Schmitt, N. and D. Schmitt. 2012. *A reassessment of frequency and vocabulary size in L2 vocabulary teaching. Plenary speech*. Available online: <http://journals.cambridge.org> (Accessed 25 March 2014).
- Schmitt, N., C.J.W. Ng and J. Garras. 2011. The word associates format: Validation evidence. *Language Testing* 28(1): 105–126.

Schmitt, N., D. Schmitt and C. Clapham. 2001. Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing* 18(1): 55–88.

Walters, J.D. 2012. Aspects of validity of a test of productive vocabulary: Lex30. *Language Assessment Quarterly* 9(2): 172–185.

Wang, J., S.L. Liang and G.C. Ge. 2008. Establishment of a medical Academic Word List. *English for Specific Purposes* 27(4): 442–58.

Waring, R. 1997. Graded and extensive reading. Questions and answers. *The Language Teacher* 27(5): 9–12.

Webb, S. 2005. Receptive and productive vocabulary learning: The effect of reading and writing on word knowledge. *Studies in Second Language Acquisition* 27(1): 33–52.

Webb, S. 2008. Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition* 30(1): 79–95.

Xue, G. and I.S.P. Nation. 1984. A university word list. *Language Learning and Communication* 3(1): 215–229.

Zareva, A., P. Schwanenflugel and Y. Nikolova. 2005. Relationship between lexical competence and language proficiency. *Studies in Second Language* 27(4): 567–595.

Zhong, H. and D. Hirsh. 2009. Vocabulary growth in an English as a foreign language context. *University of Sydney Papers in TESOL* 4: 85–113.

Zhou, S. 2010. Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners. *Asian Social Science* 6(10): 14–19.

Appendix: Collocation test

Productive Vocabulary Test

Name:

Native language:

Date:

Level of study (year):

Start hour:

Faculty:

End hour:

University:

Instruction: Complete the underlined words in the sentences below.

Example: They ma..... a beautiful couple.

They make a beautiful couple.

1. Villagers get together every year to ke..... this old tradition alive.
2. Institutions have to ex..... appropriate contexts in which to present examples of language in use for the children.
3. In order to fight against terrorism, the UN agreed on plans to res..... the export of arms to certain countries.
4. This evening, we need to ad..... the issue of legalisation of soft drugs.
5. She went on to ex..... the principle behind what she was doing.
6. We have to con..... many aspects of pollution in order to better tackle it.
7. If you do not have a regular income, you may be unable to ob..... credit.
8. It is difficult to ju..... the impact of the changes on employment patterns.
9. The latest developments will hardly af..... the perception of the crisis by the public.
10. The family will es..... temporary residence in the manor house.
11. They had to pe..... an in-depth analysis of the results.
12. Investigators are likely to ad..... a set of theories about the princess's death.
13. The schoolplanned to in..... comments from parents about the new curriculum.
14. We must make a real effort to pr..... cooperation between universities and industry.
15. They have to of..... a basic framework of ground rules for discussions.
16. Use enough gravel to fo..... a layer about 50mm thick.
17. The food shortage is likely to re..... crisis proportion.
18. She failed to co..... the task she had been set.
19. The new computer can al..... access to all the files.
20. Such a game may re..... great concentration.

21. Many developing countries hope to ac..... their goals of providing free primary education to everyone.
22. It is hoped that the new scheme will cr..... jobs in the region.
23. Society evolved to en..... a technological phase.
24. The government will re..... new statistics on the cost of living.
25. He was advised to at..... the police academy.
26. We need to ma..... contact with the organisation although it may be difficult after many years.
27. Banks will seek to re..... their exposure to risks.
28. Higher productivity has enabled them to im..... their profit margin.
29. Any surgery may de..... great precision.
30. Hotels that di..... this symbol offer activities for children.
31. Students have demonstrated that they should re..... big allocation for books.
32. The qualification should in..... my capacity to earn more.
33. It is good to co..... experts for a balanced diet.
34. Se..... the index to find the address of the data file!
35. The president aimed to co..... key ministries and reshuffled his Cabinet.
36. These criteria were used to de..... the scope of the curriculum.
37. Analysts think that the British government should pr..... aid to the area.
38. The local council will su..... new equipment for the playground.
39. You need a special password to ac..... this file.
40. Scientists should de..... technological innovations to save more energy.
41. A public outcry was needed to se..... her release from detention.
42. The results of the experiment su..... the thesis.
43. Doctors had to re..... his appendix.
44. The country needs to ra..... hard currency to pay for its oil imports.
45. We have to fo..... the safety guidelines laid down by the government.
46. Getting the right qualifications will im..... your employment prospects.
47. We could not meet because of the strike and had to ar..... a new schedule.
48. For your travel, find someone to take you to the airport or hi..... a vehicle.
49. It is the duty of the local community to pr..... accommodation for the homeless.
50. The prime minister seemed anxious to av..... controversy about these appointments.
51. Torrential rain can ca..... erosion on the hillside.
52. The management has to ac..... mediation, otherwise the strike will never be resolved.
53. It is possible to in..... further refinement on previous methods.
54. Willing volunteers needed to bu..... teams of helpers to carry everything in.
55. This is a new drug used to tr..... depression.
56. They demanded the right to ho..... peaceful assemblies.
57. It is up to the user to en..... the integrity of the data they enter.
58. It is difficult to re..... the enormity of the tragedy.
59. For the annual Thesis Award, the school had to se..... a panel of scientists from different universities.
60. His designs in..... reluctance to conform to fashion.