

This item is the archived peer-reviewed author-version of:

Towards modelled testing of productive knowledge of collocations

Reference:

Nizonkiza Deogratias, Van de Poel Kris, Van Dyk Tobie, Louw Henk.- Towards modelled testing of productive knowledge of collocations Southern African Linguistics & Applied Language Studies - ISSN 1607-3614 - 38:4(2020), p. 307-322 Full text (Publisher's DOI): https://doi.org/10.2989/16073614.2020.1858123 To cite this reference: https://hdl.handle.net/10067/1744410151162165141

uantwerpen.be

Institutional repository IRUA

Towards modelled testing of productive knowledge of collocations

Dr. Déogratias Nizonkiza

School of Journalism, Writing, and Media, University of British Columbia (UBC) — Canada and School of Languages, North-West University (Potchefstroom Campus) — South Africa E-mail: <u>deogratias.nizonkiza@ubc.ca</u>

Prof. Dr. Kris Van de Poel Department of Linguistics, University of Antwerp —Belgium and School of Languages, North-West University (Potchefstroom Campus) —South Africa E-mail: <u>kris.vandepoel@uantwerpen.be</u>

Prof. Tobie Van Dyk School of Languages, North-West University (Potchefstroom Campus) E-mail: <u>Tobie.vandyk@nwu.ac.za</u>

Dr. Henk Louw School of Languages, North-West University (Potchefstroom Campus) E-mail: <u>Henk.Louw@nwu.ac.za</u>

Abstract

Accurate use of collocations is seen as a measure of a learner's proficiency in a second/foreign language. However, research is not conclusive as to how to best test productive knowledge of collocations, and more fundamental so, there is no agreement on how vocabulary size, i.e. word comprehension, influences collocation production within constrained or unconstrained context. This paper elaborates on a battery of tests aimed at examining these issues. The tests were presented to both English as Foreign Language (EFL) and English as Second Language (ESL) students. The findings indicate that constrained and unconstrained contexts may represent different levels of productive knowledge; casting doubt on the definition of productive knowledge (cf. Laufer and Nation, 1999) collapsing the two –constrained and unconstrained– into one construct. Furthermore, both levels are influenced by the size of a learner's receptive knowledge of collocations and vocabulary size which is extended here to include productive knowledge. On the basis of these results, we argue that a new definition for "productive knowledge of collocations" is in order, where productive knowledge is categorised according to four levels, based on the amount of context provided to the learner.

Key words: collocations; productive knowledge of collocations; vocabulary size; English as Foreign/Second Language students; controlled productive knowledge; free productive Knowledge

1. Introduction

Collocations, defined by Paquot and Granger (2012, p. 136) as "lexically constrained combinations that allow for limited substitution within a particular grammatical construction (e.g., verb-object, adverb-adjective, or adjective-noun)", have been at the centre of research attention for nearly three decades now. By 'limited substitution', Paquot and Granger mean that, in English for instance, while *perform a task* or *do a task* are acceptable combinations, **make a* task is an incorrect combination. This limited substitution makes collocations challenging for second language (L2) and foreign language (FL) learners even at advanced learning stages (Laufer and Waldman 2011, Nesselhauf 2005, Wray 2002). Another reason why collocations prove to be challenging for L2 and FL learners is that they do not seem to pose a problem for comprehension (collocations are semantically transparent) and thus go unnoticed as problematic for learning by both learners and teachers (Biskup 1992, Gouverneur 2008, Henriksen and Stenius 2009, Laufer and Waldman 2011, Paquot 2008). The apparent contradiction arising here is that research evidence indicates that even though collocations pose a considerable challenge for L2/FL learners, they are important for these learners in more ways than one. It is widely accepted today that language is formulaicⁱ in nature (Pawley and Syder 1983, Wray 2002, Wray and Perkins 2000). Furthermore, both L2 and FL researchers agree that collocations constitute not only an important component of vocabulary, but they also characterise nativelike fluency (Barfield and Gyllstad 2009, Pawley and Syder 1983, Wray 2002). Collocations signal identity, reduce processing efforts (Pawley and Syder 1983, Wray 2002), and characterise maturity in speaking (Boers, Eyckmans, Kappel, Stengers and Demecheleer, 2006) and in writing (Gledhill 2000, Paquot 2008). Researchers further agree that mastery of collocations could be an indicator of overall proficiency among L2/FL learners. This relationship has been established for both receptive (Authors 2011a, Eyckmans 2009, Gyllstad 2007, 2009, Keshavarz and Salimi 2007) and productive knowledge (Authors 2011b, 2012, Boers et al. 2006, Bonk 2001, Eyckmans et al. 2004, Eyckmans 2009, Gitsaki 1999).

While this relationship seems to have been conclusively established at the receptive level with Gyllstad's (2009) standardised test, it remains inconclusive at the productive level. None of the productive tests used so far (see section 2.2 for details) has been standardised. The available

literature indicates that scholars have measured productive knowledge of collocations from two angles (Gyllstad 2007). The first one involves analysing students' essays compiled in corpora, which is known as the corpus-driven approach. The second approach is known as the experimental technique and involves selecting items and presenting them to the learners in the form of elicitation techniques: fill-in-the-blanks, translation, and cloze procedures. Corpus-driven studies make inventories of errors and categorise them in an attempt to account for their causes; while experimental studies aim to measure how well learners perform on pre-determined items. The major drawback in corpus driven studies is that unguided production might be problematic in the sense that non-natives may come up with wildly varying and deviant combinations (cf. Laufer and Nation 1999, Paquot and Granger 2012, Siyanova and Schmitt 2008), which makes it difficult to draw valid conclusions for wider applications. Experimental studies test predetermined words and are more controlled. Thus, for pedagogical purposes, they may be helpful in gaining better insights into the difficult nature of collocations (Authors 2012). However, experimental studies have been critiqued because they cannot tell what learners can produce if not prompted (Meara and Fitzpatrick 2000).

Both corpus-driven and experimental studies point to interesting results with regard to possible causes of collocation errors, their major types, and especially the extent to which knowing collocations can predict overall linguistic proficiency. However, the results from these studies do not lead to any conclusive pedagogical implications, which some scholars blame on the absence of a clear construct of what productive knowledge entails (Daller, Milton and Treffers-Daller 2007, Milton 2009, Nation 2007). For instance, Daller et al. (2007) are of the opinion that one of the challenges to measure productive knowledge in a standardised manner is a problem of definition.

In the literature, we can roughly distinguish between free productive knowledge, such as essay writing (cf. Laufer and Nation 1995) and controlled productive knowledge (cf. Laufer and Nation 1999). Free productive knowledge has been measured through a corpus-driven approach while controlled productive knowledge has been measured mainly through elicitation techniques. The problem with the corpus-driven approach is that it does not offer any means for quantifying collocations known by participants. The elicitation technique (controlled production)

4

seems to have mainly adopted Laufer and Nation's (1999) definition, which some scholars (e.g. Meara and Fitzpatrick 2000, Read 2000, Schmitt 2010) call into question especially with regard to unconstrained and constrained contexts. We depart from the definition of controlled productive knowledge, which Laufer and Nation (1999, p. 37) put in the following terms:

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.

We put this definition to the test, and the fundamental question we have is whether unconstrained and constrained contexts as defined above really refer to the same construct. We propose measuring collocations under three testing conditions (see details in section 3.2.2) which we believe represent unconstrained as well as constrained contexts. When occurring in an unconstrained context, words are presented in isolation and learners asked to produce the collocations. A context can be constrained when target words are embedded in a sentential context where the collocates have to be supplied as a whole or are prompted by the first two letters.

In this paper we attempt to answer the following questions: Do unconstrained and constrained contexts represent the same construct of productive knowledge of collocations? Furthermore, given that collocations are an aspect of vocabulary knowledge and that a larger vocabulary size entails a larger understanding of collocations (Gyllstad 2007, 2009), our question is whether this also holds for productive knowledge. More specifically, we ask the following question: Does possessing a large vocabulary size lead to knowing more collocations? This echoes Schmitt's (2010: 241) alternative approach to testing productive knowledge of vocabulary in his suggestion to test both receptive and productive knowledge simultaneously or consecutively and thus link the results for "a greater understanding" or interpret them "in an integrated manner" (Authors 2014, p. 55) to better capture the extent to which the two aspects are related. It is hoped that answering these questions will contribute towards our understanding of productive knowledge of collocations and provide some more insights into their testing which we think should be modelled.

2. Related literature

2.1. Definition of collocations

Collocations have been researched from different angles, which could be the reason why a consensus with regard to a clear and commonly agreed upon definition of collocations is yet to be reached (Eyckmans 2009, Nesselhauf 2005). One of the perspectives under which collocations have been investigated is the frequency-based approach (cf. among others, Firth 1957, Halliday 1966, Sinclair 1991). Scholars adopting this approach suggest frequency of cooccurring words as the main criterion on the basis of which collocations should be defined. This is also referred to in the literature as the statistical approach to collocations or the distributional approach (Granger and Paquot 2008, Gouverneur 2008). Lewis's (1997, p. 8) definition that collocations are "...the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency" is a good example. This approach is attributed to Firth who is regarded by many as having pioneered research on collocations and having made them known as a linguistic phenomenon (Gyllstad, 2007). Firth's (1957, p. 179) famous quotation reads as follows: "you shall know a word by the company it keeps", which is true in more ways than one.

Other proponents of this approach include Halliday (1966) who introduced the notions of "node", "collocate", and "span". These concepts respectively refer to the target word, the cooccurring word, and the space between the collocation constituents. In the following example, *Funding from the World Bank lends credibility to the project; credibility* is the node while *lends* is the collocate. In this example, the two collocation constituents follow each other. However, in the following example: *BBC backing for the scheme will enhance its credibility*; there is one orthographic word separating the collocation constituents and the span is thus two. The notions of node, collocate, and span are Halliday's addition to Firth's definition. Halliday (1966) did not, however, specify the space within which collocation constituents should occur, which Sinclair (1991), another proponent of this approach, did. He defined collocations as occurrence of two words or more within a short space in a text; a space which should be about four words.

Another popular approach to collocations is the phraseological approach (e.g., Benson, Benson and Ilson 1986, 1997, Cowie 1998, Howarth 1998, Nesselhauf 2005). Proponents of this

approach consider the syntactic categories of the co-occurring words, transparency in meaning, as well the extent to which the substitution of one collocate with another(s) is restricted (Gyllstad 2007, Nesselhauf 2005). From this perspective, collocations involve co-occurrence of words from certain syntactic classes —in a particular language— which should be acceptable. In English for instance, verbs can combine with nouns both in the verb-noun (e.g. *analyse results*) and nounverb (e.g. *results demonstrate*) combinations, but they hardly if at all combine with determiners or articles. Many scholars endorsing this view of collocations suggest that collocations should be transparent in meaning rather than opaque, even though some of them include idioms in the category of collocations. This approach is what Nesselhauf (2005) refers to as a scalar analysis of collocations. The latter suggests placing collocations on a continuum from totally free combinations to idioms. The view that is supported by many scholars adopting this approach, however, is to exclude totally free combinations and idioms from a definition of collocations.

Substitution of one collocate with another one is also an important criterion for defining collocations under this perspective. For Wray (2000, p. 474), substitution finds explanation in the fact that some lexical items "may be entirely open to any semantically plausible member of the word class, or may be subject to collocational restrictions, as with *pay attention* and *take care*, but not **take attention* or **pay care*". In Authors' (2011: 118) terms, substitution is possible although restricted, which is exemplified as follows: "In *strong wind*, *strong* can be interchangeably used with other adjectives like *terrible*, *fierce*, *moderate* etc. but not with *heavy*, which clearly indicates that substitution is possible, but restricted".

In addition to these two clearly distinct approaches, a new trend adopting a conciliatory tone, seems to be emerging; which accepts elements from the other two approaches (cf. Granger and Paquot 2008, Gyllstad 2007, Nesselhauf 2005). A definition by Nattinger and DeCarrico (1992, p. 20) can illustrate this point:

.... a collocational unit consists of a 'node' that co-occurs with a 'span' of words on either side. The span consists of particular word classes filled by specific lexical items. [...] If it is the case that the node word occurs with a span of particular words at a frequency greater than chance would predict, then the result is a collocation. The more certain the words in the span are to co-occur with the node, the more fixed and idiomatic the collocation. This definition accommodates elements from both the phraseological and frequency based approaches. The span which consists of particular word classes that are filled by lexical items refers to the syntactic categories of collocations constituents. Nattinger and DeCarrico (1992) also invoke the notion of frequency –frequency approach– which they actually suggest using to place collocations on the continuum. Given that this definition accepts the two defining criteria of collocations considered by scholars from the two approaches, it can be classified in the third approach. The latter includes collocations dictionaries, which basically use frequency as the main criterion for selecting collocations, but also consider the syntactic nature of the collocations constituents.

To conclude, the definitions provided above clearly indicate that collocations have been defined differently, which some scholars attribute to the different angles under which collocations have been researched (Eyckmans 2009, Nesselhauf 2005). This paper takes the accommodating approach and adopts the definition provided in the *Oxford collocations dictionary for students of English*, that collocations are "the way words combine in a language to produce natural-sounding speech and writing" (Lea, Crowther and Dignen 2002, p. vii).

2.2. Testing collocations

Collocations have since the 1990's been placed at the centre of research attention (Barfield and Gyllstad 2009) with regard to vocabulary and language proficiency. L2 and FL practitioners have attempted to explore growth in collocations both receptively and productively as proficiency develops using different techniques. At the receptive level, collocations have been tested using mainly cloze procedures (e.g., Keshavarz and Salimi 2007), fill-in-the-gaps format (e.g., Keshavarz and Salimi 2007), collocate matching or association tasks (e.g., Authors 2011a, Gyllstad 2007), multiple choice (e.g., Gyllstad 2007, Mochizuki 2002), and recognition tasks (e.g., Barfield 2003). All these studies, even though conducted on different participants with regard to both linguistic and sociocultural backgrounds and while using different techniques, have pointed to a strong relationship between receptive knowledge of collocations and overall L2/FL proficiency. Put differently, these results demonstrate that more proficient students understand more collocations. Based on these results, scholars argue that receptive knowledge of collocations can

reliably predict overall proficiency. Particularly, Gyllstad's (2007, 2009) tests have been validated and standardised. Results from Gyllstad's studies also reveal that learners with a larger vocabulary size are the ones who know more collocations.

At the productive level, testing knowledge of collocations has followed two tracks, i.e. corpora analysis and elicitation techniques (Barfield and Gyllstad 2009, Gyllstad 2007, Laufer and Waldman 2011). Paquot and Granger (2012, p. 130) define learner corpora as "electronic collections of texts produced by foreign or second language (L2) learners". While learner corpora are "highly heterogeneous", which makes comparing results rather difficult, they constitute the ideal source to study learners' language production of collocations and other formulaic language (Paquot and Granger 2012). They are composed of contextualised productions, mainly through argumentative essays, where learners have the freedom to choose their wording rather than being forced to use predetermined items chosen for them (Nesselhauf 2005, Paquot and Granger 2012). Even though not limited to the verb-noun type of collocations, the latter seems to have been the most widely investigated in learner corpora studies (e.g., Howarth 1998, Laufer and Waldman 2011, Nesselhauf 2005). Other types of collocations investigated include adverbadjective collocations (e.g. Granger 1998), collocations of highly frequent verbs (e.g., Altenberg and Granger 2001, Juknevičienė 2008), adjective-noun collocations (e.g., Siyanova and Schmitt 2008), and identification of all lexical collocations (e.g. Martelli 2007). The methodological approach adopted in these studies is comparing L2/FL productions to productions by native speakers of comparable level and analysing the errors they make and their possible causes. Results indicate that L2 and FL users are outperformed by native speakers of comparable class level and that most of the errors produced by L2 and FL learners are the result of transfer from the native language (L1). These results are interesting in terms of what learners can achieve and the types of mistakes they make, especially if larger corpora are concerned (Barfield and Gyllstad 2009, Laufer and Waldman 2011). However, Paquot and Granger (2012) note that they cannot be directly compared because of the different categories of learners involved and their level of proficiency which are different. More so the methods used to extract and analyse collocations are also different; what Paquot and Granger call "heterogeneity of both data and methods".

Studies adopting elicitation techniques include Bahns and Eldaw (1993), Biskup (1992),

9

Bonk (2001), Eyckmans et al. (2004), Farghal and Obiedat (1995), Gitsaki (1999), Authors (2011b, 2012, 2013). The tasks used range from cloze procedures (e.g., Bahns and Eldaw 1993, Bonk 2001, Eyckmans et al. 2004) to cued recall (e.g., Farghal and Obiedat 1995) and translation tasks (e.g., Bahns and Eldaw 1993, Biskup 1992, Farghal and Obiedat 1995) through definition tasks (e.g., Jaen 2007). These studies have also yielded more or less comparable results. Most of them, for instance, have reached the conclusion that collocational knowledge tends to grow with overall linguistic proficiency (e.g., Authors 2011b, 2012, Bonk 2001, Eyckmans et al. 2004, Gitsaki 1999), that L1 affects knowledge of L2/FL collocations (e.g., Biskup 1992), and that collocations are extremely challenging for L2 learners (e.g., Farghal and Obiedat 1995, Jaen 2007).

The relationship between overall proficiency and mastery of collocations is yet to be confirmed insofar that studies such as Bahns and Eldaw (1993) have pointed to the absence of a parallel growth between knowledge of collocations and linguistic proficiency. Furthermore, most of these studies have common methodological concerns worth pointing out. They include too few items analysed, reliability measures which are not reported (e.g., Bahns and Eldaw 1993, Biskup 1992, Farghal and Obiedat 1995), a small sample population (e.g., Jaen 2007), different types of collocations measured in one test battery (e.g., Bonk 2001, Gitsaki 1999), tests which are less workable because of their time consuming nature (e.g., Eyckmans et al. 2004), and variation of contextual information presented with test items (e.g., Authors 2011b, 2012, 2013, Eyckmans et al. 2004). Ideally these shortcomings should be eliminated.

3. The present study

3.1. Population

The present study involved L2 and FL users of English. The L2 users are South African first year students from different faculties and institutes of North-West University, Potchefstroom Campus (N = 292). They were majoring in different fields of study such as commerce, law, natural sciences, or Arts – basically from all available faculties of the university except engineering. All of them were doing a compulsory academic literacy module (AGLE 121), which is taught in English. These participants constitute a diverse population in terms of their native languages, but predominantly

speak Setswana and Afrikaans. The vast majority of them have English as their L2 and they sat the test at the beginning of the second semester, 2014.

The FL users of English are Belgians and Burundians. The Belgian participants (N = 117) were English majors from the University of Antwerp, with Dutch as their native language. They were enrolled as first year students and participated in this study at the beginning of the second semester of the academic year 2013-2014. The Burundian participants (N = 252) were English majors in their second year at the University of Burundi. They participated in this study at the beginning of the academic year 2013-2014. They speak Kirundi as their native language, French as official language, and have some basic knowledge of Kiswahili, the lingua franca of the East African region. In all the groups, no sampling was made as students were invited through their lecturers and only those who consented to participate in the study were involved.

3.2. The test battery

Four tests were used in combination: the Vocabulary Levels Test (VLT) and three collocation tests (details below).

3.2.1. The Vocabulary Levels Test

The Vocabulary Levels Test (VLT) is a receptive vocabulary test that was developed for the first time by Nation (1983). It was validated by Nation (1990), Beglar and Hunt (1999) and Schmitt, Schmitt and Clapham (2001). The test adopts a matching format wherein participants are required to match words with their definitions (see Figure 1 below for instruction in the test of Schmitt et al.). The test has been standardised as a vocabulary size measure and can be used as a placement indicator, i.e. as a proficiency measure. The test was designed by taking word frequency into account and involves five word frequency bands, i.e. 2,000-word, 3,000-word, 5,000-word, and 10,000-word bands as well as the University Word List (UWL) or Academic Word Listⁱⁱ (AWL). At each level, 18 words are randomly selected and presented with their corresponding definitions in clusters of six words and three definitions. Test-takers are instructed to match a word with its definition. The VLT developed by Schmitt et al. (2001) adopted in this study, selects 30 words at each frequency band instead of 18.

[Figure 1 should be here]

3.2.2. Collocation tests

The present study adopts a collocation test developed by Authors (2011b) which was adapted to serve the purpose of this study. We describe first the test adopted that was administered in the third position in this study. We then describe the adaptations that resulted in two tests, which were administered to participants respectively in the second and first positions. The tests are thus described in reverse order to the one in which they were administered. The test adopted was developed following frequency bands and was mapped onto Laufer and Nation (1999). It will be referred to as test three for the purpose of this study. As Authors (2011b) explain, the collocation type measured is verb-noun. The target words the test consists of, i.e. nouns, were selected from Nation's (2006) word-frequency list and the AWL (Coxhead 2000) using a systematic random sampling technique (Babbie 1990). Nation's (2006) word-frequency list is a database of words based on the British National Corpus. The first version consisted of 14 bands, but it has been updated and now consists of 25 bands. Each word band is composed of 1,000 words. The AWL consists of words (about 570) which are frequent in academic environments only and which would otherwise be considered as infrequent.

As suggested by Nation and Beglar (2007), ten words (nouns) were selected from each word band involved, which makes a total of 40 items. As the sampling technique adopted suggests, every nth word (in this case, every 100th word for Nation's database and 57th for the AWL), from a random starting point, was selected. Each time the nth word was not a noun, the next one was selected instead. Their collocates (verbs) were selected from the *Oxford collocations dictionary for students of English* (Lea et al. 2002). The procedure was to find each noun from its entry in the collocation dictionary and then list all the verbs collocating with it in the verb-noun combination. The frequency of the verbs was cross-checked from Nation's (2006) database. The idea was to consider collocates more frequent than or as frequent as the target word as Gyllstad (2007) suggests. The strength of co-occurrence of collocation constituents was also considered by running a collocation samplerⁱⁱⁱ. The verbs which were more frequent, and which had a stronger collocational strength were retained for the study.

With regard to test format, the test was modelled on Laufer and Nation's (1999) Vocabulary Levels Test, which measures controlled productive knowledge as defined in the introduction of this paper. This test measures collocation in a constrained context where a sentential context is provided and the collocations deleted but with the first two letters provided. This test requires participants to supply the missing words (verbs in this case). As appears in Laufer and Nation's (1999) definition, unconstrained and constrained tasks seem to be collapsed into one construct which makes controlled productive ability. However, we posit that unconstrained and constrained may tap into different constructs, which is actually what this study aims to test. We refer to the test which has just been described as "constrained 2". "Constrained 1" is the second test adopted in this study. It is simply an adaptation of the test described above. We retained the same sentential contexts, but we did not provide the first two letters of the word to supply. The third instrument used in this study is what we consider as "unconstrained". The test we refer to as test one in this study, retains exactly the same target words and asks participants to give their collocations in the verb-noun combination. With regard to marking, students were awarded 1pt per correct answer and 0pt per wrong or no answer and the test was marked out of 40. It is worth noting that the numbering of the tests refers to the order in which they were presented to participants in the study. Table 1 below gives a summary of the collocation tests in the order of their administration.

[Table 1 should be here]

4. Data analysis and results

4.1. Test items analysis

It is a requirement for testing experts to test the reliability of test items any test consists of. In this study the reliability of the three tests was measured by computing Cronbach's Alpha. Results show Alpha coefficients of .815 for test one, items in isolation; .840 for test two, items in sentential context; .784 for test three, items in sentential contexts with first letters provided as well. These results indicate that all the three tests can be considered as reliable as the threshold Alpha expected for a test to be reliable is .7 (cf. Pallant 2007). The Corrected Item-Total Correlation (CITC) which shows the extent to which items discriminate between test items was

also performed. It is measured on a scale from -1 to +1, which should be interpreted as the higher the figure, the better the item discriminates between test-takers. The CITC coefficients were weighed against Ebel's (1979) scale as follows: .40 and higher: definitely good items; .30 to .39: reasonably good items; .20 to .29: marginal items in need of improvement; and below .19: poor items that should be revised or eliminated (cut-off point). Results are presented in Table 2. [Table 2 should be here]

As can be seen from Table 2, in test one, only 15% of the items (6 items out of 40) do not seem to perform well as they fall below the cut-off point. In tests two and three, 35% (14 items out of 40) and 30% of the items (12 items out 40) fall below the cut-off point, respectively. What we learn from these results is that test one discriminates between participants better than the other two tests. This entails that free productive knowledge might be a better discriminator between learners with different levels of vocabulary knowledge.

4.2. Performance on collocations from unconstrained and constrained conditions

The first research question examined is the construct of productive knowledge operationalised in this study through collocations. We tested whether unconstrained and constrained contexts as defined in Laufer and Nation (1999) really refer to the same construct. Scores from the three testing conditions referred to as "unconstrained", "constrained 1", and "constrained 2" were compared by running a one-way repeated ANOVA, the results of which are presented in Table 3. It is worth reminding the readers that unconstrained means items presented in isolation, while constrained 1 and constrained 2 respectively refer to items presented in sentential context with the item to supply deleted and sentential context with the first two letters of the item to supply provided.

[Table 3 should be here]

The results presented in Table 3 indicate that the test scores vary in the following descending order: constrained 2 (both sentential context and first two letters provided), constrained 1 (only sentential context provided), and unconstrained (items in isolation). The scores are 28.69 out of 40 for constrained 2; 17.92 for constrained 1; and 13.58 for unconstrained. This order holds at the different word frequency bands. For instance, at the 2000-

word band, students achieved a score of 8.05 on constrained 2; 5.45 on constrained 1; and 4.05 on unconstrained. Overall, as shown by the Sphericity Assumed Correction test, the differences in terms of performance on the three tests are significant ($p = 0.000 \le 0.05$) both overall and at the different frequency bands.

The results were analysed further in order to test the significance of these differences by performing multiple comparisons involving the Bonferroni post-hoc test. The results are presented in Appendices A, B, C, D, and E; respectively for the 2000-word, 3000-word, AWL, 5000-word, and for overall performance. The observed differences in terms of performance on each two components compared (cf. third column entitled "Mean Difference I-J" for differences) and their related significance (cf. last column) are statistically significant.

On the basis of these findings, we posit that unconstrained (item in isolation), constrained with sentential context, and constrained with both sentential context and a clue, may represent different levels in the development of productive knowledge and should be treated as such in testing or otherwise. These findings answer the first research question about whether or not unconstrained and constrained tap into the same construct of productive knowledge: it does not seem to be the case.

4.3. Development of productive knowledge of collocations in relation with vocabulary size

The second research question the present study addresses is the extent to which productive knowledge of collocations grows in parallel with receptive knowledge of vocabulary, i.e. vocabulary size. In order to answer this question, the vocabulary size scores were used to set students in groups following the size of their vocabulary^{iv}. Three groups were distinguished: Level 3 students are students with a vocabulary size larger than 5,000 words, Level 2 students displayed a vocabulary size as large as 5,000 words and students with a vocabulary size lower than 5,000 words, are labelled Level 1. The collocation test scores were mapped onto these levels by running a one-way ANOVA. The means and standard deviations are presented in Table 4 for the three tests while ANOVA results are presented in Table 5.

[Table 4 should be here]

As the results presented in Tables 4 and 5 show, the differences in means scores achieved by the three levels were found to be statistically significant for all three tests with F(2, 373) =52.04, p = 0.000, for test one; F(2, 315) = 100.59, p = 0.000, for test two; and F(2, 315) = 79.32, p = 0.000, for test three.

[Table 5 should be here]

As the one-way ANOVA test can only show that overall differences are statistically significant without pointing to which two levels significant differences occur, the results were analysed further by performing multiple comparisons tests involving the Scheffe post-hoc test. Results are presented in Table 6 and reveal that all the three groups performed significantly differently on all the tests. This entails that Level 3 students, who are the most advanced group, have higher productive knowledge of collocations (in both unconstrained and constrained 1 and constrained 2) than Level 2 students. Level 2 students, in turn, have higher productive knowledge than Level 1 students. These results answer the second research question about the growth of productive knowledge in relation to students' vocabulary size.

[Table 6 should be here]

4.4. Discussion

The present study attempts to put to the test the definition of productive knowledge put forward by Laufer and Nation (1999), which is operationalised here through collocations. More precisely, the present study examines whether or not unconstrained and constrained uses of a word represent the same construct of vocabulary knowledge. This study also explores the development of productive knowledge of collocations in relation to vocabulary size.

With the first research question, results indicate that students performed differently on the tests referred to as unconstrained, constrained 1 and constrained 2. As could be expected, constrained 2, (the test with both contextual information and a clue) is the test on which students performed best. It was followed by the test where students were provided with sentential context; the items presented in isolation come last in terms of performance. The higher score on constrained 2 is accounted for by the fact that students not only get the most information, but also the most restrictive information. They are also guided more than what they would be in free productive writing where only their own contextual information would be a given. The differences in terms of scores on the three tests were found to be statistically significant.

What is implied through these results is that unconstrained (items in isolation) and constrained with sentential context as well as constrained with both sentential context and a clue, do not seem to represent the same construct. Even though this may require more empirical evidence, we argue that unconstrained and constrained may indeed represent different levels of collocation development and should be treated as such in testing or otherwise. Our argument with regard to modelling the testing of productive knowledge of collocations is that we can only claim to do so if we first redefine the construct of productive knowledge. This lends empirical support to the view of Daller et al. (2007), Milton (2009), and Nation (2007) that without a clear construct of what productive knowledge entails, measuring it will remain a serious challenge.

With regard to the way to go about this redefinition, the first direction is trialling the testing of productive knowledge of collocations in relation to receptive knowledge of participants (cf. Schmitt, 2010). This brings us to the second research question addressed in this study about a possible parallel development between productive knowledge of collocations and the vocabulary size of participants. Results indicate that the levels identified by the vocabulary size are also reflected in the collocation test scores. This implies that a student with a larger vocabulary size performs better than a student with a smaller vocabulary size when it comes to productive mastery of collocations in both unconstrained and constrained contexts. In the real world, this would translate into a simple statement: more words understood might lead to more words correctly collocated and thus correctly used – i.e. more fluent, accurate language. These findings confirm previous findings on the relationship between understanding collocations and vocabulary size (Gyllstad 2007, 2009), which it now extends to productive knowledge of collocations.

Given that performance on the three tests is significantly different and that scores on each test point to a possible parallel development between vocabulary size and productive knowledge of collocations, we suggest a second option for defining productive knowledge of collocations. We argue in favour of considering putting productive knowledge of collocations on the following four-point scale: unconstrained 1, unconstrained 2, constrained 1, and constrained

17

2. We define unconstrained 1 as a situation wherein learners are presented with just a context and are asked to generate sentences. We suggest considering unconstrained 2 as presenting words in isolation and asking students to use them. Constrained 1 should be understood as presenting target words embedded in a sentential context and deleting the target collocation. Constrained 2 is when the target word is embedded in a sentential context and the target collocation is deleted, but the student is provided with the first letters of the collocation to supply as a clue. The order of development should be constrained 2, constrained 1, unconstrained 2, and unconstrained 1 (cf. Figure 2). Relating our definition to Laufer and Nation's indicates that we split their definition into two aspects, constrained and unconstrained, which should be divided into two more aspects each as well.

[Figure 2 should be here]

5. Conclusion

The present study is an attempt to put to the test the definition of productive knowledge as put forward by Laufer and Nation (1999). Productive knowledge is operationalised through collocations in this study, which were tested in three different conditions representing three aspects that stand out in the said definition. Results indicate that performance on the three aspects significantly differs and that scores on the three aspects seem to grow in parallel with vocabulary size. On the basis of these results, we conclude that the three aspects that Laufer and Nation collapse into one construct may indeed represent three different levels of productive knowledge of collocations. Therefore, we argue that any attempt to model the testing of collocations should first contribute to the definition. We have in mind trialling the scale of productive knowledge development as outlined above in a follow-up study. This implies that the testing of collocations in future should consider ranking productive knowledge according to the amount of context provided to the learner. We argue that more context makes it easier for learners to produce the correct collocates and less context makes it more difficult. Productive knowledge should not therefore be seen as a constant, but rather as a constantly changing variable, which fits with constructivist theories (Ellis 2003) of the acquisition of knowledge.

References

Altenberg B, Granger S. 2001. The grammatical and lexical patterning of MAKE in native and nonnative student writing. *Applied Linguistics* 22(2): 173–195.

Babbie ER. 1990. Survey research methods. California: Wordsworth Company.

Barfield A. 2003. Collocation recognition and production: Research insights. Tokyo: Chuo University.

- Barfield A, Gyllstad H (eds). 2009. *Researching collocations in another language: Multiple interpretations.* Basingstoke: Palgrave Macmillan.
- Benson M, Benson E, Ilson R. 1986. *The BBI combinatory dictionary of English*. Amsterdam: John Benjamins Publishing Company.
- Benson M, Benson E, Ilson R. 1997: *The BBI dictionary of English word combinations* (revised edition). Amsterdam: John Benjamins Publishing Company.
- Biskup GD. 1992. L1 influence on learners' rendering of English collocations: A Polish/German empirical study. In: Arnaud PJL, Béjoint H (eds), *Vocabulary and applied linguistics*. London: Macmillan. pp. 85–93.
- Boers F, Eyckmans J, Kappel J, Stengers H, Demecheleer M. 2006. Formulaic sequences and perceived oral proficiency: Putting a Lexical Approach to the test. *Language Teaching Research* 10 (3): 245–261.
- Bonk WJ. 2001. Testing ESL learners' knowledge of collocations. In: Hudson T, Brown JD (eds), *A focus on language test development: Expanding the language proficiency construct across a variety of tests*. Honolulu: University of Hawaii Second Language Teaching and Curriculum Center. pp. 113–142

```
Cowie AP (ed.). 1998. Phraseology. Theory, analysis, and applications. Oxford: Clarendon Press. Coxhead A. 2000. A new academic word list. TESOL Quarterly 34 (2): 213–239.
```

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

Ebel RL. 1979. Essentials of education measurement. New Jersey: Prentice Hall.

- Ellis NC. 2003. Constructions, chunking, and connectionism: The emergence of second language structure. In: Doughty C, Long MH (eds), *Handbook of second language acquisition*. Oxford: Blackwell. pp. 33–68.
- Eyckmans J. 2009. Toward an assessment of learners' receptive and productive syntagmatic knowledge. In: Barfield A, Gyllstad H (eds), *Researching collocations in another language: Multiple Interpretations.* Basingstoke: Palgrave Macmillan. pp. 139–152.
- Eyckmans J, Boers F, Demecheleer M. 2004. *The Deleted-Essentials Test: An effective and affective compromise.* Available online: www.hltmag.co.uk/nov04/mart04.htm (Accessed 18 May 2014).
- Farghal M, Obiedat H. 1995. Collocations: A neglected variable in EFL. International Journal of Applied Linguistics 33 (4): 315–332.
- Firth JR. 1957. A synopsis of linguistic theory, 1930–1955. In: Firth JR (ed.), *Studies in linguistic analysis*. Oxford, UK: Blackwell. pp. 1–32.
- Gitsaki C. 1999. Second language lexical acquisition: A study of the development of collocational knowledge. San Francisco: International Scholars Publications.

Gledhill CJ. 2000. Collocations in science writing. Tubingen: Gunter Narr Verlag.

- Gouverneur C. 2008. The phraseological patterns of high-frequency verbs in advanced English for general purposes: A corpus-driven approach to EFL textbook analysis. In: Meunier F, Granger S (eds), *Phraseology in foreign language learning and teaching.* Amsterdam: John Benjamins Publishing Company. pp. 223–243.
- Granger S. 1998. Prefabricated patterns in advanced EFL writing: Collocations and formulae. In: Cowie AP (ed.), *Phraseology: Theory, analysis and applications*. Oxford: Oxford University Press. pp. 145–160.
- Gyllstad H. 2007. Testing English collocations. Unpublished PhD dissertation. Lund University.
- Gyllstad H. 2009. Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In: Barfield A, Gyllstad H (eds), *Researching collocations in another language*: *Multiple interpretations*. Basingstoke: Palgrave Macmillan. pp. 153–170.
- Halliday M. 1966. Lexis as a linguistic level. In: Bazell C, Catford J, Halliday M, Robins R (eds), *In memory of J.R. Firth*. London: Longman. pp. 148–162.

Henriksen B, Stenius LS. 2009. Processes in the development of L2 collocational knowledge: A challenge for language learners, researchers and teachers. In: Barfield A, Gyllstad H (eds), *Researching collocations in another language: Multiple interpretations*. Basingstoke: Palgrave Macmillan. pp. 224–231.

Howarth P. 1998. Phraseology of second language proficiency. Applied Linguistics 19 (1): 24–44.

- Jaén MM. 2007. A corpus-driven design of a test for assessing the ESL collocational competence of university students. *International Journal of English Studies* 7 (2): 127–147.
- Juknevičienė R. 2008. Collocations with high frequency verbs in learner English: Lithuanian learners vs native speakers. *KALBOTYRA* 59 (3): 119–127.
- Keshavarz MH, Salimi H. 2007. Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics* 17: 81–92.
- Laufer B, Nation ISP. 1995. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16 (3): 307–322.
- Laufer B, Nation ISP. 1999. A vocabulary size test of controlled productive ability. *Language Testing* 16 (1): 33–51.
- Laufer B, Waldman T. 2011. Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning* 61 (2): 647–672.
- Lea D, Crowther J, Dignen S (eds). 2002. *Oxford collocations dictionary for students of English*. Oxford: Oxford University Press.
- Lewis M. 1993. *The Lexical Approach: The state of ELT and the way forward*. Hove: Language Teaching Publications.
- Lewis M. 1997. *Implementing the Lexical Approach: Putting theory into practice*. Hove: Language Teaching Publications.
- Lewis M. 2000. *Teaching collocations: Further development in the Lexical Approach*. Hove: Language Teaching Publications.
- Martelli A. 2007. *Lexical collocations in learner English: A corpus-based approach*. Alessandria: Edizioni dell'Orso.
- Meara P, Fitzpatrick T. 2000. Lex30: An improved method of assessing productive vocabulary in an L2. *System* 28 (1): 19–30.

Milton J. 2009. *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

Mochizuki M. 2002. Exploration of two aspects of vocabulary knowledge: Paradigmatic and collocational. *Annual Review of English Language Education in Japan* 13: 121–129.

Nation ISP. 1983. Testing and teaching vocabulary. *Guidelines* 5(1): 12–25.

Nation ISP. 1990. Teaching and learning vocabulary. Rowley: Newbury House.

- Nation ISP. 2006. How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63 (1): 59–82.
- Nation ISP. 2007. Fundamental issues in modeling and assessing vocabulary knowledge. In: Daller H, Milton J, Treffers-Daller J (eds), *Modelling and assessing vocabulary Knowledge*. Cambridge: Cambridge University Press. pp. 35–43.

Nation ISP, Beglar D. 2007. A vocabulary size test. The Language Teacher 31 (1): 9–13.

- Nattinger JR, DeCarrico JS. 1992. *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nesselhauf N. 2005. *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.
- Pallant J. 2007. SPSS survival manual. Buckingham: Open University Press.

Palmer H. 1933. Second interim report on English collocations. Tokyo: Kaitakusha.

- Paquot M. 2008. Exemplification in learner writing: A cross-linguistic perspective. In: Granger S, Meunier F (eds), *Phraseology in foreign language learning and teaching*. Amsterdam: John Benjamins Publishing Company. pp. 101–119.
- Paquot M, Granger S. 2012. Formulaic language in learner corpora. *Annual Review of Applied Linguistics* 32: 130–149.
- Pawley A, Syder FH. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: Richards, JC, Schmidt RW (eds), *Language and communication*. London: Longman. pp. 191–227.

Read J. 2000. Assessing vocabulary. Cambridge: Cambridge University Press.

Schmitt N. 2010. *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.

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

Sinclair J. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press.

- Siyanova A, Schmitt N. 2008. L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review* 64 (3): 429–458.
- Wray A. 2000. Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics* 21 (1): 463–489.

Wray A. 2002. Formulaic language and the lexicon. Cambridge: Cambridge University Press.

Wray A, Perkins M. 2000. The functions of formulaic language: An integrated model. *Language and Communication* 20 (4): 1–28.

ⁱ Collocations belong to this big category referred to as formulaic language, a term adopted by Wray (2002).

ⁱⁱ The UWL is the first academic vocabulary list developed by Hu and Nation (1984) while the AWL is an updated version of academic vocabulary developed by Coxhead (2000).

ⁱⁱⁱ The collocation sampler is available online at: <u>http://www.collins.co.uk/Corpus/CorpusSearch.aspx</u>. It was accessed in May 2009.

^{iv} Laufer and Ravenhorst-Kalovski's (2011) formula was used. They suggest that the score at the 1,000-word band is the same as that at the 2,000-word band and that the score at the 4,000-word band is obtained by averaging scores at the 3,000- and 5,000-word bands (Laufer and Ravenhorst-Kalovski 2010: 21). The size at the 10000-word band for which Laufer and Ravenhorst-Kalovski do not suggest any formula was estimated using Schmitt et al.'s (2001) suggestion that a word band is mastered when students achieve a score of 24 out 30, which represents 80 percent.