Measuring improvement in first-year ESL university students' productive knowledge of verb-noun collocations through integrated productive oriented teaching strategies

Mercia Morris

orcid.org/0000-0001-8849-3118

Dissertation accepted for the degree *Magister Artium in Applied Language Studies* at the North-West University

Supervisor:	Dr K Van den Berg
Co-supervisor:	Dr D Nizonkiza
Assistant Supervisor:	Prof CA Breed

Graduation: July 2022

Student number: 21975957

ACKNOWLEDGEMENTS

This dissertation is dedicated to my parents and to Nic. To my parents, Gert and Yvonne – thank you for teaching me the value of hard work and perseverance. To Nic – thank you for your love, support, and patience.

I would like to thank the following:

1. My supervisor and mentor, Dr Karien Van den Berg – for your unwavering dedication and for believing in me when I could not believe in myself. You have made me the teacher and researcher I am today.
2. My co-supervisors, Prof Adri Breed and Dr Déogratias Nizonkiza – for your invaluable insight and guidance. This dissertation would not have been possible without you.
3. Prof Johann Van der Walt – for your guidance during the initial stages of this study.
4. My friends, Dr Louise Strydom and Dr Janien Linde – for your advice, your ears, and occasionally your shoulders.
5. Dr Lande Botha and Mrs Maryka Van Zyl – for sharing your knowledge and assisting with *WordSmith*.
6. My colleagues in the English Department – for your support throughout this journey.
7. Finally, my cat, Bilbo – for keeping me company in the early morning hours and reminding me to take breaks by sitting on my keyboard every few hours.

I would also like to thank the Research Unit and the Department of Arts and Culture for financial assistance.

ABSTRACT

Research has firmly established that vocabulary plays a crucial role in both receptive and productive skills. Several studies have shown that learners who are able to use complex vocabulary, particularly multiword items, tend to come across as being more proficient and are closer to reaching native-like proficiency, both in speaking tasks (Boers, Eyckmans, Kappel, Strengers & Demecheleer, 2006; Strengers, Boers, Housen, and Eyckmans, 2010, 2011; Boers, Demecheleer, Coxhead & Webb, 2014) and writing tasks (e.g., Crowther, Dignen, & Lea, 2002; Dai & Ding, 2010). Researchers are thus in agreement on the benefits for L2 and FL learners to acquire "considerable numbers of multiword items" (Boers, 2020:143; also see Hatami, 2015), emphasising its status as "fundamental to the way language is used, processed and acquired in both the L1 and L2" (Martinez & Schmidt, 2015: 299). Formulaic language is thus central to communicative competence (Nation, 2001; Schmitt, 2004; Wray, 2002). According to Nation (2001:317-318), mastering language chunks in relation to fluency shows the benefits of focussing on formulaic language in L2 teaching. However, language learners seem to have difficulty especially in acquiring multiword units, collocations, lexical bundles, and idioms (e.g., Abel, 2003; Irujo, 1986; Kecskés, 2000). However, vocabulary is a complex phenomenon and it is difficult to develop a sense for multiword units (Martinez et al. 2015). As a result, many scholars consider teaching collocations as an integral part of teaching vocabulary (e.g., Barfield & Gyllstad, 2009; Basal, 2019; Boers & Lindstromberg, 2012; Chen, 2011; Nesselhauf, 2003).

The same applies to ESL learners at South African universities (Scheepers, 2017:109). Understanding what to count when identifying multiword items (MWI) to measure competence in this regard, is crucial in order to accurately determine a learner's level of proficiency. In addition, attempts to support students in building their vocabulary adequately and appropriately through formal instruction must be informed by a clear definition of vocabulary. However, a lack of clear guidelines on what to count when identifying collocations in student writing for measurement purposes effectively limits research in this important area. Defining MWI can be challenging since such definitions depend greatly on what words or multiword items are understood to be, as well as what exactly is meant by knowing a word. Likewise, the results of a study also depend on how the object of interest – i.e., vocabulary and specifically collocations – is defined and assessed for the purpose of that particular study. This study aimed to define collocational knowledge within the framework of a psycholinguistic approach to language, with particular focus on form-meaning mapping. Having defined the notion of collocations within the context of the present study, the researcher compared the proficiency of verb-noun

collocations of two cohorts of participants. The cohorts consisted of a control group and an experimental group. Following a quasi-experimental design, the two cohorts were assessed repeatedly using a productive version of the Vocabulary Levels Test that was adapted to assess collocations. The tests were based on the Academic Collocations List (Ackermann & Chen, 2013) and compiled using a variety of published, peer reviewed articles. Participants were also required to complete writing tasks that were used to conduct a collostructional analysis to determine which verb-noun collocations occurred frequently in their writing, and whether participants made errors in their use of collocations. The experimental group received treatment in the form of teaching interventions based on awareness-raising techniques. The results from the control group were compared with that of the experimental group pre- and post-treatment to determine whether the techniques employed in the intervention phase resulted in statistically significant improvement. Results indicated that the intervention was successful. The results from the present study could be used to inform future research on measuring and teaching collocations, as well as the process utilised to select which collocations are necessary to teach.

Keywords: awareness-raising techniques, collocations, vocabulary assessment, vocabulary knowledge, collostructional analysis, form-meaning mapping.

OPSOMMING

Navorsing het onteenseglik bewys dat woordeskat ’n belangrike rol speel in beide reseptiewe en produktiewe taalvaardighede. Verskeie studies wys daarop dat leerders wat komplekse woordeskat – veral multiwoordeenhede – gebruik, geneig is om meer taalvaardig te wees en nader daaraan is om moedertaalvlakvaardigheid te bereik in beide mondelinge (Boers, Eyckmans, Kappel, Strengers & Demecheleer, 2006; Strengers, Boers, Housen, and Eyckmans, 2010, 2011; Boers, Demecheleer, Coxhead & Webb, 2014) en geskrewe take (e.g., Crowther, Dignen, & Lea, 2002; Dai & Ding, 2010). Navorsers stem saam dat dit voordelig is vir tweedetaal- en vreemdetaalleerders om ’n groot aantal multiwoordeenhede (Boers, 2020:143; sien ook Hatami, 2015) aan te leer, om dit ’n fundamentele aspek is van hoe taal in beide ’n L1 en L2 gebruik, verwerk en aangeleer word (Martinez & Schmidt, 2015: 299). Formuletaal speel dus ’n sentrale rol in kommunikatiewe bevoegdheid (Martinez & Schmidt, 2015: 299). Nation (2001:317-318) meen dat die bemeestering van taalstukke korreleer met vlotheid, wat weer verder daarop wys dat dit voordelig is om op formuletaal in tweedetaalonderrig te fokus. Leerders blyk egter veral te sukkel met die leer van multiwoordeenhede, kollokasies, leksikale bundels, en idiome. Woordeskat is egter ’n komplekse verskynsel en dit is moeilik om ’n gevoel vir multiwoordeenhede te ontwikkel (Martinez et al., 2015). As gevolg hiervan beskou baie kenners onderrig in kollokasies as ’n belangrike deel van woordeskatonderrig (e.g., Barfield & Gyllstad, 2009; Basal, 2019; Boers & Lindstromberg, 2012; Chen, 2011; Nesselhauf, 2003).

Dieselfde aspekte is relevant vir toepassing op leerders wat Engels as ’n tweedetaal aan Suid-Afrikaanse universiteite studeer (Scheepers, 2017:109). Begrip om multiwoordeenhede te identifiseer om bevoegdheid te meet, is deurslaggewend om ’n leerder se vlak van vlotheid op akkurate wyse te bepaal. Daarby moet pogings om studente te ondersteun om voldoende woordeskat op te bou deur formele onderrig, onderlê word met ’n duidelike definisie van woordeskat. ’n Tekort aan duidelike riglyne om kollokasies te identifiseer in studente se skryfwerk as maatstaf, beperk navorsing in hierdie belangrike area. Om kollokasies te identifiseer hang grotendeels af van hoe hierdie hulle gedefinieer word, sowel as wat dit presies beteken om ’n woord te ken. Eweneens hang die resultate van ’n studie af van hoe dít wat van belang is – d.i. die woordeskat en spesifiek kollokasies – vir doeleindes van die spesifieke studie gedefinieer en beoordeel word. Hierdie studie poog om kennis van kollokasies te definieer binne die raamwerk van ’n psigolinguistiese benadering tot taal, met spesifieke fokus op vorm-betekenispare. Nadat die idee van kollokasies binne die konteks van die huidige studie gedefinieer is, het die navorser die vaardigheid van werkwoord-naamwoordkollokasies

vergelyk op twee groepe deelnemers. Die twee groepe het bestaan uit ’n kontrolegroep en ’n eksperimentele groep. Na gelang van ’n kwasi-eksperimentele ontwerp is die twee groepe herhaaldelik geassesseer met die gebruik van ’n aangepaste weergawe van die Vocabulary Levels Test wat aangepas is om die aktiewe gebruik van kollokasies te assesseer. Die toetse is gebaseer op die Engelse Akademiese kollokasielys (Academic Collocations List) van Ackermann & Chen (2013) en saamgestel deur die gebruik van verskeie gepubliseerde en geakkrediteerde artikels. Deelnemers is vereis om skyftake uit te voer wat gebruik is om ’n kollekseemanalise uit te voer om vas te stel watter kollokasies dikwels voorgekom het in hulle skryfwerk. Hierdie analise het ook inligting verskaf oor die tipe kollokasiefoute wat deelnemers gemaak het. Die eksperimentele groep het onderrigintervensie ontvang wat gebaseer is op bewustheidsopbrengstegnieke. Die uitkoms van die kontrolegroep is vergelyk met die eksperimentele groep se skryfwerk van voor- en ná intervensies. Sodoende kon vasgestel word of die tegnieke wat gebruik is in die intervensiefase statistiese betekenisvolle verbetering toon. Die uitslag het aangetoon dat die intervensie suksesvol was. Die uitslag van die huidige studie kan gebruik word in toekomstige navorsing rondom die onderrig van kollokasies. Hierdie studie bied ook insigte in die proses wat gebruik kan word om kollokasies uit te kies wat belangrik is om te bemeester.

Sleutelterme: woordeskat, onderrigintervensie, kollokasies, kollekseemanalise, vorm-betekenispare, assessering, bewustheidsopbrengstegnieke

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Problem statement and motivation

In the last few decades, much emphasis has been placed on the importance of vocabulary in English as Second and/or Foreign Language (ESL/EFL) learners' development. It has been established that there is a link between receptive vocabulary size (how many words a learner knows) and vocabulary knowledge (how well a learner knows a word), and a learner's level of proficiency and their productive vocabulary use, which enables a learner to communicate fluently and accurately (Schmitt, 2010). Quite a number of studies have established a relationship between vocabulary knowledge and overall proficiency (see among others Meara & Buxton, 1987; Meara, 1996; Nation 1990; Nation & Beglar, 2007). Vocabulary development is also often parallel to the development of overall proficiency (Nation, 1990, Nizonkiza, 2011). Furthermore, research evidence indicates that having adequate knowledge of more complex structures such as formulaic sequences, proverbs, similes and collocations, helps second language learners to come across as being more proficient (Boers *et al.*, 2006; Boers *et al.*, 2014). Particularly, knowledge and use of formulaic sequences and collocations are important and useful in academic discourse.

Collocations can be defined as words that co-occur naturally and frequently, for example, *fully* aware and *make a decision*. Erman and Warren (2000) state that written discourse might consist of formulaic sequences by as much as 52.3%. This also applies to academic writing. Adequate knowledge of formulaic sequences therefore plays a vital role in overall language proficiency and academic achievement (Jones & Haywood, 2004). However, learning vocabulary in a second language (L2) is a complex process and formulaic sequences and multi-word units seem to take a longer time to acquire (Boers *et al.* 2014) and, according to Forsberg (2010), at times only advanced L2 learners display adequate knowledge of these structures. Verb-noun collocations (e.g. *conduct research*, *draw a conclusion*) seem to be particularly problematic (Laufer & Waldman, 2011; Barfield & Gyllstad, (ed.), 2009; Nesselhauf, 2005). Wei also recognises that collocation is an important but undervalued aspect of productive vocabulary, as learners need to learn a wide variety of ways in which words collocate with each other (1999:1).

Following the importance of collocations and their problematic nature, researchers have more recently focused their attention on the learning and more specifically, the explicit teaching of

collocations in ESL/EFL courses (Barfield, 2009; Boers *et al*. 2006, Webb & Kagimoto, 2009). Such scholars include Chan and Liou (2005), Boers *et al.* (2006), and Ab Manan, Jaganathan and Pandian (2014) to name but a few. Chan and Liou used five web-based exercises aimed at improving Chinese ESL learners' use of verb-noun collocations and found that there was a significant increase (2005). Boers *et al*. (2006) compared the overall oral proficiency of a control group and experimental group after receiving explicit instruction on collocations, and the experimental group scored higher than the control group. Ab Manan *et al*. (2014) also used an experimental design to investigate the benefit of direct instruction of academic formulas (DIAF) in a group of students at a Malaysian university. The experimental group showed a significant improvement in the quality of their academic writing. While several studies have been done to determine the effects of explicitly teaching collocations, similar studies have not been attempted at universities in South Africa. However, studies at one South African university show that the problematic nature of collocations is evident.

Studies at the North-West University focusing on productive vocabulary knowledge in academic writing suggest that first-year students reaching the 3000-word band, or even higher, still tend to fall short of expectations when it comes to productive knowledge of verb-noun collocations and that they are only able to master verb-noun collocations from the 2000-word band (Nizonkiza *et al.*, 2013; Nizonkiza & Van Dyk, 2015). Nizonkiza and Van Dyk (2015) found that receptive knowledge is not problematic, while Nizonkiza *et al.* (2013) found that collocations in particular are problematic. Furthermore, research conducted at the North-West University using a productive collocation test and a vocabulary levels test, indicates that first-year university students have a higher level of receptive vocabulary compared to their productive vocabulary in an academic context (Nizonkiza, 2016). Van Dyk *et al.* (2016) found that there is a strong correlation between the productive knowledge of collocations and academic literacy of first-year student taking an Academic Literacy course. The combined findings of these four studies suggest that it is imperative that students at the North-West University improve their productive knowledge of collocations to enhance their overall academic achievement. Thus, more explicit production-oriented teaching of verb-noun collocations could improve learners' productive knowledge (Nizonkiza *et al.,* 2013; Barfield, 2009, Van Dyk *et al*., 2016). This may be done following one, or a combination of two main approaches to teaching collocations: the awareness-raising approach and the attention-drawing approach (Nizonkiza, 2012). These two approaches are both based on Nation's (2001) three psychological conditions for learning vocabulary: noticing, retrieving, and generation.
Various studies have found that receptive vocabulary is often larger than productive vocabulary (Laufer & Paribakht, 1998; Webb, 2008, among others). If this is the case, Laufer and Nation

(1995) suggest that it is necessary to implement instruction that would give learners the opportunity to put their vocabulary knowledge to productive use. According to Lee and Muncie, explicit target vocabulary instruction appears to scaffold learner writing, which is a central skill for success in academic achievement (2006:296). They also attribute an improvement in writing to the teacher's use of interactive elicitation of vocabulary and a writing frame, and specific instruction to learners to use target vocabulary. Zhong and Hirsh (2009) also found that students' productive knowledge can be increased through productive classroom tasks. Based on the finding that the majority of Japanese English Foreign Language (EFL) learners focus on receptive rather than productive vocabulary learning, Al-Murthada (2014) concludes that it is crucial to expose learners to productive vocabulary learning strategies.

This suggestion is in line with Nation (2001), who recommends that vocabulary learning programs should include an explicit, intentional learning component, as well as a component focused on maximising exposure to target vocabulary, in order to allow maximum engagement with the target vocabulary. Schmitt (2008) outlines Nation's four suggested learning strands (meaning-focused input, meaning-focused output, language-focused learning, and fluency development) and suggests that this approach "gives balanced attention to learning new information about lexical items, and then provides for consolidation and enhancement of that knowledge" (p. 343).

It has been established by Nizonkiza *et al*. (2013), Nizonkiza and Van Dyk (2015) and Nizonkiza (2016) that productive use of collocations are problematic at the North-West University's Potchefstroom Campus. This study therefore aimed to investigate the extent to which controlled productive knowledge of verb–noun collocations could be improved by explicit awareness-raising teaching strategies for first-year students at the North-West University, enrolled for an ESL course aiming to improve their English language proficiency (ENLS 111). The purpose of the module is to develop students' knowledge and understanding of the fundamentals of the English language and acquire the skills required to interpret and produce a text, such as an academic essay. The students taking ENLS 111 were therefore suitable participants for this study. The present study drew on Nation's approach to introduce selected collocations to first-year ESL students, providing input and giving opportunity for productive use of relevant lexical items. The selection of collocations comprised all of the verb-noun collocations identified in the Academic Collocation List (ACL) developed by Ackermann and Chen (2013). To incorporate Nation's (2001) proposed learning strands, a purposefully integrated blended learning approach was used for the explicit teaching instruction, specifically following the awareness-raising technique as outlined by Nizonkiza (2012).

Though South African-based studies have examined collocations (e.g. Nizonkiza et al. 2013; Nizonkiza & Ngwenya 2015), an intervention study specifically focusing on productive knowledge of verb-noun collocations is yet to be conducted in this context. Results of this study will be of comparative value within the international body of literature in this regard. In addition, the proposed methodology provides more defined insight into the productive knowledge of verb-noun collocations of first-year ESL students (at the North-West University) in terms of the strength and accuracy of the selected verb-noun collocations than what can be achieved by frequency counts alone. Furthermore, insight into the proposed productive oriented approach to teaching collocations (Nizonkiza, 2012) were gained and recommendations were made concerning the duration of intervention required or preferred in order to bring about a change in this type of knowledge; and awareness-raising techniques and tasks to be implemented for classroom practice to help improve learners' ESL proficiency and academic writing. Finally, the study may also serve as basis for future research in continued tracking of students' collocation use in ESL writing.

1.2 Research goals

In replicating international research, the primary aim of this study was to determine the extent to which ESL first-year university students' controlled productive verb-noun collocation knowledge could be improved by means of explicit awareness-raising techniques targeting selected verb-noun collocations and maximising learner engagement with the target language. In order to reach the primary goal of the proposed study, the researcher attempted to answer the following sub-questions:

1. What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?
2. To what extent does intervention with explicit teaching of verb-noun collocations result in improving their mastery among first-year ESL students at the North-West University? This question is answered by determining whether
    a. there is any statistically significant difference between the results of the pre-test, post-test and delayed post-test of the experimental group.

1.3 Method of research

1.3.1 Survey of scholarship

Relevant literature on ESL proficiency, collocations, collexeme strength, vocabulary learning and teaching, the role of vocabulary in ESL writing skills, and measuring knowledge of collocations was reviewed. This included literature on vocabulary knowledge, defining collocations, measuring collocational strength, the current trends in vocabulary teaching, approaches to teaching collocations/vocabulary and the importance of accurately using verb-noun collocations. Literature on language testing, the reliability of current tests and technological developments in measuring vocabulary use was also reviewed.

1.3.2 Empirical research

In this section the details of the methodology are discussed. This includes details on the data collection and analyses, participants of the study, the instrumentation utilised, as well as the procedure followed to obtain data for the present study.

1.3.3 Research design

The study followed quantitative design comprising a collostructional analysis and a two-group experimental design with a pre-test, two post-tests, and a delayed post-test. Data were collected from a non-randomized control group and experimental group, comprising 600 participants in total in order to determine the extent to which the controlled productive knowledge of verb-noun collocations could be improved in first year English Second Language (ESL) students taking an English for Specific Purposes course (ENLS 111). For the purposes of the study, only quantitative data were collected to evaluate participants' performance over a period of six weeks. Only quantitative data were collected to enable the researcher to statistically measure the performance of the participants, and accurately compare the 2017 and 2018 cohorts.

1.3.3.1 Data collection and analysis

1.3.3.1.1 Participants

The participants constituted two groups of first-year English second language university students enrolled for the ENLS 111 module, aiming to improve ESL proficiency, at the Potchefstroom Campus of the North-West University. The control group constituted the 2017 cohort of students, and the experimental group constituted the 2018 cohort. The complete population, i.e. the whole group of ENLS 111 students for 2017 and 2018, were used for the

purpose of this study (n=624). Using a total population sampling technique allowed for deep insights and analytic generalisations. Relevant statistical tools were used to control for variance. The decision is discussed in detail and motivated in Chapter 4.

1.3.3.1.2 Instrumentation

A two-pronged approach was used to address the sub-questions. This called for different instruments, as discussed in detail in Chapter 4. First, a collostructional analysis was conducted. Second, a series of pre- and post-tests were administered.

The collostructional analysis was conducted to answer research question 1. A collostructional analysis as designed by Stefanowitsch and Gries (2003) and Gries and Stefanowitsch (2004a; 2004b) was employed to determine the strength of the verb-noun collocations selected from the ACL. This application is based on the Fischer exact test (Pedersen, 1996) and is useful for determining effect size in terms of the extent to which a construction is typically associated with a lexical item (Gilquin, 2010:195). For the purpose of this study, it was used to determine the strength between the verb-noun collocations selected from the Academic Collocation List (Ackermann & Chen, 2013) as used in ESL students' writing. See section 4.3.1.1 for details on the processes followed to analyse the essays, and section 5.2.1 for an interpretation of the results.

A series of tasks was used to track potential improvement of the controlled productive knowledge of ENLS 111 students over a period of six weeks. The tasks were administered on eFundi, an online platform used by the university, and on the same timeline in 2017 and 2018. Each task will have a time limit of one hour. The task series comprised a pre-test exercise (T_1), four intermittent exercises (t_i, t_{ii}, t_{iii} and t_{iv}), a post-test exercise (T_2), a second post-test exercise (T_3) and a delayed post-test exercise (T_4). The post-test exercise was administered twice at three-week intervals, followed by the delayed post-test exercise which was administered early in the second semester (see *Procedure* for a timeline and 4.3.1.2 for a detailed discussion and motivation*)*.

These exercises were marked and served as part of general class participation only. During the six weeks the 2017 participants received no intervention, whereas the 2018 participants were exposed to explicit teaching and awareness-raising techniques as outlined by Nizonkiza (2012) to improve the participants' use of verb-noun collocations. The extent of the intervention's impact was then measured (see below) and conclusions were drawn concerning

the duration of impact needed; the success of the proposed teaching method; and implications for future teaching and course development.

The format of the exercises, as based on Laufer and Nation's Productive Levels Test (1999) which is still considered the closest instrument to a standardized vocabulary test available (Meara, 1996), comprised gap-filling and cloze tests. The item selection was based on the ACL by Ackermann and Chen (2013). The three main exercises (T_1, T_2, T_3) had the same format, but different items were selected for each instance to control for student memory playing a role in taking the tests. Exercises (t_i and t_{ii}) were administered intermittently during the intervention phase, aiming to collect continuous data for more careful tracking of potential improvement. The content of these tests varied slightly to reflect the variety of intervention techniques (discussed in detail in *Phase 3*, under *Procedure*).

1.3.3.1.3 Procedure

The procedure entailed three phases: the collostructional analysis, followed by the experimental phase and a final phase in which the final statistical calculations and comparisons were made.

Phase 1

Four collostructional analyses were done in order to determine the strength of the ACL verb-noun collocations used by participants: two to determine the strength of the verb-noun collocations used in two essays of the 2017 cohort ESL students' writing (unintervened); and two to determine the strength of the selected constructions as used in the 2018 cohort ESL learner writing.

Collocations that measure very high in strength are typically conventionalised. For example, *conduct research* is more common and conventionalised than an expression such as *make research*. The strength of the selected verb-noun collocations in the ACL was measured to in participants' essays to identify more and less conventional verb-noun collocations. The verb-noun collocations that can be deemed highly conventional (i.e. collocate very strongly) in the ACL could be expected to be evident in a proficient ESL learners' writing. The analyses were conducted to confirm whether participants have mastered the identified verb-noun collocations sufficiently. If the strength of the same verb-noun collocations in the 2017 cohort writing compared well with the ACL list, the suggestion is that participants had actually mastered these collocations sufficiently, and in fact may not need additional instruction. This was not the case for the present study. The fourth analysis then served to confirm whether the explicit

intervention resulted in stronger verb-noun collocations of the selected items. An increase in strength of the selected verb-noun collocation from the 2017 to the 2018 cohort ESL student writing would suggest that the 2018 intervention was successful and support findings from literature cited. Results that compare well to that of the ACL list indicated mastery of the selected verb-noun collocations. These findings served to support that of the pre- and post-tests, and informed recommendations for teaching practice.

Phase 2

This phase of the study consisted of two parts entailing the completion of the experimental procedure for the control group (2017) and for the test group (2018), followed by a third stage for data analysis. The timeline for the procedure was as follows:

Timeline:	Beginning 1st wk	1st wk May	end of wk1	wk 2	end wk 2	wk 3	end wk 3	wk 4	end wk 4	wk5	end wk 5	wk 6	end wk6	early 2nd semester
Control group 2017	T1	xx	ti	xx	tii	Intervention I3	T2	xx	ti	xx	tii	xx	T2 repeat	T3
Experimental group 2018	T1	Intervention I1	ti	Intervention I2	tii	Intervention I3	T2	Intervention I4	ti	Intervention I5	tii	Intervention I6	T2 repeat	T3

Table 1-1: Procedure timeline

For both cohorts, the pre-test was administered in the first week following the university holiday in April to ensure a continuous period of at least 6 weeks before an examination period started. The procedure for each cohort is outlined below:

2017 – Control Group procedure

Data gathered during this phase served as benchmark against which to compare the effect of the planned interventions for the 2018 cohort. The 2017 control group procedure commenced in April 2017 and comprised only a series of formative tests. The pre-test (T_1) for the control group was administered, intermittent tests t_i and t_{ii}, the post-test administered twice (T2 and repeated T2), and the delayed post-test (T3) six weeks later. No intervention took place during this phase.

2018 - Experimental Group procedure

This phase commenced in April 2018. The same outlined procedure was followed as in 2017, but with the inclusion of interventions I1 to I6. The tests were administered as in 2017 and according to the same timeline.

The intervention took place over a six-week period from April to May. The purpose of the interventions was to introduce participants to the concept of collocations, and raise awareness of how collocations are used. This was achieved through explicit teaching of the selected verb-noun collocations in class by the course lecturer, classroom activities and exercises as well as via the eFundi online platform available to students at the North-West University, to ensure optimal integration and engagement with the target language following a blended learning approach. The target vocabulary in each intervention was taken from the ACL as mentioned above. Details of the intervention phase are outlined in section 4.3.1.2.

Phase 3: Statistical analysis and comparison of results

The NWU Statistical Consultation Services informed the research design and process for this study, and provided support in conducting the relevant analyses for the study.

Phase 1 analyses were conducted as described and the results will be used for supporting purposes here.

The repeated post-test and use of intermittent exercises in addition to the pre-, post- and delayed post-test exercises served to increase the number of observation points to provide a clearer picture of any developmental pattern that may be evident between the pre-test – post-test and post-test – delayed post-test phases. In this way reliability of the results was increased. In addition, it served to inform the duration required for the intervention, i.e. comparing the extent of development after three weeks' intervention, and after six weeks' intervention.

In terms of analyses, variance within participants as people, as well as variance between the tests needed to be controlled. Hierarchical Linear Measures (HLM) Hancock & Mueller, 2010; Woltman *et al.*, 2012) was applied to control for variance in the groups, with person dependence as the primary unit of analysis. This analysis allows for missing data and is useful in applications to the full survey group to ensure the intragroup and intergroup comparability. Woltman *et al.* (2012:53) explain: "HLM simultaneously investigates relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels than other existing analyses". As it was anticipated

for this study that the independent variable would impact the dependent variable, the effect size of the treatment was determined by means of Cohen's D value (Cohen, 1988; Sink & Mvududu, 2010), based on the HLM averages and variance.

The results for the three collostructional analyses (Phase 1) were considered in comparison to the results of the experimental procedure (Phase 2) in order to inform final conclusions and recommendations.

1.4 Ethical concerns

Ethical clearance was sought from the Ethical Commission at the North-West University. Ethical clearance was granted and the ethics number received is NWU-00592-17-S8. As this was an experimental study, the aim was to bring about a change in participants' knowledge. Ethical implications could therefore arise and precautionary measures were taken to address potential issues.

Participant identities remained anonymous and a coding system was used for identification purposes. Data was only included for those participants who had given signed consent (see Annexure A for the consent form that participants received). No negative impact was expected or incurred as the aim was to improve participants' knowledge of controlled productive verb-noun collocations. Furthermore, the tests were administered as class exercises and contributed to class participation only and did not influence semester marks. However, this point was not stressed to participants to ensure they take the tasks seriously, as per normal classroom and homework activities. In addition, all participants within respective cohorts were treated equally: none of the 2017 participants received the treatment, while all the 2018 participants did, regardless of whether they have signed consent or not. Using participants from two year-groups provided an alternative to the problematic design of exposing one part of a group, or some classes in one year-group to an intervention, while withholding the intervention from others. Adding, revising, or removing teaching activities from one year-group to the next is common practice and may occur to varying degrees from one year to the next. Such changes could for example result from different lecturers teaching the module, selecting a different prescribed text, the content being revised to suit students' needs more closely, or because the course is being recurricularised.

1.5 Programme of Study

Following Chapter 1, Chapter 2 consists of an in-depth literature review detailing the types and role of vocabulary knowledge in ESL proficiency and academic success, and the concept of collocations, collocations as part of vocabulary knowledge and its role in ESL proficiency. The main purpose of Chapter 2 is to conceptualise the nature of vocabulary knowledge and collocational knowledge, in order to inform a working definition of collocations and theoretical approach taken in this study. This chapter further explores the importance of form-meaning mapping in collocational knowledge. Chapter 3 focuses on various teaching approaches and techniques relevant to learning and teaching vocabulary and collocations, including explicit and implicit teaching, productive vs receptive oriented approaches, the four learning strands (Nation, 2001), and specific approaches to teaching collocations. Chapter 3 further aims to determine which collocations are important in ESL teaching. Chapter 4 provides an overview of the research design and the empirical procedures followed. In this chapter, collostructional analyses will be considered in detail, as well as experimental design. Chapter 5 consists of a discussion of the empirical findings. The conclusion and recommendations for further research follow in Chapter 6.

CHAPTER TWO: CONCEPTUALISING COLLOCATIONAL KNOWLEDGE

2.1 Introduction

Vocabulary has long been recognised as one of the essential building blocks of language learning (Boers, 2020; Horiba, 2012; Milton & Treffers-Daller, 2013; Nation, 2001; Nizonkiza & Van den Berg, 2014), and it has been established that vocabulary is of particular importance when it comes to academic success (Gardner, 2013; Scheepers, 2018), and overall language proficiency (Iwashita *et al.* 2008; Meara & Buxton, 1987; Meara, 1996; Nation 1990; Nation & Beglar, 2007; Schmitt *et al.*, 2011; Staehr, 2009) in both a first (L1) and second language (L2). Horiba (2012), in a comparative study between Chinese- and Korean-speaking L2 learners and L1 speakers of Japanese, noted that text comprehension was poorer in learners with lower-level vocabulary knowledge, and that learners "who have richer word knowledge are more likely to encounter more familiar words in a text" (p.108).

Vocabulary is not only essential, it is also complex in that it contains lexical units that consist of two words or more (McCarthy, 1990:6), i.e., meaning is not necessarily confined to a single word. Research has firmly established that vocabulary plays a crucial role in both receptive (reading and listening) and productive (speaking and writing) skills. Several studies have shown that learners who are able to use complex vocabulary, particularly multiword items, tend to come across as being more proficient and are closer to reaching native-like proficiency, both in speaking tasks (Boers *et al.*, 2006; Stengers *et al.*, 2011; Boers *et al.*, 2014) and writing tasks (e.g., Crowther *et al.*, 2002; Dai & Ding, 2010).

One aspect of vocabulary knowledge that has received much attention over the last two decades, is formulaic language. This came as a result of the work of Pawley and Syder (1983), Nattinger and DeCarrico (1992), and Lewis (1993), who emphasise the importance of formulaic sequences. Based on the evidence above, it seems that researchers are in agreement about the benefits for L2 and FL learners to acquire "considerable numbers of multiword items" (Boers, 2021:143; Hatami, 2015), emphasising its status as "fundamental to the way language is used, processed and acquired in both the L1 and L2" (Martinez & Schmidt, 2015: 299). Formulaic language is thus central to communicative competence (Nation, 2001; Schmitt, 2004; Wray, 2002). According to Nation (2001:317-318), mastering language chunks in relation to fluency shows the benefits of focussing on formulaic language in L2 teaching.

However, language learners seem to have difficulty, especially in acquiring multiword units, collocations, lexical bundles, and idioms (Abel, 2003; Irujo, 1986; Kecskés, 2000). Hill (2000) also indicates that most students lack competence in the use of collocations, despite their value (p.50). Recently, Nguyen and Webb (2016:331) confirmed this, reporting that Vietnamese learners knew about 33% of the targeted figurative collocation meaning, despite having a relatively high language proficiency in English. That is because vocabulary is a complex phenomenon and it is difficult to develop a sense of feeling for multiword units (Martinez *et al.*, 2015). As a result, many scholars consider teaching collocations as an integral part of teaching vocabulary (e.g., Barfield & Gyllstad, 2009; Basal, 2019; Boers & Lindstromberg, 2012; Chen, 2011; Nesselhauf, 2003). Scheepers (2017:109) also indicates that undergraduate students at a South African university have difficulty with high-frequency verb-noun collocations.

Attempts to support students in building their vocabulary adequately and appropriately through formal instruction must be informed by a clear definition of vocabulary. Likewise, having an adequate definition of what vocabulary knowledge entails, and understanding what to count when counting multiword items (MWI), is crucial in order to accurately determine a learner's level of proficiency. However, defining vocabulary knowledge and, by extension, MWI, can be challenging, since such definitions depend greatly on what words or multiword items are understood to be, as well as what exactly is meant by knowing a word. The results of a study such as the present one also depend on how the object of interest – i.e., vocabulary and collocations in particular – is defined and assessed for the purpose of that particular study.

The purpose of this chapter is to explore the notion of vocabulary knowledge and MWI in order to present a working definition to inform the present study. The chapter is divided into three main sections. Section one provides a discussion of the nature of vocabulary knowledge and the dimensions of knowing a word, and distinguishing vocabulary size and depth (deep-word knowledge). Receptive and productive knowledge are also explored in relation to vocabulary size and depth. Various definitions of vocabulary knowledge are taken into consideration in order to conceptualise vocabulary knowledge. Section two aims to define collocations and collocational knowledge, and investigates the importance of collocations as part of vocabulary knowledge and in English Second Language (ESL) teaching. Section three explores the concept of the mental lexicon and form-meaning mapping in relation to vocabulary and collocational knowledge in order to inform a clear definition for the purpose of the present study and its methodology.

2.2 What is a word?

Before the concept of word *knowledge* can be defined, the following question must be answered: What exactly is a *word*?

Words are defined on different levels. They can be defined as *tokens* (items) or *types*. Tokens amount each word in a text, and repeated words are counted each time they are used (Biber *et al*, 2002; Schmitt, 2000). For example, the sentence *The children are playing on the playground*, contains seven tokens (items), because there are seven individual words. Counting tokens enables language teachers and researchers to determine the number of words in a text. Words can also be defined as *types*. Here the focus is on the number of *different* words, and repeated words are counted only once. For example, the sentence *The children are playing on the playground* contains six types of words, and 'the' is only counted once. However, there is not always consensus on what constitutes a different word type (Treffers-Daller *et al.*, 2016:306) and this poses a challenge when it comes to determining the size of a learner's vocabulary. The crucial question to ask in this instance is whether words such as *play* and *playing* would fall into the same category. In other words, do derivatives and inflections of words count as *items* related to the same word or different *types* of words.

There have been many developments in this particular area of research over the last few decades. Research shows that morphological knowledge is crucial for developing learners. Nagy *et al.* (1991), for example, found that knowledge of morphological structures, specifically derivational affixes, still develop into readers' teenage years (p.15). Several researchers (Engber, 1995; Laufer, 1991; Jarvis, 2002) include inflected forms of a word as a tokens of the same type. Others (Durán *et al.*, 2004) adopt the same view, but exclude fused forms such as *fell* (past tense of *fall*) as separate types. Durán *et al.* (2004) exclude fused forms because they do not add to lexical diversity. In contrast, Yu (2010) considers all inflected forms as different types, while Laufer and Nation (1995) include all inflected and derivational forms of a word as a word family unit up to Level 3 of Bauer and Nation's (1993:254) scale (including affixes such as *-able*, *-er*, *-ish*, *-less*, *-ly*, *-ness*, *-th*, *-y*, *non-*, and *un-*). Table 2.1 below provides a brief overview of Bauer and Nation's (1993) word families scale. Level 1 is not included, since "every different form is counted as a different word, so *develop*, *develops*, etc. are all different words at Level 1" (Bauer & Nation, 1995:254, authors' italics).

Word families			
2	develop develops developed developing	wood wood's woods wooded	bright brighter brightest
3	developable undevelopable developer(s) undeveloped	woody woodiest woodier woodiness	brightly brightish brightness
4	development(s) developmental developmentally		
5	development-wise semi-developed anti-development	wooden	brighten
6	redevelop predevelopment	anti-wood	

Table 2-1: Additions to a word family at different levels of inflection and affixation (Bauer & Nation, 1995:254)

While establishing the number of tokens and types a learner is able to distinguish can give an indication of variation in L2 learners' vocabulary, such counts provide little evidence of the extent or depth of a learner's vocabulary knowledge (Schmitt, 2001:2).

In order to determine learner vocabulary proficiency, *word families* are taken into consideration. Word families include a base word (or root) and all its inflections, and its most common derivatives, which means that, although words have "different *orthographic* (written) shapes, they are closely related in meaning" because they have the same root (Schmitt, 2001:2). Schmitt (*Ibid.*) uses the following example to illustrate:

> *Walked, walking and walks are inflections of the root word walk. However, if the affixes change the word class of a stem, the result is derivative. Thus stimulative (adjective) and stimulation (noun) are derivates of stimulate (verb).*

Word families then consist of words that are similar in meaning, with derivatives belonging to different classes of lexical words. It is important to note that different words can only be included in a word family if they have the same base form. For example, words such as *spectate* and *special* are not related because *spec* is not a root word, but *special* and *specialise* form part of the same family, just as *spectate* and *spectator* do. Should a learner know the

word family of a certain number of words, it provides a more accurate indication of that learner's knowledge of vocabulary. Bauer and Nation (1993:253) suggest that,

> *the important principle behind the idea of a word family is that once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort.*

Having discussed the types of words a learner can know, the next question is whether knowing a word entails either: a) a learner simply passively recognises a word and the meanings associated with it, or b) a learner being able to actively use that word and all, or at least most, of its associated meanings. The following section considers the notion of receptive and productive knowledge in relation to vocabulary size and depth.

2.3 Vocabulary knowledge

Lexical competence is multidimensional (Larsen-Freeman, 2007; Read, 2004), and – due to the complexity of vocabulary acquisition, learning, and development – several definitions have been put forward in an attempt to define what exactly vocabulary knowledge entails. Nation (2013) recognises that "words are not isolated units of the language, but fit into many related systems and levels" and therefore "there are many degrees of knowing" a word (p.44).

This idea has its roots in the theory that vocabulary knowledge can be separated into two dimensions, namely breadth and depth. Breadth of knowledge relates to the size of a learner's vocabulary and is concerned with how many words a learner knows, while depth (or deep-word knowledge) refers to how well a learner knows a word (Anderson & Freebody, 1981:20). Vocabulary size and depth are closely related, as there is a causal relationship between how many words a learner knows and the quality of language they are able to produce. Depth, however, is more dependent on breadth and a learner may recognise or understand many words receptively, but remain unable to produce them accurately and appropriately. Receptive vocabulary knowledge precedes productive vocabulary knowledge (Nation, 2013), and receptive learning and use are considered to be easier than productive learning and use (Ellis & Beaton, 1993; Nation, 2013). Results from studies measuring receptive vocabulary knowledge therefore cannot be assumed to be indicative of productive vocabulary knowledge (Siyanova-Chanturia, 2017:292-293).

2.3.1 Receptive vs productive vocabulary knowledge

Receptive and productive vocabulary knowledge are defined in terms of what are considered receptive and productive language skills (Nation, 2001; Palmer, 1921). Reading and listening are categorised as receptive skills where a learner uses a passive vocabulary in order to comprehend language input, whereas writing and speaking require active use of language in order to produce output (Laufer, 1998a; Meara, 1990a; Palmer, 1921) – hence the term productive knowledge. According to Nation (2001:24-25), receptive vocabulary use involves "perceiving the form of a word while listening and reading and retrieving its meaning," whereas productive vocabulary use "involves wanting to express meaning through speaking and writing, and retrieving and producing the appropriate written or spoken form". Productive knowledge can further be divided into two types: controlled and free. Laufer and Nation (1999) define controlled productive knowledge 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" (p.37). Free productive knowledge is defined as "the ability to use a word at one's free will" (*Ibid.*).

It is important to note that several terms have been used to refer to receptive and productive knowledge (cf. Meara, 1990a; Corson, 1995; and Lee & Muncie, 2006). Meara (1990a) distinguishes active and passive vocabulary in terms of word association. For Meara (1990a) words are connected in networks of association: passive vocabulary "consists of items which respond only to external stimuli; active vocabulary does not require an external stimulus, but can be activated by other words" (p. 153). Meara (1992:69) further argues that, because of these associations, L1 and L2 should view vocabulary as a structure, instead of a collection of individual words. While there is certainly merit to Meara's (1990a, 1992) arguments, Nation (2001:47; 2013:38) criticises this view, stating that "language use is not only associationally driven, but, more basically, is meaning driven." He argues that "being able to actively name an object using a second language (L2) word can be externally stimulated by seeing the object without necessarily arousing links to other L2 words" (*Ibid.*).

Corson (1995) also refers to passive and active knowledge. However, Nation (2013:38) argues that Corson's (1995) view of vocabulary knowledge is determined by use, and not by degree. For Corson (1995: 44-45), passive vocabulary includes active vocabulary as well as three other types of vocabulary: words that are only partly known, low-frequency words that are not readily available for use, and words that are avoided. According to Corson (1995), then, a speaker of

a language might have knowledge of words that they do not use, because there is no motivation to. For example, a speaker might know the meaning of a curse word while at the same time choosing not to use it.

Each of the views discussed above have different implications for both vocabulary teaching and testing. While vocabulary teaching and testing are discussed in detail in Chapter 3, it is important to note at this stage that the distinction between receptive and productive knowledge (as well as free and controlled productive knowledge) is crucial when measuring a learner's vocabulary knowledge.

Receptive and productive knowledge are not only measured differently (Lee & Muncie, 2006), but results also have different implications in terms of what can be said about a learner's comprehension and use of a word. For example, in the event that a vocabulary test (whether it be to assess receptive or productive knowledge) has an item that is considered a curse word, a learner might opt not to answer the question (Corson, 1995). The results would then indicate that the learner does not know the word, except if they explicitly indicate that they did not want to provide an answer. Meara's (1990a; 1992) word association view also poses challenges in terms of testing. Meara and Fitzpatrick's (1999) Lex30 test, for example, is used to assess free productive vocabulary knowledge by assessing whether test-takers are able to produce words that are associated with the test item. While this test is also discussed in more detail in Chapter 3, it is necessary to mention the effect that Meara's (1990a; 1992) shortcomings might have on results here. Target items' associations are pre-selected and might not include all possible associations of each item. For example, the word *baggage* will have different associations in different contexts, and importantly, for different test-takers. One person might associate *travel* with *baggage*, while another might think of *emotional issues*.

For the purpose of the present study, Nation's (2001; 2013) view of vocabulary knowledge is adopted, since this view – like Meara's (1990a; 1992) – takes into account that "words are not isolated units of the language, but fit into many related systems and levels" (Nation 2013:44). It also asserts that "there are many degrees of knowing" a word (Ibid.) and that knowledge of a word is not only determined by whether a learner *uses* a word, but whether or not a learner knows the *meaning* of a word. In order to further explore this view, it is necessary to define vocabulary size and depth as they are linked to receptive and productive knowledge. The following two sections explore these terms in order to conceptualise vocabulary knowledge.

2.3.2 Vocabulary size

Research on vocabulary size dates back to the 1920s with authors such as Thorndike (1924, cited in Nation, 2013:528), and Seashore and Eckerson (1940) emphasising the importance of vocabulary size for understanding written texts. While vocabulary size has been used to accurately predict L2 learners' proficiency levels and academic performance in South Africa (Nizonkiza *et al.*, 2013; Scheepers, 2016) and abroad (Meara & Buxton, 1987; Meara, 1996; Nation 1990; Nation & Beglar, 2007), there is no consensus on how large a vocabulary is sufficient to enable learners to communicate effectively, because there is no consensus on the size of the English lexicon is (e.g. Nagy & Anderson, 1984; Goulden *et al.*, 1990).

In addition, while it might be useful to determine the number of words in a language, words do not function in single units or in isolation. Goulden *et al.* (1990) found that *Webster's Third New International Dictionary* (1963) contained 54 000 word families, excluding proper names and alternative spellings. It is unlikely that native speakers know all these word families, and so ESL learners cannot be expected to learn the entire English vocabulary (Schmitt, 2001). The question is: How many words should a proficient learner know to be considered proficient?

Laufer (1997) established that the threshold for minimum comprehension when reading a text is knowledge of 95% of the words in that text. This translates to approximately 3 000 word families. However, more recent research indicates an even higher threshold for sufficient comprehension at 98-99% (Hu & Nation, 2000). This percentage amounts to 8 000 to 9 000 word families (Nation, 2006).

While an adequate vocabulary size is necessary for receptive tasks, productive tasks require deep -word knowledge as learners are required to not only understand language input, but also to provide comprehensible output. This requires a learner to fully know a word family and to be able to use each word in a word family accurately within different contexts. This entails accurately using words that frequently co-occur with items in a word family. For example, a learner should possess knowledge of collocates that occur with each word. The word *run*, for example, collocates in the following ways: *run to, run into, run in*. In the first instance, you *run to someone for help*. The second example could be used to indicate that *you've run into trouble* or that *you've run into someone*, i.e., went to a place and saw a person you know there by chance. Here, it is also important to distinguish between the negative and positive connotations that *run into* has in different contexts. The last example could be used to say that something is common in a particular situation, for example, *red hair runs in my family*.

Seashore and Eckerson (1940) argue that a learner's vocabulary size is dependent on individual learner differences, which, according to Anderson and Freebody (1981:82-84), could include general knowledge and mental aptitude. Anderson and Freebody (*ibid*.) distinguish three views that attempt to explain the causal relationship between vocabulary size and language use. These ideas include the instrumentalist view, the aptitude view, and the knowledge view.

The instrumentalist view assumes that good vocabulary knowledge enables good reading comprehension. The aptitude view states that a learner's mental abilities would affect all areas of learning, and therefore also the size of their vocabulary because they are able to learn vast quantities of words. The knowledge view states that word knowledge plays a crucial part in vocabulary size. An example would be if a learner has knowledge of aviation, it would be easier for that learner to learn new words within a context where aviation is discussed. Mezynski (1983:255) later suggests that all three of these views must be combined to enable a learner to use language effectively. This is known as the access view. Mezynski (1983) argues that language skill is dependent on how easily a learner can access known vocabulary, and so language fluency is taken into account for the first time. This fluency of use could be linked to vocabulary size, but more importantly, with deep-word knowledge. If a learner knows a word such as *tolerate*, but does not know its derivatives *tolerable* and *tolerant*, the learner has limited knowledge of the levels on which *tolerate* can function and thus would not be able to use, and/or potentially understand the derivatives. In order to assess whether a learner knows a word, they should be able to use it to communicate optimally within various contexts. This requires deep-word knowledge. The following section is devoted to exploring this concept in detail.

3.2.3 Deep-word knowledge

As discussed in the previous section, the number of words a learner knows could give an indication of receptive vocabulary knowledge, but it does not give any indication of the level of that knowledge. Hulstijn (2002:12) distinguishes lower- and higher-order language knowledge, arguing that lexical units represent stable form-meaning mapping, combining lower-order formal representations with higher-order cognitive representations. Thus, knowing a word is central to communicating complex concepts in various contexts. It is here where deep-word knowledge is essential, and how it is defined is fundamental to an understanding of overall linguistic development in L1 and L2 (Elgort & Nation, 2011:89). Nation (1990) identifies several

types of word knowledge that could indicate that a learner knows a word on different levels. These types include meanings of words, the written form of a word, the spoken form of a word, the grammatical behaviour of a word, collocations of the word, the register, associations, and frequency of a word. For example, a learner should know which verb tense is the correct one to use within a given context (grammatical behaviour). A learner should also know which words are not appropriate in a certain situation that requires a more formal register, for example, the difference between the words *give* and *provide*. These types can further be divided into three different categories, namely form, meaning, and use, depending on whether a word is understood receptively or used productively (Nation, 2013:48). See Table 2.2 (Nation, 2013:49), which illustrates the different types of word knowledge related to vocabulary depth.

The form of a word has to do with the way in which a word is written or pronounced. In terms of receptive skills, this would mean that a learner can recognise the word form in written or spoken input. Using word forms accurately on a productive level would mean that a learner knows how to spell and pronounce a word in order to produce comprehensible output. It is also here where knowledge of word families is important. For example, a learner should know the difference between the words *tolerant* and *tolerable* in order to communicate meaning accurately and appropriately in context. Should a learner not fully know the difference between those two words, communication could be hindered, or the learner could be misinterpreted. For example:

1. *The boy is naughty, but tolerable* means that even though the boy is naughty, his behaviour can be endured.
2. *The boy is tolerant* means that the boy is able to endure or accept something he might not agree with.

In terms of meaning, learners should be able to understand the input they might receive in written or spoken form, and infer from the context of the input which meanings are associated with a word and which are not, in that communicative context. A learner should then also be able to use that word within different contexts to convey various meanings that might be associated with that word. For example, should a word be a homograph of another (a word that is spelt the same, but pronounced differently and has another meaning), a learner should be able to discern which meaning to associate with a word based on the context. An example of such a word would be *bass*, which could mean a freshwater fish or a low-pitched sound, and the meaning would depend on the context and pronunciation.

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concepts and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this word make us think of?
		P	What other words could be used instead of this one?
Use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or type of words occur with this one?
		P	What words or type of words must we use with this one?
	constraints on use	R	Where, when and how often would we meet this word?
		P	Where, when and how often can we use this word?

Table 2-2: What is involved in knowing a word? (Nation, 2013:49)

The word referring to a freshwater fish is pronounced /bas/, while the one referring to a low-pitched sound is pronounced /beɪs/. If a learner truly knows the meaning of a word, they could also be expected to produce synonyms and antonyms for that word without difficulty, or refer to it collectively, as in the following examples:

i. The bass seemed to like the new bait. Fish were pulling at the rods all along the riverbed.

ii. *While he strummed the bass, she was hitting the top notes like an expert soprano*.

iii. *Ha! Slapping, the ketchup of the bass world!* (Randy Jackson)

iv. *Johnny played guitar; Jenny played bass* (*Pop goes the World* by Men Without Hats)

In *i*, bass is a type of fish, and the substitution of the word *fish* for bass is clear. In *ii*, bass refers to a musician playing bass guitar which is signalled by the word *strummed*, while a female musician is singing in a high voice. *Bass* and *soprano* are juxtaposed, but they are not used as direct opposites, because the meaning in context is different. Example *iii* is another reference to musical bass, but in this case, bass does not necessarily refer to a bass guitar, but rather to the bass music line, which could also be interpreted by drums, a bass violin or bass clarinet. In example *iv*, we have two guitar players distinguished by the reference to *guitar* and *bass.*

Usage of a word concerns a deeper level of understanding, as a learner should know which words can co-occur with certain vocabulary items, for example, knowing that they should say *make a decision* and not *do a decision*. In order to master this level of word knowledge, a learner must know the grammatical functions a word can serve, as well as grammatical constraints of using it. Consider for instance:

v. *A decision has to be made.*
vi. *They have to make a decision.*
vii. *A decision he must make.*
viii. *She a decision must make.*

Both *v* and *vi* are grammatical and commonly used, while *vii* is also grammatically correct, though used mainly for emphasis. However, *viii* is ungrammatical, though apparently not that much different to *iii*, apart from the position of the sentence subject in relation to object and verb.

Form, meaning, and use have been associated with what Nation (1990, 2001) respectively calls precise word knowledge, comprehensive word knowledge, and network word knowledge. For the purposes of this study, it is argued that vocabulary knowledge would develop from precise to comprehensive knowledge, until finally reaching network word knowledge. These assumptions posited by Nation (1990, 2001) follow from proposals by Richards (1976), Anderson and Freebody (1981) and Mezynski (1983).

For example, Richards (1976:83) suggests eight assumptions that cover various aspects of what knowing a word entails. These relate to Nation's (2013:49) categories. These assumptions are:

1. Native speakers' vocabulary knowledge continues to expand throughout their adult life (as opposed to grammatical competence which is relatively stable);

Furthermore, a learner should know:

2. the probability of encountering a word/lexical item in speech or print;
3. the limitations on the use of a word;
4. the syntactic behaviour of a word;
5. the underlying form of a word and its possible derivations;
6. the network of associations between a word and other words in the language;
7. the semantic value of a word;
8. the various meanings associated with a word.

While Richards' (1976) and Nation's (1990, 2001) assumptions have formed the cornerstones of how vocabulary knowledge is defined, some researchers have critiqued the role these assumptions have played in defining vocabulary knowledge (cf. Meara, 1996a; Henriksen, 1999).

An important point of criticism offered by Meara (1996a), Henriksen (1999), and Meara and Wolter (2004) is that, while these aspects provide a general description of what lexical competence entails, the grouping of knowledge components and learning processes is random, which indicates the need to "discuss the nature of and interrelationships among various aspects of lexical competence and learning and production processes" in order to develop "a unified theoretical construct of lexical competence and a model of vocabulary development" that aims to be more precise and standardised (Henriksen, 1999:304). Furthermore, Meara (1996a:46) notes that Richards' (1976:83) assumptions serve as a useful checklist to determine what it means to know a word, but it is difficult to practically test each item on the list when dealing with a large lexicon. These assumptions relate to features of individual words only, and can therefore not be used to make assumptions regarding a person's entire vocabulary when assessing vocabulary knowledge.

In this regard, Meara (1996a:46) recognises Nation's list as an improvement on Richards'. Nation's (1990, 2001) list extends Richards' assumptions in that it considers how each item is related to receptive and productive knowledge. However, what seems to be lacking is guidance toward operationalising the construct of vocabulary knowledge. No model is presented, for

example, to show how the features relate to one another or how they interrelate across a lexicon network (see Meara, 1996a:47, as discussed below).

Nation's (1990; 2001) approach still leaves much to be desired in terms of the *process* of learning a word, and this makes determining when word knowledge develops from receptive to productive knowledge challenging. It is also problematic when it comes to *testing* vocabulary knowledge. Meara and Wolter (2004:86-87) point out that the notions put forward by Richards (1976) and, subsequently, by Nation (1990; 2001; 2013) do not account for vocabulary *development*. Using the items on Richards' (1976:83) list and Nation's (1990; 2001, and later 2013:49) table, does not allow for many test items, and therefore limits the results one can obtain. It is therefore necessary to not only be concerned with the different types of word knowledge to determine when a word is known, but also with the *quality* of that knowledge in terms of associations between words.

In order to test this quality, Meara (1996a), Meara and Fitzpatrick (1999), and Meara and Wolter (2004) propose determining how a learner's vocabulary is structured or organised in terms of how one word might be associated with other words. Meara states that there is a need for a "characteristic that is able to distinguish between someone who has a large vocabulary because they have just learned a long list of words, and someone whose vocabulary is more structured than that" (Meara, 1996a:46-7). This is related to Richards' (1976:83) idea of network associations between an individual word and other words in a language, but Meara (1996a:47) suggests developing a model of exactly how "words are associated [and] the properties of the networks that these associations form". That would involve placing a word at the centre of a "complex web of associations for native English speakers, some paradigmatic, some syntagmatic, some situational, some emotional, but all of them contributing in different ways to the meaning" of that word (*Ibid.*).

Following Meara's (1996) distinction between vocabulary size and organisation, Henriksen (1999) develops a more comprehensive definition that takes into account the complexity of language learning, especially learning lexical items. Henriksen (1999) argues for a third dimension of vocabulary knowledge, namely the receptive-productive continuum. According to Henriksen (1999), receptive vocabulary knowledge can be defined as the kind of knowledge a learner needs to understand input in the form of reading or listening, while learners need productive vocabulary knowledge in order to produce comprehensible output when speaking or writing. Henriksen (1999:304) proposes three dimensions of vocabulary knowledge: (a) partial-precise knowledge, (b) depth of knowledge, and (c) a receptive-productive dimension.

The partial-precise dimension is related to vocabulary size as defined by Meara (1996a) and Nation (1990). The second dimension, depth of knowledge, generally refers to Read's (1993) notion of "quality of the learner's vocabulary knowledge". As mentioned, vocabulary knowledge is multi-faceted and multidimensional. Henriksen (1999) expands on this idea by distinguishing paradigmatic, syntagmatic, and analytic relations of a word. Paradigmatic relations refer to synonyms and related meanings of a word, while syntagmatic relations refer to collocations of a given word, and are related to Nation's (1990, 2001, 2013) idea of comprehensive word knowledge. An example of a syntagmatic relation is knowing the combination *make a decision* instead of saying *do a decision*. Syntagmatic relations are the focus of this study.

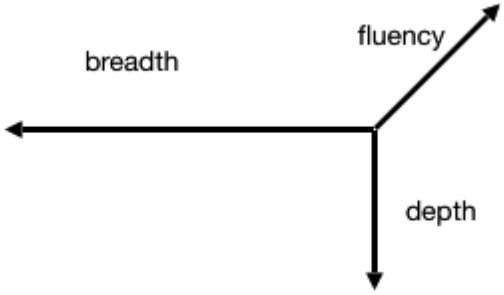


Figure 2-1: The lexical space: dimensions of word knowledge and ability (Daller et al., 2007:7)

Henriksen's (1999) view is supported by more recent research. Jiang suggests that knowing an L2 word includes knowing "semantic differences from its L1 translation and other semantically related L2 words" (2004:427). Both Jiang (2004) and Hulstijn (2002) suggest that lexical knowledge goes beyond merely knowing how a word changes form in different contexts; knowing a word can be conceptualised as having pragmatic knowledge and sociocultural competence. Erman, Lundell and Lewis (2016:111-112) emphasise that high levels of language proficiency are closely linked with high levels of idiomaticity: "Being highly proficient in a language means not only knowing how to combine words according to their selectional and morpho-syntactic restrictions but also knowing what linguistic expressions in what situations and contexts" (2016:111). This has been found to be the case in several studies (see Columbus, 2010; Ellis *et al.*, 2008).

Thus, knowing a word entails placing a word within a broader context of society and knowledge of the world, human attitudes, beliefs and their social behaviour in a given language community (Elgort & Nation, 2011:91). Sinclair (2004:173) notes that learners should be able to make connotational connections between lexical items, for example, the verb *cause* often occurs in

a negative context, such as *cause an accident*. He refers to this as semantic prosody. Daller *et al.* (2007:7) also emphasise the importance of viewing a learner's vocabulary knowledge as a three-dimensional lexical space, recognising different aspects of knowing a word as presented in Figure 2.1.

The horizontal axis represents lexical breadth, or vocabulary size, to indicate how many words a learner knows, regardless of how well they know them. This is associated with receptive knowledge, and relates to form and meaning as introduced by Nation (2001). The vertical axis is concerned with lexical depth and resembles how much a learner knows about a word, including comprehensive and network word knowledge. Lexical depth would then include associations, grammatical functions, collocations and limitations to how a word can be used (Nation, 2001; Richards, 1976). The horizontal and vertical axes each represent elements of traditional assumptions about vocabulary knowledge.

However, the final axis represents fluency, indicating "how readily and automatically a learner is able to use the words they know and the information they have on the use of these words" (Daller *et al.*, 2007:8). It is here where a learner's lexical competence is put to use, and where the transition between receptive knowledge and productive knowledge becomes clear. This transition is especially important when dealing with collocational knowledge because collocations play a crucial role in reaching native-like proficiency (cf. Boers *et al.*, 2006; Boers *et al.,* 2014; Martinez & Schmidt, 2015; Boers 2020). In order to better understand this transition and the role of collocations in vocabulary knowledge, section 2.6 explores the concept of collocations; considers collocational knowledge as it relates to vocabulary knowledge; and investigates how to go about identifying which collocations that must be taught.

2.4 Collocational Knowledge

It has been established that knowing many multiword units is essential for language fluency because it reduces processing effort and therefore language is more readily available. This is especially true of collocations (Barfield & Gyllstad, 2009; Boers *et al.*, 2006; Nesselhauf, 2003, 2005), since learners who are competent in using them generally come across as being more proficient and native-like (Barfield & Gyllstad, 2009; Boers *et al.*, 2006:2-3; Crossley *et al.*, 2015; Erman *et al.*, 2016; Hsu & Chiu, 2008; Keshavarz & Salimi, 2007). Nation (2013:84) notes that collocations indicate patterns in language and being aware of those patterns can reduce the learning burden of certain words. However, like vocabulary knowledge, collocations and other multiword units are difficult to define and identify (Henriksen, 2013:30). Henriksen (2016:29)

provides an overview of formulaic sequences, which include chunks of lexical items such as idioms (*we'll cross that bridge when we get there*), figurative expressions (*break a leg*), pragmatic formulas (*see you soon*), discourse markers (*give me a minute*), lexicalised sentence stems (*as a result of this…*), and collocations (*heavy rain, strong wind*).

Defining formulaic sequences is difficult because they are diverse and serve different purposes (Schmitt & Carter, 2004:3). Collocations form part of a larger group of expressions that are also called *formulaic sequences* (see Erman *et al*, 2016) and multiword items (Martinez *et al*. 2015; Siyanova-Chamturia, 2017; Wray, 2002). Schmitt and Carter (2004:3) note that formulaic sequences can be used to "express a message or idea, functions [such as declining assistance], social solidarity, and to transact specific information in a precise and understandable way". A wide range of terminology has been used to refer to these expressions, such as *prefabricated units*, *prefabs*, *phraseological units*, *(lexical) chunks*, *multiword units*, *collocations*, *conventionalised forms*, and *prefabricated routines* (Nesselhauf, 2005:1; Wray 2002:9), to name but a few. For the purposes of this study the term *collocation* will be used as a subset of formulaic sequences.

Several definitions of collocations have been suggested. Nesselhauf (2005:1) describes them as expressions that are "made up of more than one word and are lexically and/or syntactically fixed to a degree" (Nesselhauf, 2005:1). Wray (2002:9) defines formulaic language as

> a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.

For the purposes of the present study, only collocations will be considered as a category of multiword units that function as a one unit that contains meaning. Defining the term *collocation* has proven difficult because various definitions have been proposed over the last several decades. These definitions all subscribe to one of three views of collocation: the frequency-based approach, the phraseological approach, or more recently, a combination of the two (Granger and Paquot, 2008; Nesselhauf, 2005).

2.4.1 The frequency-based approach

The frequency-based approach started developing with Firth in the 1950s, and was later expanded by Halliday (1966) and Sinclair (1991). According to this view, collocation is "the

occurrence of two or more words within a short space of each other in a text" (Sinclair, 1991:170) and when words "co-occur more often than their respective frequencies and the length of the text in which they appear would predict" (Sinclair 1974:21), they are seen as significant collocations. In other words, not all words that co-occur within a short span can be categorised as important combinations of words. For example, many nouns would co-occur with article or determiners (*the+n* or *my+n*), but that co-occurrence in itself does not mean that *the+n* is a collocation.

Halliday (1966, cited in Halliday, 2002:168) introduced concepts used to describe collocations such as 'node' (the main word in a collocation), 'collocate' (words that co-occur with the node), and 'span' (the distance between the node and the collocate). Sinclair (1991:170) later proposed that the optimal span is approximately four words to the left as well as to the right of the node. According to the frequency-based approach, "the degree of likelihood of two words co-occurring in a corpus within a given span of discourse can be quantified through one of the available measures of collocational strength such as the mutual information (MI) score" (Boers & Webb, 2018:78). The higher the score, the more likely words are to co-occur.

2.4.2 The phraseological approach

There has been some disparity in terms of how collocation within the frequency-based approach is understood. This is largely due to how researchers define what constitutes a word, and whether derivations and inflections of a lexical item are considered individual forms, and therefore each collocate that occurs with each different form would be a collocation in its own right (Nesselhauf, 2005:13). However, if collocations are considered to be a relationship between lexemes of a given lexical item, "*a strong argument, he argued strongly, the strength of the argument, his argument was strengthened* would all be considered instances of the collocation" (Nesselhauf, 2005:13, author's emphasis).

It stands to reason then, that frequency alone cannot be used to identify collocations. The second approach to collocations is the phraseological approach with Cowie and Howarth as its main proponents (Nizonkiza *et al.*, 2013:167). Cowie (1998:145-146) considers collocations as a kind of word combination that is different from idioms and what he calls 'free combinations', and he is concerned with the syntactic function of collocations and their constituents, as well as transparency and substitutability. Transparency refers to whether or not the combination of words has a literal or non-literal meaning, whereas substitution refers

to “whether and to what degree the substitution of the elements is restricted” (Nesselhauf, 2005:14).

Cowie (1994) distinguishes four types of word combinations: free combinations, restricted collocations, figurative idioms, and pure idioms. Free combinations include combinations such as *drink water* or *eat food*. All elements in the combination have literal meanings, and substitution is specified on semantic grounds, while restricted collocations have at least one element with a non-literal meaning. In terms of substitution, the limitations are arbitrary and only some substitution is possible. For example, the combination *perform a task* is possible, so is *complete a task*, but not *make a task*. Figurative idioms on the other hand are combinations that have figurative meaning, but literal interpretations. An example would be *do/make a U-turn* and substitutability is seldom possible. Pure idioms such as *take the bull by the horns* leaves no room for substitutability and the combinations are purely figurative.

For the purposes of this study, only restricted collocations with limited substitutability were considered. This decision was made based on evidence that these collocations seem to be easier for learners to acquire. According to Lesniewska (2006:98), fixed multiword units (e.g., idiomatic expressions) may be less problematic for learners than multiword units with subtler restrictions, such as collocations. Howarth (1998:42) notes that the challenge with collocations lies in the fact that some combinations are free, and some are restricted. Furthermore, collocations are challenging because they are not entirely ‘free’, but not absolutely ‘fixed’ either (Howarth, 1998:160).

The phraseological approach requires that all elements of a collocation are syntactically related, and therefore only some combinations are possible. However, the restriction placed on collocations once again depends on the approach taken. Hausmann (1997, cited in Nesselhauf, 2005:17) only considers combinations that “appear in a pre-defined set of syntactic relations” such as adjective-noun, verb-noun, adverb-verb, and adverb-adjective combinations. Hausmann also “only allows the combination of two lexical elements” and would not allow prepositions to be part of those combinations (Nesselhauf, 2005:17).

This introduces an important distinction made by Benson, Benson and Ilson (1997) between grammatical and lexical collocations. Lexical collocations are usually limited to collocations consisting of lexical words: nouns, (main) verbs, adjectives and adverbs, whereas grammatical collocations include lexical words as well as functional words such as prepositions. They also include combinations with a *that* clause and the *to* + infinitive.

For this study, the focus is on lexical collocations. Below is a summary of the lexical collocations listed by Benson *et al.* (1997:xxxi-xxxiv):

1. L1: verb-noun/pronoun for example, inflict a wound, set a record;
2. L2: verbs essentially meaning to eradicate/nullify in combination with a noun, for example, reject an appeal. These collocations are also known as EN collocations.
3. L3: adjective-noun combinations, for example, strong tea or a sweeping generalisation. These collocations also include nouns that are used as adjective, for example, house arrest.
4. L4: noun-verb collocations where the verb names an action characteristic of the person or thing designated by the noun, for example, blizzards rage and adjectives modify.
5. L5: collocations that indicate a unit that is associated with a noun, for example, a pack of dogs.
6. L6: adverb-adjective collocations that would include hopelessly addicted or utterly disappointed.
7. L7: verb-adverb collocations. Examples are apologise humbly or affect deeply.

However, for the purposes of this study, lexical collocations L1 and L2 will be combined to only form a single category: verb-noun collocations. The motivation for combining L1 and L2 from the list above comes from Nesselhauf's (2003) definition of collocations, specifically related to verb-noun collocations.

Central to Nesselhauf's (2003) definition is the notion of a classification of combinations called *restricted sense*, taking into account the arbitrary constraints on verb-noun combinations. This relates to Cowie's (1994) distinction of free combinations, collocations, and idioms. In each of these categories, combinations are more or less restricted and substitution of elements of the combinations is not possible to the same degree. Nesselhauf (2003) applies this idea specifically to verb-noun collocations, stating that in a combination, the verb is "the basis for the distinction of free combinations and collocations" because "if the sense in which the verb is used is restricted, the combination is a collocation (rather than a free combination)" (p. 226).

A potential problem related to verb-noun collocations is that verb forms often consist of phrasal verbs that include prepositions, for example, *bear down on something/someone*, and excluding those combinations because they contain prepositions could lead to a less comprehensive

picture of verb-noun collocations as a whole. To illustrate, consider the verbs *see*, *watch*, and *look*. The word see can be combined with most nouns, such as *see a car*, *see a person*, *see a movie.* All of these combinations are possible with *watch*, but there is a restriction to some degree because of the semantic properties of watch. However, in the case of *look*, the verb has to be followed by a preposition, for example, *look for [something]*. Here, Henriksen's (2013) definition of collocations provides some guidance. She defines collocations as "frequently recurring two-to-three word syntagmatic units which can include both lexical and grammatical words" (p. 30). This view also allows for a combination of the frequency-based and phraseological approaches.

In other words, given the complex nature of collocations, it seems reasonable to include lexical units that have a statistically significant re-occurrence rate, and those units should be semantically and syntagmatically related to form meaning in a particular context. This is especially important, since several studies have found that verb-noun collocations in particular are challenging for L2 learners, both in spoken and written language (e.g. Barfield, 2003; Bonk, 2001; Laufer & Waldman, 2011; Nesselfhauf, 2003; Nizonkiza *et al*., 2013, Nizonkiza, 2017; Revier & Henrisksen, 2006; Webb & Kagimoto, 2010) and there seems to be a need to address this issue further and focus more on explicitly teaching verb-noun collocations. Following these findings, the present study focusses only on verb-noun collocations.

2.4.3 The combined approach

The third approach to identifying collocations is a combination of the frequency-based and phraseological approach to both defining and extracting collocations from texts. As discussed, the frequency-based approach to identifying collocations could pose a challenge in that combinations could be identified that have no semantic relation or syntactic meaning (see section 2.4.1). However, the phraseological approach alone also falls short when it comes to identifying collocations. Given the complex nature of collocations as part of a larger group of multiword units, one cannot rely on linguistic intuition alone in terms of restrictedness of word combinations. In recent years, researchers have turned to a combination of frequency-based corpus data extraction and the phraseological approach. An example of such a study includes Al-Thubaity and Baazeem (2017), who developed a collocation extraction tool (*Musaheb*) for Modern Standard Arabic. *Musaheb* is able to extract *n*-gram collocations that can be limited to a 2-gram language model, but the collocates of specific word types may also be extracted to up to 15 positions from the node. In addition, it also allows the user to choose one of eight different collocation strength measures to create a collocation list (Al-Thubaity & Baazeem,

2017:25-26). An extraction tool such as this one addresses what is lacking in the frequency-based as well as the phraseological approach, as it uses both statistical data and syntactic and semantic relations to determine the likelihood of a word combination being a collocation. While it does not eradicate the possibility of an extraction tool identifying a node and collocate that are not true collocations, it does contribute to the field of teaching collocations in a significant way. Information extracted from collocations tools such as *Musaheb* have played a role in creating dictionaries (Kallas *et al.*, 2015), data-driven language learning (Rahimi & Momeni, 2012), second language acquisition (Demir, 2017), and computational linguistics (Schultz & Wilker, 2016), to name a few. These extraction tools have enabled researchers to accurately identify collocations that are most frequently used and, therefore, most relevant for teaching.

Thus, the present study follows the combined approach to identify and define collocations. It adopts the definition presented in the *Oxford collocations dictionary for students of English* (2009) that collocations are "common word combinations such as 'bright idea' or 'talk freely' – [that] are the essential building blocks of natural-sounding English". However, according to Paquot and Granger (2012:136), collocations are "lexically constrained combinations that allow for limited substitution within a particular grammatical construction (e.g. verb-object, adverb-adjective, or adjective-noun)". There are also constraints in terms of register and context. On the one hand, researchers cannot fully rely on extraction tools used to identify collocations based on frequency only. On the other hand, they cannot fully rely on the phraseological approach, since human error could play a significant role in misidentifying collocations. For example, the words *play a video game* might co-occur frequently, but might not be a collocation *per se*, while *play a role*, is a recognised collocation in various registers. Thus, it is necessary to take a more theory-based approach to the combined approach on defining and identifying collocations.

The combined approach also ties in with the psycholinguistic approach to language acquisition, which focusses on the "cognitive processes that support the acquisition and use of language" (De Bot & Kroll, 2010:124). In other words, it is assumed that language use involves cognitive processes that draw from various faculties of the human brain to attach abstract meaning to the physical form and sound of language. The cognitive processes relevant for vocabulary development and use are 1) how words are recognised receptively and 2) how abstract thought is turned into lexical items that form meaning, and are then used productively. In order to explain this process, one can apply Levelt's 'Speaking' model (1989, 1999). While Levelt's (1989; 1999) focus is on speaking, and this study is concerned with written language, both

these skills are categorised as productive language skills (see section 2.3.1). The basic principles of the 'Speaking' model are therefore relevant.

The model aims to describe this process in three components of processing (Levelt, 1989:11). First, a language user 'conceptualises' a communicative intention which is turned into something that can be expressed linguistically. The second component is the 'formulator'. This component entails single words and meanings that are turned into sentences. These sentences are translated into sounds by the articulator, which is the third component. This process is of particular importance with regard to vocabulary, "especially as language production is largely lexically based" (De Bot & Kroll, 2010:126). According to Levelt (1989), formulation processes are lexically driven (p.181). When language is produced, lexical items have to be selected from what is known as a mental lexicon. The mental lexicon includes all aspects of a word: meaning, syntax, phonology, and morphology. Below is an illustration of Levelt's (1989:182) depiction of the mental lexicon of a lexical item.

Figure 2-2: Internal structure of the mental lexicon. (Adapted from Levelt, 1989:182)

Accessing information about a lexical item in order to use it to express meaning, entails accessing all aspects of knowing the word and 'activating' that knowledge. Through this activation, syntactic procedures are triggered and that, in turn, leads to sentence formation. For sentences to be formed and produced, items in the mental lexicon have to be activated on two levels: the lemma and the lexeme. According to Levelt (1989), "semantic information in a lemma specifies what conceptual conditions have to be fulfilled in the message for the lemma to become activated" (p.189), while the morphological and phonological structures of the mental lexicon are "composed of smaller lexical elements… [T]he process by which the speaker retrieves this information…suggests the existence of a multi-level organization of word-form properties" (p.284). This is known as the lexeme.

Thus, the lemma refers to the meaning of a lexical item and the lexeme categorises that lexical item according to its morpho-phonological properties. For the purpose of this study, the researcher adopted a psycholinguistic approach, and focussed on form-meaning connections. The researcher therefore adopted this approach in both proposed teaching intervention strategies, and data collection and selection. Since meaning is the main focus of this study, certain errors that participants might make are accepted as correct. For example, all of the statements below will be accepted:

1. Yesterday I made a decision.
2. Yesterday I make a decision.
3. Yesterday I making a decision.
4. Yesterday I makes a decision.

The argument for accepting all these statements as correct is that each sentence contains the word *make* as a verb. It could be argued that each participant is able to recognise the associative meaning between the words *make* and *decision*. Thus, all inflections of the verb form will be accepted. However, should a participant use a derivative of the verb or misspell a word and inadvertently use a completely different word, that item will not be accepted. For example, if the word *paid* (pay) is misspelled as *payed*, the item will not be counted.

2.5 Conclusion

The aim of this chapter is twofold. On the one hand, it defines vocabulary knowledge in order to inform a definition of collocations. This is done by determining 1) what a word is; 2) when a word is known; and 3) the dimensions of knowing a word. Receptive and productive knowledge are defined, and the importance of developing the receptive-productive continuum is emphasised in relation to developing fluency. Accordingly, a combination of Henriksen's (1999) and Nation's (2013) view was used to inform the definition of vocabulary knowledge for the purpose of this study.

On the other hand, the aim is to determine the most effective way to identify collocations, since they are a crucial aspect of overall language proficiency and processability. It is established that vocabulary knowledge entails knowing words on a pragmatic and syntagmatic level which includes being able to use words within restricted contexts. Therefore, the definition of collocations adopted for the present study relies heavily on a combination of the frequency-based and phraseology approach, as set out in Henriksen (2013:30).

In short, the researcher's definition of vocabulary knowledge and collocations may be summarised as follows: in order to fully know a word, a learner must know that a word can function on various levels, depending on the context. A word may take on various forms such as having inflections or derivations, and these forms have limited functions and meanings. In terms of collocations, learners should be aware that the limitations on usage of words are not only determined by grammatical structure of language, but also by syntagmatic form-meaning connections. These connections are not automatically known but learnt as receptive comprehensible input becomes productive comprehensible output.

While it is not possible to address the full vocabulary of a group of learners, researchers can focus on areas that have proven to be problematic. As discussed in 2.4.2, verb-noun collocations seem particularly difficult for second and foreign language learners. Thus, the present study focusses on the controlled productive use of verb-noun collocations within an academic context. Chapter 3 aims to address this issue by discussing the teaching and testing of vocabulary, with a particular focus on collocations.

CHAPTER 3: TEACHING AND ASSESSING VOCABULARY

3.1 Introduction

As established in Section 2.1, there is a causal link between vocabulary knowledge, overall language proficiency (see Meara & Buxton, 1987; Meara, 1996; Nation, 1990; Nation & Beglar, 2007) and academic success (Gardner, 2013; Scheepers, 2016) in both L1 and L2 performance. Chapter 2 further explored different types of word knowledge, distinguishing size and depth, and receptive versus productive knowledge (Corson, 1995; Laufer, 1998; Lee & Muncie, 2006; Meara, 1990a; Nation, 2001; 2013; Palmer, 1921). This distinction is especially important when aiming to determine how well a learner knows and can use a word or phrase, specifically in relation to collocations (Boers *et al.*, 2006; Boers *et al.*, 2014). While vocabulary plays a pivotal role in language use, it has been established that learners often lack the ability to fully comprehend language and effectively utilise it because of low levels of word knowledge (Horiba, 2012:108). Consequently, the current research is concerned with the role of formal instruction to improve vocabulary knowledge.

For the purpose of the present study, this chapter provides an overview of strategies that have been employed to teach and test knowledge of lexical items, especially in relation to collocations. Various strategies are considered to determine the most effective one for teaching multi-word units in an attempt to inform the empirical component of the present research. This chapter, therefore, addresses the sub-question of the role that instruction plays in vocabulary learning and acquisition, as well as the teaching strategies that could be employed to teach vocabulary – and specifically collocations – most effectively. The chapter is divided into three sections: the first section outlines form-focussed and meaning-focussed instruction, section two explores strategies employed in teaching vocabulary with a specific focus on collocations, while section three addresses key issues in the kinds of tasks utilised to test knowledge of vocabulary and collocations. The chapter concludes with an outline of the theoretical stance taken in relation to the instruction and assessment of collocations used to inform the methodology of the present study.

3.2 Form-focussed vs meaning-focussed instruction

Against the backdrop of Krashen's distinction (1981; 1982; 1985) between learning and acquisition, the role of formal instruction was debated extensively (see e.g. Ellis, 1996; Li, 2010; Norris & Ortega, 2000). If it is assumed that only unconsciously acquired language can be spontaneously retrieved, it stands to reason that formal instruction cannot have an effect on language acquisition, as it takes place through naturalistic exposure to meaning-focussed communication. According to Savingnon (1983:vi), learners learn language automatically as a result of "the opportunity they are given to interpret, to express, and to negotiate meaning in real-life situations." The assumption that language acquisition occurs naturally is strengthened by Krashen's input hypothesis (1981; 1982; 1985), which assumes that a learner will acquire language forms through comprehensible input, and through what he termed 'i+1'. Krashen's argument stems from the assumption that acquisition of a second language follows the same natural order as that of a first language, and that learners need input that provides examples of language forms that naturally develop in sequence of one another. Acquisition then takes place when a learner understands language containing 'i+1', and this will happen automatically when communication is successful (cf. e.g. Alahmadi, 2019). "Acquisition requires meaningful interaction in the target language – natural communication – in which speakers are concerned not with the form of their utterances but with the messages they are conveying and understanding" (Krashen, 1989:22).

While research on Second Language Acquisition (SLA) in the 80s and early 90s challenged the importance and relevance of instruction in language acquisition and learning (see Felix, 1981; Krashen, 1982, 1985; Prabhu, 1987; Schwartz, 1993), the effectiveness of formal instruction in second and foreign language learning is no longer contested (Ellis, 2001; Nassaji, 2016; Norris & Ortega, 2000; Spada, 1997). The idea that attention to form should be an integral part of language classrooms within a communicative context "developed as a reaction against a communicative approach which advocates the exclusive use of meaning-focussed activities in language classrooms" (Nassaji, 1999:385-386). Thus, there was a need to incorporate formal instruction into an approach that was initially only meaning-focussed. Nassaji (1999) states that theoretical perspectives on language learning and teaching have since changed and that "[n]ew perspectives advocate a principled, form-focussed approach to L2 learning, arguing that a totally message-based approach is inadequate for the development of an accurate knowledge of language" (p.386). This shift was informed by various proponents, including Schmidt (1990; 1992), whose Noticing Hypothesis suggests that language learning

is a conscious process during which learners are aware of the forms of linguistic features in the input they receive. Swain (2000) argues that learners who make an effort to communicate their messages effectively develop second language proficiency better than those who do not. Other researchers found that, while learners who did not receive any instruction on grammatical forms scored high on comprehension and fluency, they did not achieve high levels of accuracy (Harley & Swain, 1984; Lapkin *et al.*, 1991; Alahmadi, 2019).

The key difference between communicative language teaching (CLT) and earlier methods of teaching is how language use is viewed. Savignon (2006:673) notes that in CLT, “it is the engagement of learners in communication [that allows] them to develop their communicative competence.” This is in contrast to earlier views that focussed on communication without form.

Howatt (1984:279) distinguishes between a weak and a strong interpretation of CLT:

> The “weak” version… stresses the importance of providing learners with opportunities to use their English for communicative purposes and, characteristically, attempts to integrate such activities into a wider program of language teaching… The “strong” version of communicative language teaching, on the other hand, advances the claim that language is acquired through communication, so that it is not merely a question of activating an existing but inert knowledge of the language, but of stimulating the development of the language system itself. If the former could be described as “learning to use” English, the latter entails “using English to learn it".

However, Savignon (2006:676) notes that it is a common misconception that CLT is strictly meaning-focussed, since it

> does not exclude metalinguistic awareness or conscious knowledge of rules of syntax, discourse, and social appropriateness... CLT is properly seen as an approach or theory of intercultural communicative competence to be used in developing materials and methods appropriate to a given context of learning.

Regardless of whether one adopts a strong or weak version of CLT, it is important to recognise that the reference to "form" should not necessarily be understood to mean grammatical form only, and that there are various ways of incorporating focus on form. Ellis (2016:408) states that ‘form’ can refer to lexical, grammatical, and pragmalinguistic properties of language. Form-focussed instruction is not only concerned with grammatical forms, but can be expanded to all aspects of vocabulary as well (Laufer & Girsai, 2008:695). Consequently, the notion is considered here to contextualise the focus of this chapter.

Two main distinctions have been made with regard to types of form-focussed instruction (FFI): Focus on Form (FonF) (Long, 1991:44) and Focus on Forms (FonFs) (Long & Robinson, 1998). Long (*Ibid.*) defines FonFs as the structural and synthetic approach to language teaching where learners' exposure to language is isolated and de-contextualised. This type of instruction "assumes that the target L2 forms can and need to be taught one by one in a sequence externally orchestrated according to linguistic complexity" (Norris & Ortega, 2000:420). A learner's attention is drawn to "formal properties of a linguistic feature which appears to cause trouble on that occasion, is learnable given the learner's internal developmental state, and is likely to be useful in future communication" (*Ibid.*). However, Spada (1997), Williams (2005) and Ellis (2001; 2015; 2016), among others, expand on the definition of FonF. Doughty and Williams (1998b:4), for example, distinguish between FonF and FonFs by stating that:

> Focus on formS and a focus on form are *not* polar opposites in the way that 'form' and 'meaning' have often been considered to be. Rather a focus on form *entails* a focus on the formal elements of language, whereas focus on formS is *limited* to such a focus, and focus on meaning *excludes* it… the fundamental assumption of focus-on-form instruction is that meaning and use must be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across. (Authors' italics)

Nation (2001; 2007; 2013) prefers to use the term language-focussed learning, as this eliminates the assumption that the focus should only be on grammatical aspects of language, and allows for independent *learning* in contrast to assuming that this process can take place through *instruction* only. This means that language-focussed learning includes all aspects of language. This would include levels of vocabulary knowledge that relate specifically to collocational knowledge, as discussed in Chapter 2.

The role of FFI within the context of CLT has therefore been reassessed and L2 FFI is considered to be potentially effective "if it is provided appropriately" (Nassaji, 2016:35). Nassaji (1999:386) notes that more recent perspectives of CLT "advocate a principled, form-focused approach to L2 learning, arguing that a totally message-based approach is inadequate for the development of an accurate knowledge of language". The present research therefore departs from the assumption that L2 FFI is potentially valuable to improve vocabulary and collocational knowledge. In turn, such knowledge is likely to support ESL proficiency development.

More recently, however, the focus has shifted to which FFI is most beneficial (Nassaji, 2016) and how FFI can be successfully integrated into a communicative setting. These questions are considered to inform the present study.

Nation (2001; 2007; 2013) recommends equally incorporating four strands in language courses that contribute to creating an optimal learning environment, given that certain conditions are met. These strands can be implemented to aid learning all features of language, including vocabulary. Similar cases have been made regarding vocabulary instruction, specifically when dealing with technical vocabulary and collocational knowledge (see Nation, 2001, 2013; Nizonkiza, 2012; Nizonkiza *et al.*, 2013; Nizonkiza & Van de Poel, 2019). In language-focussed vocabulary learning and teaching, learners' attention can be drawn "to lexical items (single words and multi-word units) within a communicative task environment, if these lexical items are necessary for the completion of a communicative, or an authentic language task" (Laufer & Girsai, 2008:695). Spada (1997:82) also suggests that FFI is most beneficial to SLA when it is incorporated in a communicative setting, and the four strands suggested by Nation (2001; 2007; 2013) allow for this to be done effectively in vocabulary teaching.

The progression from purely meaning-focussed teaching to integrated meaning-focussed and form-focussed teaching has led to much research on the effects of formal vocabulary instruction (Allen, 1983; Gairns & Redman, 1986; Nation, 1991; 2001; Lewis, 1997), the assessment of vocabulary (Meara & Buxton, 1987; Nation, 1990; 2001; 2007; Read, 1997; 2000; Meara *et al.*, 2007), and the validity and efficacy of tests used to assess vocabulary knowledge (Fitzpatrick, 2007; Meara, 2005; Richards & Malvern, 2007; Van Hout & Vermeer, 2007). Each of these aspects is discussed in the remainder of this chapter, along with an outline of the four strands of language courses (Nation, 2001; 2007; 2013) to be applied in vocabulary teaching.

3.3 Vocabulary Teaching Strategies

As far as vocabulary-teaching strategies are concerned, two main approaches have been adopted to implement form-focussed instruction: explicit and implicit teaching techniques (Nassaji, 2016). Within these categories, FFI instruction can either be isolated or integrated, deductive or inductive. While other strategies, such as input enhancement (see White, 1998; Williams & Evans, 1998; Lee, 2007; Winke, 2013), processing instruction (PI) (see Banati, 2005; VanPatten & Cadierno, 1993; VanPatten & Oikkenon, 1996), interactional or corrective feedback (Locktman, 2002; Lyster & Mori, 2006; Nassaji, 2015; Oliver & Mackey, 2003), consciousness-raising tasks (Nassaji & Fotos, 2010), or incidental FonF (Loewen, 2003; Ellis *et al.*, 2001) have been utilised to teach form, each of these can be placed within the broader framework of the approaches and categories listed above, and they can be both explicit and implicit. Table 3.1 illustrates FFI techniques on a matrix, showing where they can be placed in

terms of the approach and category they fall under. As illustrated, almost every strategy, whether isolated or integrated, deductive or inductive, can be both explicit and implicit.

This section first provides definitions of the strategies mentioned above, as well as a brief overview of studies that have investigated the effects of various types of instruction in both linguistic and lexical features of ESL and EFL learners. Focus is then directed to explicit and implicit strategies employed specifically in vocabulary teaching, including general vocabulary and collocations. The section concludes with a synthesis of strategies employed in the present study in an attempt to improve first-year university ESL learners' productive knowledge of academic verb-noun collocations.

	Explicit	**Implicit**
Isolated	Interactional/corrective feedback	Interactional/corrective feedback
Integrated	Input enhancement Processing instruction Interactional/corrective feedback	Input enhancement Interactional/corrective feedback
Inductive		Incidental FonF
Deductive	Input enhancement	Input enhancement

Table 3-1: Matrix of vocabulary teaching approaches and strategies

3.3.1 Explicit vs implicit teaching and feedback strategies

Scott (1990:779) defines an explicit approach to teaching grammar as one that "insists on the value of deliberate study of grammar rule in order to recognise linguistic elements efficiently and accurately", while an implicit approach "suggests that students should be exposed to grammatical structures in a meaningful and comprehensible context in order that they may acquire, as naturally as possible, the grammar of the target language". In other words, explicit teaching strategies draw learners' attention to consciously learning language forms, while implicit learning can occur through repetition without deliberate effort on the part of the learner.

Many studies have investigated whether explicit strategies are more effective than implicit strategies and vice versa. For example, Scott (1989) compared two groups of French learners and found that the group receiving short-term explicit teaching on relative pronouns and the

subjunctive outperformed the group that received implicit teaching on a written test. They did not, however, outperform the group receiving implicit instruction on an oral test. Carroll, Roberge and Swain (1992) as well as Carroll and Swain (1992) examined the effects of corrective feedback, with the latter study indicating that explicit corrective feedback performed better than other feedback groups. De Graaff (1997) also found that explicit instruction assisted the acquisition of morphologically and syntactically complex features of an artificial language in 56 Dutch undergraduate students. In a meta-analysis of 49 studies published between 1980 and 1998, Norris and Ortega (2001) found that, compared to implicit FFI, explicit FFI was more beneficial and had lasting effects. In a 2012 study, Nazari also found that explicit teaching strategies have a better effect on improving the EFL learners' productive use of the present perfect tense than implicit strategies in elementary level adult learners. Tasks in Nazari's (2012) study included direct explanation of grammar rules, target-item extraction from sentences, completing exercises, and translation (p. 158). There are several other studies that show similar results when explicit teaching or feedback strategies were compared with implicit strategies (e.g. Alanen, 1995; Robinson, 1996; Laufer & Girsai, 2008; Spada & Tomita, 2010).

In summary, the discussion of implicit and explicit teaching strategies presented here suggests that explicitly teaching language elements is most beneficial for the improvement of learners' language knowledge and use. In the next section, integrated and isolated teaching strategies are considered.

3.3.2 Integrated vs Isolated teaching and feedback

Spada and Lightbown (2008) draw a distinction between isolated and integrated FFI. They define isolated FFI as "the provision of instruction in lessons whose primary purpose is to teach students about a particular language feature" (p.187). This could be to prepare a learner for a communicative task when "the teacher believes that students are unlikely to acquire the feature during communicative activities without an opportunity to learn about the feature in a situation where its form and meaning can be made clear" (*Ibid.*). In contrast, integrated FFI "occurs in classroom activities during which the primary focus remains on meaning, but in which feedback or brief explanations are offered to help students express meaning more effectively or more accurately within the communicative interaction" (Spada & Lightbown, 2008:187). In other words, explicit instruction only occurred in the feedback section of tasks, and not during input. The results of this study indicated that both isolated and integrated instruction were beneficial, depending on the language feature being learned, the learner characteristics, and the learning conditions. Spada and Lightbown (2008:200) note that isolated instruction might be more

beneficial in a setting where learners share the same L1 and need to overcome similar problems brought on by L1 interference, whereas integrated instruction could be more useful in helping develop fluency and automaticity.

Furthermore, in a study investigating the effects of integrated and isolated FFI compared with incidental vocabulary learning of adult intermediate proficiency students, File and Adams (2010) found that both instructional strategies resulted in improvement. Valeo (2013) compared the effects of meaning-focussed, content-based instruction versus FFI, with results indicating that – while both strategies resulted in improvement – FFI seemed to benefit learners more.

Both integrated and isolated teaching strategies can be used in implicit and/or explicit language teaching. Based on the brief discussion presented here, it is evident that research results vary as to which one is more beneficial. However, the varying results may also be attributable to the fact that different types of language elements were targeted. This could suggest that integrated and/or isolated teaching strategies may be more useful to improve some aspects of language than others. Hence, it could be argued that the nature of different aspects of language should be taught differently. For example, grammatical structure and vocabulary might not be equally effective if the same teaching strategy is employed. The present study aims to improve controlled productive collocational knowledge through explicit integrated strategies, since these have been shown to be more beneficial than implicit isolated strategies.

3.4 Vocabulary learning and teaching strategies

As with general language instruction, vocabulary instruction can be divided into implicit or explicit learning and teaching activities (Hulstijn, 2001:266). Incidental (or implicit) vocabulary learning has been defined as “learning without an intent to learn, or as the learning of one thing, for example vocabulary, when the student’s primary objective is to do something else” (Laufer & Hulstijn, 2010:10), whereas intentional/deliberate learning involves actively learning a word in order to make a connection between the form and meaning of a target word (Hulstijn, 2003:360). Various studies (mentioned below) have found that the average retention rate of intentional vocabulary learning is higher than that of incidental learning (Hulstijn, 2003; Schmitt, 2008), and that intentional learning triggered both the representational and functional aspects of lexical knowledge (Elgort, 2011). According to Biemiller and Boote (2006), for example, explicit instruction of vocabulary-building strategies is most effective as it may extend the learning process beyond the classroom.

Researchers have also advocated the need for teaching collocations explicitly as part of vocabulary instruction, and have explored techniques to achieve this (see Barfield & Gyllstad, 2009; Laufer & Girsai, 2008; Nizonkiza, 2012; Nizonkiza *et al,* 2013; Van Dyk *et al.*, 2016). While the value of employing explicit vocabulary teaching strategies is largely uncontested, there are limitations to a purely explicit learning and teaching environment (Schmitt, 2000). Moreover, the question that remains to be answered is *how* best to teach vocabulary, and perhaps more importantly, *which* lexical items to teach. This is especially important when dealing with technical and subject-specific vocabulary, as well as collocations.

As mentioned in Section 3.1, Nation (2001; 2007; 2013) suggests implementing a language course that includes four different strands in order to ensure optimal learning of language features, including vocabulary. These strands are meaning-focussed input, meaning-focussed output, language-focussed learning, and fluency development. Each of the strands utilises different teaching and learning strategies that focus on developing both receptive and productive skills. The purpose of the following section is to outline the objectives of each strand in terms of the teaching and learning of vocabulary and collocation, and to outline strategies that have been successfully employed to teach vocabulary and collocations.

3.4.1 The four strands

Nation (2001; 2007; 2013) suggests including various aspects of language learning and teaching techniques to provide a framework in which learners are provided with "an appropriate balance of opportunities for learning" (Nation, 2007:2), focussing on both receptive and productive skills. Each strand (meaning-focussed input, meaning-focussed output, language-focussed learning, and fluency development) is dependent on every other strand in the framework; none of the strands can function in isolation. Each strand is discussed in turn below.

3.4.1.1 Meaning-focussed input

Learning through meaning-focussed input is mostly incidental, and constitutes input through listening and reading. Studies have shown that incidental acquisition of new vocabulary is possible through reading in both an L1 (Freebody & Anderson, 1983; Nagy *et al.*, 1985) and L2 (Horst, 2005; Zahar *et al.*, 2001). Nagy *et al.* (1985) note that native speakers learn most vocabulary through natural incidental exposure, as they have ample opportunity to do so. This is not the case for ESL/EFL learners. In order for incidental vocabulary learning to take place in ESL/EFL, Nation (2002:38) suggests meeting three requirements:

> First, the unknown vocabulary should make up only a very small proportion of the tokens, preferably around 2 per cent… Second, there needs to be a very large quantity of input, preferably one million tokens or more per year. Third, learning will be increased if there is more deliberate attention to the unknown vocabulary through the occurrence of the same vocabulary in the deliberate learning strand of the course.

Learners should be exposed to aural input that meets these three requirements. Nation (2007:3) notes that "[t]ypical activities in this strand include extensive reading, shared reading, listening to stories, watching TV or films, and being a listener in a conversation". In terms of FFI strategies, input enhancement would also be useful to ensure that learners are exposed to target vocabulary through both extensive and intensive reading activities. Hulstijn (2001: 275) suggests that retention of new information is determined by the quality and frequency of information processing activities such as "elaboration of a word's form and meaning, plus rehearsal".

However, in order for incidental vocabulary learning to be possible, extensive reading is essential (Schmitt, 2000; Nation, 2001) to ensure that learners are repeatedly exposed to unknown words. When exposed to written input, L2 learners must know approximately 98% of running words of a text for them to guess word meanings from context (Hirsh & Nation, 1992; Hu & Nation, 2000; Nation, 2006; Schmitt *et al.*, 2011).

While extensive reading can result in incidental vocabulary learning, this is not always the case. Scheepers (2008), for example, investigated the effect of an extensive reading programme on the vocabulary of grade 7 learners. Results indicated that learners showed a lack of vocabulary knowledge, even at the end of the study, suggesting that reading input alone is not sufficient, and that explicit vocabulary instruction is necessary in order to develop vocabulary knowledge.

There is evidence that collocations can be learned incidentally. Webb *et al.* (2013) investigated the effects of repetition on the learning of collocations on Taiwanese university EFL students. Participants were simultaneously exposed to 18 collocations while reading and listening to graded readers. Participants were exposed to one of four versions of readers containing varying encounters of the collocations (1, 5, 10 and 15). Receptive knowledge was assessed in a four-part test immediately following the reading and listening sessions. The tests assessed knowledge of receptive and productive, as well as receptive and productive meaning of the target collocations. Webb *et al.* (2013) found that exposing participants to collocations 15 times in a graded reader resulted in higher retention of new collocations. However, as the study did

not include a pre-test of participants' productive lexical knowledge, the interpretation of the results is limited. Macis (2008) conducted a similar study to determine whether lexical knowledge of the figurative meanings of duplex collocations (collocations with both a lexical and figurative meaning (Macis & Schmitt (2017)) could be enhanced by reading a semi-authentic novel that contained 38 target items. The study examined the meaning-recall level and how repetition affected lexical knowledge by conducting one-on-one interviews, a pre-test, post-test and delayed post-test. While results showed that each participant's knowledge of the target collocations was enhanced, the results were not always statistically significant.

While collocations (and other multi-word units) fulfil basic communicative and social needs (Wray, 2002:175), as well as communicative functions (Warren, 2005), Henriksen (2013) notes that "some collocations are not salient and therefore not noticed as readily as other units by the L2 learner" (p.41). Ying and O'Neill (2009:185) suggest that L2 learners might not notice collocations in input because they lack awareness of collocations as lexical units. This is also true for learners who are fairly advanced (Henriksen & Stæhr, 2009). This suggestion is supported by Das Neves Seesink (2007).

Based on a case study investigating the effect of teaching academic vocabulary collocations on the writing development of six students in an Intensive English Programme, Das Neves Seesink (2007:161) found that the concept and its importance had to be explained to learners to clarify the concept. Nizonkiza and Van de Poel (2019:15) also suggest that, while formulaic language is believed to be implicit in nature and should be taught implicitly, sufficient exposure to target language in L2 or foreign-language contexts is not possible, and therefore explicit teaching is necessary. A case has therefore been made for explicitly teaching collocations and other dimensions of formulaic language (cf. Lewis, 1993; Barfield, 2009; Barfield and Gyllstad, 2009; Henriksen, 2013; Nizonkiza *et a*l., 2013; Van Dyk *et al*., 2016).

Based on this discussion of meaning-focussed input as an aspect of vocabulary and collocational learning, it can be suggested that only meaning-focussed input could result in learning vocabulary, specifically collocations. However, while studies have shown that incidental learning is *possible*, it is less effective than deliberate learning. The following section considers the effect that meaning-focussed output might have on vocabulary and collocation knowledge.

3.4.1.2 Meaning-focussed output

The objective of this strand is to provide opportunity for learners to produce language in order to "develop their knowledge of language through speaking and writing activities where their main attention is focused on the information they are trying to convey" (Nation, 2001:2). These activities are useful in learning vocabulary as they help learners focus on words they may not have to focus on when listening or reading. Such activities "include talking in conversations, giving a speech or lecture, writing a letter, writing a note to someone, keeping a diary, telling a story and telling someone how to do something" (Nation, 2007:4). Nation (2002:39) also suggests three ways in which vocabulary enhancement can take place in this strand:

First, activities can be designed, such as those involving the use of annotated pictures or definitions, which encourage the use of new vocabulary. Second, speaking activities involving group work can provide opportunities for learners to negotiate meanings of unknown words with each other… Third, because the learning of a particular word is a cumulative process, using a partly known word in speaking or writing can help strengthen and enrich knowledge of the word.

Other productive activities could include tasks where learners provide interactional and corrective feedback to their peers. The kinds of activities used in both meaning-focussed input and output strands depend on the context and type of vocabulary that is learnt. For example, the activities would be different in an academic context where subject-specific vocabulary is used.

Laufer and Nation (1995) suggest that it is necessary to implement instruction that would give learners the opportunity to put their vocabulary knowledge to productive use. According to Lee and Muncie, explicit target vocabulary instruction appears to scaffold learner writing, which is a central skill for success in academic achievement (2006:296). They also attribute an improvement in writing to the teacher's use of interactive elicitation of vocabulary and a writing frame, and specific instruction to learners to use target vocabulary. Zhong and Hirsh (2009) also found that students' productive knowledge can be increased through productive classroom tasks. Based on the finding that the majority of Japanese English Foreign Language (EFL) learners focus on receptive rather than productive vocabulary learning, Al-Murthada (2014) concluded that it is crucial to expose learners to productive vocabulary learning strategies. In a systematic statistical synthesis of the effects of output stimulus tasks on L2

incidental vocabulary learning in 12 studies, Haung (2010) found that using output stimulus tasks to learn vocabulary was more beneficial than a read-only input approach.

However, productive vocabulary knowledge often lags behind receptive knowledge in relation to general vocabulary (Laufer 1998; Webb 2008) and academic vocabulary (Zhou, 2009). There is a need for language-focussed learning – in this instance, with an emphasis on vocabulary learning. The following section explores language-focussed vocabulary learning.

3.4.1.3 Language-focussed learning/Deliberate vocabulary learning

According to Nation, language-focussed learning "involves the deliberate learning of language features such as pronunciation, spelling, vocabulary, grammar and discourse" (2007:6; also see Spada, 1997, and Norris & Ortega, 2000). This strand focusses mainly on improving receptive skills to result in better output. Nation (2001, 2002) identifies four main strategies that could result in learning vocabulary, and by extension, collocations, viz. guessing from context, deliberately studying word cards, using word parts, and dictionary use. These and other strategies have been used in experimental studies (see Fordyce, 2014; Loewen, 2015; Nassaji, 2009; Nassaji & Fotos, 2010) to teach vocabulary explicitly.

In one example, Almari and Rogers (2018) compared the effects of written vs visual instruction and timing of word-recognition activities on the retention of academic vocabulary in seventy Arabic learners of English. Their results showed that explicit teaching is more beneficial than implicit-only conditions, and that using visual aids in that condition, while pre-teaching target vocabulary, is more effective than using written input and post-teaching vocabulary items. Szudarski (2012) compared the effect of implicit and explicit strategies on the acquisition of verb-noun collocations through reading-only vs. reading followed by treatment. Participants were EFL learners and were assigned to one of the groups. Participants in the reading-only group were exposed to target collocations by reading texts in which each new collocation was repeated 6 times. The reading-plus-treatment group completed the reading exercises, but were also asked to complete vocabulary activities that focussed explicitly on collocational patterns. Sessions occurred once per week for three weeks, and lasted 45 minutes. The collocational knowledge of the reading-only treatment group improved significantly in terms of verb form recall and verb form recognition, but not in overall collocation form recall.

Some of the strategies mentioned in this section can be applied to teaching formulaic language. For example, learners could be exposed to meaning-focussed input by means of various

reading activities, as well as tasks that require meaning-focussed output. However, a general approach to teaching vocabulary such as those discussed in 3.4.1.2 and 3.4.1.3 are not necessarily the best approaches to teaching specific vocabulary such as collocations. Granger and Meunier (2008) and Nizonkiza and Van de Poel (2014), among others, state that the main questions in teaching collocations is not whether to teach them explicitly, but *how* to teach them effectively and, as mentioned earlier, *which items* to teach. Collocation teaching strategies are discussed in more detail in section 3.4.1.5.

3.4.1.4 Developing fluency

The final strand of a language-teaching course is aimed at developing fluency. This strand should include reading, listening, speaking, and writing, and should aim to help learners "make the best use of what they already know" (Nation, 2007:7). In other words, learners are exposed to language and content that they are largely familiar with. Nation (2002:43) notes that the "conditions needed for fluency development involve a large quantity of familiar material, focus on the message and some pressure to perform at a higher than normal level".

Fluency activities can be divided into two main approaches: the first approach relies mainly on repetition of reception or production of the same material, and the second approach relies on activities that require learners to make connections and associations with words that they know.

Studies of fluency development in both first (Kuhn & Stahl, 2003) and second language readers (Chung & Nation, 2006) have shown to be beneficial. Kuhn and Stahl (2003:17) found that activities focussed on improving fluency has a positive effect, and that these assisted activities were more effective than unassisted activities. For second language readers, Chung and Nation (2006) found that timed activities were beneficial in assisting learners to improve fluency. Fluency development activities typically follow the 4/3/2 technique, where activities are repeated in decreasing time frames (four minutes, then three minutes, then two) and studies have shown that fluency as well grammatical accuracy and complexity increased (see Arevart & Nation, 1991; Nation 1989).

Despite evidence that fluency development is beneficial, Nation (2007) notes that fluency development is often neglected in language courses "possibly because it does not involve the learning of new language items and thus is not seen as moving the learners forward in their knowledge of the language" (p.8).

3.4.1.5 Considerations for collocation teaching strategies

As mentioned in section 3.4.1, it is crucial that a balanced approach is taken when implementing the four strands to enable learners to notice vocabulary items, retrieve the different levels of word knowledge in input and output tasks, and finally, to generate meaningful communication. Nation (2001:401) provides a table to summarise the four strands of a language course, as well as the conditions needed for learning and examples of activities:

Strand	Conditions for learning	Examples of activities
Meaning-focussed input	1. Focus on the message 2. Include a small number of unfamiliar items 3. Draw attention to the new items	1. Extensive graded reading 2. Listening to stories 3. Working with familiar content
Language-focussed learning	1. Focus on language features 2. Do deliberate repeated retrieval of items	1. Learning from word cards 2. Grammar exercises 3. Read difficult text
Meaning-focussed output	1. Focus on the message 2. Include a small number of new items	1. Communication activities 2. Research and write
Fluency development	1. Focus on the message 2. Work with completely familiar material 3. Work at a higher than normal speed 4. Do a large quantity of language use	1. Repeated reading 2. Repeated speaking on familiar topics 3. Graded reading

Table 3-2: The four strands of a language course (Nation, 2001:401)

As indicated in section 3.4.1.3, collocations can be taught within the same framework, but results are not always significant. Van Dyk *et al*. (2016) suggest adopting a productive-oriented approach to teaching collocations. This may be done following one, or a combination of two main approaches to teaching collocations: the awareness-raising approach and the attention-drawing approach (Nizonkiza, 2012). Ying and O'Neill (2009:184) suggest teaching collocations through the AWARE-approach. They define it as

> a process-oriented learning approach, [that] is based on two central beliefs: One is that learning will be more effective if learners are made aware of language and language learning at three levels: noticing the particular language features that they need to learn, developing an awareness of learning strategies, and a metacognitive awareness of reflecting on their learning process and content. The other central belief is that effective learning is more likely to take place when learners

see the significance of learning what they learn and are given opportunities to exhibit what they have learned.

The attention-drawing technique (Peters, 2009) is based on similar theoretical idea. However, this technique differs from the AWARE approach in that it only draws attention to lexical items or collocations through highlighting the items in a text. The key in both approaches is to make learners aware of collocations that they do not know, and to actively engage them in learning more about those items. This can be achieved by combining techniques suggested by Nation (2001; 2002) and activities aimed at awareness raising. Nizonkiza (2017), for example, achieved positive results when using the awareness-raising technique and an adapted version of McCarthy and O'Dell's (2005) collocation web model as intervention techniques to teach verb-noun collocations to students majoring in English at a university in Burundi, ultimately, improving collocational knowledge results in developing fluency. However, Nizonkiza and Van de Poel (2019:17) state that, while "the role of collocations in L2 teaching and learning contexts has been widely acknowledged" and "the calls to explicitly teach collocations are increasing in number", a gap that still needs to be bridged is *which* collocations to teach.

The AWARE-approach draws from psycholinguistic principles. The focus is on three main levels of language awareness (Ying & O'Neill, 2009:183):

1. *Noticing*, which Schmidt (1992) considers the initial level of awareness of specific linguistic features of language input. Schmidt and Frota (1986) claim that learners who "notice most learn most" (p.317). This is critical for the next stage.
2. *Cognitive awareness* – on this level, cognitive strategies are employed to allow for deep processing and internalising of noticed language features (Ying & O'Neill, 2009:183). Following this stage, learners may reach the final stage of language awareness:
3. *Metacognitive awareness* – Little (1997) defined this higher level of awareness as an awareness that learners reach when they "develop a psychological relation to their learning content and process" (p.94). Hence, the learner has more control over language learning and use.

According to Ying and O'Neill (2009:184), the learner is central in each of these dimensions as they notice the importance of the linguistic features they are learning, are actively involved in the learning process, and can make use of cognitive strategies and develop metacognitive awareness of both the content and learning process to make their learning more effective. Similar results were obtained in Barfield (2009).

Another popular approach to teaching collocations is psycholinguistic in nature and suggests teaching collocations as intact units rather individual items. Proponents of this approach argue that presenting learners with collocations as individual items poses a challenge at the time of assembling them for productive use in speaking or in writing (Boers *et al.*, 2014a; Boers *et al.*, 2017; Boers & Lindstromberg, 2008; 2009; Boers *et al.*, 2014b; Boers *et al.*, 2012; Wray, 2002).

Boers *et al.* (2014b) investigate the reasons for slow L2 collocation acquisition. They identify lack of attention and lack of exposure in a short time as two main factors that hamper collocation acquisition in L2 learners (*ibid.*, p.45-49). Another factor that hinders acquisition is failing to retain collocations once they have been learned (*ibid.*, p.50). Their study comprised of 36 undergraduate language major taking English as a foreign language. Participants were exposed to a set of 22 collocations through dictation twice, and asked to write down the collocations. Results showed that participants experienced difficulty retaining the collocations that they were exposed to incidentally. Boers et al. (2014b: 56) note that there are many impediments that contribute to learning collocations. These include learning new word strings and noticing collocations. Furthermore, “when learners do notice collocations, they still face the task of retaining them in memory” (*Ibid.* p.56).

3.4.1.6 Which collocations should we teach?

While it is apparent that collocations and other formulaic sequences are crucial to L2 proficiency and success, what is not apparent are which items to teach and test. Even though structures such as verb-noun collocations have been identified as particularly problematic (Nesselhauf, 2003), it is still difficult to determine which verb-noun collocations to teach.

In an attempt to overcome the problem of finding the ‘right’ words to teach, researchers and teachers have for decades been attempting to establish which vocabulary items to teach. As a result, several word lists have been compiled based on the vocabulary items identified as important, especially for L2 learners. These lists include the General Service List (West, 1953), the University Word List (Xue and Nation, 1984), Coxhead’s Academic Word List (1998) and the New Academic Word List (AWL) (Coxhead, 2000). These lists have been utilised by learners and teachers, and they have been used to inform pedagogy. The AWL (Coxhead, 2000) is one of the most widely used lists in English for Academic Purposes (EAP) (Akcermann & Chen, 2013:235).

Various methods have been used to identify items on each list, and, for the first few lists at least, word frequency was the main criterion. Items were taken from relevant sources such as academic journals and textbooks to ensure that the items were suitable for an academic register. However, these lists only focus on single-word items; lists containing multi-word units were considered only as recently as 2009, with Durrant (2009) considering an academic collocation list, and Simpson-Vlach and Ellis (2010) compiling an Academic Formulas List (AFL). Martinez and Schmitt (2012) compiled the PHRASE list in an attempt to narrow the scope of previous lists, such as the AFL (Simpson-Vlach & Ellis, 2010). The Academic Collocation List (ACL) (Ackermann & Chen) was finally developed in 2013.

Whereas traditional approaches to collocation research relied on expert intuition to identify collocations (Ackermann & Chen, 2013; Boers 2020; Martinez & Schmitt, 2012), more recent approaches combine both phraseological and frequency-based approaches (as discussed in Chapter 2). The nature of multi-word units and collocations means that items could not be selected merely on how frequently one or both (in the case of collocations) occurred. The frequency of recurrence of both words as a word combination had to be considered. This means that computational and statistical analyses are necessary to provide an accurate depiction of which multi-word units are used in academic writing.

Simpson-Vlach and Ellis (2010) compiled a list of the most frequently used academic formulas, using "target corpora of academic discourse [that] included 2.1 million words each of academic speech and academic writing" (2010:491). Texts from various disciplines were included in the corpus. Upon identifying the most frequent n-grams, a Mutual Information (MI) test narrowed the results down to only significant combinations of words. MI is a statistical measure "designed to assess the degree to which words in a phrase occur together more frequently than would be expected by chance" (Simpson-Vlach, 2010:493). MI was used to eliminate phrases that were not useful. For example, the most frequent n-gram identified by frequency only was *on the other*, while with MI it was *due to the fact that*. The n-gram *on the other* cannot be used on its own, even if it was a frequent n-gram in the corpus. Similarly, Martinez and Schmitt (2012) used selection criteria that "revolved around high frequency, meaningfulness, and relative non-compositionality" (p. 304). In other words, the list consist of phrases that are frequently encountered by L2 learners, are meaningful, and whose meanings are not fixed.

Ackermann and Chen (2013) relied on statistical measures (MI score and *t*-score) and expert review to compile the ACL. Items were identified using the written component of the Pearson International Corpus of Academic English (PICAE). Ackermann and Chen (2013:237) define

collocation as “a single word that tends to co-occur in the span of ± 3 words from the reference word, co-occurring at least five times in total across at least five different texts with a Mutual Information score of at least 3 and a t-score of at least 2”. The initial list was refined by identifying items that met the criteria. The first list contained more than 130 000 entries, which were reduced to 2 468. The review and refinement process included eliminating combinations that were not lexical collocations through part-of-speech tagging (POS-tagging). Items were manually vetted by the researchers to exclude entries that were linguistically incomplete, combinations with a high degree of fixedness, adverbs referring to time of frequency, common transparent adjectives, concrete geographical references, and combinations that are often hyphenated.

The next step in refining the list was to have it reviewed by experts to determine which items had to be included and excluded. Two main considerations were 1) Whether the entry could be regarded as an appropriate collocation for teaching and/or learning purposes, and 2) Whether the collocation was pedagogically relevant.

The next step was systemisation, where the following changes were made to each collocation:

1. Listing collocations in their base form;
2. Harmonising entries that appear in British and American English, with British English as the preferred form;
3. Adding definite and indefinite articles to verb + noun collocations in line with dictionary conventions;
4. Adding optional copula *be* to adverb + verb past participle combinations if they could be used as a predicate;
5. Adding dominant prepositions to collocations.

The ACL is a useful list when teaching collocations in academic writing and has been used to identify some of the most important collocations. It is also the list that was used as a starting point for the present study because it is currently the only list of its kind. However, the researcher suggests that there are several points of critique that could be raised with regard to the process followed as well as the final list. The process followed to compile the list was not sufficient for the purposes of this study, as the collocates suggested for each word are too limited. The benchmark for optimal span is about four items to the left and right positions of the node (Sinclair, 1991). The author argues that there are several structures in language that could result in the node and its collocate being moved further apart. For example, the passive voice in English allows for the noun to be moved further away from the verb, and it will also

change the position of the noun in relation to the verb, i.e. the noun will move from the right to the left of the verb. If a sentence contains an embedded clause that separates the node and the collocate, the span will increase as well. For example, a collocation such as *make a decision* could occur in a sentence like *This was a decision that he struggled to make*. In this case, the node and collocate are more than four positions apart. Therefore, it could be argued that a span of three would not be sufficient to identify all possible collocations in a corpus, as it is below the benchmark. Despite having considered possible shortfalls in the compilation of the ACL, it is currently the most complete list available for the most frequently occurring academic collocations. Therefore, for the purpose of the present study, the ACL was used as a starting point; it identifies the most frequent verbs and (most) of their noun collocates.

In their study, Nizonkiza and Van de Poel (2019) aimed to identify the most frequent collocations as a possible starting point. Their results showed that, of the eight categories of lexical collocations (as discussed in Chapter 2, section 2.4.2), adjective-noun and verb-noun collocations are the most frequent. They do, however, recommend a thorough assessment of all types of collocations (Nizonkiza & Van de Poel, 2019:23). Several studies have been conducted in order to improve receptive and/or productive knowledge of collocations, including evaluating the effects of spacing techniques (Farvardin, 2019), video captioning (Teng, 2019) and presenting and organising collocations (Akpinar and Bardakçı, 2016). However, similar studies have not been conducted in a South African context. The present study therefore aims to contribute to research by determining the effects of awareness-raising teaching techniques on the controlled, productive use of verb-noun collocations of first-year university students on the Potchefstroom Campus of the North-West University in South Africa.

This section has provided an outline of teaching strategies that can be employed to teach vocabulary – both implicitly and explicitly, integrated or in isolation, or inductively and deductively. Four major components of language teaching were also explored: meaning-focussed input, meaning-focussed output, deliberate vocabulary learning, and fluency development. Each of these components plays an integral role in improving vocabulary knowledge. Based on the discussion presented, it is evident that each of the components is essential for improving vocabulary knowledge. Learners need to be exposed to meaningful input in order to produce meaningful output. Meaningful input can be supplemented with deliberate learning and teaching of target items, while meaningful output forms part of the activities necessary to develop fluency. One aspect that remains to be explored in this chapter is how to test vocabulary and collocational knowledge. The following section provides a detailed discussion of vocabulary assessment.

3.5 Assessing Vocabulary Knowledge

Over the last 30 years, the assessment of vocabulary knowledge has received much attention in SLA research (Webb & Sasao, 2013), and various tests have been designed to assess this knowledge (Sasao & Webb, 2017). The formats of these tests range from gap-fill exercises, multiple-choice tests, and tests that require test-takers to match words with their correct meaning (Read, 2000). Multiple-choice tests and matching tests have been especially useful because they are easy to write and score, and they are time effective (Read, 2000). Tests that are well formulated can be used to accurately measure learner's levels of reading comprehension, vocabulary size, progress over a period of time, and to point out weaknesses in a learner's knowledge (Read, 2000:3). The purpose of this section is to provide an overview of vocabulary assessment, considering the following factors: a) the need for assessing vocabulary knowledge, b) fundamental issues with vocabulary assessment, c) measures used to assess vocabulary, and d) the validity of the measures used. These factors will be discussed in the sections that follow to identify the most suitable collocation assessment practices in order to inform the methodology of the present study.

3.5.1 The need for assessing vocabulary

As outlined in 2.1, vocabulary knowledge is essential for effective communication, and there are many aspects to knowing a word. As researchers' understanding of lexical knowledge and competence expanded, tools used to test this competence developed. Measuring vocabulary size has been a longstanding area of research, and is useful for both native speakers and second language learners. Read (2007:107) states that "since vocabulary size is closely associated with reading comprehension ability, vocabulary tests have traditionally had a significant role in research on reading development and in literacy programmes."

Read (2000:3-4) distinguishes two major approaches to assessing vocabulary knowledge: the discrete-point approach, and the communicative approach. The discrete-point approach "involves designing tests to assess whether learners have knowledge of particular structural elements of language". Discrete-point tests assess learners' knowledge of particular word forms, structural patterns, phonological differences between words, etc. However, this approach has been criticised, and the validity of tests that fall within this stream, has been questioned. Read (2000:3-4) identifies five major criticisms against this approach: 1) Scores of these tests cannot be used to make a general statement about a learner's vocabulary; 2) Since these tests focus on single structures and meanings of single words, they cannot be used to

accurately determine a learner's overall proficiency; 3) Standard discrete-point tests assess receptive rather than productive knowledge; 4) The context in which words are used is not taken into consideration because they are tested in isolation; and 5) Learners can compensate for lack of lexical knowledge in communication situations by guessing meaning from context.

Along with developments in the understanding of language competence and communicative language teaching (CLT), the communicative approach to language testing became dominant (Read, 2007:106). A learner's ability to successfully complete a particular task became a means of assessing various aspects of language competence, including vocabulary knowledge (see Bachman and Palmer, 1996; McNamara, 1996). Bachman and Palmer (1996:68) identified two main components of testing language ability. The first component is language knowledge, which includes knowledge of vocabulary, grammar, pronunciation, and spelling. The second component, strategic knowledge, entails the ability to use that knowledge effectively and accurately to communicate. This framework addresses some of the main criticisms against the validity of discrete-point tests, as it considered both organisational and pragmatic knowledge.

Organisational knowledge includes grammatical knowledge and textual knowledge, i.e. how utterances and texts are organised. Pragmatic knowledge includes functional knowledge and sociolinguistic knowledge, i.e. how organisational knowledge is related to communicative goals within a particular context. However, even though Bachman and Palmer's (1996:68) framework drew attention to communicative ability, vocabulary knowledge is categorised as an aspect of grammatical knowledge, and not tested separately. Vocabulary knowledge in and of itself is essential to overall language proficiency, and much research has been dedicated to defining the multifaceted and complex nature of vocabulary knowledge (Richards, 1976; Nation, 1990, 2001; Meara, 1996; Henriksen, 1999).

Chapter 2 further showed that lexical knowledge is dimensional (Meara, 1996; Henriksen, 1999), comprising size, depth, and receptive-productive knowledge. It is thus necessary to treat vocabulary as a separate aspect of language knowledge that contributes significantly to strategic knowledge. As such, measuring vocabulary knowledge specifically has become common practice and essential in determining language proficiency, with tests designed to assess size and/or depth (cf. Daller *et al.*, 2007; Meara, 2002; Nation, 1983; Nation, 1990; Nation & Beglar, 2007; Nation & Laufer 1995, 1999; Read, 1993; 1998; 2000; Zareva *et al.*, 2005, among others). Examples of tests that have contributed greatly to research and development of vocabulary assessment include the Eurocentres Vocabulary Size Test (EVST)

(Meara & Jones, 1988) the Vocabulary Levels Test (VLT) (Nation, 1990; Schmitt *et al.*, 2001), Read's Word Associates Test (WAT) (1993; 1998), Paribakht and Wesche's Vocabulary Knowledge Scale (VKS) (1997), and the Academic Word Levels Test (AWLT) (Webb & Sasao, 2013). However, certain aspects of vocabulary assessment remain problematic. The following section deals with fundamental issues in vocabulary assessment, and considers factors such as test design and validity, whether the test is designed to assess size or depth, and receptive or productive knowledge. The Vocabulary Levels Test (VLT), Word Associates Test (WAT), Vocabulary Knowledge Scale (VKS), and the Academic Word Levels Test (AWLT) are considered. These tests have been most widely used, are well established, and have been designed to measure what students know about words (Webb & Sasao, 2013:263).

3.5.2 Fundamental issues in vocabulary assessment

One of the major factors that contribute to the complexity of vocabulary assessment is determining whether to test size or depth, and whether the focus is receptive or productive. This decision depends mainly on the reason for testing, which again influences the test design. In this section, tests to assess size and depth are discussed separately to provide a clear outline of test design, validity, and to provide examples of studies that have incorporated each of the tests. These tests are discussed separately because they are used to assess different aspects of vocabulary knowledge. Vocabulary size tests count the number of words a learner knows, while vocabulary depth tests aim to determine how well a learner knows a word. In other words – can the learner use a word accurately on different levels?

3.5.2.1 Measuring vocabulary size

A vocabulary size test typically requires a large sample of words from word frequency lists, and test-takers are required to indicate whether or not they know the meaning of the target words. One such test is the VLT (Nation, 1983, 1990).

The VLT was designed for classroom use to enable teachers "to develop a suitable vocabulary teacher and learning programme for their students" (Read, 2000:118). Test items include word-definition matching, where test-takers are expected to match words to their definitions. The test consists of five parts, each measuring knowledge of vocabulary within a frequency band: 2000, 2001-3000, 4001-5000, and 9001-10 000 word families, and words from the University Word List. Schmitt *et al.* (2001) made useful improvements to the original version of the test, such as increasing reliability by increasing the number of items for each level from 18 to 30.

Coxhead's Academic Word List (2000) was used to evaluate academic vocabulary knowledge. Test-takers were not required to complete the entire test, because the scores are assigned for each individual level of the test and, as a result, accurate assumptions can be made about a learner's vocabulary size. Webb and Sasao (2013) note that the highest level of frequency not mastered would give an indication of where vocabulary learning should be focussed.

Read (1988) aimed to determine the validity of the VLT by analysing the results produced by 81 students (ranging from low-intermediate to advanced) taking a three-month intensive English course for academic purposes. The test was administered shortly after the beginning of the course, and test results showed a consistent pattern of declining scores from the highest to the lowest frequency level. Similarly, Beglar and Hunt (1999) produced a revised version of the VLT at the 2000-word and University word levels to assess Japanese learners' English vocabulary knowledge. Results confirmed that the test is a valid measure to estimate vocabulary size.

While the VLT has proven to be a valuable tool to accurately measure vocabulary size (Read, 2000), there are certain limitations. Webb and Sasao (2013:265) suggest that there are two major limitations: the test does not include a 1000-word level, and the lists used for the frequency levels are outdated. They motivate the first limitation by emphasising that the "relative value of vocabulary is a function of frequency; words that are encountered and used more often have greater value to learners because they will have a greater impact on comprehension and production" (*Ibid*). Secondly, they emphasise that the General Service List (West, 1953) was used to create the first 2000-word level, while frequency criteria from Thorndike and Lorge (1944) and Kučera and Francis (1967) were used to create the lists for the 3000, 5000, and 10 000-word levels.

In order to address these limitations, the New Vocabulary Levels Test (NVLT) was developed. The NVLT includes only 1000, 2000, 3000, 4000, and 5000-word levels, since words from each of these levels have been proven most useful, especially for second language learners. The word lists were created by Nation (2012) and items on the lists were derived using two sets of criteria. The 1000 and 2000-word lists were compiled using a greater proportion of spoken than written text to ensure that the high frequency words were suitable for courses designed for teaching English as a foreign language. The 3000 to 5000-word lists were compiled using word frequency rankings from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). Further, the NVLT does not include a measurement of academic vocabulary for two reasons: first, Schmitt *et al.*'s (2001) version of the VLT is a valid measure

of the Academic Word List and, second, a separate Academic Word Levels Test (AWLT) has been developed (Webb & Sasao, 2013). The AWLT aims to provide a more valid measure of knowledge of academic vocabulary, drawing from Coxhead's (2000) Academic Word List (AWL). The AWL consists of 10 sublists. The AWLT consists of only five levels, simultaneously measuring two sublists. This ensures a manageable learning goal of 120 words for a course and it "provides a practical tool that can be used to measure progress in learning the AWL" (Webb & Sasao, 2013:268). These tests could be useful to measure receptive collocational knowledge, but they are not necessarily an option for assessing productive collocational knowledge.

While tests that consist of word-definition matching mainly measure receptive knowledge, Laufer and Nation (1995; 1999) aimed to measure productive vocabulary knowledge by developing the Lexical Frequency Profile (LFP). The 1995 study saw the initial application of the LFP to measure learner's ability to use words in their own speech and writing, in order to calculate the relative percentages of high and low-frequency vocabulary, as well as academic vocabulary. Laufer and Nation (1999) used a sentence-based gap-filling test consisting of the target items on the VLT to measure relative size of receptive and productive vocabulary. The format of this test allows for the testing of collocations as well as general vocabulary knowledge, because it can be utilised to test deep-word knowledge and vocabulary size. In terms of collocations, the test can be adjusted in such a way that nodes and their collocates can be assessed by requiring participants to fill in either of the two components. For example, to assess knowledge of verb-noun collocations, the following format could be used:

Q: The engineer had to co___ research before she could start the project.
A: The engineer had to conduct research before she could start the project.

In this example, a clue as to what the node (verb) is, is provided, while the noun collocate is used to provide additional information, and in itself is also a clue.

A major consideration with the tests discussed above is that they are mainly used to measure vocabulary size. In other words, they accurately measure whether or not a learner knows a word, but not the quality of that knowledge. This poses problems when aiming to assess productive vocabulary proficiency, which is essential for academic success. Alternative testing methods are necessary in order to do so. The next section details ways of testing vocabulary depth.

3.5.2.2 Measuring vocabulary depth

Read (2007:113) states that “compared to [the] robust initiatives to develop and apply various tests of vocabulary size, there has been rather less progress in measuring quality (or depth) of vocabulary knowledge”. There seems to be consensus among researchers that more attention should be paid to this aspect of vocabulary assessment. Nizonkiza and Van den Berg (2014), for example, argue that focussing on size alone “means that vocabulary knowledge is being investigated and measured in terms of only one part of a whole” and that this emphasis “poses problems for validity in terms of construct-underrepresentation” (p.46).

In an early attempt to address this long-standing issue, Read (1983; 1998) developed the Word Associates Test (WAT). The WAT was an important innovation in vocabulary assessment, as it provided a means to measure knowledge of multiple aspects of vocabulary knowledge to determine how well words are known, instead of only requiring test-takers to know form-meaning connections of words. Test-takers are presented with cue words and eight words to choose from to identify which words are related to the cue word. Four of the eight words are correct associations, and learners are expected to select the correct items. The WAT also measures meaning-knowledge based on synonyms (Gyllstad & Schmitt, 2019), and has been applied widely in second language research. For example, Qian (1999) compared results from the VLT and the WAT to determine size and depth respectively, and found that the WAT accounted for a significant amount of the variance in reading comprehension scores. Bogaards (2000) reviewed work on the measurement of depth and reported on a study investigating the validity of using a simplified version of the WAT (Meara’s Euralex French Tests).

While research on vocabulary depth has proven valuable, much is lacking in terms of measuring (productive) knowledge of more complex form-meaning connections such as multi-word units and collocations (Webb & Sasao, 2013:270). Gyllstad and Schmitt (2019) note that “there is still no consensus on the best ways to measure FL [formulaic language] and no test which has been recognized as a standard measurement” (p. 175). According to Gyllstad and Schmitt (2018), factors that contribute to the problems associated with assessing formulaic language are 1) “formulaic language is made up of numerous disparate categories, which all have their own particular characteristics” (p. 175), and 2) there are a large number of formulaic sequences. They also state that “[t]hese two factors work against both the identification of the target population and the representative sampling of items from that population” (Gyllstad & Schmitt, 2018:175). The two sections that follow discuss 1) studies aimed at identifying

collocations to teach (and, by extension, assess), and 2) measures that have been employed to assess collocational competence.

3.5.3 Measuring productive collocational competence

There have been attempts to measure collocational knowledge and determine the rate of acquisition of collocations in ESL/EFL learners, but to date no standardised test has been developed. While various research methods have been used to investigate L2 learners' collocational knowledge, use and development, some challenges remain.

In terms of measuring collocational competence, Biskup (1992), for example, found that EFL learners with an L1 (Polish) that is more distant from English made fewer collocational errors than learners whose L1 (German) was closer to English. Nesselhauf (2003) confirmed these findings in a study that examined verb-noun collocation errors made by advanced learners – L1 interference seemed to account for a large proportion of errors (p. 235). Laufer and Girsai (2008) found that explicitly contrasting L1 and L2 collocations in translation activities seemed to minimise L1 interference on learner's use of collocations. Siyanova and Schmitt (2008) have identified three general elicitation tools utilised to collect data from learner production: 1) online written and oral tasks; 2) offline elicitation tools such as productive translation tasks (Biskup, 1992; Webb & Kagimoto, 2011), cloze tests (Durrant, 2008; Revier, 2009; Prentice, 2010) and association tasks, and receptive multiple-choice and judgement tasks (Gyllstad, 2007; Barfield, 2009; Fitzpatrick, 2012); and 3) online reaction tasks (Columbus, 2010; Millar, 2011; Wolter & Gyllstad, 2011). Webb and Kagomoto (2011) found that collocations that share the same collocate seem to be easiest to learn, because fewer word associations need to be memorised.

Webb and Sasao (2013) suggest that the AWLT could be adapted to measure collocational knowledge and compare scores to more traditional tests of form-meaning connection. The WAT (Read, 1993; 1998) has also been used to test collocational knowledge. However, Gyllstad, Vilkaitė and Schmitt (2015) suggest that the multiple-choice format inevitably leads to problems, as test-takers could guess the answers without knowing the correct combinations. Furthermore, the test can accommodate only a limited number of target items.

Three tests that are quite similar in design are the COLLEX and COLLMATCH (Gyllstad, 2009), and the Discriminating Collocations Test (DISCO) (Eyckmans, 2009), all of which use a recognition format to test receptive collocational knowledge. The COLLEX and COLLMATCH tests "were designed to measure advanced Swedish learners' (upper secondary school and

university) receptive recognition knowledge of English verb+noun word combinations" (Gyllstad & Schmitt, 2019). The tests were developed and administered to create assessments that are reliable and produce valid scores (Gyllstad, 2007). The COLLEX test uses a three-option, multiple-choice format that contains 50 English word combinations in a decontextualised format, and test-takers have to select the item that is frequent and natural. The COLLMATCH test is a yes-no test that specifically assesses collocational knowledge. Test-takers must judge whether or not a word combination is frequent and natural in English. COLLMATCH test items are also decontextualised. Gyllstad (2007) administered both these tests along with the VLT to 307 participants, and found that both collocation tests produced very similar scores, and the scores also correlated with the VLT scores.

Eyckmans (2009) administered the DISCO test, which required test-takers to identify two-word combinations that are English idiomatic expressions. The test was based on a receptive vocabulary size test. This was a longitudinal study, which measured English proficiency and phrasal knowledge by administering pre and post-tests in an instructional setting amounting to 60 hours of input-driven instruction. Results indicated that both oral proficiency and phrasal knowledge improved, suggesting that "the cognitive processes involved in taking the test – distinguishing salient Verb + Noun combinations from non-salient ones – relate to the learner's set of stored mental representations that are ready for productive use" (p.150).

Revier's (2009) CONTRIX differs from COLLMATCH, COLLEX and DISCO, in that Revier (2009) claims that it assesses aspects of productive knowledge of verb-noun collocations, even though the test employs a receptive format. Test-takers are required to select items to fill three gaps in a sentence. The combinations that need to be selected in each case comprise a verb, an article, and a noun. Revier (2009) argues that the test assesses productive knowledge because "test takers must not only create (i.e., produce) meaning by combining lexical constituents, but they must also grammatically encode the noun constituent for determination" (p. 129). However, Gyllstad and Schmitt (2019) argue that "the initial pilot only investigated differences in scores between learners of different proficiency levels, and differences in scores between transparent, semi-transparent, and non-transparent collocations" and, therefore, "there was no evidence that the test provides an indication of productive collocation mastery as an independent construct" (p. 179).

Finally, Martinez's (2011) PHRASE test uses four-option multiple-choice format, containing the target item, and a short, non-defining sentence to provide context. Items are taken from a PHRASE List (Martinez & Schmitt, 2012), and test-takers should complete the gaps in a

sentence by using the options provided. Gyllstad and Schmitt (2019) argue that, since items are taken from a finite list, the "percentage correct on the test can be interpreted as the percentage known on the whole PHRASE List" (p. 181). This implies that scores can be interpreted in relation to overall size, unlike scores achieved for COLLEX, COLMATCH or DISCO. Thus, "future tests of FL [formulaic language] may need to focus on much more constrained, and thus identifiable, subsets of FL, in order to make the resulting scores more meaningful" (Gyllstad & Schmitt, 2019:181).

Henriksen (2013:44) notes that, while these studies have been valuable, comparing the results is problematic for various reasons. There are multiple study aims and approaches, and different aims mean that different research instruments are utilised (*Ibid*.) In addition, Given the fact that collocations are defined differently (as discussed in Chapter 2), and developing a standardised tool to measure collocations would be subject to consensus among researchers. Therefore, while measurement tools such as COLLEX, COLLMATCH, DISCO and PHRASE are available, the present study made use of more conservative tools to measure the development of controlled productive knowledge of verb-noun collocations of first-year students taking an English for Specific Purposes course at North-West University. An adapted version of the Vocabulary Levels Test was used, with a multiple-choice cloze test format and gap-fill items with the first one or two letters provided. Participants were presented with sentences that contain verb-noun combinations. The verb or noun is missing in each sentence, and test-takers had to fill in the correct word based on the context of the sentence.

3.7 Conclusion

This chapter established the need for explicitly teaching collocations as part of vocabulary in L2 learners. Various vocabulary-teaching strategies were discussed from the four strands suggested by Nation (2001; 2007; 2013).

However, studies indicated that, while strategies to teach general vocabulary might prove valuable when teaching multi-word units, the nature of formulaic language lends itself to being taught differently. The AWARE approach was identified as the most effective way to explicitly teach collocations.

This chapter also investigated various vocabulary-knowledge assessment tools to identify the most appropriate tool with which to measure productive collocational knowledge. As mentioned, research on vocabulary depth has proven valuable, but much is lacking in terms of

measuring (productive) knowledge of more complex form-meaning connections. Furthermore, there is no consensus on how to measure productive use of formulaic language. While tests such as COLLEX, COLLMATCH, DISCO, and PHRASE have shown to be useful in testing collocational knowledge, these tests have not yet been standardised or fully validated, and further research is necessary.

The present study therefore used a productive version of the VLT to measure development of collocational knowledge in first-year university students receiving integrated explicit treatment. Several studies have utilised some version of the VLT in an attempt to measure productive collocational knowledge. One such study was conducted by Nizonkiza (2017). Nizonkiza (2017) aimed to improve academic vocabulary of English majors at the University of Burundi taking a course built on a collocation-based syllabus, in an attempt to improve academic writing abilities. The awareness-raising approach (Ying & O'Neill, 2009) as well as an adapted version of McCarthy and O'Dell's (2005) Collocation Web Model were used to teach collocations. Collocation knowledge was measured using Nizonkiza's (2014) collocation test, which is based on Laufer and Nation's (1999) VLT. Participants completed pre and post-tests, and open writing tasks. Results indicated that the scores obtained for the post-test were significantly higher, and that overall academic writing skills and productive knowledge of collocations improved. The present study utilised strategies similar to those of Nizonkiza (2017) to improve controlled productive use of verb-noun collocations. The AWARE approach was utilised in various ways to aid in teaching collocations. The main focus of the teaching strategies were to enable participants of the study to notice verb-noun collocations, engage with content that enhanced learning processes, and allow learners to actively engage with the content in order to create a deeper level of awareness of the significance of the content. The researcher also used the format of the VLT (Laufer & Nation, 1999), because it allows for measurement of productive collocational knowledge, as suggested in section 3.5.2.1.

The next chapter, Chapter 4, provides details on the methodology of the present study.

CHAPTER 4: METHOD OF RESEARCH

4.1 Introduction

This chapter provides a detailed description of the research method followed to determine the extent to which first-year university students' controlled productive knowledge of verb-noun collocations could be improved by means of explicit teaching. An overview of the empirical process is provided here. The empirical process was completed in three phases to address the three research questions. Each of phases is discussed in detail below.

The aim of this study was to determine the extent to which first-year university students' controlled productive knowledge of verb-noun collocations could be improved by means of explicit teaching. To determine this, the following sub-questions were posed:

1. What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?
2. To what extent does intervention with explicit teaching of verb-noun collocations result in improving their mastery among first-year ESL students at the North-West University? This question is answered by determining whether
    a. there is any statistically significant difference between the results of the pre-test, post-test and delayed post-test of the experimental group.

This study followed a quantitative design comprising, firstly, a collostructional analysis and, secondly, a two-group quasi-experimental design with a pre-test, post-test and delayed post-test. Data were collected from a non-randomised control group and experimental group comprising 624 participants in total in order to determine the extent to which the productive knowledge of verb-noun collocations could be improved in first year English Second Language (ESL) students taking an English for Specific Purposes course (ENLS 111). For the purposes of the study, a quantitative design was considered appropriate in order to evaluate participants' performance over a period of six weeks. Performance was measured after a six-week period because this timeline allowed for continuous intervention during one semester while participants completed vocabulary assessments.

4.2 Participants

The participants constituted two groups of first-year English second language university students enrolled for the ENLS 111 module, aiming to improve ESL proficiency, at the

Potchefstroom Campus of the North-West University. The control group constituted the 2017 cohort of students (n=349), and the experimental group constituted the 2018 cohort (n=275). The complete population, i.e. the whole group of ENLS 111 students for 2017 and 2018 was used for the purpose of this study. Using a total population sampling technique allowed for deep insights and analytic generalisations. Relevant statistical tools were used to control for variance. These will be discussed in more detail below where relevant.

4.3 Data collection and analyses

A two-pronged approach was used to address the sub-questions. This required different instruments to collect various forms of quantitative data. The selection is discussed and motivated below.

4.3.1 Instrumentation

First, a collostructional analysis was conducted. Second, a series of pre and post-tests was administered. Finally, a statistical analysis was conducted to determine whether the intervention received by the 2018 cohort had an effect on their productive knowledge of verb-noun collocations. The methods and instrumentation utilised to address each of these questions are discussed in separate sections below.

4.3.1.1 Q1: What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?

A collostructional analysis, as designed by Stefanowitsch and Gries (2003) and Gries and Stefanowitsch (2004a; 2004b), was employed to determine the strength of the verb-noun collocations selected from the ACL. This application is based on the Fisher exact test (Pedersen, 1996) and is useful for determining effect size in terms of the extent to which a construction is typically associated with a lexical item (Gilquin, 2010:195). This first step was included in order to ascertain to what degree participants could use verb-noun collocations freely within an academic context. Participants were not specifically instructed to use certain verb-noun combinations, and therefore the collostructional analysis could provide an objective perspective of what knowledge participants had of verb-noun collocations that they indeed used. A higher collocational strength between a verb and a noun would indicate that many participants had mastered the use of the construction, and that intervention may not be

necessary for that particular item. This step thus has the advantage of aiding researchers in identifying which verb-noun collocations might be problematic and identify them as items that need to be taught more explicitly.

For the purpose of this study, the collostructional analysis was used to determine the strength between the verb-noun collocations selected from the Academic Collocation List (Ackermann & Chen, 2013) as used in ESL students' writing. In total, there are 112 verb-noun collocations on the ACL. The verb 'have' and its inflections were not included in this analysis because it was mostly used as an auxiliary verb in participants' texts. All the other 111 verbs were included and were present in at least one group of texts. In order to analyse participants' writing, the researcher followed the procedure below:

a) First, students were instructed to write two short essays during the course of the semester. The essays were between 400-600 words long, and had to be written in an academic register. These essays were submitted in *MS Word* format on an online platform at the NWU (eFundi) on two occasions, with both cohorts submitting one essay before taking any of the tests, and one after the first post-test had been completed. Both cohorts were restricted to writing about the following topics:

    1. Argue for or against the following statements:
        i. Fake news is problematic.
        ii. Social media should be regulated by the government.
        iii. Marijuana should be legalised for private use.

    Providing specific topics meant that verb-noun combinations were limited. In other words, the researcher could not expect to find all of the verb-noun collocations in student writing, and this would not necessarily be an indication that participants did not have knowledge of those collocations. For example, there might be a higher likelihood of encountering *collect data* than *perform a task*. This is a limitation to the study. However, limiting the possible combinations could provide an indication of the vocabulary size and proficiency of the groups in the sense that it could indicate that not all participants were on the same level. This was especially apparent with errors that participants made. For example, with the verb *change* typical collocates would be *an attitude*, or *a perception*, but one participant used the phrase *change someone's conception**. Errors such as these indicate more serious issues with vocabulary knowledge, especially if found in many texts. These errors as well as the most problematic verbs are discussed in more detail in Chapter 5.

b) The texts were collected by the researcher to create a corpus for each essay. There were four corpora in total: two for the control group (*CG_SE1* and *CG_SE2*) and two for the experimental group (*EG_SE1* and *EG_SE2*). The essays were then edited to ensure that:

    1. All personal participant details were removed;
    2. All reference lists were removed in order to have an accurate word count of the written essays, and no unnecessary details that could affect the collostructional analysis;
    3. The essay files were renamed using the coding mentioned above. For example, a participant's essays from the 2017 cohort were renamed *CG_SE1_001* and *CG_SE2_001*.

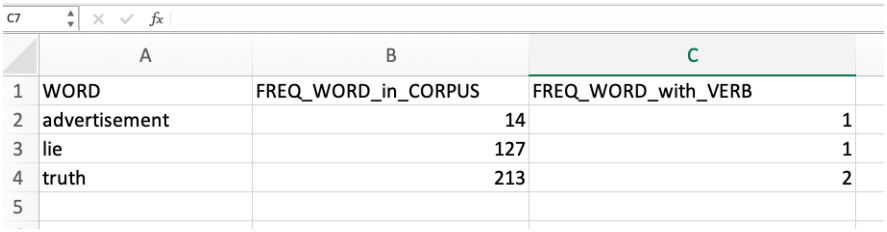
c) For the same reason stated under ii), spelling errors were corrected. The corrections included changing American English spelling to British English spelling. Should two words such as *specialise* and *specialize* appear in the corpus, the programme used to analyse the texts would count the words as two separate words, and this would affect the word count, making the corpus seem larger than it actually was.

d) Next, in order to create a corpus for each group of essays, the researcher converted the texts from *MS Word* to *TXT* format and uploaded the files to *Wordsmith Tools version 6* (Scott, 2011). *Wordsmith* was then used to create a wordlist to retrieve an accurate word count for each corpus. This is crucial because the precise frequencies are necessary for each of the verbs and nouns for which the collostructional strength needed to be determined.

e) After the wordlists were created, the ACL (Ackermann & Chen, 2013) was used to search for verbs used in participants' writing, and a list was made of the relevant nouns that were used with the verbs. In this instance, 'relevant' means that the noun had to be suitable to the academic register. In other words, if a participant used an expression such as *play a game*, 'game' would not be included in the analysis. However, in cases such as *play a part* or *play a role*, both nouns were used because they are appropriate within the context.

f) Once all the relevant nouns were extracted, a list was compiled in MS Excel-format. In this document, two distinctions were made with regard to the noun frequencies. First, the frequency of the noun as it occurs with the verb in question and, second, the frequency of the noun within the corpus as a whole. This enables researchers to determine how frequently the noun co-occurs with a particular verb compared

with how frequently it co-occurs with other verbs as well. Below is an image of the MS Excel file.

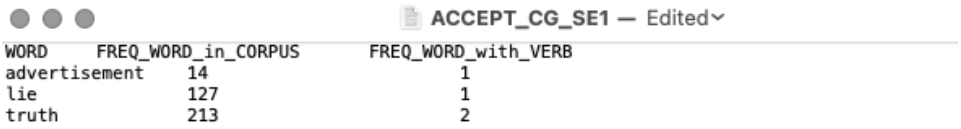


	A	B	C
1	WORD	FREQ_WORD_in_CORPUS	FREQ_WORD_with_VERB
2	advertisement	14	1
3	lie	127	1
4	truth	213	2
5			

Figure 4-1: MS Excel file (accept)

In Figure 4.1, Column A represents the noun that occurred with *accept*, Column B shows the frequency of the noun in the entire corpus, and Column C indicates how many times the noun co-occurred with accept.

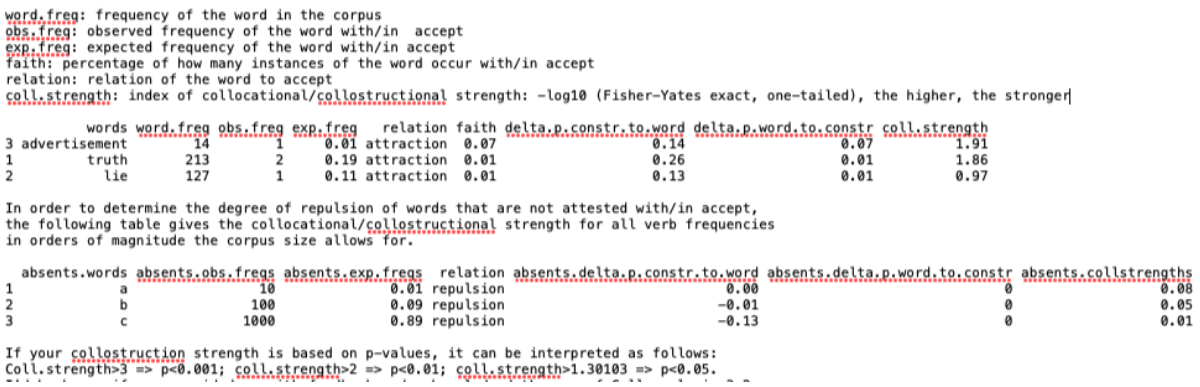
g) After the MS Excel-document is complete, it is converted to an MS DOS Formatted file. The file is then ready to be analysed using the Fisher-Yates exact test (Pedersen, 1996) in order to conduct a collostructional analysis using a software programme called *R* (The R Foundation for Statistical Computing, 1997-). This programme was used to conduct a collostructional analysis as designed by Stefanowitsch and Gries (2003) and Gries and Stefanowitsch (2004a; 2004b), and the collostructional strength was determined for each noun and the verb it co-occurred with. This procedure was repeated for each corpus and the results were compared in order to determine the initial proficiency of the first-year students at the NWU taking the ESP course. Below are two figures (4.2 and 4.3) to show the MS DOS TXT input and TXT output files, along with an explanation of the process to measure collostructional strength.



ACCEPT_CG_SE1 — Edited

```
WORD     FREQ_WORD_in_CORPUS      FREQ_WORD_with_VERB
advertisement    14                       1
lie              127                      1
truth            213                      2
```

Figure 4-2: MS DOS TXT input file (accept)



```
word.freq: frequency of the word in the corpus
obs.freq: observed frequency of the word with/in  accept
exp.freq: expected frequency of the word with/in accept
faith: percentage of how many instances of the word occur with/in accept
relation: relation of the word to accept
coll.strength: index of collocational/collostructional strength: -log10 (Fisher-Yates exact, one-tailed), the higher, the stronger

          words word.freq obs.freq exp.freq   relation faith delta.p.constr.to.word delta.p.word.to.constr coll.strength
3 advertisement        14        1     0.01 attraction  0.07                   0.14                   0.07          1.91
1         truth       213        2     0.19 attraction  0.01                   0.26                   0.01          1.86
2           lie       127        1     0.11 attraction  0.01                   0.13                   0.01          0.97

In order to determine the degree of repulsion of words that are not attested with/in accept,
the following table gives the collocational/collostructional strength for all verb frequencies
in orders of magnitude the corpus size allows for.

  absents.words absents.obs.freqs absents.exp.freqs  relation absents.delta.p.constr.to.word absents.delta.p.word.to.constr absents.collstrengths
1             a                10              0.01 repulsion                          0.00                              0                  0.08
2             b               100              0.09 repulsion                         -0.01                              0                  0.05
3             c              1000              0.89 repulsion                         -0.13                              0                  0.01

If your collostruction strength is based on p-values, it can be interpreted as follows:
Coll.strength>3 => p<0.001; coll.strength>2 => p<0.01; coll.strength>1.30103 => p<0.05.
```

Figure 4-3: R TXT output file (accept)

Figure 4.3 indicates the results of the collostructional analysis of the word *accept* in the control group's first essay. The topic of this essay was *Fake news is problematic*. In the table, there are six columns that are important for interpreting the results of the current study. These are 1) *words*, 2) *word frequency (word.freq*), 3) *observed frequency (obs. freqs)*, 4) *expected frequency (exp. freqs)*, 5) *relation* and 6) *collostructional strength (coll.strength)*.

The first column, *words*, indicates which words were used with *accept*. Results show that *advertised, truth* and *lie* co-occur with *accept* in the student data. The second column, *word. freq*, shows how many times the word occurred in the corpus. The word 'advertised' occurs 14 times. The third column, *obs. freq*, gives an indication of how many times the word occurred with *accept*. *Advertisement* occurred only once with *accept* in this particular corpus. The fourth column, *exp. freq*, indicates how many times the word is expected to occur with *accept*, based on the Fisher-Yates exact calculation. The fifth column, *relation*, shows that there is an attraction between the words, because the observed frequency is higher than the expected frequency. If the observed frequency is lower than the expected frequency, it would indicate that there is a repulsion rate between the words. Gries *et al.* (2005) explain this phenomenon as follows: "[*fall*] is significantly repelled by the intransitive-motion construction because it occurs in this construction less often than expected on the basis of pure chance" (p. 647). This indicated that the words are not expected to co-occur. While the repulsion strength could indicate errors, it could also indicate that combinations are not frequent in the corpus, even if the combination is correct. This results in the collostructional strength – the last column. The higher the score, the stronger the relation between the words. In this case, *accept* shows a high attraction rate to advertisement, and a low attraction rate to *lie*.

The results of this section of the study are discussed in Chapter 5.

4.3.1.2 Q2: To what extent does intervention with explicit teaching of verb-noun collocations result in improving their mastery among first-year ESL students at the North-West University?

In section 3.4.5.1, *Considerations for Collocation Teaching Strategies*, two cognitive approaches to teaching collocations were discussed. These approaches were the attention-drawing technique (Bishop, 2004; Laufer & Gersai, 2005; Boers *et al.*, 2006) and the awareness-raising approach (Ying & O'Neill, 2009). For the present study, the awareness-raising approach was utilised by employing strategies used in studies that had previously shown successful retention of collocations.

The teaching intervention used in the present study constituted three steps and took place during allotted teaching time for the module ENLS 111. After each step, participants provided their answers, and the researcher as well as other participants provided feedback.

1. **Reading exercise to identify verb-noun collocations**

   Participants were provided with texts from *Speed Readings for ESL Learners 3000 BNC* (Millet, 2017). This source was written by the Victoria University of Wellington's School of Linguistics and Applied Language Studies. It contains twenty 400-word passages, each of which is followed by ten comprehension questions. This source was selected based on two criteria: 1) The ACL (Ackermann & Chan, 2013) consisted of verbs from the 3000-word frequency band; and 2) the nouns on the list mostly fell into the 4000-word frequency band. According to research, university students should be able to comprehend 4000-5000 word families in order to understand lectures at undergraduate level (see Nation, 2006; Schmitt 2008, 2010). In a study investigating the size of students' vocabulary, Nizonkiza and Van Dyk (2015) found that students at the North-West University mostly had a large enough vocabulary to enable academic success. However, a percentage of the participants did not reach the minimum requirements in terms of productive vocabulary falling in the 2000- and 3000-word bands. Nizonkiza and Van Dyk (2015) suggest graded readers to improve vocabulary comprehension (p.165).

   Once participants of the present study were provided with a text from Millet (2017), they were given 10 minutes to read the text. They were then asked to write down any expressions that they thought were verb-noun collocations. Examples of verb-noun collocations were provided as a guideline.

2. **The collocation web model**

   The second part of the intervention exercises consisted of an adapted version of McCarthy and O'Dell's (2005) collocation web model. The researcher identified verbs from each passage that were included in the ACL to use as a starting point for the exercise. The node was placed in the centre of the web model, and participants were asked to fill out nouns that could be considered collocates, based on clues that were provided. See figure 4.4 below for details.

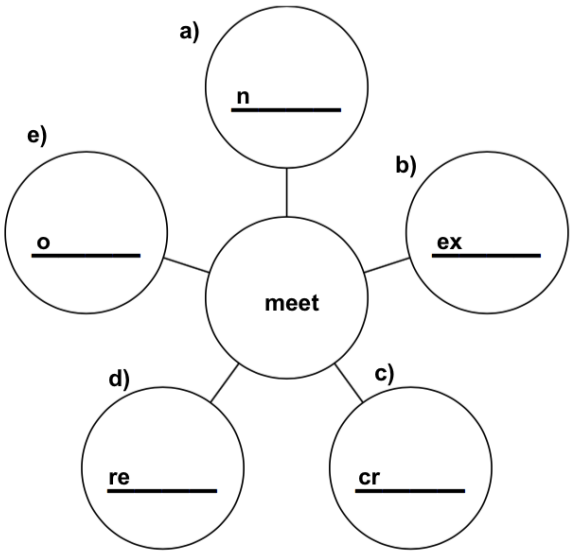


Figure 4-4: Example of collocation web model used in interventions

3. **Productive writing exercise**

   Step three of the exercise involved a short productive writing task in which participants had to choose two collocations and write down a sentence for each. This final step was included to enable participants to practise using the verb-noun collocations productively in a controlled exercise.

This intervention was repeated six times. It was evident during the first intervention that participants were not aware of the concept 'collocation', and most of the time was used to explain the concept. Even after an explanation and examples, it became apparent during the second intervention that participants still struggled with the concept. Participants could not identify collocations, and some identified phrases that were parts of sentences that did not contain any collocations. For example: *has a mixed economy*, *are exported to 90 countries*, and *find new places*. It was thus necessary to adjust the initial teaching invention to increase

exposure to and awareness of collocations. The researcher introduced additional materials such as collocations dictionaries, online dictionaries, and corpus data to illustrate what collocations are. The fundamental teaching approach (awareness-raising) was the same throughout the intervention phase, but the presentation of the content varied slightly based on learning styles, such as visual learning styles and more analytic learning styles. This enabled the researcher to ensure interventions that could be more pedagogically sound in that they accommodated a greater number of participants' learning styles. An example of one of the intervention exercises can be found in Annexure B.

To gauge the extent to which intervention may have resulted in collocations development, the following question was answered: Is there any statistically significant difference between the results of the pre-test, post-test and delayed post-test of the experimental group?

To this end, the productive version of the Vocabulary Levels Test (Laufer & Nation, 1999) was used. This test is discussed in detail in section 3.5.2.1, and was used to assess participants' progress over a number of weeks. The test items were selected from published and peer-reviewed articles from a variety of academic fields that included Economic, Agricultural, Educational, Legal, Health, and Communication Studies. Published articles were used in order to ensure that the test items were based on authentic language output. Most of the participants taking the course were enrolled in the School of Communication Studies and the Law Faculty. Participants would therefore have been exposed to these registers and would have been required to use the registers accurately. A complete list of articles used for the tests is available on request. This phase of the study consisted of two parts entailing the completion of the experimental procedure for the control group (2017) and for the test group (2018), followed by data analysis. The timeline for the procedure was as follows:

Timeline:	Beginning 1st wk	1st wk May	end of wk1	wk 2	end wk 2	wk 3	end wk 3	wk 4	end wk 4	wk5	end wk 5	wk 6	end wk6	early 2nd semester
Control group 2017	T1	xx	ti	xx	tii	Intervention I3	T2	xx	ti	xx	tii	xx	T2 repeat	T3
Experimental group 2018	T1	Intervention I1	ti	Intervention I2	tii	Intervention I3	T2	Intervention I4	ti	Intervention I5	tii	Intervention I6	T2 repeat	T3

Table 4-1: Testing timeline

For both cohorts, the pre-test was administered in the first week following the university holiday in April to ensure a continuous period of at least 6 weeks before exams started.

2017 Control Group procedure

Data gathered during this phase served as benchmark against which to compare the effect of the planned interventions in the following phase.

This phase commenced in April 2017 and comprised only a series of formative tests. The pre-test (T_1) for the control group was administered, intermittent tests (t_i, t_{ii}, t_{iii} and t_{iv}) and the post-test was administered twice (T2 and T3). The delayed post-test (T4) six weeks later. No intervention took place during this phase.

2018 Experimental Group procedure

This phase commenced in April 2018. The same outlined procedure was followed as in 2017, but with the inclusion of interventions I1 to I6. The same exercises were administered as in 2017 and according to the same timeline.

The repeated post-test and use of intermittent exercises in addition to the pre-, post-, and delayed post-test exercises served to increase the number of observation points to get a clearer picture of any developmental pattern that may be evident between the pre-test, post-test, second post-test, and delayed post-test phases. In this way, reliability of the results was increased. In addition, it served to inform the duration required for the intervention, i.e.

comparing the extent of development after three weeks' intervention, and after six weeks' intervention.

In terms of analyses, variance within participants as people, as well as variance between the tests needed to be controlled. Hierarchical Linear Measures (HLM) (Hancock & Mueller, 2010; Woltman *et al.*, 2012) were applied to control for variance in the groups, with person dependence as the primary unit of analysis. This analysis allows for missing data and is useful in applications to the full survey group to ensure the intragroup and intergroup comparability. Woltman *et al*. (2012:53) explains: "HLM simultaneously investigates relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels than other existing analyses". As it was anticipated for this study, the independent variable affected the dependent variable, and the effect size of the treatment was determined by means of Cohen's D value (Cohen, 1988), based on the HLM averages and variance. The results of the rests are discussed in detail in Chapter 5.

The results of the constructional analysis and the two-group quasi-experimental design with a pre-test, post-test and delayed post-test were also considered carefully. Inferences were drawn based on the findings achieved in the preceding stages described above. These inferences and recommendations are discussed in Chapter 5. This information serves to motivate and guide future classroom practices related to teaching collocations, as well as future research on the matter.

CHAPTER 5: PRESENTATION AND INTERPRETATION OF RESULTS

5.1 Introduction

In this chapter a detailed account of the findings of the present study is provided. The structure of this chapter follows that of section 4.3, addressing each of the research questions separately, along with the findings that resulted for that phase of the study.

The aim of this study was to determine the extent to which first-year university students' controlled productive knowledge of verb-noun collocations could be improved by means of explicit teaching. To determine this, the following sub-questions were posed:

1. What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?
2. To what extent does intervention with explicit teaching of verb-noun collocations result in improving their mastery among first-year ESL students at the North-West University? This question is answered by determining whether
    a. there is any statistically significant difference between the results of the pre-test, post-test and delayed post-test of the experimental group.

This study followed a quantitative design comprising firstly a collostructional analysis and secondly a two-group quasi-experimental design with a pre-test, post-test and delayed post-test. Data were collected from a non-randomised control group and experimental group comprising 624 participants in total in order to determine the extent to which the controlled productive knowledge of verb-noun collocations can be improved in first year English Second Language (ESL) students taking an English for Specific Purposes course (ENLS 111). For the purposes of the study, only quantitative data were collected in order to evaluate participants' performance over a period of six weeks for both groups.

5.2 Interpretation of results

A two-pronged approach was used to address the sub-questions. Results of the collostructional analysis are reported and discussed first. Then the results are reported and discussed for the of pre-, post- and delayed post-test administered. Finally, the results are

reported for statistical analysis conducted to determine whether the intervention received by the 2018 cohort had an effect on their productive knowledge of verb-noun collocations.

5.2.1 Collostructional analysis of participants essays

Q1: What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?

To answer Question 1, a collostructional analysis as designed by Stefanowitsch and Gries (2003) and Gries and Stefanowitsch (2004a; 2004b) was employed to determine the extent to which participants in the study had mastered the verb-noun collocations selected from the ACL (Ackermann & Chen, 2013) before the intervention and could master the verb-noun collocations after the invention phase. The intervention phase is discussed in detail in sections 4.3.1.2 and 5.2.2.

The collostructional analysis is based on the Fisher-Yates exact test (Pedersen, 1996), which is useful for determining effect size in terms of the extent to which a construction is typically associated with a lexical item (Gilquin, 2010:195). This first step was included in order to ascertain the degree to which participants freely mastered the verb-noun collocations on the ACL within an academic context. The results of the analyses provided insight into which collocations participants could master, and what types of errors were made. This information is potentially valuable to inform future interventions aimed at the kinds of difficulties that first year students may encounter with mastering verb-noun collocations. The essays for both cohorts were edited and analysed according to the method discussed in section 4.3.1.1 in order to compile a list of verb-noun collocations used by participants based on the 112 verb-noun collocations on the ACL. The analysis and the editing process of the essays indicated that participants typically made the following types of errors: a) using an informal academic register instead of a formal or academic register; b) using verbs as nouns; c) using words with similar meaning in the same sentences with the same verb (as part of the same clause); d) verb tense and agreement errors; and e) using incorrect noun collocates with verbs. A complete list of these errors can be found in Appendix F, but some examples are discussed here. While errors a), b), and d) were found throughout participant writing, they were not limited to verbs used in the analyses and show a broader picture of the writing of the participants used in the study. However, errors c) and e) were identified in the case of the verbs that were analysed and indicate a specific problem area with the use of verb-noun collocations for these cohorts. Each type of error is briefly discussed below.

a) Incorrect register

Results indicated that participants showed a preference to more informal vocabulary. Some examples of a) include expressions such as 'find a way' as opposed to 'find a means'. Similar errors occurred with nouns such as *money vs profit* and *money vs income*, as well as *measures vs rules*. In certain cases, the collostructional analysis of the second essays, which were written after the intervention had taken place, indicated that participants started using more formal expressions, particularly in the case of 'make money' and 'make a profit'. More participants used 'profit' as the preferred collocate with *make* to indicate financial gain than before the intervention. This could signal a greater awareness and attentiveness to using the more formal phase which is more appropriate in the academic context.

b) Using verbs as nouns and vice versa

The collostructional analyses also showed that participants in the study had great difficulty with nominal derivational suffixes. In other words, they did not know which suffixes to add to a word to turn it into a noun. For example, participants used expressions such as 'leads to a reduce in corruption' (participant EG238) and 'If the Government succeeds in regulating social media it will make a mock of our Democracy' (participant CG172) instead of 'lead to a reduction' and 'make a mockery'. Participant CG_002 also used the word *comfusement* in the sentence 'Not only was this information wrong, it spread like wild fire, causing heavy confusement'. When referring to the publishing process, a participant used the phrase 'the publishment of something'. *Publishment* is used in a very particular context when referring to a printed work that is for sale or when referring to the act of publishing a text and would not typically be used within the context of publishing a news article, for example, and was inappropriate for the context of the essay.

Other verb/noun errors included using the words 'affect' and 'effect' interchangeably as a noun or a verb. This was also the case for the words *belief vs believe* and *extent vs extend*. Examples from the data include 'it will effect people', and 'they belief something'. Participant CG_36 used the sentence 'people sometimes belief in the madness that is been spread'.

c) Using words with similar meaning

Errors included using expressions such as ‘to gain financial or political gain’, or ‘gain attraction’ instead of ‘gain attention’. Both examples illustrate uncertainty as to the meaning and use of the problematic lexical items. In the former example, the word *gain* is repetitively in both verb and noun form. Should the adjectives *financial* and *political* be removed, the phrase would be *to gain gain*. This is a problematic combination and the noun *gain* more likely collocates with the verb *to achieve* One would rather say 'to achieve financial and political gain' or 'to gain financially and politically'. In the latter case, the verb *draw* would have been better suited with *attention*. This example shows that there is not a clear understanding of the difference in meaning between *attention* and *attraction*. Another example from the data ‘is to make a fortune of profit’ where *profit* and *fortune* have similar meanings within the particular context, but *fortune* is less formal than *profit* in an academic context.

This phenomenon could indicate that participants lack knowledge of the meaning of the words that they used, and that they have difficulty identifying synonyms or near-synonyms of those words.

d) verb tense and agreement errors

Results from the analysis also showed that participants had difficulty using verbs correctly as they made errors related to verb tenses and agreement. Some examples from participants’ texts include the following: *something is been given*; *have/has became*; *news is been provided*; *that been said* (that being said); and *everyone has went through it*. For example, participant EG124 used the phrase ‘misinformation can be reduces to people or community’. Participant CG_36 also made verb errors (as well as errors listed under c). An example is ‘people sometimes belief in the madness that is been spread’. A complete list of these errors are presented in Appendix E.

e) incorrect noun collocations

In order to provide a detailed discussion of errors related to incorrect noun collocations, the verbs *make* and *cause* are discussed in detail below.

The ACL contains 112 verb-noun collocations. Each of the verbs were analysed in all four corpora that were compiled using participant essays. Of the 112 verbs, *make* and *cause* yielded the most verb-noun collocations and contained many of the errors discussed above and illustrated below. Even though only two of the verbs are discussed as a way of exemplifying

the errors that participants made, summaries of the analyses of all 112 verbs are available on request. Below, a table is provided for each of the two verbs – *make* and *cause* – following the same format as Figure 4.3 (section 4.3.1.1). The nouns of each table are organised alphabetically. The items that are printed in bold indicate noun collocates that are erroneously used with *make* and *cause*. These erroneously used noun collocates also allow the researcher to identify items that should most likely be included in an intervention course, because there are several errors of the same verb-noun collocation across both cohorts. This indicates that participants did not know the construction before the intervention, and continued to have difficulty with the construction upon completion of the intervention phase. In this way the collostructional analyses provide insights as to which collocations should be taught more explicitly. This is a key question in current research (see section 3.4.1.6 for a detailed discussion).

Below, Tables 5.1 and 5.2 are discussed simultaneously in order to illustrate the method used to compare results from both cohorts. The collostructional strength is significant at 1.3. However, for the purpose of this study all collocates that occurred with the verbs were included. The main consideration for these analyses was not whether participants used verb-noun collocations that were significant according to the method employed, but to identify all the collocates that were used. Therefore, collocates with a collostructional strength below 1.3 were also included. Based on these results, inferences could be made with regard to the noun collocates that participants mastered, the type of register participants preferred to use in their writing, as well as errors that occurred.

SUMMARY OF ANALYSIS OF *MAKE*

NOUN	CG_SE1					CG_SE2					EG_SE1					EG_SE2				
	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength
A_LIVING	6	5	0.43	A	5.00	12	4	0.73	A	2.34	4	3	0.29	A	2.86	8	4	0.43	A	3.31
ACCUSATION	33	5	2.34	A	1.09															
ALLEGATION	57	5	4.05	A	0.42											1	1	0.05	A	1.27
ARGUMENT																44	1	2.37	R	0.51
ARTICLE	**341**	**1**	**24.22**	**R**	**9.71**															
ASSUMPTION						1	1	0.06	A	1.21	7	7	0.50	A	8.03	5	1	0.27	A	0.62
ATTEMPT						45	1	2.75	R	0.64	14	1	1.00	R	0.13					
CASE	92	1	**6.53**	R	**2.05**															
CHOICE	13	2	0.92	A	0.63	13	3	0.80	A	0.73	12	4	0.86	A	2.09	9	1	0.48	A	0.41
CLAIM	24	3	1.70	A	0.62						21	1	1.50	A	0.26	6	1	0.32	A	0.55
COMMENT	12	1	0.85	A	0.23	33	4	2.02	A	0.85	8	1	0.57	A	0.35					
CONCLUSION	**101**	**7**	**7.17**	**R**	**0.24**	**64**	**2**	**3.91**	**R**	**0.62**	**69**	**3**	**4.94**	**R**	**0.58**	**69**	**1**	**3.72**	**R**	**0.97**
CONNOTATION																1	1	0.05	A	1.27
CONTRIBUTION											3	1	0.21	A	0.70	4	1	0.22	A	0.70
DAMAGE						**43**	**2**	**2.43**	**R**	**0.30**	**47**	**1**	**3.36**	**R**	**0.85**					
DEBATE																**40**	**1**	**2.15**	**R**	**0.45**
DECISION	37	14	2.63	A	7.03	33	10	2.02	A	4.77	47	32	3.36	A	25.62	25	16	1.35	A	14.35
DIRECTIVE						**3**	**1**	**0.18**	**A**	**0.76**										
DISCUSSION						**20**	**1**	**1.22**	**R**	**0.19**										
DISTINCTION	5	1	0.36	A	0.51						1	1	0.07	A	1.15					
EARNING(S)											3	1	0.21	A	0.70					
EFFECT	**155**	**1**	**11.01**	**R**	**3.89**															
ERROR											10	1	0.72	A	0.28					
EXCEPTION	1	1	0.07	A	1.15											1	1	0.05	A	1.27
EXPERIMENT																**1**	**1**	**0.5**	**A**	**1.27**
FORTUNE	2	2	0.14	A	2.30						1	1	0.07	A	1.15	1	1	0.05	A	1.27
GUIDELINE	**2**	**1**	**0.14**	**A**	**0.86**															
IDEA	**44**	**1**	**3.12**	**R**	**0.77**															
IMPACT	**477**	**6**	**33.87**	**A**	**9.02**						**122**	**8**	**8.73**	**R**	**0.31**	**62**	**3**	**3.34**	**R**	**0.24**
INCOME											8	1	0.57	A	0.35					
INCONVENIENCE	**1**	**1**	**0.07**	**A**	**1.15**															
INFERENCE	5	5	0.36	A	5.75						4	4	0.29	A	4.59					
INFORMATION	**780**	**1**	**55.39**	**R**	**24.49**						**470**	**1**	**33.65**	**R**	**14.12**					
INVESTIGATION	18	2	1.28	A	0.43															
JUDGEMENT						5	1	0.31	A	0.57	10	2	0.72	A	0.80					
MEASURE						**34**	**1**	**2.08**	**R**	**0.43**										
MISTAKE	2	1	0.14	**A**	0.86	3	1	0.18	A	0.76	11	1	0.79	A	0.25	7	3	0.38	A	2.34
MOCK						**1**	**1**	**0.06**	**A**	**1.21**										
MOCKERY	2	1	0.14	**A**	0.86	4	1	0.24	A	0.65										
MONEY	120	38	8.52	A	15.13	74	15	4.53	A	4.50	143	51	10.24	A	22.82	83	11	4.47	A	2.34
OPINION	**155**	**1**	**11.01**	**R**	**3.89**						**92**	**1**	**6.59**	**R**	**2.08**					
PLEA																3	2	0.16	A	2.08
POINT						52	1	3.18	R	0.79	49	2	3.51	R	0.51					
PRESUMPTION	1	1	0.07	A	1.15															
PROFIT	29	13	2.06	A	7.62	4	2	0.24	A	1.69	36	7	2.58	A	1.90	6	3	0.32	A	2.56
PROGRESS	5	2	0.36	A	1.36											10	4	0.54	A	2.87
PROVISION						**3**	**1**	**0.18**	A	**0.76**	**2**	**2**	**0.14**	**A**	**2.29**					
REMARK	2	1	0.14	A	0.86															
REPORT						**43**	**1**	**2.63**	**R**	**0.60**	**64**	**1**	**4.58**	**R**	**1.30**					
RESEARCH	**90**	**4**	**6.39**	**R**	**0.65**	**31**	**1**	**1.90**	**R**	**0.37**										
REVENUE											17	2	1.22	A	0.46	26	1	1.40	R	0.23
SPECULATION	**11**	**1**	**0.78**	**A**	**0.26**															
STAND	**8**	**2**	**0.57**	**A**	**0.97**															
STATEMENT	62	12	4.40	A	2.93	62	12	3.79	A	3.51	50	3	3.58	A	0.29	18	1	0.97	A	0.20
STIR	8	1	0.57	A	0.35															
STUDY	**49**	**1**	**3.48**	**R**	**0.89**											**104**	**1**	**5.60**	**R**	**1.68**
SURVEY	**37**	**1**	**2.63**	**R**	**0.60**	**8**	**1**	**0.49**	**R**	**0.40**										
USE OF STH	37	13	2.63	A	6.13	98	14	5.99	A	2.69	45	11	3.22	A	3.61	479	9	25.79	A	4.25

Table 5-1: Summary of collostructional analysis of MAKE

Table 5-2: Summary of collostructional analysis of CAUSE

SUMMARY OF ANALYSIS OF *CAUSE*

NOUN	CG_SE1					CG_SE2					EG_SE1					EG_SE2				
	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength	Freq.	Obs. Freq.	Exp. Freq.	Attraction/ Repulsion	Coll. Strength
ALARM																				
ARGUMENT						40	2	0.94	A	0.61										
ATROCITY	2	1	0.08	A	1.12															
ATTENTION											**97**	**1**	**4.60**	**R**	**1.29**					
CHANGE																24	2	1.00	A	0.58
CHAOS	30	11	1.15	A	8.19	18	5	0.42	A	4.34	20	11	0.95	A	9.58					
COMMOTION											2	1	0.09	A	1.03					
COMPLICATION																3	1	0.13	A	0.92
CONCERN											17	2	0.81	A	0.72					
CONFLICT						20	4	0.20	A	2.97	37	17	1.76	A	12.86					
CONFUSION	72	35	2.77	A	30.28	19	7	0.45	A	6.85	49	29	2.32	A	25.84					
CONTROVERSY	10	1	0.38	A	1.27	5	2	0.12	A	2.28	4	3	0.19	A	3.39					
DAMAGE	31	18	1.19	A	17.56	43	20	1.01	A	21.27	46	24	2.18	A	19.65	28	11	1.17	A	8.22
DECREASE						22	2	0.52	A	1.03										
DETRIMENT											6	1	0.28	A	0.60					
DIFFICULTY	5	1	0.19	A	0.75											13	1	0.54	A	0.37
DISAGREEMENT						3	1	0.07	A	1.16										
DISMAY	1	1	0.04	A	1.41															
DISRUPTION	2	2	0.08	A	2.83	5	4	0.12	A	5.83	2	2	0.09	A	2.65	1	1	0.04	A	1.38
DISTRAUGHT											**1**	**1**	**0.05**	**A**	**1.32**					
DISTRESS	2	1	0.08	A	1.12						7	5	0.33	A	5.34	5	3	0.21	A	3.17
DISTRUST	7	1	0.27	A	0.62															
DISTURBANCE						3	2	0.07	A	2.79										
DIVISION	3	3	0.12	A	4.25															
DOUBT	28	1	1.08	R	0.15															
DOWNFALL	5	1	0.19	A	0.75															
DRAMA	8	4	0.31	A	3.88															
DROP						1	1	0.02	A	1.63	3	1	0.14	A	0.87					
EFFECT											**95**	**3**	**4.51**	**R**	**0.48**	**368**	**14**	**15.35**	**R**	**0.37**
FRICTION											4	4	0.19	A	5.30					
FRIGHT											**1**	**1**	**0.05**	**A**	**1.32**					
GROWTH																**21**	**1**	**0.88**	**A**	**0.23**
HARM	54	18	2.08	A	12.24	27	12	0.64	A	12.59	49	20	2.32	A	13.82	34	16	1.42	A	13.24
HAVOC	4	4	0.15	A	5.67	2	1	0.05	A	1.33	10	6	0.47	A	5.71	1	1	0.04	A	1.38
HOSTILITY	1	1	0.04	A	1.41															
HYSTERIA	6	1	0.23	A	0.68						4	3	0.19	A	3.39					
IMPLICATION											**14**	**2**	**0.66**	**A**	**0.85**	**9**	**1**	**0.11**	**A**	**0.50**
INCIDENT																5	2	0.21	A	1.80
INSTABILITY											8	2	0.38	A	1.28					
ISSUE	145	3	5.58	R	0.73	226	3	5.33	R	0.67	77	2	3.65	R	0.55	54	5	2.25	A	1.14
LOSS	12	2	0.46	A	1.12	10	1	0.24	A	0.67	29	2	1.38	A	0.39	46	11	1.92	A	5.71
MISPERCEPTION	7	1	0.27	A	0.62															
MISTRUST	**9**	**3**	**0.35**	**A**	**2.40**	**2**	**1**	**0.05**	**A**	**1.33**										
MISUNDERSTANDING	14	4	0.54	A	2.80	6	1	0.14	A	0.87	10	1	0.47	A	0.41					
OUTRAGE	7	1	0.27	A	0.62	6	1	0.14	A	0.87	3	3	0.14	A	3.98					
PARANOIA	4	2	0.15	A	2.08															
PROBLEM	289	22	11.12	A	2.77	213	17	5.03	A	5.01	254	38	12.05	A	9.82	188	23	7.84	A	5.57
REDUCTION						2	2	0.05	A	3.26										
RIOT	3	2	0.12	A	2.37															
RISK																**73**	**1**	**3.05**	**R**	**0.73**
STIR											4	2	0.19	A	1.90	1	1	0.04	A	1.38
STRESS											11	3	0.52	A	1.88					
SUCCESS	**8**	**1**	**0.31**	**A**	**0.57**															
TERROR	5	2	0.19	A	1.86						4	2	1.90							
THREAT																**13**	**1**	**0.54**	**A**	**0.37**
TROUBLE	14	4	0.54	A	2.80						11	3	0.52	A	1.88					
TURMOIL	3	2	0.12	A	2.37	1	1	0.02	A	1.63	3	2	0.14	A	2.19					
UNCERTAINTY											19	7	0.90	A	4.81					
UNSTABLENESS											**1**	**1**	**0.05**	**A**	**1.32**					
UPROAR	8	3	0.31	A	2.56	7	2	0.17	A	1.97	4	2	0.19	A	1.90	1	1	0.04	A	1.38

5.2.1.1 Discussion of *MAKE* and *CAUSE*

As mentioned in the introduction of this section, the items in bold are nouns that do not typically occur with the verb *make*.

For many of these nouns, the observed frequency in the corpus is lower than the expected frequency based on the data provided in terms of the overall frequency of the noun, and how many times it occurred with the verb. This results in a repulsion (compare Gries *et al.* 2005:656; 646), even though the collostructional strength is high. In other words, there is a very low likelihood of these words co-occurring (see section 4.3.1.1 for a detailed discussion). The results can apparently be attributed to the fact that these were instances where participants used the verb-noun erroneously. For instance, one example that surprised the researcher was an error in the CG_SE1 and EG_SE1 group, viz. 'make information'. This error only occurred once in both of the mentioned corpora, and in the essay on *Fake news is problematic*. In the same essay, one participant in CG_SE1 (CG_98) used the expression 'malicious articles are made to cause dismay'. Here, the *make* was used instead of *write*. These are examples of error *e*, because the incorrect noun collocate was used with *make*.

The most common error made by participants from both cohorts is 'make a conclusion'. This is another example of using the incorrect noun collocate with *make* (i.e. error *e*). The noun *conclusion* would more frequently be used with the verbs *draw* and *reach*, and less frequently with *come to*. However, while this error occurred in all four of the corpora, it occurred less frequently in the second essay of each cohort. The control group (CG) did not receive specific instruction on how to use the word *conclusion* with appropriate verbs. However, the control group did encounter instances of verb-noun collocations including *conclusion* in the test exercises that were used as the main source for statistical analysis to determine whether there was an improvement in the experimental group. This could indicate that participants learned the correct combination implicitly through exposure to the test items.

By implication, erroneously used noun collocates such as these allow the researcher to identify items that should most likely be included in an intervention course, because there are several errors of the same verb-noun collocation across both cohorts. In other words, participants have not mastered these items. Similar examples from Table 5.1 are *damage*, *research*, *study*, and *survey*. To illustrate further, *research* and *study* are discussed.

Results from the 2017 cohort show that fewer participants used the expression ‘make research’ than in previous accounts. In essay 1 (CG_SE1) *research* occurred 90 times in the corpus as a whole, and was used with *make* four times. In essay 2 (CG_SE2) *research* occurred 31 times, but only co-occurred with *make* once in that corpus. Results also showed that a participant in the experimental group used the expression ‘make a study’ after the intervention phase. Participants were not only exposed to the verbs that *study* is a collocate to in interventions, but also in the test items throughout the study. Items such as this one are crucial to take notice of, because it can inform the selection of verb-noun collocations to teach in future interventions. It is, however, necessary to note here that many of these nouns showed a repulsion to *make*. In other words, the words are not likely to co-occur and occurred in combination with *make* less frequently than expected. These kinds of errors could indicate that participants have difficulty making the form-meaning connection of the collocation on a syntagmatic level with these particular items (see section 2.3.3 for a detailed discussion).

Results indicate that the experimental group preferred more formal nouns to collocate with expressions related to income. For example, in both essays, the control group preferred to use ‘make money’ instead of ‘make a profit’. This is not the case with the experimental group. In the last essay of the EG-cohort, only ‘make a profit’ and ‘make revenue’ are used, instead of the less formal ‘make money’. This indicates that some participants became more aware of the academic register throughout the study.

However, a common error made by participants from both cohorts was using words with similar meaning in the same sentence or phrase – i.e. tautology. For example, ‘make a fortune of money/profit’. The expression ‘make comments and remarks’ was also present in the corpus. This was also evident with *cause*. Corpus analyses showed that participants often used expressions such as ‘cause havoc and drama or chaos’. This indicates that participants are not familiar with synonyms of nouns such as *havoc* and *chaos*, especially when informal vocabulary is used in conjunction with formal vocabulary. This error also occurred with the word *address*. Participants often used expressions such as ‘address a problem or issue’. Another example is ‘change a view and opinion’ – these words are close synonyms and the repetition is unnecessary. It does not serve to clarify or add essential information.

The tables above indicate that numerous noun collocates were correctly used with *make* and *cause*. However, throughout both corpora, some frequent errors occurred where participants used the incorrect form of a word. For example, one participant in the CG-cohort used ‘make a mock of online users’ instead of the correct form ‘mockery’. This kind of error also occurred

with *cause* in the expression ‘cause distraught’ instead of the correct ‘cause distress’. These types of errors indicate that participants have difficulty distinguishing verbs and nouns (and other parts of speech), and this is quite a common error in first years taking the ENLS 111 course[1]. These types of errors could be one of the reasons why verb-noun collocations in particular (for this group) are difficult to master.

The aim of this section was to address the question as to the extent of collocational competence of the 2017 and 2018 cohorts. The collostructional analysis provided insight into what participants knew in terms of verb-noun collocations. Results confirmed that verb-noun collocation are problematic, at least for the population in question. Several types of errors were highlighted, and examples of the most common errors were provided. Based on the results of the analyses, the conclusion can be reached that a collostructional analyses of participant writing could provide insight into the level of their collocational knowledge, but more crucially, it provided insight into the frequent errors as well as items that consistently proved to be challenging for participants. A collostructional analysis would therefore be a useful tool to utilise in teaching interventions focussed on collocational knowledge. Results from the present study can be used as a starting point to design awareness raising techniques to help first-year university students master verb-noun collocations that are suitable to the academic context.

5.2.2 Teaching intervention: Experimental group

Q2: To what extent does intervention with explicit teaching of verb-noun collocations result in improving their mastery among first-year ESL students at the North-West University?

In section 3.4.5.1, *Considerations for Collocation Teaching Strategies*, two cognitive approaches to teaching collocations were discussed. These approaches were the attention-drawing technique (Bishop, 2004; Laufer & Gersai, 2005; Boers *et al.*, 2006) and the awareness-raising approach (Ying & O’Neill, 2009). For the present study, the awareness-raising approach was utilised by employing strategies used in studies that had previously shown successful retention of collocations using this approach.

Each exercise in the teaching intervention used in the present study constituted three steps and took place during allotted teaching time for the module ENLS 111. After each step

[1] This observation has been made by the researcher who has been teaching the course for several years.

participants provided their answers, and the researcher as well as other participants provided feedback. Initially, the researcher allowed 30 minutes of the allotted class time to complete each intervention exercise, dividing each phase into intervals of 10 minutes.

1. Reading exercise to identify verb-noun collocations

In this section, the details of the reading exercises used in the interventions are discussed.

Participants were provided with texts from *Speed Readings for ESL Learners 3000 BNC* (Millet, 2017). Each text consisted of 400 words, followed by ten comprehension questions (See Appendix C for a list of references to the reading passages and Appendix D for the questions). During this stage of the intervention exercises, it became apparent that participants experienced several difficulties. Part of the ENLS 111-course focussed on comprehension and reading strategies. Initially, participants were given 10 minutes to scan through each text, after which they were asked to identify and write down verb-noun combinations that seemed significant within an academic context (See Annexure C for examples. The verb-noun collocations are marked in bold). Participants were given an example of what such a combination might look like, for example *make a decision*. They were then asked to read through the text more carefully a second time and revise their list. After identifying possible significant expressions, participants provided feedback and the verb-noun collocations that were present in the text were discussed during the session. This enabled participants to identify possible errors.

During this phase of the exercise, it became apparent that participants needed more than 10 minutes to scan through the texts (often up to almost 20 minutes). This resulted in a longer intervention session during class time overall. During this phase, it also became apparent that participants had difficulty identifying verb-noun combinations as a whole, and had very little knowledge of verb-noun collocations that are common and occur frequently in speech and writing. Below is an example of the expressions that were identified in some of the intervention exercises:

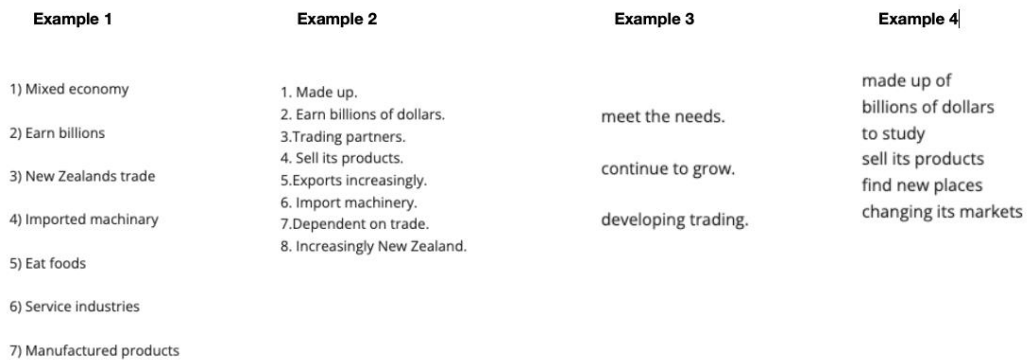


Figure 5-1: Examples of verb-noun combinations identified by participants

While some of the examples in Figure 5.1 are relevant (*meet needs*, *sell products*), most of the examples provided by participants are incorrect. For example, mixed economy is an example of an adjective-noun collocation, and (within the context of the text) *exports increasingly* is an incomplete phrase. In most cases, the examples provided are random sections of clauses found in the text. For details on the text, please consult Appendix D.

2. The collocation web model

The second part of the intervention exercises consisted of an adapted version of McCarthy and O'Dell's (2005) collocation web model (See Appendix C for the list of references to reading texts). For each intervention, the researcher selected three verbs that were present in the text and created a collocation web model. The node was placed in the centre of the web model, and participants were asked to fill out nouns that could be considered as collocates, based on clues that were provided. Below is a diagram that was provided in one of the exercises that were based on the same text as the example answers given by participants in Figure 5.1. The answers of one of the participants are also provided as an example. Feedback was provided by both the participants and the researcher to identify as many verb-noun collocations that were possible with a particular verb.

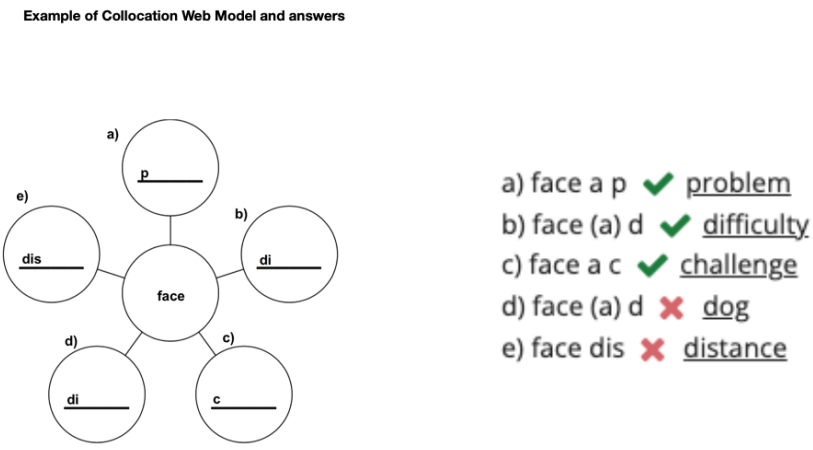


Figure 5-2: Collocation Web Model and example answers

In Figure 5.2, items *a-c* are correct. Items *d* and *e* are incorrect. Participants were explicitly instructed to complete the Collocation Web Model using academic vocabulary, and ensuring that they would be able to use the expression in a sentence that has clear meaning. The combination *face a dog* is highly unlikely to occur at all except in a very particular context in which a person might be confronted with an aggressive dog, and therefore even less likely to occur in academic writing. The same is true for *face distance*.

In figure 5.3, the verb/noun errors discussed in section 5.3.1 are evident. This figure shows the answers to the web model illustrated in figure 5.2, as well as the sentences that were produced by the participant.

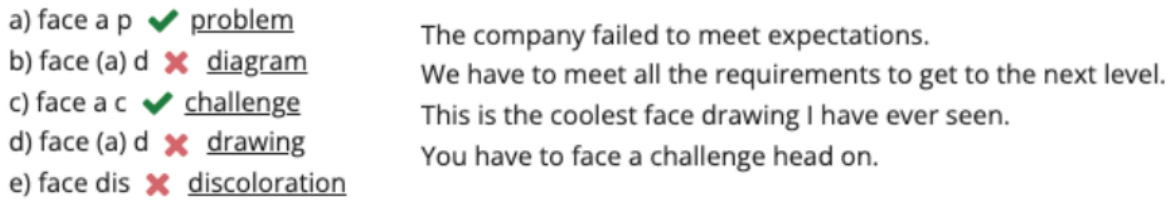


Figure 5-3: Examples of participant answers to collocation web model and productive writing exercise

In this example, the participant had difficulty distinguishing *face* as a verb and a noun in some instances. Item d) the participant used the expression 'face a drawing'. From the example sentences it is evident that the participant misinterpreted the meaning of *face* in this exercise, based on the sentence *This is the coolest face drawing I have ever seen*. This was a common occurrence in the intervention exercises. However, there are instances where this participant

formed correct verb-noun collocations, such as ‘face a problem’ and ‘face a challenge’. This is also evident from the three other sentences in which the collocations are used accurately.

5.2.2.3 Productive writing exercise

Step three of the exercise involved a short productive writing task in which participants had to choose two collocations that they had already come up with during step two of the exercise and write down a sentence for each. Participants were asked to use two of the collocations they provided in the collocation web model in separate sentences. This final step was included to enable participants to practise using the verb-noun collocations productively in a controlled exercise. This also enable the researcher to gain a better understanding of how participants used the expressions in productive language. Since there were multiple verbs in each text, participants were allowed to use any of the verbs with their correct collocates. Below is an example of the answers provided by the same participant as in Figure 5.2.

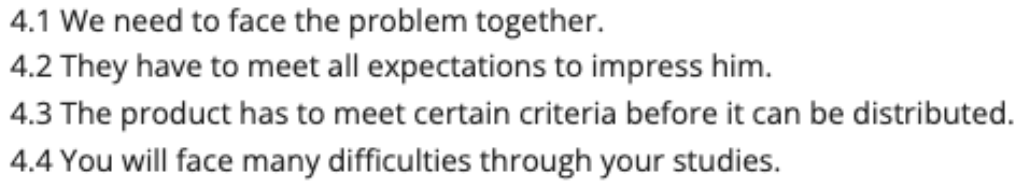


Figure 5-4: Example sentences in productive exercise

In Figure 5.4, the participant used the collocations ‘meet needs/criteria’ and ‘face a problem/difficulty’ correctly. Figure 5.2 illustrates that the participant could correctly identify and produce some verb-noun collocations with *face*.

These types of errors gradually decreased in number during the intervention period.
This intervention was repeated six times, with one exercise each week. It was evident during the first intervention that participants were not aware of the concept ‘collocation’, and the researcher used most of the allocated time explaining what collocations are to the participants. Even after an explanation and examples, it became apparent during the second intervention that participants still struggled with the concept. Participants could not identify collocations, and some identified phrases that were parts of sentences that did not contain any collocations. For example: *has a mixed economy*, *are exported to 90 countries*, and *find new places*. There are similar examples in Figure 5.1. It was thus necessary to adjust the initial teaching invention to increase exposure to and awareness of collocations.

The researcher introduced additional materials such as collocations dictionaries, online dictionaries, and corpus data to illustrate what collocations are and how they are used. Participants were instructed to consult an online dictionary and an online collocation dictionary to complete step two of the intervention exercise. This enabled participants to investigate which nouns are commonly used with particular verbs that they encountered in the exercises and provided an opportunity to be exposed to examples of sentences in which the collocations were used. The researcher also showed participants examples of verb-noun collocations found in concordances in COCA (Corpus of Contemporary American English, Davies, 2008-). This also enabled participants to be exposed to authentic language input that included examples of verb-noun collocations. These techniques were employed in order to facilitate the initial intervention exercise design and focussed on increasing awareness of verb-noun collocations.

At this stage, it might be useful to discuss errors made by participants in the tests overall. Participants made errors similar to those discussed section 5.2.1. Common errors included verb-tense errors and word choice errors. For example, several participants used the expression 'payed attention' instead of 'paid attention'. In this instance, the error was not accepted as correct because *payed* and *paid* have different meanings, and *payed* is not merely a spelling error. From the test results it was also evident that the word *face* proved difficult for participants. Many participants could not identify the nouns that should have been used with the word, and in the case where *face* was the missing item in a test, participants had difficulty with verb tenses and concord

At this point in the study it was not yet evident whether the intervention phase had a significant effect on the participants' controlled productive use of verb-noun collocations. In order to fully address this question, data from the tests had to be analysed. The next section therefore focusses on the results of the test data.

5.3.2 Discussion of the controlled productive test results

For this study, it was deemed important to examine the extent to which test results were significant and thus give an indication of collocations development because of intervention. The following question guided the analysis: Is there any statistically significant difference between the results of the pre-test, post-test and delayed post-test of the experimental group?

A productive version of the Vocabulary Levels Test (Laufer & Nation, 1999), discussed in detail in section 3.5.2.1 was used to assess participants' progress over a number of weeks. Below, the results of the pre-test, post-test 1, post-test 2, and delayed post-test (30 items) and the shorter tests between interventions (15 items) are discussed respectively.

5.3.2.1 Results for the pre-test, post-test and delayed post-test

Each of the tests consisted of 30 items, divided into two types of questions. The bulk of the test constituted a gap-fill exercise in which participants were required to fill in the missing noun based on the verb provided in a sentence. Section two of the test constituted a cloze-test design. The motivation for the selection of this type of test is discussed in detail in section 3.5.3, *Measuring productive collocational competence*. Participants were provided with a paragraph with missing items, and were given a multiple-choice question to select the most appropriate missing item to complete the passage. These tests are provided in Annexure D. Below are the results of these tests for both cohorts.

			95% Confidence Interval	
Group	Test	Mean	Lower Bound	Upper Bound
CG	Pre-test	**17.12**	16.584	17.662
	Post-test 1	**15.16**	14.568	15.765
	Post-test 2	**17.00**	16.314	17.696
	Delayed post-test	**19.46**	18.421	20.502
EG	Pre-test	**15.93**	15.189	16.670
	Post-test 1	**14.29**	13.594	14.992
	Post-test 2	**17.63**	17.015	18.257
	Delayed post-test	**23.24**	22.567	23.927

Table 5-3: Means of pre-test, post-test 1, post-test 2, and delayed post-test on verb-noun collocations

Table 5.3 above shows the results of the pre-test, post-test 1, post-test 2, and delayed post-test. The CG and the EG achieved a means of 17.12/30 and 15.93, respectively on the pre-test. Both groups have comparable means, which is also the case for post-test 1 with the CG and EG achieving means of 15.16 and 14.29 respectively. The means remain seemingly comparable on port-test 2, with a slight advantage for the EG, with a mean of 17.63, compared to CG's mean of 17.00. While this is a rather small difference, it does indicate a steady improvement over time, especially when comparing the size of the cohorts, and the overall performance of the EG at the end of the intervention. While this difference is rather small, it became larger on the delayed post-test in which the EG cohort scored a significantly higher mean (23.24 compared to 19.46).

This is more clearly visible in Figure 5.5 below. In the graph, the bold line indicates the results for the control group, while the light dotted line indicates the results for the experimental group.

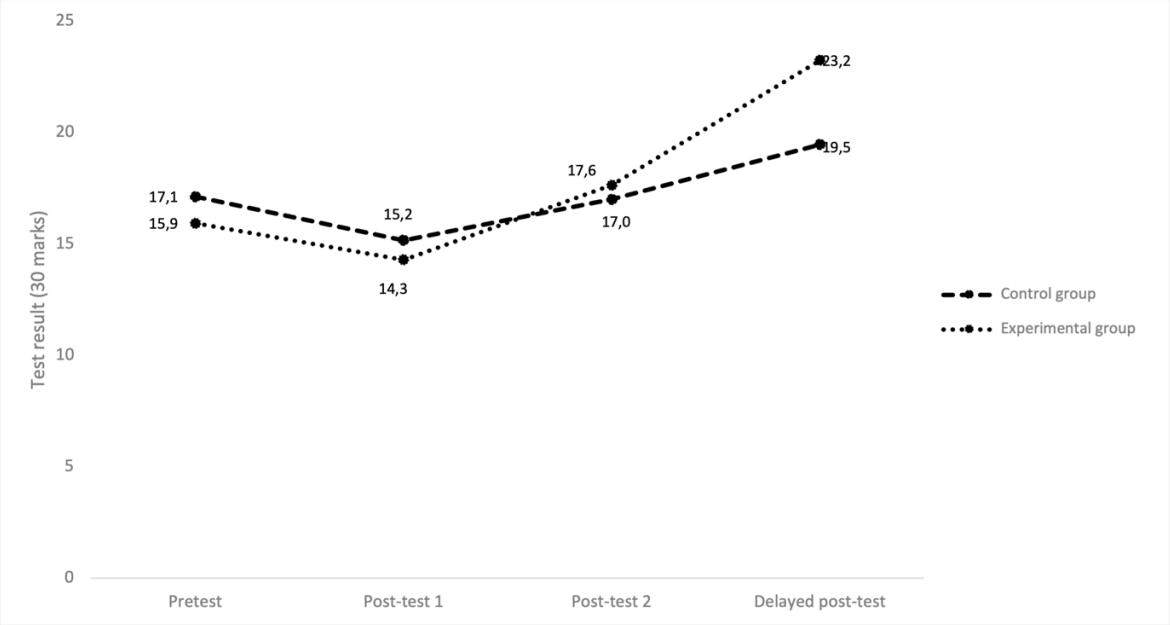


Figure 5-5: Pre-test, post-test 1, post-test 2, and delayed post-test results

The significance of these results was tested using SPSS (Bonferroni, 1936), and the results are presented in Table 5.4 below. These differences were statistically significant for the post-test at the 0.05 level, as reported in the column 'Mean Difference (I-J)'.

Test			Mean Difference (I-J)	Std. Error	df	**Sig**
1	CG	EG	**1.193***	0.467	1356.374	**0.011**
	EG	CG	**-1.193***	0.467	1356.374	**0.011**
2	CG	EG	**0.874**	0.469	1353.344	**0.063**
	EG	CG	**-0.874**	0.469	1353.344	**0.063**
3	CG	EG	**-0.631**	0.474	1356.012	**0.183**
	EG	CG	**0.631**	0.474	1356.012	**0.183**
4	CG	EG	**-3.786***	0.634	1376.485	**0.000**
	EG	CG	**3.786***	0.634	1376.485	**0.000**

Table 5-4: Mean difference with Bonferroni correction

Furthermore, when determining Cohen's value (1992) for the mean difference between the two cohorts, results showed that the difference was statistically significant and that the EG-cohort improved in their use of verb-noun collocations gradually.

Test	Effect size
Pre-test	0.26
Post-test 1	0.19
Post-test 2	0.14
Delayed post-test	0.82

Table 5-5: Cohen's (1992) effect size on the mean difference between CG and EG

From Table 5.5, it is evident that the effect size between the pre-test and the delayed post-test is significant. While the difference between the CG and EG effect size is small for the pre-test, with 0.26, that of 0.82 for the delayed post-test indicates a large difference, therefore making it statistically significant.

Mean scores were also compared within each cohort to track possible changes. Results are presented in Table 5.4. The means of each test were compared in a two-way calculation, and indicated that the mean difference in each case was consistent. In other words, the CG and EG results were compared in one calculation, and the inverse was compared in a separate calculation and the results were similar in each case. As presented in Figure 5.5, the means of the CG pre-test and delayed post-test are 17.1 and 19.5 respectively. However, the EG results of the same tests are 15.9 and 23.2. These mean scores along with the statistical analyses indicate that there was a significant improvement in the controlled productive knowledge of verb-noun collocations in this particular group.

The same measures were applied to the shorter intervention/intermittent tests (see section 4.3.1.3 for a detailed discussion) and yielded the results presented in Tables 5.6, 5.7 and 5.8. The tests are available in Annexure E. Table 5.6 indicates the mean scores of each test. The intermittent tests were named as follows: *ti*, *tii*, *tiii* and *tiv*.

Group	Test	Mean	95% Confidence Interval	
			Lower Bound	Upper Bound
CG	Ti	8.60	8.31	8.89
	Tii	9.27	8.98	9.57
	Tiii	6.57	6.25	6.89
	Tiv	7.59	7.28	7.90
EG	Ti	7.78	7.46	8.10
	Tii	9.13	8.80	9.47
	Tiii	6.08	5.76	6.39
	Tiv	7.46	7.11	7.81

Table 5-6: Means of ti, tii, tiii, and tiv on verb-noun collocations

The significance of these results was tested using SPSS (Bonferroni, 1936), and the results are presented in Table 5.7 below.

Test			Mean Difference (I-J)	Std. Error	df	**Sig**
ti	CG	EG	0.82	0.22	1629.592	**0.00**
	EG	CG	-0.82	0.22	1629.592	**0.00**
tii	CG	EG	0.14	0.22	1666.170	**0.53**
	EG	CG	-0.14	0.22	1666.170	**0.53**
tiii	CG	EG	0.49	0.23	1694.002	**0.03**
	EG	CG	-0.49	0.23	1694.002	**0.03**
tiv	CG	EG	0.13	0.23	1729.117	**0.57**
	EG	CG	-0.13	0.23	1729.117	**0.57**

Table 5-7: Mean difference with Bonferroni correction (ti, tii, tiii, tiv)

However, when applying statistical measures determining Cohen's value (1992) for the mean difference between the two cohorts, the results indicated a much lower effect size for *ti*, *tii*, *tii*, and *tiv*, as is evident in Table 5.8.

Test	**Effect size**
ti	0.33
tii	0.06
tiii	0.20
tiv	0.05

Table 5-8: Cohen's (1992) effect size on the mean difference between CG and EG (ti, tii, tiii, tiv)

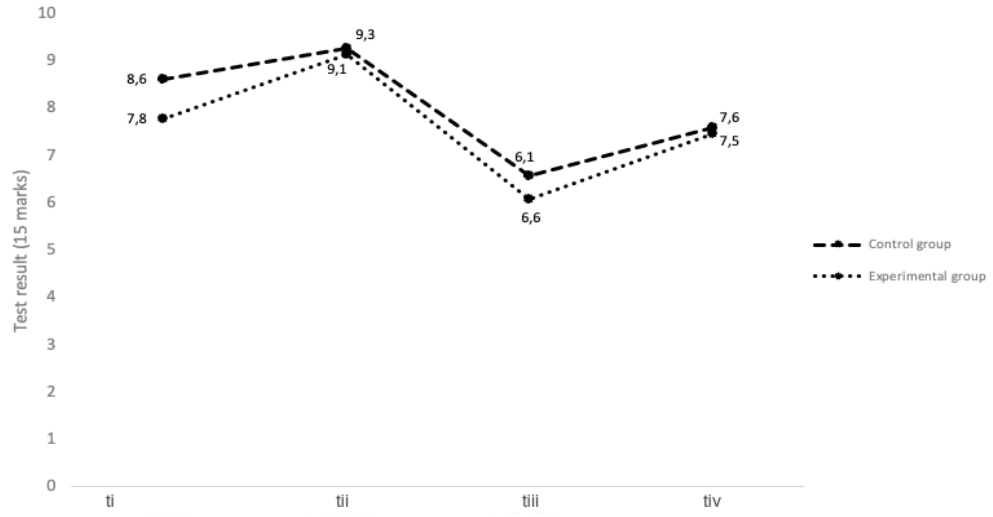


Figure 5-6: Results for ti, tii, tiii and tiv

From the results in the tables above and in Figure 5.6, it is evident that there is not a significant difference between the CG and EG performance in the intermittent tests. However, the EG performed poorer than the CG in both *ti* and *tiii*. A possible reason for this could be that the EG initially had difficulties grasping the concept of verb-noun collocations, and therefore had difficulties completing the tests accurately. In the following section common errors that occurred in the test results are discussed.

5.3 Conclusion

Based on the results discussed above, it can be concluded that: a) the pre-test results of the CG and the EG show that there is little difference between their controlled productive use of verb-noun collocations. While the results for post-test 1 and post-test 2 are not significant, the results for the delayed post-test show that the EG performed significantly better than the CG in retaining knowledge of verb-noun collocation after an intervention period of six weeks. From the results it is evident that the intervention exercises to raise awareness of the use of verb-noun collocations was successful.

Given the duration of intervention in this study —six weeks— and the results obtained, especially the statistically significant differences on post-test 2, it might be interesting to reflect on the question of the length of intervention needed to achieve sustained improvement in controlled productive verb-noun collocation knowledge. Addressing such a question requires

further research, but the absence of a significant difference between on post-test 1 compared to the statistically significant difference on post-test 2 is an indication that a longer intervention could potentially lead to bigger gains and possibly sustained improvement. While the present study does not answer this question, it lays a solid foundation for addressing it and recommends it for further research.

The researcher therefore suggests that the time of exposure to collocations during an intervention session is another important variable that should be carefully considered. This was an issue, and a serious limitation, in the present study.

One of the limitations experienced was a time limit during which interventions could take place. The ENLS 111-course is not designed to specifically teach collocations. The researcher had limited time available during allotted class time, and this proved to be challenging. Participants experienced difficulty with understanding the concept of collocations, and their reading speed affected the exercises in that the researcher had to spend more time than initially planned on the reading section of the intervention exercises. However, the results indicated than even limited explicit exposure to awareness-raising techniques improved controlled productive collocational knowledge and that that knowledge was retained well.

The results of this study can be used to inform and guide future classroom practices related to teaching collocations, as well as future research on the matter. This is discussed in more detail in Chapter 6.

As presented in the above sections, all the questions examined in this study have been answered. The results of the study indicated that participants taking a first-year ESL course had an average level of mastery of academic verb-noun collocations. The control group mastered verb-noun collocations at 57% in the pre-test, and 65% in the delayed post-test. However, the experimental group initially scored an average of 53% for the pre-test, compared to 77% for the delayed post-test. Throughout the study, their knowledge increased and issues such as informal vocabulary use seemed to improve with participants from the EG group showing a greater tendency to use formal vocabulary in their writing.

After results of the constructional analysis and the two-group quasi-experimental design with a pre-test, post-test and delayed post-test were considered, the researcher can confidently conclude that the study was successful, and that the explicit teaching of verb-noun collocations through the use of awareness-raising techniques is recommended to improve the controlled

productive knowledge of collocations among first-year university students. While this study is not exhaustive and has limitations, it is believed these results can be used to inform and guide future classroom practices related to teaching collocations, as well as future research on the matter. This is discussed in more detail in Chapter 6.

CHAPTER 6: CONCLUSION

6.1 Introduction

This study aimed to measure improvement in first-year ESL university students' productive knowledge of verb-noun collocations through integrated productive-oriented teaching strategies. To conclude this study, this chapter synthesises the main aspects discussed in the preceding chapters, and evaluates the outcome of the study, discusses its limitations, and makes recommendations for further research.

6.2 Measuring improvement in first-year ESL university students' productive knowledge of verb-noun collocations through integrated productive-oriented teaching strategies

It has been established that vocabulary is of particular importance when it comes to academic success and overall language proficiency in both a first (L1) and second language (L2) – see section 2.1. Research has also firmly established that vocabulary plays a crucial role in both receptive (reading and listening) and productive (speaking and writing) skills (section 2.1). However, vocabulary is complex because meaning is not necessarily confined to a single word, and lexical units can consist of two words or more. As discussed in section 2.1, these multi-word items often fall under the category of formulaic language and is considered central to communicative competence. Mastering language chunks lessens the cognitive load of processing an L2 and improves fluency as language is more readily available. This suggests the benefits of focussing on formulaic language in L2 teaching. Collocations are particular patterns in language and being aware of those patterns can reduce the learning burden needed to learn them.

However, as discussed in Chapter 2.1, formulaic language, especially collocations, is difficult to master. As a result, many scholars, including the present researcher, consider teaching collocations as an integral part of teaching vocabulary (Barfield & Gyllstad, 2009; Basal, 2019; Boers & Lindstromberg, 2012; Chen, 2011; Nesselhauf, 2003). Despite the importance of collocations, they are difficult to define and therefore often difficult to teach. The challenges identified in 3.4.1.5 and 3.4.1.6 motivated the present research.

Chapter 2 therefore focussed on defining vocabulary knowledge as well as defining collocations conceptually in terms of their significance for the present study. It also considered

the role of receptive and productive vocabulary knowledge. In this regard, the first conclusion that was drawn was that knowing a word entails placing a word within a broader context of society and knowledge of the world, human attitudes, beliefs and their social behaviour in a given language community (Elgort & Nation, 2011:91; see section 2.3.3). Sinclair (2004:173) notes that learners should be able to make connotational connections between lexical items, for example, the verb *cause* often occurs in a negative context, such as *cause an accident*.

The second conclusion drawn in regard to what it means to know a word, is that a single approach cannot be used to identify collocations due to their complex nature (see section 2.4.1). Three main approaches have been employed in recent decades: the frequency-based approach, the phraseological approach, and a combination of the two. All three of these approaches have value, but not one is comprehensive enough for the purpose of defining collocations for the present research. Consequently, the researcher concluded that a combined approach is best suited to identify collocations for the purpose of this study (see section 2.4.1).

The researcher conceptualises the notion of collocation as follows to inform the present research: in order to fully know a word, a learner must know that a word can function on various levels, depending on the context. A word may take on various forms such as having inflections or derivations, and these forms have limited functions and meanings. In terms of collocations, learners should be aware that the limitations on usage of words are not only determined by grammatical structure of language, but also by syntagmatic form-meaning connections. These connections are not automatically known but learnt as receptive comprehensible input becomes productive comprehensible output.

Since learners often lack the ability to fully comprehend language and effectively utilise it because of low levels of word knowledge, current research (see White, 1998; Williams & Evans, 1998; Locktman, 2002; Oliver & Mackey, 2003; Lyster & Mori, 2006; Lee, 2007; Ellis *et al.*, 2010; Nassaji & Fotos, 2010, Winke, 2013; Nassaji, 2015;) is concerned with the role of formal instruction to improve vocabulary knowledge. Chapter three therefore discussed various aspects of teaching and assessing vocabulary knowledge. The purpose of Chapter three was to determine the role that instruction plays in vocabulary learning and acquisition, as well as which teaching strategies could be employed to teach vocabulary – and specifically collocations - most effectively. Various vocabulary-teaching strategies were considered in 3.4 and results of several studies indicated that the most effective approach is an integrated approach in which vocabulary is taught explicitly (see 3.3). In terms of teaching collocations, a

review of literature indicated that the AWARE-approach was identified as the most effective (section 3.4.1.5).

Vocabulary assessment has also received much attention in recent decades, specifically assessing collocational knowledge. Various assessment methods were considered in 3.6. This review revealed that a productive version of Laufer and Nation's (1999) Vocabulary Levels Test is a reliable method to assess controlled productive vocabulary knowledge. Laufer and Nation (1999) used a sentence-based gap-filling test consisting of the target items on the VLT to measure relative size of receptive and productive vocabulary. The format of this test allows for the testing of collocations as well as general vocabulary knowledge, because it can be utilised to test deep-word knowledge and vocabulary size. In terms of collocations, the test can be adjusted in such a way that nodes and their collocates can be assessed by requiring participants to fill in either of the two components. Therefore, the format of this test was adopted for the pre-test, post-test and delayed post-test, as well as for all the intermittent tests, employed for the purpose of the present research.

Having defined vocabulary knowledge conceptually and identified an appropriate format for measuring this knowledge, the researcher operationalised measures to explore the controlled productive knowledge of verb-noun collocations of first-year students at the North-West University empirically.

As a first step, the initial productive collocational knowledge was determined by utilising collostructional analyses. The details of the analysis are discussed in section 4.3.1.1 to address research question 1 - *What is the level of controlled productive verb-noun collocation proficiency of first-year ESL students at the North-West University?* Results as reported in section 5.2.1 revealed which collocation items participants could use well, and which word combinations were challenging. The analyses also allowed the researcher to determine whether there was any difference in vocabulary use after teaching intervention took place, as per results reported in 5.2.3.

In order to answer research question two, awareness-raising teaching interventions followed, and were crucial in determining whether it was possible to achieve a statistically significant difference between the control group and the experimental group, based on the results of the pre-test, post-test 1, post-test 2, and the delayed post-test, as reported in 5.2.3. The specific teaching technique was therefore successful in improving collocational knowledge. However, a variety of approaches were used to raise awareness of collocations, and the researcher could

not only rely on a single method for in this study. See sections 4.3.1.2 and 5.2.2 for a detailed discussed on the methods used. The researcher can therefore conclude that employing a single awareness-raising method in a teaching intervention might not be the best practice, especially if participants do not have knowledge of the concept or importance of collocations.

The results of the study indicate that it is indeed possible to improve the controlled productive collocational knowledge of first-year students at the North-West University through integrated, explicit teaching of verb-noun collocations. These results were achieved by employing awareness-raising techniques that included exercises that required both receptive and productive skills (discussed in sections 3.4.1.5 and 4.3.1.2). These exercises included reading texts in an attempt to identify verb-noun collocations in each text, completing gap-fill exercises to ascertain whether participants could provide noun collocates for verbs encountered in the texts, and sentence-writing exercises to encourage productive use of those verb-noun collocations.

However, these results cannot be generalised beyond the context of this study. Limitations to the study and recommendations for further research are therefore discussed in section 6.3.

6.3 Limitations of the study and recommendations for further research

In this section several limitations are discussed, and recommendations are made for further research.

Firstly, these results cannot be generalised or applied to first-year students at every university in South Africa. The participants were limited to students at the Potchefstroom Campus of the North-West University. However, the study could be replicated in any ESL course taken by first-year university students in South Africa.

This resulted in very limited instruction opportunity, because the ENLS 111-course did not include collocations as part of the syllabus. In other words, the time of exposure to collocations during an intervention session is an important variable that should be carefully considered. This was an issue, and a serious limitation, in the present study because the ENLS 111-course is not designed to specifically teach collocations. The researcher had limited time available during allotted class time, and this proved to be challenging. Participants experienced difficulty with understanding the concept of collocations, and their reading speed affected the exercises in that the researcher had to spend more time than initially planned on the reading section of

the intervention exercises. It is the researcher's belief that the results could have been different should ideal conditions be met.

With regard to the testing phase, some limitations included the participation of participants. The tests were not compulsory, and participants could therefore opt not to complete them. This resulted in missing data items for some of the participants. It also resulted in participants not taking the tests seriously and rushing through the tests. Collocations and vocabulary development were not part of the official course, and this resulted in many participants not completing the tests as diligently as they completed other assessments. While this may not have affected the results of the study because of measures taken during the data analysis process —for example, allowing for missing data and to ensure intragroup and intergroup comparability (see section 4.3.1.3) — it might benefit the teaching of collocations if they were an integral component of the course. Replicating the current study in such a setting might yield different results.

Another suggestion for further research is therefore to replicate the study in a context in which collocations form an integral part of the course design, and compare the results to the current study. Furthermore, the current study adopted an integrated approach wherein different techniques were combined, which makes it impossible to determine the relative contribution of each of them. For the purpose of modelling the teaching of collocations, it might be interesting to conduct sub-studies that implement one technique at a time and therefore compare their contribution. This could help determine which technique that may be more efficient and the extent this is the case.

Based on the errors that were made by participants (discussed in section 5.2 and 5.3), further investigation into the word forming process of ESL learners might prove useful. Webb and Sasao (2013) developed an Academic Word Levels Test (AWLT) that could be useful in combination with the methods used in the present study. Webb and Sasao state that the test "provides a practical tool that can be used to measure progress in learning the AWL" (2013:268). Using assessment techniques that focus specifically on the development of vocabulary knowledge on different levels provide researchers with a clear view of specific problem areas in vocabulary knowledge. The AWLT focusses on the development of morphological knowledge and could provide insights into learner vocabulary that were identified with the collostructional analyses of the present study. This would better inform teaching techniques and which vocabulary items (or aspect of deep-word knowledge) researchers and teachers should pay special attention to.

In section 3.4.1.6, the researcher identified which existing word list would be most useful to inform the selecting of verb-noun collocations that are necessary to include in formal instruction. The Academic Collocation List (ACL, Ackermann & Chen, 2013) was selected. This list was selected because it was the most recent and most complete list at the time of this study.

However, the researcher identified possible drawbacks to using this list. Ackermann and Chen (2013:237) define collocation as "a single word that tends to co-occur in the span of ± 3 words from the reference word, co-occurring at least five times in total across at least five different texts with a Mutual Information score of at least 3 and a t-score of at least 2". The process followed to compile the list was not sufficient for the purposes of this study, as the collocates suggested for each word are too limited. The benchmark for optimal span is about four items to the left and right positions of the node (Sinclair, 1991). It is the opinion of this researcher that a more complete list should be compiled. There are several structures in language that could result in the node and its collocate being moved further apart. For example, the passive voice in English allows for the noun to be moved further away from the verb, and it will also change the position of the noun in relation to the verb, i.e., the noun will move from the right to the left of the verb. If a sentence contains an embedded clause that separates the node and the collocate, the span will increase as well. It might therefore be necessary to widen the span when compiling a list, to ensure more accurate and comprehensive data. Furthermore, the ACL was limited in the number of noun collocates provided with each of the verbs. Searches in other corpora such as COCA Academic (Davies, 2008-) yielded more combinations with the same verbs. In order to accurately identify which items are most useful and relevant to teach based on the frequency of those items, several corpora need to be taken into account to compile a new list that is more comprehensive.

6.4 Discussion of results

After a brief overview of the results from this study and related limitations and suggestions for further studies presented above, this section offers a discussion of the results especially in relation to previous research.

With the first research question, this study has demonstrated that common verb-noun errors made by ESL learners can be identified through a collostructional analysis. The analysis also provided insight into which combinations participants had already mastered before

intervention. Moreover, the analysis proved useful to measure improvement of the use of collocations over a period of time.

With the second research question about the extent to which intervention leads to measurable gains in controlled productive knowledge of collocations, this study shows that controlled productive knowledge of collocations improved among ESL students at North-West University following intervention. This result supports previous findings, for example Barfield (2009) and Webb and Kagimoto (2009, 2011) in a Japanese context, Laufer and Girsai (2009) in an Israeli context, Nizonkiza (2017) in a Burundian context, Ying and O'Neill (2009) in a Chinese context, that reached similar conclusions. This study extends these previous findings to a South African context, which is its merit. It can therefore be concluded that awareness-raising teaching techniques are efficient and should be recommended for teaching collocations.

With the third research question, the researcher was able to quantify the overall improvement of the experimental group. From the results it was evident that the teaching interventions employed were successful, and that measurable improvement was indeed possible.

6.5 Conclusion

This study aimed to measure whether it would be possible to improve the controlled productive knowledge of verb-noun collocations of first-year ESL university students through the use of integrated productive-oriented teaching strategies. The study contributes to current research on collocations in that it addressed a well-researched area of vocabulary knowledge, and sought to provide a solution to a known problem, i.e. that ESL and EFL learners have difficulty learning and using collocations, specifically verb-noun collocations. Through collecting data in the form of essays and test scores, and through utilising explicit awareness-raising teaching methods, the following conclusions can be reached.

Test scores and statistical analyses indicate that first-year university students' controlled productive knowledge of verb-noun collocations could be improved through explicit awareness-raising techniques. This further establishes that explicit teaching techniques are beneficial for vocabulary acquisition, and are specifically beneficial when teaching verb-noun collocations. The results of the present study therefore contribute to the field in solidifying the foundation towards modelling the teaching of collocations in order to achieve sustained collocational knowledge. The study further contributes to the question of *which* collocations to teach (section

3.4.1.6) in that the researcher was able to identify possible areas in which participants were lacking in collocational knowledge through the collostructional analyses (see section 5.2.1).

Therefore, the techniques used in this study could be used for further research and can be recommended for developing teaching tools in university courses to improve mastery of collocations.

BIBLIOGRAPHY

Abel, B. 2003. English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, 19, 329–358.

Ab Manan, N.A., Jaganathan, P. & Pandian, A. 2014. The benefits of formula instruction in enhancing EAP learners' academic writing performance. *Journal of English Language and Literature*, 2(1):141-153. http://dx.doi.org/10.1016/j.jeap.2013.08.002 Date accessed: 24 September 2019.

Ackermann, K. & Chen, Y. 2013. Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes,* 12(4): 235-247.

Akbarian, I. 2010. The relationship between vocabulary size and depth for ESP/EAP learners. *System,* 38: 391-401.

Akpinar, K.D. & Bardakçi, M. 2016. Retention of collocations: the effect of presenting and organising collocations. *The Buckingham Journal of Language and Linguistics*, 9:129-151

Alanen, R. 1995. Input enhancement and rule presentation in second language acquisition. (*In* Schmidt, R., *ed.*, Attention and awareness in foreign language acquisition. p. 259–302 University of Hawaii, Honolulu).

Allen, V. F. 1983. *Techniques in Teaching Vocabulary*. New York: Oxford University Press

Alamri, K. & Rogers, V. 2018. The effectiveness of different explicit vocabulary-teaching strategies on learners' retention of technical and academic words. *The Language Learning Journal*, 46(5), 622-633. DOI: 10.1080/09571736.2018.1503139 Date of access: 19 Sep. 2020.

Al-Murthada, M. 2014. The relationship between productive vocabulary tasks and accuracy, fluency, and motivation of Japanese EFL learners. *VERB*, 3(1): 2-4.

Anderson, R. C., & Freebody, P. 1981. Vocabulary knowledge. (*In* Guthrie J.T., *ed*. Comprehension and Teaching: Research Reviews. p. 77-117. Newark: International Reading Association).

Arevart, S., & Nation, I. S. P. 1991. Fluency improvement in a second language. *RELC Journal, 22*(1): 84-94.

Bachman, L. F. and Palmer, A. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.

Bahrick, H.P., & Bahrick, P.O. 1964. A re-examination of the interrelations among measures of retention. *Quarterly Journal of Experimental Psychology*, 16(4): 318-324.

Barfield, A. 2009. Following individual L2 collocation development over time. (*In*: Barfield A. & Gyllstad, H., *eds.*, Researching collocations in another language. p. 208-223. New York: Palgrave Macmillan).

Barfield, A. & Gyllstad, H, *eds.* 2009. *Researching collocations in another language*. New York: Palgrave Macmillan.

Bauer, L., & Nation, I. S. P. 1993. Word families. *International Journal of Lexicography,* 6(4): 253-279.

Beglar, D., & Hunt, A. 1999. Revising and validating the 2000 word level and the university word level vocabulary tests. *Language Testing, 16*(2):131-162.

Benson, M., Benson, E. and Ilson, R. 1997. *The BBI Dictionary of English Word Combinations*. Amsterdam: John Benjamins.

Benati, A. 2005. The effects of processing instruction, traditional instruction and meaning-output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9(1): 67–93.

Biemiller, A., & Boote, C. 2006. An effective method for building meaning vocabulary in the primary grades. *Journal of Educational Psychology,* 98(1): 44-62.

Biskup, D. 1992. L1 influence on learners' renderings of English collocations: a Polish/German empirical study. (*In* Arnaud, P. J. L.& Bejoint, H., *eds.*, Vocabulary and Applied Linguistics. p. 85-93. London: Macmillan.)

Boers, F. 2020. Factors affecting the learning of multiword items. (*In* Webb, S. *ed.*, The Routledge Handbook of Vocabulary Studies. p. 134-157. New York: Routledge).

Boers, F., Eyckmans, J., Kappel, J., Stengers, S., & Demecheleer, M. 2006. Formulaic sequences and perceived oral proficiency: Putting the Lexical Approach to the test. *Language Teaching Research*, 10(3):245-261.

Boers, F., Demecheleer, M., Coxhead, A. & Webb, S. 2014. Gauging the effects of exercise on verb–noun collocations. *Language Teaching Research*, 18(1)*:*54-74.

Boers, F., & Lindstromberg, S. 2012. Experimental and intervention studies on formulaic sequences in a second language. Annual Review of Applied Linguistics, 32(3): 83-110.

Bogaards, P. 2000. Testing L2 vocabulary knowledge at a high level: the case of the Euralex French Tests. *Applied Linguistics, 21*(4), 490-516.

Carroll, S., Roberge, Y. & Swain, M. 1992. The role of feedback in adult second language acquisition: Error correction and morphological generalization. *Applied Psycholinguistics* 13(2): 173–189.

Carroll, S. & Swain, M. 1993. Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition* 15(3): 357–386.

Chan, T., & Liou, H. 2005. Effects of Web-based Concordancing Instruction on EFL Students' Learning of Verb-Noun Collocations. *Computer Assisted Language Learning*, 18(3): 231-250.

Cobb, T. 2000. One size fits all? Francophone learners and English vocabulary tests. *Canadian Modern Language Review,* 57(2): 295-324.

Cohen, J. 1988. Statistical power analysis for the behavioural sciences. 2nd ed. Hillsdale: Lawrence Erlbaum.

Corson, D. J. 1995. Using English Words. Dordrecht: Kluwer Academic Publishers.

Cowie, A.P., *ed*. 1998. *Phraseology. Theory, Analysis, and Applications.* Oxford: Clarendon Press.

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.

Crowther, J., Dignen, S., & Lea, D., *eds.* 2002. *Oxford collocations dictionary students of English*. Oxford, England: Oxford University Press.

Dai, Z., & Ding, Y. 2010. Effectiveness of text memorization in EFL learning of Chinese students. (*In* Wood, D. *ed*., Perspectives on formulaic language: Acquisition and communication. p. 71–87. New York, NY: Continuum.)

Dale, E. 1965. Vocabulary measurement: techniques and major findings. *Elementary English,* 42(8):895-901.

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

Davies, M. 2008–. The Corpus of Contemporary American English (COCA). https://www.english-corpora.org/coca/

de Bot, K. & Kroll, J.F. 2010. Psycholinguistics. (*In* Schmitt, N., *ed*. An Introduction to Applied Linguistics. 2nd ed. p. 124-142. London: Hodder Education).

de Graaf, R. 1997. The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition* 19(2):249–276

Doughty, C. & Williams, J. 1998. *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press.

Durán, P., Malvern, D., Richards, B., & Chipere, N. 2004. Developmental trends in lexical diversity. *Applied Linguistics,* 25(2): 220-242.

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.

Elgort, I. 2011. Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, 61(2): 367-413.

Elgort, I. 2013. Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing,* 30(2): 253-272.

Elgort, I., & Nation, P. 2010. Vocabulary learning in a second language. (*In* Seedhouse, P., Walsh, S. & Jenks C., *eds*. Conceptualising 'Learning' in Applied Linguistics. p. 89–104. Basingstoke: Palgrave Macmillan.)

Ellis, N. C., & Beaton, A. 1993. Factors affecting foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology,* 46A, 3: 533-558.

Erman, B., & Warren, B. 2000. The idiom principle and the open choice principle. *Text,* 20(1):87-120.

Engber, C. A. 1995. The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing,* 4(2): 139-155.

Farvardin, M. T. 2019. Effects of spacing techniques on EFL learners' recognition and production of lexical collocations. *Indonesian Journal of Applied Linguistics*, 9(2):395-403. doi: 10.17509/ijal.v9i2.20237

Felix, S. 1981. The effect of formal instruction on second language acquisition. *Language Learning* 31(1):87–112.

File, K., A. & Adams, R. 2010. Should vocabulary instruction be integrated or isolated? *TESOL Quarterly* 44(2):222–249

Firth, J.R. 1957. *Papers in Linguistics 1934–1951*. London: Oxford University Press.

Fordyce, K. 2014. The differential effects of explicit and implicit instruction on EFL learners' use of epistemic stance. *Applied Linguistics* 35(1):6–28.

Forsberg, F. 2010. Using conventional sequences in L2 French. *IRAL*, 48(1):25-51.

Gairns, R. & Redman, S. 1986. *Working with Words: A Guide to Teaching and Learning Vocabulary.* Cambridge: Cambridge University Press.

Gardner, D. 2013. *Exploring Vocabulary: Language in Action*. Oxford: Routledge.

Goulden, R., Nation, P., & Read, J. 1990. How large can a receptive vocabulary be? *Applied Linguistics,* 11(4): 341-363.

Granger, S. & Meunier, F., *eds*. 2008. *Phraseology. An interdisciplinary perspective*. Amsterdam: John Benjamins.

Granger, S., & Paquot, M. 2008. Disentangling the phraseological web. (*In* Granger, S. & Meunier, F., *eds*. 2008. *Phraseology. An interdisciplinary perspective*. Amsterdam: John Benjamins).

Gries, S.Th. & Stefanowitch, A. 2004a. Extending collostructional analysis: A corpus-based perspective on 'alterations'. *International Journal of Corpus Linguistics*, 9(1): 97-129.

Gries, S.Th. & Stefanowitch, A. 2004b. (*In* Achard & Kemmer, *eds*. Language, Culture and Mind. Stanford CSLI Publications. P. 225-236.)

Gilquin, G. 2010. *Corpus, cognition and causative construction.* Amsterdam: John Benjamins.

Gyllstad, H. 2007. Testing English collocations: Developing receptive tests for use with advanced Swedish learners. Lund: 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. p 153–170. New York: Palgrave Macmillan).

Gyllstad, H., & Schmitt, N. 2018. Testing formulaic language. (*In* Siyanova-Chanturia, A. & Pellicer-Sánchez, A., *eds*. Understanding Formulaic Language. p. 124-142. London: Hodder Education).

Gyllstad, H., Vilkaité, L., & Schmitt, N. 2015. Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL International Review of Applied Linguistics*, 166(2): 278-306.

Halliday, M.A.K. 1966. Lexis as a Linguistic Level. (*In* Bazell, C.E., J.C. Catford, J.C., M.A.K. Halliday, M.A.K. & Robins, R.H., *eds*. In Memory of J.R. Firth. .p. 148–162. London: Longman.)

Hancock, G.R., & Mueller, R.O. 2010. The reviewer's guide to quantitative methods in the social sciences. New York: Routledge.

Hatami, S. 2012. Review of Researching and Analyzing Vocabulary. *TESOL Quarterly,* 46(4): 868-869.

Henriksen, B., & Stæhr, S.,L. 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. p. 224-231. New York: Palgrave Macmillan.)

Henriksen, B. 2013. Research on L2 Learners' Collocational Competence and Development — A Progress Report. (*In* Bardel, C., Lindqvist, C. & Laufer, B., *eds.*, L2 Vocabulary Acquisition, Knowledge and Use. New Perspectives on Assessment and Corpus Analysis. Eurosla Monographs Series 2. p. 29-56. European Second Language Association.)

Hirsh, D., & Nation, P. 1992. What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2): 689-696.

Howatt, A. P. R. 1984. A History of English Language Teaching. Oxford University Press, Oxford.

Horiba, Y. 2012. Word Knowledge and Its Relation to Text Comprehension: A Comparative Study of Chinese- and Korean-Speaking L2 Learners and L1 Speakers of Japanese. *The Modern Language Journal*. 96(1):108-12.

Horst, M. 2005. Learning L2 vocabulary through extensive reading: a measurement study. *Canadian Modern Language Review, 61*(3): 355-382.

Hsu, J.Y., & Chiu, C. 2008. Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL Journal*, 10(1):181–204.

Hu, M., & Nation, I.S.P. 2000. Vocabulary density and reading comprehension. *Reading in a Foreign Language,* 13(1): 403-430.

Hulstijn, J. 2001. Intentional and incidental second-language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. (*In* Robinson, P., *ed*. Cognition and Second Language Instruction. p. 258-286. Cambridge: Cambridge University Press).

Hulstijn, J. 2002. Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research,* 18(3): 193-223.

Jarvis, S. 2002. Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing,* 19(1): 57-84.

Irujo, S. 1986. A piece of cake: learning and teaching idioms. *ELT Journal,* 40(3): 236-242.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. 2008. Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics,* 29(1): 24-49.

Keshavarz M.H., & Salimi H. 2007. Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics* 17(1): 81–92.

Kecskés, I. 2000. A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics*, 32 (5):605-625. DOI: 10.1016/S0378-2166(99)00063-6

Krashen, S. 1981. *Second language acquisition and second language learning*. Oxford: Oxford University Press.

Krashen, S. 1982. *Principles and practice in second language acquisition*. New York: Pergamon Press.

Krashen, S. 1985. *The input hypothesis: Issues and implications*. Oxford: Pergamon Press.

Kucera, H., & Francis, W. N. 1967. *A computational analysis of present-day American English*. Providence, R.I.: Brown University Press.

Kuhn, M.R., & Stahl, S.A. 2003. Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1): 3-21.

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

Laufer, B. & Girsai, N. 2008. Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics* 29(4):694–716.

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

Laufer, B. & Nation, P. 1999. A vocabulary-size test of controlled productive ability. *Language Testing,* 16(1): 33-51.

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

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

Lee, S.H. & Muncie, J. 2006. From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly,* 40(2): 295-320.

Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Massachusetts: MIT Press.

Lewis, M. 1993. *The Lexical Approach*. Hove, UK: Language Teaching Publications.

Lesniewska, J. 2006. Collocations and second language use. Instytut Filologii Angielskiej. Studia Linguistica Universitatis Iagellonicae Cracoviensis, 123.

Lightbown, P. & Spada, N. 2008. Form-focused instruction: Isolated or integrated? *TESOL Quarterly* 42, 181-207.

Little, D. 1997. Language Awareness and the Autonomous Language Learner. *Language Awareness* 6(2 and 3): 93–104.

Lochtman, K. 2002. Oral corrective feedback in the foreign language classroom: How it affects interaction in analytic foreign language teaching. *International Journal of Educational Research* 37 (3): 271–283. DOI: 10.1016/S0883-0355(03)00005-3

Loewen, S. 2015. *Introduction to instructed second language acquisition*. New York: Routledge

Long, M. 1991. Focus on form: A design feature in language teaching methodology. (In de Bot, K., Ginsberge, R & Kramsch, C., *eds*. Foreign language research in cross-cultural perspective. p. 39–52. Amsterdam: John Benjamins).

Long, M. & Robinson, P. 1998. Focus on form: Theory, research and practice. (*In* Doughty, C & Williams, J., *eds*. Focus on form in classroom second language acquisition. p. 15–41. Cambridge: Cambridge University Press).

Lyster, R. & Mori, H. 2006. Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition* 28(2): 269–300.

Martinez, R., & Schmitt, N. 2012. A phrasal expressions list. *Applied Linguistics* 33(3): 299-320. doi:10.1093/applin/ams010 Date accessed: 13 April 2018.

McCarthy, M. 1990. Minding your words: two important areas of vocabulary learning. *English Studies (British Council)*, 5, 6-7.

McCarthy, M. and F. O'Dell. 2005. *English collocations in use*. Cambridge: Cambridge University Press.

McNamara, T. F. 1996. *Measuring second language performance.* London, UK: Longman.

Meara, P. 1990a. A note on passive vocabulary. *Second Language Research,* 6(2): 150-154.

Meara, P. 1990b. Some notes on the Eurocentres vocabulary tests. (*In* Tommola, J., *ed*. Foreign Language Comprehension and Production. Turku: AFinLa Yearbook. p. 103-113).

Meara, P. 1996. The Dimensions of Lexical Competence. (*In* Brown, G., Malmkjaer, K. & Williams, J., *eds.* Competence and Performance in Language Learning. p. 35-53. Cambridge: Cambridge University Press).

Meara, P. & Buxton, B. 1987. An alternative to multiple choice vocabulary tests. *Language Testing,* 4(2): 142-154.

Meara, P. & Wolter, B. 2004. VLinks: Beyond vocabulary depth. *Angles on the English-speaking World,* 4: 85-96.

Mezynski, K. 1983. Issues concerning the acquisition of knowledge: effects of vocabulary training on reading comprehension . *Review of Educational Research,* 53(2): 253-279.

Millet, S. 2017. Speed readings for ESL learners 3000 BNC. *English Language Institute Occasional Publication No. 26*.

Milton, J., & Treffers-Daller, J. 2013. Vocabulary size revisited: the link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4(1):151-172.

Nagy, W. E., Herman, P., & Anderson, R. C. 1985. Learning words from context. *Reading Research Quarterly, 20*(2): 233-253.

Nassaji, H. 1999. Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *Canadian Modern Language Review* 55 (3):385–402.

Nassaji, H. 2015. *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. London: Bloomsbury.

Nassaji, H. 2016. Research Timeline: Form-focused instruction and second language acquisition. *Language Teaching,* 49(1): 35–62. doi:10.1017/S0261444815000403

Nassaji, H. & Fotos, S. 2010. *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context.* New York: Routledge, Taylor & Francis Group.

Nation, I.S.P. 1983. Testing and teaching vocabulary. *Guidelines,* 5(1): 12-25.

Nation, I.S.P. 1990. Teaching and Learning Vocabulary. Rowley, Mass.: Newbury House.

Nation, I.S.P. 1991. *Language Teaching Techniques.* Wellington: English Language Institute, Victoria University of Wellington.

Nation, I.S.P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. 2002. Best practice in vocabulary teaching and learning. (*In* Richards, J.C. & Renandya. W.A., *eds.* Methodology in Language Teaching. p. 267-272. Cambridge: Cambridge University Press).

Nation, I.S.P. 2007. The four strands. *Innovation in Language Learning and Teaching,* 1(1): 1-12.

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

Nation, P. 1990. Teaching and learning vocabulary. Rowley, MA: Newbury House.

Nattinger, J.R. & DeCarrico, J.. 1989. Lexical phrases, speech acts and teaching conversation. *AILA review*, *6*, pp.118-139.

Nattinger, J., & DeCarrico, J. 1992. *Lexical phrases and language teaching*. Oxford: OUP.

Nesselhauf, N. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24(2): 223-242. doi:10.1093/applin/24.2.223

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

Nguyen, L. T. C., & Nation, I.S.P. 2011. A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal,* 42(1): 86-99.

Nizonkiza, D. 2011. The relationship between lexical competence, collocational competence, and second language proficiency. *English Text Construction*, 4(1): 113–146.

Nizonkiza, D. 2012. Quantifying controlled productive knowledge of collocations across proficiency and word frequency levels. *Studies in Second Language Learning and Teaching*, 2(1): 67-92.

Nizonkiza, D. 2016. First-year university students' receptive and productive use of academic vocabulary. *Stellenbosch Papers in Linguistics*, 45(1):169-187.

Nizonkiza, D. & Van den Berg, K. 2014. Dimensional Approach to Vocabulary Testing: What Can We Learn from Past and Present Practices? *Stellenbosch Papers in Linguistics,* 43(1): 45-61.

Nizonkiza, D. & Ngwenya, T. 2015. Challenges of testing deep word knowledge of vocabulary: Which path to follow? *Journal for Language Teaching*, 49(1):223-253. http://dx.doi.org/10.4314/jlt.v49i1.9 Date accessed: 12 Sep. 2016.

Nizonkiza, D. & Van de Poel, K. 2014. Teachability of collocations: The role of word frequency counts. *Southern African Linguistics and Applied Language Studies* 32(3):301-316. doi:10.2989/16073614.2014.997061 Date accessed: 28 Oct. 2019.

Nizonkiza, D., Van Dyk, T.J. & Louw, H. 2013. First-year university students' productive knowledge of collocations. *Stellenbosch Papers in Linguistics Plus* 42(1):165-181. doi:10.5842/42-0-143. Date accessed: 18 Apr. 2016.

Nizonkiza, D. & Van Dyk, T.J. 2015. Academic literacy of South African higher education level students: Does vocabulary size matter? *Stellenbosch Papers in Linguistics*, 44(1):147-174.

Norris, J. M. & Ortega. L. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50(3):417–528.

Oliver, R. & Mackey, A. 2003. Interactional context and feedback in child ESL classrooms. *The Modern Language Journal,* 87(4):519–533.

Palmer, H. E. 1921. *The Principles of Language Study*. London: George G. Harrap & Co.

Pawley, A., & Syder, F. 1983. Two puzzles for linguistic theory: Nativelike selection and native-like fluency. (*In* Richards, J. & Schmidt, R., *eds*. Language and communication. p. 191– 226. London: Longman).

Pedersen, T. 1996. *Fishing for exactness*. Proceedings of the SCSUG 96 in Austin, TX, 188-200.

Peters, E. 2009. Learning collocations through attention-drawing techniques: a qualitative and quantitative analysis. (*In* Barfield, A. & Gyllstad, H., *eds.* Researching collocations in another language: Multiple interpretations. p. 194–207. New York: Palgrave Macmillan).

Prabhu, N. S. 1987. *Second Language Pedagogy*. Oxford: Oxford University Press.

Pawley, A., & Syder, F.H. 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. (*In* Richards, J.C. & Schmidt, R. W., *eds*. Language and Communication. p. 191-225. London: Longman.)

Read, J. 1988. Measuring the vocabulary knowledge of second language learners. *RELC Journal,* 19(2), 12-25.

Read, J. 1993. The development of a new measure of L2 vocabulary knowledge. *Language Testing,* 10(3): 355-371.

Read, J. 1995. Refining the word associates format as a measure of depth of vocabulary knowledge. *New Zealand Studies in Applied Linguistics,* 1: 1-17.

Read, J. 2007. Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies, 7*(2): 105-125.

Read, J. 2013. Research timeline: Second language vocabulary assessment. *Language Teaching,* 46(1): 41-52.

Read, J. A. S., & Nation, I.S.P. 2006. An investigation of the lexical dimension of the IELTS speaking test. *IELTS Research reports,* 6, 207-231

Revier, R.L. 2009. Evaluating a new test of whole English collocations. (*In* Barfield, A., & Gyllstad, H., *eds.* Researching collocations in another language: Multiple interpretations. p. 125-138. New York: Palgrave Macmillan.

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

Robinson, P. 1996. Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition* 18(1):27–67.

Sasao, Y. & Webb, S. 2017. The word part levels test. *Language Teaching Research,* 21(1): 12-30. DOI:10.1177/136216881558608 Date accessed: 10 Sep. 2019.

Scheepers, R.A. 2006. The effects of immersion on Grade 7 learners' vocabulary size: is incidental learning of vocabulary enough? Journal for Language Teaching, 40(2):1-20.

Scheepers, R. 2016. The Importance of Vocabulary at Tertiary Level. *Journal for Language Teaching*, 50(1): 53-77.

Scheepers, R. 2017. South African students' use of delexical multiword units: The trouble with high frequency verbs. *Stellenbosch Papers in Linguistics*, 47:89-114. doi:10.5842/47-0-263 Date of access: 25 Feb. 2019.

Scheepers, R. 2018. Probing the depths: Can both size and depth of vocabulary predict academic performance? *Per Linguam,* 34(2):52-68. http://dx.doi.org/10.5785/34-2-812 Date of access: 20 Sep. 2020.

Schmitt, N. 2008. *Teaching vocabulary*. Pearson Education handout.

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: Palgrave Macmillan.

Schmidt, R.W. & Frota, S. 1986. Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. (*In* Day, R., *ed.* Talking to Learn: Conversation in Second Language Acquisition. p. 237-326. Rowley, Mass.: Newbury House).

Schwartz, B. 1993. On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition* 15(2):147–163.

Scott, M. 2011. *WordSmith Tools version 6*. Liverpool: Lexical Analysis Software.

Scott, V. 1989. An empirical study of explicit and implicit teaching strategies in French. *The Modern Language Journal* 73(1): 14–22.

Seashore, R.H. & Eckerson, L.D. 1940. The measurement of individual differences in general English vocabularies. *Journal of Educational Psychology,* 31(1): 14-38.

Simpson-Vlach, R. and N.C. Ellis. 2010. An academic formulas list: New methods in phraseology research. *Applied Linguistics* 31(4): 487-512. doi:10.1093/applin/amp058 Date of access: 7 Apr. 2020.

Sinclair, J. M. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Siyanova-Chanturia, A. 2017. Researching the teaching and learning of multi-word expressions. *Language Teaching Research*, 21(3): 289-297. doi:10.1177/1362168817706842 Date of access: 20 Sep. 2019.

Spada, N. 1997. Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching* 30: 73–87.

Spada, N. & Lightbown, P. 2008. Form-focused instruction: Isolated or integrated? *TESOL Quarterly* 42(2): 181–207.

Spada, N. & Tomita, Y. 2010. Interaction between type of instruction and type of language feature: A meta-analysis. *Language Learning* 60(2): 263–308.

Staehr, L. S. 2009. Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4): 577– 607.

Stefanowitch, A., & Gries, S.Th. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics,* 8(2): 209-243.

Stengers, H., Boers, F., Housen, A., & Eyckmans, J. 2011. Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *IRAL*, *49(4)*: 321-343. DOI: 10.1515/iral.2011.017

Thorndike, E. L., & Lorge, I. 1944. *The Teacher's Word Book of 30,000 Words*. New York: Teachers College Columbia University.

Teng, F. 2019. The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers and Education*, 142. doi:10.1016/j.compedu.2019.103655 Date of access: 18 Mar. 2021

Valeo, A. 2013. The integration of language and content: Form-focused instruction in a content-based language program. *The Canadian Journal of Applied Linguistics* 16(1): 25–50.

Van Dyk, T., Louw, H., Nizonkiza, D. & Van de Poel, K. 2016. Productive knowledge of collocations may predict academic literacy. *Per Linguam*, 32(2): 66-81.

VanPatten, B. & Cadierno, T. 1993. Explicit instruction and input processing. *Studies in Second Language Acquisition* 15(2)*:* 225–244.

VanPatten, B. & Oikkenon, S. 1996. Explanation vs. structured input in processing instruction. *Studies in Second Language Acquisition* 18(4): 495–510.

Warren, B. 2005. A model of idiomaticity. *Nordic Journal of English Studies,* 4(1): 35-54.

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

Webb, S. & Kagimoto, E. 2009. The Effects of Vocabulary Learning on Collocation and Meaning. *TESOL Quarterly*, 43(1): 55-77.

Webb, S. & Sasao, Y. 2013. New Directions in Vocabulary Testing. *RELC*, 43(3): 263-277.

Wei, Y. 1999. *Teaching Collocations for Productive Vocabulary Development*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, New York, NY, 9-13 March. http://files.eric.ed.gov/fulltext/ED457690.pdf. Date of access: 13 Sep. 2016.

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

White, J. 1998. Getting the learners' attention: A typographical input enhancement study. (*In* Doughty, C. & Williams, J., *eds*. Focus on form in classroom second language acquisition. p. 85–113. Cambridge: Cambridge University Press).

Williams, J. & Evans, J. 1998. What kind of focus and on which forms? (*In* Doughty, C. & Williams, J., *eds*. Focus on form in classroom second language acquisition. p. 139–155. Cambridge: Cambridge University Press).

Winke, P. 2013. The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition,* 35(2): 323–352.

Woltman, H., Feldstein, A.J., MacKay, J.C. & Rocchi, M. 2012. An introduction to hierarchical linear modelling. *Tutorial in Quantitative Methods for Psychology,* 8(1): 52-69. doi: 10.20982/tqmp.08.1.p 052

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

Xue, G., & Nation, I. S. P. 1984. A university word list. *Language Learning and Communication, 3*(2), 215-229.

Ying Y. & O'Neill M. 2009. Collocation learning through an 'AWARE' approach: Learner perspectives and learning process. (*In* Barfield A. & Gyllstad, H., *eds*. Researching Collocations in Another Language: Multiple Interpretations. p. 181–193. New York: Palgrave Macmillan).

Zahar, R., Cobb, T., & Spada, N. 2001. Acquiring vocabulary through reading: effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(3): 541-572.

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

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

Zhou, A. A. 2009. What adult ESL learners say about improving grammar and vocabulary in their writing for academic purposes. *Language Awareness, 18*(1): 31-46.

ANNEXURE A: INFORMED CONSENT

Dr K Van den Berg
Supervisor
Faculty of Arts
English Department
North-West University
018 299 1554

Ms M Morris
Researcher
Faculty of Arts
English Department
North-West University
018 299 1490

INFORMED CONSENT

You are invited to take part in a research project conducted by **M Morris** in the Faculty of Arts: English Department at the North-West University. The results of this research will be published in the form of a Master's thesis. It is important that you read and understand the following general principles, which are applicable to all participants in our research projects. I trust that you will be able to take part in this research, and I wish to thank you in advance.

SECTION A: GENERAL PROJECT INFORMATION

1. TITLE OF THE PROJECT:

Measuring improvement in first-year ESL university students' productive knowledge of verb-noun collocations through integrated productive oriented teaching strategies.

2. INSTITUTION/SUBJECT GROUP:

North-West University, English Department

3. **NAMES AND CONTACT DETAILS OF PROJECT STAFF:**

You are most welcome to contact the Supervisor of this project for enquiries, help and complaints related to the project or your participation in the project. If you need any help, feel uncertain or have any questions regarding the project, or if you experience any unwanted effects of the project interventions, feel victimised or have any other complaints related to the project, or wish to terminate your participation to the project, you may contact these persons at any time. You are welcome to refer general enquiries to the Researcher.

	Supervisor	Researcher
Title, Initials and Surname:	Dr K Van den Berg	Ms M Morris
Contact details:	018 299 1554	018 299 1490
Email address:	karien.Hattingh@nwu.ac.za	21975957@nwu.ac.za

4. **RESEARCH QUESTION AND PURPOSE OF RESEARCH:**

RESEARCH QUESTION:

To which extent can ESL first-year university students' controlled productive knowledge of verb-noun collocations be improved by implementing awareness-raising teaching techniques?

PURPOSE OF RESEARCH:

The primary aim of the study is to determine the extent to which ESL first-year university students' controlled productive verb-noun collocation knowledge can be improved by means of explicit awareness-raising techniques targeting selected verb-

noun collocations and maximising learner engagement with the target language. The focus is on vocabulary development of first-year ESL students.

5. PROCEDURES

The procedure for this study will not interfere with, or add to your normal ENLS 111 activities. Some of the eFundi-based exercises included as part of the module will serve to inform the researcher as to students' controlled productive knowledge of verb-noun collocations. There will be eight exercises, and performance and improvement will be measured across all eight exercises.

Two groups of students will be involved: a control group and a test group. You will not know which group you belong to. The test group will receive an instructional treatment that the control group will not. If results are significantly positive, control group candidates will be afforded the opportunity to receive this same input.

Your module mark will not be affected directly by your performance on the exercises in question.

What is expected of you as participant? To make available your data in the form of ENLS 111 exercises (written, classroom, and electronic-based) and marks achieved (this may be based on electronic-, peer-, self- and/or lecturer-evaluations) for the purposes of this research project.

SECTION B: GENERAL PRINCIPLES OF RESEARCH

1. PARTICIPATION

Participation in the research project is voluntary, and participants have the right to withdraw from this project at any time.

2. POTENTIAL RISKS INVOLVED:

No negative impact is expected as the aim is to measure the impact of certain interventions, or the lack thereof, on students' controlled productive knowledge of verb-noun collocations.

3. POTENTIAL BENEFITS:

Participants may benefit from this project in that it may contribute to vocabulary development, which could potentially benefit participants' academic performance.

4. PAYMENT FOR PARTICIPATION:

Participants will receive no payment for taking part in this project.

5. CONFIDENTIALITY:

Participant identities will remain anonymous and a coding system will be used for identification purposes. Data will only be included for those participants who have given signed consent, and the data collected will not be used for any purposes other than for this project.

6. TO THE SIGNATORY OF THE CONSENT CONTAINED IN SECTION C OF THIS DOCUMENT:

Should you have any questions or concerns relating to this project, please contact Dr Van den Berg at karien.Hattingh@nwu.ac.za or 018 299 1554.

a) Your participation in this project is completely voluntary and you can choose to withdraw from the project at any time, without stating reasons, and you will in no way be harmed by so doing,
b) There are no risks involved in taking part in this project,
c) You will not derive any benefit personally from your participation in the project, although the knowledge that may be gained by means of the project, may benefit aspects of your academic performance,
d) All participants will remain anonymous.

7. DECLARATION:

As Project Head, I, Ms M Morris, confirm to participants that the above information is complete and correct.

_____ _____

Signature: Researcher Date signed

Signed at: _____

SECTION C: INFORMED CONSENT

TITLE OF PROJECT: *Measuring improvement in first-year ESL university students' productive knowledge of verb-noun collocations through integrated productive oriented teaching strategies.*

PARTICIPANT

I, _____ (full names and surname) have read the preceding premises in connection with the project, as discussed in ***Section A*** and ***Section B*** of this informed consent form, and have also heard the oral version thereof and I declare that I understand it. I have also initialed every page of ***Section A*** and ***Section B***. I was given the opportunity to discuss relevant aspects of the project with the Project Head and I hereby declare that I am taking part in the project voluntarily.

_____ _____

Signature: Participant Date signed

Signed at: _____

ANNEXURE B: INTERVENTION EXERCISE (EXAMPLE)

Verb-Noun Collocations Exercise 3

This assessment follows the same format as the Vocabulary Exercises we completed in class.

Ensure that you read every question carefully.

You will have 45 minutes to complete the exercise.

You can take this quiz only once.

Part 1

1

Read *The New Zealand Economy* (attached).
Identify and write down verb-noun collocations from the text.

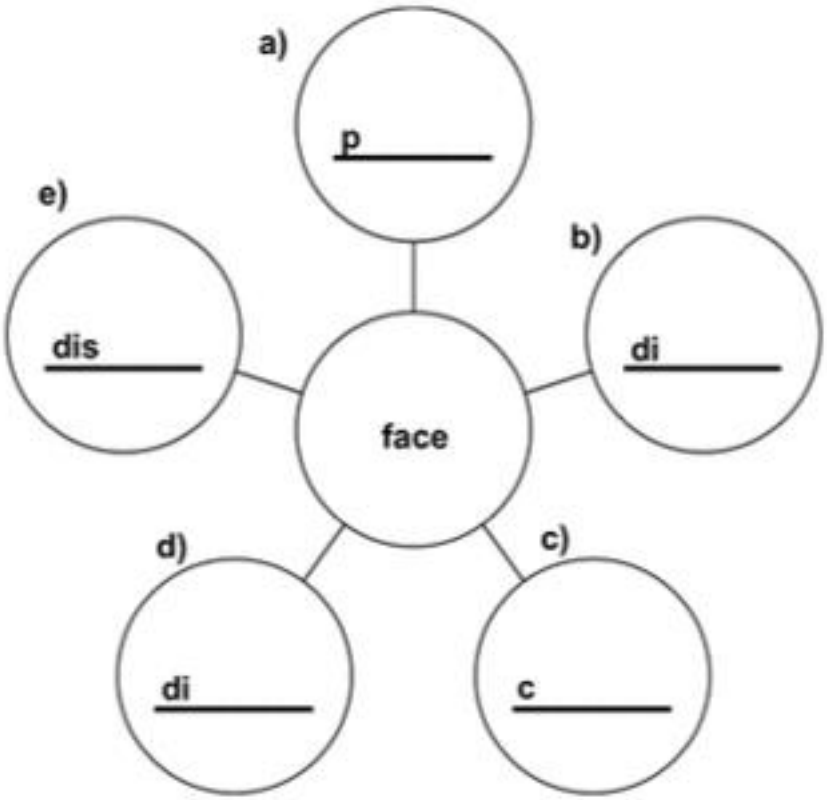
e.g. 1. fall into a category
2. solve a problem

Attachments
Verb-Noun Collocations (Exercise 3).pdf

Answer Point Value: 1.0 points
Model Short Answer: .....

2

Attachments



Complete the diagram by filling in the gaps below.
Use the centre word as a starting point, then fill in all the gaps around the centre word.
The centre word is a **verb**. Only enter **nouns** in the gaps.
Use the first letter(s) on each line as a clue.

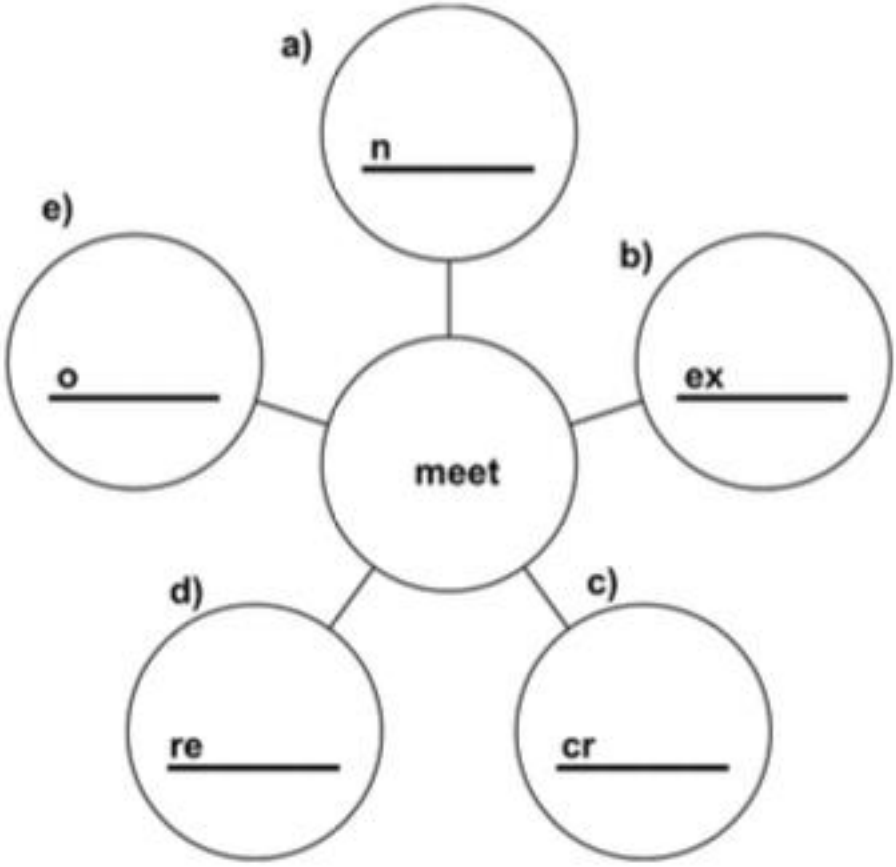
(These words are not necessarily in 'The New Zealand Economy')

a) face a p____
b) face (a) d____
c) face a c____
d) face (a) d____
e) face dis____

Answer Point Value: 5.0 points
Answer Key: problem|roblem, dilemma|difficulty|ilemma|fficulty, challenge|hallenge, dilemma|difficulty|ilemma|fficulty, crimination|discrimination

3

Attachments



Complete the diagram by filling in the gaps below.
Use the centre word as a starting point, then fill in all the gaps around the centre word.
The centre word is a **verb**. Only enter **nouns** in the gaps.
Use the first letter(s) on each line as a clue.

(These words are not necessarily in 'The New Zealand Economy')

a) meet the n____
b) meet ex____
c) meet cr____
d) meet re____
e) meet o____

Answer Point Value: 5.0 points

Answer Point Value: 5.0 points
Answer Key: need|needs|eed|eeds, expectation|expectations|pectation|ectations, criteria|iteria, requirements|requirement|quirements|quirement, objective|objectives|bjective|bjectives

4

Choose four (4) of the collocations from your diagrams, and make one (1) sentence with each.

4.1

4.2

4.3

4.4

Answer Point Value: 4.0 points
Model Short Answer: --------

ANNEXURE C: EXAMPLE TEXT – INTERVENTIONS

Endangered Species

S Millet - English Language Institute Occasional Publication No.22

There are millions of species (different kinds of animals and plants) on earth. Scientists have classified about 1.7 million of them. However, it is now thought that there are between ten million and one hundred million species in the world. Each one has its own unique **part to play** in keeping the natural balance of the world.

Over millions of years, countless animals and plants have died out and been replaced by others. Should we be worried about this? Isn't this a natural process? Why have people become so concerned about endangered species in recent years?

People are concerned about protecting endangered species because people are the reason why they are endangered. With rapidly increasing world population, people are taking land that animals once used. Consequently there is nowhere for some species to

live and they become endangered. Think of the panda in China and the kiwi in New Zealand. Scientists believe that rain forests like the Amazon in South America may contain half of the world's species. However, every year large areas of rainforest are destroyed to provide land for farms and to produce crops like palm sugar. With every square mile of land destroyed, more plants and animals are lost.

Nobody knows how many species have died out recently, but they do know the reason why they have. Human beings. It is our responsibility to save endangered species.

There are four main reasons why people should save and protect endangered species. The first reason is that all species have a right to share the earth, not just humans. Another reason is that all species are unique and interesting and they make our world a richer and better place to live. Third, humans can **make useful products** from other species, such as medicines from

plants so it is important for us to **protect** our own **interests** by protecting the environment we live in. The fourth reason is that we need other species. We cannot exist by ourselves. If we change the natural balance of the world, we may destroy ourselves.

People and governments are learning more about the importance of saving endangered species, and are **taking steps** to **solve the problem**. For example the New Zealand Government has **initiated programmes** to help save the kiwi and other endangered species.

The world would be a much poorer place without kiwi and pandas. We should preserve our environment for future generations.

ANNEXURE D: PRE-TEST, POST-TEST 1 AND 2, DELAYED POST-TEST

Name: _____
Score: _____ / _____

Pre-test

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take the verb tenses and subject-verb agreement into consideration. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last three questions), please write the word that you select as the answer, and not only the letter corresponding to the answer.

A. I have read the instructions.

B. .

C. .

2

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Wastewater use is increasingly seen as an option to m____ the growing needs for water.

3

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Research f____ that users' engagement with social media contributes to tension and jealousy beyond the offline relationship dynamics.

4

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Irrigation with untreated and inadequately treated wastewater pr____ a serious public health risk to farmers and consumers who are exposed to a range of contaminants.

5

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

One important skill for economists is the ability to a____ data.

6

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Exposure was determined using interviews to c____ information on contact with wastewater, personal hygiene practices, use of personal protective measures, and other relevant information.

7

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Social media affordances complicate individuals' ability to control who r____ information about themselves because they are encouraged to have an ever-growing number of "friends."

8

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Informal employment became widespread among both migrant workers and urban workers. In response to these conditions, workers increasingly e____ in protests.

9

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

The relationship between parent and child po____ unique challenges within a social media platform.

10

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

The Constitution requires that the absentee be treated in a manner that t____ his or her age into consideration.

11

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

We use this question to c____ an analysis of the contract status of the same worker before and after the law.

12

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Finally, the survey pr____ limited information on the effects of the provisions for open-ended contracts that upset opponents of the law more than any others.

13

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Conflicting results have been f____ in existing studies on the economic impacts of smoking bans.

14

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Although participants can largely communicate in English on day-to-day topics, their mother tongue p____ an important role in stressful circumstances like health-related situations.

15

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Last year, investors in the stock market of China p____ more attention to environmental issues.

16

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

In an attempt to find a solution to this problem, the researchers c____ an exploratory literature review on learner absenteeism.

17

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Theoretically-based studies can o____ insights into the subtle, striking effects that new media exert on young adult women.

18

Fill in the gap in the sentence, using the first letter of the missing verb as a clue.

Information and communication technologies can enh____ teaching and learning.

19

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text.

Social media are digital platforms for networking and interaction between people, where users 1.____ and comment on content. Social media such as Facebook, Twitter, and Snapchat are incredibly popular; for example, Facebook reported 1.44 billion active users in March 2015 (Facebook, 2015). Social media provide many 2. ____ to users, including connecting with friends and family and building personal and professional networks. However, there are also many privacy concerns that have not been adequately considered, let alone addressed. These questions are of concern to policymakers, therapists and health care professionals, parents, educators, technology companies, and the general public.

Social media pose a privacy paradox. On the one hand, they are increasingly prominent in users' lives and are the repositories of an extensive trove of personal information freely shared by those users. They also 3. ____ some privacy advantages over face-to-face communication, including the ability to edit one's words before sharing them and the ability to hide emotions that might otherwise be communicated nonverbally. On the other hand, research 4. ____ that most users have concerns about their privacy, even as they share more and more information on social media. This concern for privacy appears to have increased over time.

1. a. produce
   b. make
   c. generate

2. a. benefits
   b. services
   c. disadvantages

3. a. give
   b. provide
   c. enable

4. a. recommends
   b. suggests
   c. provides

20

China's new Labour Contract Law 5. ____ effect in January 2008 and required firms to give migrant workers written contracts, strengthened labour protections for workers and contained penalties for firms that did not follow the labour code. This article 6. ____ survey data of migrant workers in the Pearl River Delta before and after the law, and a retrospective question on when workers received their first labour contract to 7. ____ the effects of the law on labour outcomes. The evidence shows that the new law 8. ____ the percentage of migrant workers with written contracts, which in turn raised social insurance coverage, 9. ____ the likelihood of wage arrears and 10. raised the ____ that workers had a union at their workplace.

5. a. took
b. gave
c. made

6. a. presents
b. conducts
c. uses

7. a. interpret
b. analyse
c. assess

8. a. improved
b. increased
c. impacted

9. a. lowered
b. influenced
c. reduced

10. a. likelihood
b. data
c. percentage

21

Choose the correct option for each missing word in the texts below, then fill in your answer in gap in the text.

At the practice and training level, it is important to include educational programs that 11. ____ the issues related to cultural and linguistic biases, and 12. ____ opportunities for practitioners to interact with individuals of different backgrounds and linguistic abilities to practice facilitative communication strategies. This is particularly key given that the provision of bilingual or L1 services is not always feasible. Therefore, it is vital to educate practitioners about the pitfalls of L2 communication, and particularly regarding the increased chance of 13. ____ biased judgments about nonnative speakers, and about individuals whose cultures are different from one's own.

11. a. involve
    b. compare
    c. address

12. a. aim
    b. provide
    c. give

13. a. making
    b. doing
    c. choosing

Name: _____
Score: _____ / _____

Post-test 1

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last two questions), please write the word that you select as the answer, and not the letter.

I have read the instructions carefully.

- True
- False

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The goal of education can be seen as achieved when a student either ch____ his attitude or causes somebody else to change.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

An issue to be di____ with great care is which one of these interpretations is to be prioritised within various Education for Sustainable Development contexts.

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

In its public usage, to apply a label to someone is to implicitly m____ a moral judgment.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Xinjiang's Hotan prefecture c____ a total area of 248,100 km2.

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The difficulties en___ by teachers implementing education for sustainable development.

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We pr___ considerable evidence for our argument using newly collected data on human-made walls separating states from 1800 to 2014.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The survey was c___ in the native language of the students, administered online and promoted via an email to students.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Research in Britain has begun to ex___ the emotional difficulties faced by Early Childhood Education and Care practitioners in forming relationships with children and parents.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The threat p___ by terrorism is one to be taken seriously.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

There is a tendency to f___ on the emotional aspects of the child.

12

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Competence is about technocratic efficiency, the right workforce to ac___ the right outcomes.

13

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We could u___ these data to create border province dyads for all countries.

14

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Falling prices of low carbon technologies has generated political pressure against policies designed to re____ carbon emissions.

15

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The researcher ad____ an evolutionary economic approach to analyse barriers that are rooted in the complex interconnections between choices that consumers must navigate on a daily basis.

16

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Maternal and child health advocates are also critical to help turn the tide and m____ common-sense policies like paid family leave a reality.

17

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Sport Education is a student-centred model that has the potential to provide trainees with a holistic education through nonnegotiable features in which they t____ on a number of roles.

18

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The nature of the first author's teaching cr____ an environment where the teachers could share their lived experience in the focus groups.

19

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

One conclusion to d____ from the reflection upon the distinction presented above, is that if the expression is going to be substantial and meaningful, critical values education has to open up for a moral reasoning.

20

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Faculty members that un____ this research need to engage systematically with existing theoretical and conceptual work.

21

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We s____ researchers' arguments that teachers who are teaching through Sport Education need both content and pedagogical content knowledge.

22

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

A master, who had already de____ his competence in a craft or trade, would engage a number of apprentices, who would be provided with food and lodging.

23

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

Psychical Education teachers' learning agenda includes learning about the specific content being taught, learning about learning and learning about teaching. Hence, being successful as a Physical Education Teacher Education (PETE) faculty member is not only concerned with having teachers 22.1 ____ attention to the subject matter being taught but also the practices 22.2 ____ in presenting the subject matter. One researcher terms this the "pedagogical turn". That is, learning what is being taught while at the same time questioning, examining and learning about the way in which it is actually being taught. PETE faculty members need to provide 22.3 ____ to the thoughts and actions that shape teachers practice and unpack their teaching in ways that 22.4 ____ them access to the pedagogical reasoning, uncertainties and dilemmas of practice that are inherent in understanding teaching as being problematic.

22.1 a. give
b. pay
c. make

22.2 a. employed
b. given
c. combined

22.3 a.information
b. access
c. benefits

22.4 a. gives
b. allocates
c. allows

24

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

It is necessary to increase the intensity of science research and technology integration of assembling complete sets, continue to 23.1 ____ research of low-carbon agricultural technology. We know that Hotan is a serious lag on the regional and national decoupling stages, and that it cannot 23.2 ____ its agricultural carbon emission reduction goals by 2030. However, Hotan lacks an incentive mechanism to 23.3 ____ the development of low carbon agriculture.

23.1 a. carry out
b. achieve
c. carry

23.2 a. make
b. conducts
c. achieve

23.3 a. interpret
b. promote
c. assess

25

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

One of the two stated objectives of the new "Research Note" section of Political Psychology is to 24.1 ____ short reports that highlight novel methodological approaches. Toward that end, we call readers' attention to the "flanker task," a research protocol widely 24.1 ____ in the study of the cognitive processes involved with detection, recognition, and distraction. The flanker task has increasingly been modified to study social traits, and we believe it has untapped value in the area of political psychology. Here we describe the flanker task—discussing its potential for political psychology—and illustrate this potential by presenting results from a study correlating political ideology to flanker effects.

24.1
a. present
b. give
c. have

24.2
a. conduct
b. employ
c. collect

Name: _____
Score: _____ / _____

Post-test 2

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last four questions), please write the word that you select as the answer, and not the letter.

A. I have read the instructions.

B. .

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

It will be necessary to de____ methods for estimating changes in carbon stocks in grasslands.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

It appears that 'gadget' children who en____ in sedentary rather than active leisure activities particularly with their parents are more likely to report high screen time and have an unhealthy weight.

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Chi squared statistics and forced entry logistic regression was used to id____ factors associated with obesity in children.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Does this structure m____ at least the current requirements of companies, not to mention the future requirements of employers?

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Their data sh____ that this did not affect children with specific language impairment.

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Species migration induced by climate change may c____ even greater stresses.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Those who did not f____ i____ this most active category were grouped together into 1 comparison group.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We did not formally assess articulation because we felt that the judgements of speech and language therapists who work with the children each day and who had full access to each child's medical history were superior to any judgment that we could m____ on the basis of a five minute assessment using a standardised task.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The maps created in this study pr____ useful overviews of the spatial distribution of grasslands and their soil organic carbon stocks in Japan.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

In view of the key function of teachers in the formal education system's ability to ac____ environmental education goals, Environmental Education is a necessary component of preservice teacher training programs.

12

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The mechanism by which high screen time in____ the likelihood of obesity in children is unclear.

13

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The findings re____ that these differences were not significant.

14

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Before we co____ the implications of these findings, it is important to consider whether the results reflect methodological flaws in the design of the study, rather than substantive discoveries.

15

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

These factors create the context in which environmental educators work, so they should serve as guiding factors and co____ to the development and implementation of effective Environmental Education.

16

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Data were pr____ as adjusted odds ratios (OR) of the likelihood of specified outcomes.

17

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The method of procedure does not in any way a____ the final outcome of the case.

18

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The soil profile data u____ in this study were based on the 1959–1978 soil survey.

19

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Before any action of this kind is taken the whole issue should be carefully co____.

20

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Teachers p____ a key role in developing environmental literacy in future generations.

21

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The results obtained for this category may pr____ insight into the influence of parents' education as one of many socioeconomic factors that creates a household atmosphere receptive and responsive to public agenda issues that receive media exposure.

22

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

Text 1
To 22.1 ____ more accurate and comprehensive information of these changes, it is important to develop the unified grassland categorization system and to 22.2 ____ soil organic carbon data for several types of grasslands

22.1. a. make
b. obtain
c. receive

22.2. a. get
b. gain
c. collect

23

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

The spatial distribution of grasslands was determined using the vector data of the vegetation and wetland maps of Japan, based on data from the National Survey on the Natural Environment 23.1 ____ by the Environment Agency. The survey has been conducted about every 5 years since 1973, in order to 23.2 ____ nationwide information needed to promote and implement policies to protect the natural environment.

23.1. a. conducted
b. done
c. observed

23.2. a. present
b. gather
c. use

24

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

At the practice and training level, it is important to include educational programs that 24.1 ____ the issues related to cultural and linguistic biases, and 24.2. ____ opportunities for practitioners to interact with individuals of different backgrounds and linguistic abilities to practice facilitative communication strategies. This is particularly key given that the provision of bilingual or L1 services is not always feasible. Therefore, it is vital to educate practitioners about the pitfalls of L2 communication, and particularly regarding the increased chance of 24.3. ____ biased judgments about nonnative speakers, and about individuals whose cultures are different from one's own.
24.1. a. involve
b. compare
c. address

24.2. a. aim
b. provide
c. give

24.3. a. making
b. doing
c. choosing

25

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

The authors 25.1 ____ a method where the learner gets the concepts in his mother tongue and is initiated to convert the same in English. Here the status of the stock is sufficient, but the supply lacks free flow. On receiving the first version of the converted matter from the vernacular language to English, the authors carefully identify the problem areas of the learner. A thorough assessment is 25.2 ____ in the presence of the learner with an attempt to rectify the overall defects in the writing.
After this, a second attempt is made and a considerable improvement is experienced. By this time, the learner 25.3 ____ some knowledge of writing sentences of his own.
25.1. a. experimented
b. employed
c. gathered

25.2. a. provided
b. explored
c. made

25.3. a. gains
b. understands
c. makes

Name: _____
Score: _____ / _____

Delayed post-test

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last three questions), please write the word that you select as the answer, and not the letter.

A. I have read the instructions carefully.

B. .

2

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

For the first time in a long time we were encouraged to ga____ information through other methods than the usual desk research and eld research via interviews.

3

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Online wiki documentation for the Celestia OSSD project de____ this process and stresses its importance as a form of “apprenticeship“.

4

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Facilitating social circumstances in____ the likelihood of design solutions being accepted by open source software development communities.

5

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

We co____ research by experiencing the task of playing ourselves.

6

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Addiction to social media is the issue that has received high concerns in society because people nowadays can ha____ more frequent access to social media from portable devices.

7

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Some scholars have suggested that social media could be used as a tool to help people re____ stress.

8

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Since a comic uses visuals to co____ messages to the readers, the language it's used could be described as a "visual language".

9

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

These advantages mostly rely on the manga's ability to put the reader inside the story, to be immersed within as manga itself had de____ specific and distinctive visual storytelling techniques to achieve this.

10

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

To test whether a threat has deterrence effect, people's behaviour should be compared before and after the threat is po____ or becomes credible.

11

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

According to a China Internet Network Information Center (CNNIC) report pu____ in January 2017, over 90 percent of Chinese online users adopt at least one social media application.

12

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

We refer to social media as web-based communities where users create personal profiles; sh____ information, ideas, and personal messages; and communicate with other users.

13

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Under this new welfare regime, socialised and medicalised solutions to social problems are usurped by penalised solutions and the receipt of help often requires that one as____ the role of victim or perpetrator.

14

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

The goal of this research is to ga____ a structured insight into the collaborative design process within OSSD projects.

15

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Self-efficacy determines the choice of activities to en____ in, how much energy to spend doing the activities and how to maintain it.

16

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

The developer has second thoughts whether the proposed colour palette me____ the technical requirements for colour management in the printing process.

17

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Existing scientific knowledge does not of____ solid insights into the causes for design methodologies.

18

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

In conventional organisations, members communicate and act in adherence to recognised roles and plans within a well structured production environment where defined methods, goals, schedules, etc. pl____ important roles in achieving production goals.

19

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

They must filter through other communications in order to ex____ relevant information.

20

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

New recognition of children exposed to domestic violence as crime victims al____ children access to state funds and associated services for which they would not otherwise be eligible.

21

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

A 2001 study co____ by Nancy Lemon finds that implementation of the rebuttable presumption was inconsistent and that courts were unaware of the code and confused about what type of finding triggered the presumption.

22

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

It is important to note that the toolkits also ha____ the potential to serve as entrepreneurial goods that return monetary value to the design students.

23

Fill in the gap in each of the following sentences with a verb, using the first letter as a clue. When writing your answer, include the letters given as a clue.

Empowering the end-user to ma____ substantive contributions to the interpretation process facilitates her gaining more control over the value generated.

24

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

Despite some research limitations, results from this study provide significant contribution to social media research. Importantly, the authors have 24.1 ____ additional evidence to the literature on the impact of social media addiction, which is an area that still lacks empirical investigation at the present time. Specifically, this research focuses on the impact of social media addiction on mindfulness and coping strategy, which are factors that were left unexplored. In addition, by focusing on employee samples and workplace outcomes, this research is able to extend the context of social media research to 24.2 ____ its impact in the workplace environment. The overall results 24.3 ____ additional support to extant studies that found the negative consequences of social media usage.

24.1. a. produced
b. provided
c. generated

24.2. a. assess
b. search
c. interpret

24.3. a. give
b. provide
c. enable

25

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

This work also contributes to the literature on personal credit approvals, which 25.1 ____ different statistical methods, such as logistic regression, decision tree, nearest neighbour, and neural networks, to distinguish between "good" and "bad" borrowers. The results of our study 25.2 ____ that social media-related variables can be added to the classification models as account-specific covariates to 25.3 ____ the performance of the models.

25.1. a. takes
b. uses
c. makes

25.2. a. present
b. conduct
c. suggest

25.3. a. interpret
b. improve
c. assess

26

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

There's been a paradigm shift in design from focusing on aesthetic worth to focusing more on the interplay of form and function to 26.1 ____ social responsibility and to pursue social change through innovation. As a result, the discipline needs models for how to educate responsible designers who see design not only as a commercial enterprise but more importantly as a catalyst for social change, and are able to innovate visual technologies that 26.2 ____ social problems that are wicked by nature, and are far more complex and interdisciplinary than merely problem-solving how to aestheticise a client's content.

26.1. a. undertake
b. assume
c. change

26.2. a. include
b. address
c. discuss

ANNEXURE E: TI, TII, TIII, TIV

Name: _____
Score: _____ / _____

ti

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last two questions), please write only the word that you select as the answer, and not the letter of the option you choose.

A. I have read the instructions.

B. .

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Researchers have argued that children should ac____ a range of skills early in life, long before formal schooling begins.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Does self-awareness of cultural and linguistic biases c____ the attitudes and practices of health/law professionals?

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Accented speech can take longer for listeners to process. This can cause listeners to d____ conclusions about the speaker based on the origin or the mere existence of an accent rather than the content of the utterance.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Data were c____ in person from a selected sample of the research population from each district.

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Having public conversations with friends c____ the impression that the user is popular and well-liked.

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

There is always a need to put a greater emphasis on professional improvements through continuous training of teachers to m____ the challenging demands of the profession of teaching.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

When people feel excluded from a group, even a group of strangers, they typically ex____ distress and a need to belong.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Exposure was determined using interviews to c____ information on contact with wastewater.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The collected data were an____ with the help of Statistical Package for Social Sciences.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Early mastery of skills fa____ learning and development at later ages and also makes learning more likely to continue.

12

Choose the correct option for each missing word in the texts below, then fill in your answer in gap in the text. Write the whole word and not only the letter.

Research should 11.1. ____ a combined inter- and transdisciplinary approach if it seeksto advance theory and to 11.2. ____ to practical progress and development of new policies. Such an approach can 11.3. ____ a systematic and comprehensive theoretical framework for the definition and analysis of problems and challenges in a given field, in that it can integrate concepts and methods from different academic disciplines as well as questions, perspectives and priorities of different stakeholders, including policy-makers and practitioners.

11.1. a. choose
b. adopt
c. generate

11.2. a. improve
b. contribute
c. enhance

11.3. a. give
b. make
c. provide

13

Choose the correct option for each missing word in the texts below, then fill in your answer in gap in the text. Write the whole word and not only the letter.

Online interventions are 12.1 ____ an increasingly critical role in health campaigns designed to reach young adults, and campaigns are more likely to change attitudes if they 12.2 ____ into account relevant theories, understand the audience, and tailor the message to particular audience characteristics.

12.1. a. contributing
b. making
c. playing

12.2. a. consider
b. take
c. use

Name: _____
Score: _____ / _____

tii

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last two questions), please write the word that you select as the answer, and not the letter.

A. I have read the instructions carefully.

B. .

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Pickering found that the majority of teachers in her study felt safe to r____ controversial issues within their teaching team.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

This pro____ further evidence that young children need specific opportunities for skills to be taught and practised.

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Last year, Newman pu____ an article on the Montessori approach for the local teacher education community.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

All sessions were conducted on an individual basis, and lasted 15 minutes or until any event selected had started and re____ its conclu-sion.

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

A mixed method approach was used to ob____ information relating to perceptions, understanding, and level at which climate change has been taught in Nigerian universities.

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Centres are required to sh____ evidence of reflection through a self-review process.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

A recent narrative review su____ this argument.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Children com____ tasks that included physical activity assessment at 19 months, 3.5 and 5 years.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

There seems to be a need to en____ the teaching of and learning about climate change.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

An estimated 50-250 million Africans will f____ increased water stress by 2020.

12

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

The results from this study also 12.1 ____ that students in the environmental sciences (i.e., geography) have more (50.01%) knowledge of climate change than students in the humanities and other faculties (Fig. 1). This implies that the high level of understanding of climate change as shown in Table 3 were mainly students with class experience on climate change, who graduated from environmental sciences. The results from the focus group discussion further show that apart from environmental sciences departments, there were few related courses taught in other faculties, which had little or no climate change content in their curriculum. This may be because faculties of environmental sciences are the major university units that 12.2 ____ research on environmental issues, such as climate change. It also shows that the majority of universities in Nigeria do not offer classes that 12.3 ____ the issue of climate change.

12.1
a. presented
b. revealed
c. addressed

12.2
a. do
b. make
c. conduct

12.3
a. address
b. talk about
c. initiate

13

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

Dr Montessori 13.1 ____ an experimental approach; they 13.2 ____ her method and 'field-based curriculum and materials' as being continually tested by trial and error by children and teachers in countries around the world.

13.1
a. develop
b. took
c. made

13.2
a. described
b. praised
c.implemented

Name: _____
Score: _____ / _____

tiii

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last two questions), please write the word that you select as the answer, and not the letter.

Remember that you are allowed to use a dictionary to find the correct answers.

A.
I have read the instructions.

B.
.

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

A qualitative approach was deemed most appropriate for our goal to de____ strategies and procedures to enable us to consider experiences from the informants' perspectives.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We would like to underline the potential that hospital pedagogy has to co____ to the development of a hospital education.

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Answering these questions would require evaluation evidence that shows how preschool programs' design a____ children's outcomes in both the short and long run.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The first goal was to id____ the key issues in measuring the economic returns from preschool programs.

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We em____ qualitative research methods to examine university students' written reactions to assignments

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We p____ data from a cohort of 421 students enrolled at ACU in the combined degree course from 2009 until 2012.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

These university students g____ insight and appreciation for the music theme's connection to academic content after the course.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Researchers f____ a number of challenges when using benefit-cost analysis to evaluate a preschool program.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We elected to f____ these specific procedures.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The central authority establishes curriculum bases, and then the autonomous communities develop its contents to m____ their specific requirements.

12

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

A second challenge is related to the fact that preschool programs are expected to 11.1 ____ outcomes throughout the life course. Thus a full accounting of potential benefits would require projecting outcomes from preschool participation into the future, beyond the point of the last follow-up. To connect outcomes measured at younger ages with expected outcomes at older ages, we need to 11.2 ____ assumptions about the causal relationships through time.

11.1. a. produce
b. affect
c. change

11.2. a. make
b. give
c. provide

13

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

Motivated by policy proposals to expand preschool access at the state and federal levels, several studies have been 12.1 ____ to 12.2 ____ such estimates. Table 3 summarizes the key features of five state studies, all conducted in the mid-2000s. These studies all project the benefits from state-level universal preschool programs—in some cases for a one-year program, in other cases for a two-year program. The studies 12.3 ____ the impact of increasing access to high-quality preschool programs, relative to current enrolment levels.

12.1. a. conducted
b. made
c. researched

12.2. a. present
b. estimate
c. provide

12.3. a. interpret
b. consider
c. look at

Name: _____
Score: _____ / _____

tiv

Part 1

1

Read each question carefully. When filling in the gaps in each sentence, take verb tenses and subject-verb agreement into consideration. Write the whole word, and include the letters given as clues in your answer. e.g. My garden is filled with f_____ flowers. Answer: fragrant. For the questions where there are multiple choice options (the last two questions), please write the word that you select as the answer, and not the letter.

A. I have read the instructions.

B. .

2

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

These issues should be ad____ by investigations emerging from sustained research.

3

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The Web of Science g____ access to multiple databases that reference cross-disciplinary research.

4

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The first year is usually 12 months of full-day classroom education mirroring topics co____ in the first two years of medical school.

5

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

The Research Section hosted several research conferences at____ by many leading clinical and basic science researchers.

6

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

This review article explores research that has been c____ o____ with respect to knowledge translation in agriculture.

7

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Students are expected to perform to the level of their education and continuously en____ performance through expansion of the knowledge and refinement of skills.

8

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

We need to de____ and evaluate the most effective methods for facilitating physical therapist acquisition.

9

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Kolb's LSI is a forced-choice instrument using short phrases to id____ students' preferred ways of learning without segregating the students into defined groups.

10

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Education for the Physician Assistant (PA) undergoes continuous modification to m____ the expanding requirements of modern medicine.

11

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Literature reviews can c____ a wide range of subjects at various levels of completeness and comprehensiveness.

12

Fill in the missing verb in the sentence, using the first letter(s) as a clue.

Studies used alternative terms such as ''extension'' and ''technology transfer'' to de____ the process of knowledge transfer.

13

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

These experiences 13.1 ____ opportunities for students to 13.2 ____ and receive feedback and interact with faculty.

13.1. a. provide
b. make
c. generate

13.2. a. benefit
b. give
c. consider

14

Choose the correct option for each missing word in the text below, then fill in your answer in gap in the text. Write the only missing word. Do not write the letter.

The initial survey was distributed on paper and collected using an anonymous drop-box. Subsequent surveys were 14.1 ____ using an Internet-based survey engine without 14.2 ____ information pertaining to participant IP addresses.

14.1. a. produced
b. conducted
c. done

14.2. a. presenting
b. getting
c. obtaining

ANNEXURE F: PARTICIPANT ERRORS

Annexure E:
Participant errors

Error type	Spelling	Incorrect plural	Incorrect PoS	Non-word	Nominal derivational suffix	Inflectional suffixes
1	instagram					
2	twitter					
3	rumored					
4	radifications					
5	wheter					
6	inbodyment					
7	eachother					
8	humanbeen					
9	devios					
10	humanrace					
11	unlimted					
12	polititian					
13	realize					
14	rumors					
15	realizing					
16	confusement				confusement	
17	i					
18	Wich					
19	civilized					
20	organization					
21	organized					
22	fueled					
23	Webopedia					
24	publically					
25	characterizing					
26	networpubliclyking					
27	recognized					
28	immerged					
29	Nasional					
30	socialize					
31	now a days					

32	modern day					
33	percieved					
34	globalization					
35	premiumize					
36	departement					
37	interpereted					
38	nececcarily					
39	quickle					
40	posession					
41	cociane					
42	sensualize					
43	cliënts					
44	where as					
45	behavior					
46	could of					
47	people's life's					
48	it's		it's			
49	mesmerized					
50	minimize					
51	favor					
52	harboring					
53			affect			
54	sosial					
55	dillema					
56	impak					
57	wikipideia					
58	editted					
59	there fore					
60	presiving (perceiving)					
61	thim (them)					
62	legidement					
63	compiniese					
64	responsebilty					
65	piont					

66	strugle					
67	diffrensiate					
68	juicsy					
69	stitistics					
70	fals					
71	upperhand					
72	mearlly					
73	generaly					
74	ruine					
75	lifes					
76	rol					
77	serounding					
78	targetting					
79	Afrika					
80	againts					
81	exsploted					
82	propoganda					
83	conclution					
84	aposses					
85	greates					
86	Whatsapp					
87	scrutinize					
88	civilized					
89	Gaurdian					
90	polarization					
91	unignorably					
92	analyzed					
93	recognize					
94	jeopardize					
95	tipe					
96			(make us) worrying			
97	whitch					
98	honorable					
99			mislead (misleading)			

100			enlights	enlights		
101	weaponize					
102	criticized					
103	influnce					
104			maybe (may be)			
105	wach					
106	fakenes					
107				unsurety (unsureness)		
108	tention					
109	sensationalizing					
110			believes (beliefs)			
111	verrifyig					
112	electronical			electronical (electronic)		
113	im					
114	mas-distractions					
115	humor					
116	who`s		who's (whose)			
117	prioritize					
118	South-Africa					
119	instill					
120	favor					
121	contradictive			contradictive		
122	negativly					
123	availabile					
124	never the les					
125	favoring					
126			has became			
127			is been said			
128	were		were (where)			
129	defense					
130	flavors					
131	clickbaiting					
132	unrealiable					
133	dishonor					

134	utilizing					
135	utilization					
136	offense					
137	personalize					
138	maximize					
139	biasedness					
140	indimidate					
141	ligitimate					
142	endeavors					
143	secretely					
144	exucutive					
145	jounrnalists					
146	posseses					
147	qulafications					
148	disclipned					
149	wat (what)					
150			varies (various)			
151	publishment					
152			that been said			
153	demoralize					
154			have proved			
155	publicize					
156	used the believe (used to believe)					
157	unhuman					
158	satarical					
159	constution					
160	(lives of) other's		other's			
161	fakeisms					
162	chail (jail)					
163			affect (effect)			
164	neighborhood					
165	summarize					
166	restuarants					
167	assortiated					

168	desparately					
169	belieavable					
170	news are been casted		news are been casted			
171	news is been provided		news is been provided			
172	individuals is been affected		individuals is been affected			
173	systematized					
174	violance					
175	enviromment					
176	extend		extend (extent)			
177	meausure					
178	sugestions					
179	propogandists					
180	contextualization					
181	allogations					
182	ther's reputation		ther's (reputation)			
183	falls (false)					
184	koo (coup)					
185	formaer					
186	emed					
187	privelage					
188	rummors					
189	leaugue					
190	bussines					
191	uncreditable					
192	goverment					
193	ruïne					
194	ment (meant)					
195			have became			
196	supress					
197	bread winner					
198	visualizations					
199	specialize					
200	ingenuine					
201	neutralize					

202	persausive					
203	aswell					
204	meduim					
205	allected (elected)					
206	socialble					
207	personalize					
208	fith (fifth)					
209	maximize					
210	celebrate (celebrity)					
211	decieved					
212	decieving					
213	becuase					
214	descripted			descripted (described)		
215	devided					
216	there is been		there is been			
217	presedenit					
218	alowed					
219	coutry					
220	artickels					
221	probleme					
222	vals					
223	theis					
224	articales					
225	belived					
226	ammedeatly					
227	goverment					
228	paragrav					
229	allot					
230	allso					
231	givven					
232	adeer					
233	oppening					
234	bracking					
235	oppen					

236	reuld					
237	socila					
238	surculate					
239	publick					
240	panick					
241	criticizing					
242	is been given		is been given			
243	unnessacery					
244	anyone to writes something		(anyone) to writes (something)			
245	demoralizes					
246	polarization					
247	polarized					
248	deformation (defamation)					
249	unethicality's					
250	offenses					
251	monitering					
252	cauld					
253	organization					
254	itweb (the web)					
255	uppon					
256	offence is been made					
257	it's					
258	abovementioned					
259	idolize					
260	offis					
261	coheir					
262	playes					
263	inmportant					
264	differenciate					
265	HIV- infected					
266	wheter					
267	goes (go)					
268	Altough					
269	emostions					

270	behavior					
271	monotoring					
272	imprisonated					
273	imporant					
274	someone;s					
275	publically					
276	defense					
277	twenty- four					
278	atleast					
279	cliënts					
280	now a days					
281	There am I for					
282	then (than)					
283	inforce					
284	un acceptable					
285	Sosial					
286	adolosent					
287	sosialmedia					
288	sould					
289	surten					
290	beoming					
291	sicial					
292	goverment					
293	soccial					
294	acording					
295	exsessive					
296	regulatoin					
297	diffrent					
298	clossing					
299	ress					
300	oppinion					
301	throung					
302	negitive					
303	kommunication					

304	Cina					
305	speach					
306	surtend					
307	exctend					
308	priventing					
309	Conclution					
310	nesisary					
311	rumors					
312	South- Africa					
313	must rather been seen					
314	There is to many uncontrolled					
315	can not					
316	implimenting					
317	regualtion					
318	dangours					
319	intsgram					
320	deplicted					
321	curiousity					
322	disasteress					
323	thruth					
324	forture					
325	preposterious					
326	fueling					
327	Insitute					
328	alienize					
329	publically					
330	SOLEMENELY					
331	have took place over the years					
332	lesquerella					
333	ites					
334	are been monitored					
335	happenning					
336	scandels					
337	imagies					