LEXIS AND THE UNDERGRADUATE: ANALYSING VOCABULARY NEEDS, PROFICIENCIES AND PROBLEMS

by

PATRICIA ANNE COOPER

submitted in fulfilment of the requirements for
the degree of

MASTER OF ARTS

in the subject

LINGUISTICS

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF E H HUBBARD

JANUARY 1999

Abstract

Although vocabulary is widely acknowledged as a key to second language proficiency, the precise role of the second-language student's English vocabulary at the level of tertiary education is currently being debated. This study concerns research into the relationship between students' vocabulary size and their performance at undergraduate level, with particular focus on academic vocabulary, which comprises a significant component of the vocabulary used across a range of academic disciplines. The assumption underlying this study is that students' scores on a test of vocabulary will correlate significantly with their final marks. The aim was to establish to what extent vocabulary serves as an index of academic performance. Results suggest that a measurement of academic vocabulary will reflect trends in academic performance rather than providing a clear indication of student potential.

Key terms

Vocabulary size; vocabulary testing; academic performance; reading comprehension; language proficiency; basic vocabulary; academic vocabulary; advanced vocabulary; second-language learners; undergraduate students; tertiary education.
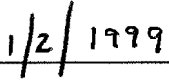
Declaration

Student number: **797-526-0**

I declare that LEXIS AND THE UNDERGRADUATE: ANALYSING VOCABULARY NEEDS, PROFICIENCIES AND PROBLEMS is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

P A Cooper

Ms P A Cooper

1/2/1999

Acknowledgements

Table of contents

Chapter 3: Research method

Chapter 4: Results

Chapter 5: Conclusion

Chapter 1

Vocabulary and academic performance

1.1   Introduction

> Words are a means of organizing experience, of marshalling ideas, and the ability to use them
> efficiently will not only improve your own capacity for assimilating new ideas and for passing
> them on and making them work for you, but will also enable you to make that capacity known
> to other people.
>
> (Wenborn, 1981:10)

Disciplines in the Arts require that scholars aspire to the roles of philosopher, orator and wordsmith. The power of words within academic debates is reflected in, among others, our ability to explain and defend a point of view. In this sense the mental lexicon[1] may be regarded as a deciding factor in influencing others of our understanding and competence, in the assessment of our academic performance and in determining our scholastic success. To what degree are undergraduate students equipped with the vocabulary skills required of them in the academic environment? Although vocabulary is acknowledged to be a key component of second language proficiency, the precise role of the second-language student's English vocabulary at the level of tertiary education needs to be considered in detail. This dissertation is motivated by the need to examine the role of vocabulary in the academic context. Given the common assumption that vocabulary is relevant to academic performance, this study attempts to establish to what degree this is the case. In the event of there being a direct relationship between vocabulary size and academic performance, it

---

[1] The terms *mental lexicon* and *lexicon* are used interchangeably in this study to refer to "the speaker's mental representation of all the semantic, syntactic and phonological specification of the lexical items in a language" (Laufer-Dvorkin, 1991:4).

may be possible to estimate students' potential performance from a measure of vocabulary size. The extent to which vocabulary size estimates academic performance will depend on the strength of the relationship between these variables. Given a significantly strong correlation it may be feasible to use a vocabulary test to identify students who require additional tuition. In other words, the vocabulary test could serve as a test of broader linguistic and cognitive skills. The study is, therefore, justified by the number of questions surrounding the issue of vocabulary size and the possibility of improving the students' performance. Simply stated, if vocabulary size does have a significant influence on academic performance, it should be feasible to design courses aimed at developing vocabulary which would facilitate students' comprehension of the reading material and improve the quality of their writing.

The purpose of this chapter is to identify the issues which comprise the background to this study. In order to do so I will advance the key questions which prompted my research. In addition, I will outline the research methods, aims, objectives and hypotheses of this dissertation, and provide a synopsis of its structure.

1.2   Identifying key questions

Without a clear understanding of the components of language proficiency, we cannot adequately affect or influence language competence. The resurgence of interest in vocabulary has led to a questioning of the degree to which the vocabulary of the language user affects overall linguistic competence. This has in turn led to specific conjectures that vocabulary size affects academic performance (cf. Anderson and Freebody, 1981; Diack, 1975; Meara and Jones, 1987). If one considers the correlation between the number of first- and second-language speakers of English who register at universities in South Africa and the relative success of each group at the end of their first year of study, it is apparent that a significantly larger proportion of mother-tongue students is successful. A number of socio-economic and political factors, rooted in the educational policies implemented during the era of the apartheid government, have contributed to this discrepancy. The

policies surrounding Bantu Education in particular have resulted in a legacy of educational deprivation for black students. These students, as second-language speakers of English, represent the 'linguistic minorities' who continue to suffer the effects of a racist segregation policy and unequal distribution of resources.

One of the factors which may contribute to the academic difficulties experienced by second-language speakers is poor linguistic competence which results in inadequate handling of the subject material (Blacquiére, 1989; Macdonald, 1990; Perkins, 1991; Pretorius, 1995; Vorster and Piper, 1995; Vorster and Reagan, 1990). A primary reason for students struggling through their undergraduate courses is an inability to cope with the reading material prescribed for each course (Pretorius, 1995:33). To what extent is this problem rooted in an inadequate vocabulary? Clearly a discrepancy between the level of vocabulary expected from the first-year student and the student's actual command of lexis places the second-language student at a significant disadvantage. This discrepancy and its detrimental consequences are all too apparent in the written assignments I have assessed in my capacity as a lecturer in the Department of Linguistics at the University of South Africa which show that the second-language student's vocabulary frequently does not meet the requirements of academic writing. This inadequacy is reflected in, among others, a range of lexical errors, including the use of semantically vacuous lexical items[2], a lack of awareness of selection restrictions which results in inappropriate collocations, semantic errors such as the use of inappropriate synonyms and the misuse of terminology (cf. Savini, 1992). The effect of such errors on academic performance becomes apparent when the response of markers to errors in the writing of second-language students is examined. Santos (1988) evaluated university teachers' reactions to academic writing in terms of comprehensibility, irritation and acceptability, and found that "lexical errors were the most serious" (1988:69). In examining the pedagogical implications of this finding, Santos

---

[2] The terms *lexical item* and *lexeme* are used interchangeably in the literature to refer to the abstract form underlying all grammatical variants. For the purposes of this study, however, the terms *lexeme* and *lemma* are used to refer to the base form of a word (or headword) which is the abstract form underlying both the inflected and derived forms in a word family. Simply stated the lexeme is "the least complex form of a word, without affixes" (Schmitt and McCarthy, 1997:327). The term *lexical item* is used more broadly to refer to all forms of words, including base forms, inflections, derivations and multi-word units which convey a single meaning. *Word* is used in the same sense as lexical item.

(1988:69) proposes that more emphasis should be placed on vocabulary improvement and lexical selection within ESL writing courses.

If we accept the premise that reading plays a significant role in academic studies, then we must recognise that an adequate knowledge of vocabulary is basic to the understanding of text. Text comprehension derives from the integration of a number of features, including micro- and macro-levels of processing, the perception of semantic unity, rhetorical conventions, connectedness and the ability to infer meaning from the text (Pretorius, 1995:37). Is vocabulary knowledge simply a matter of automatic word recognition at the micro-linguistic level? To what degree is a knowledge of content (in addition to linguistic knowledge) required to interpret the meaning of words within the text? Alternatively, to what degree does an assessment of vocabulary knowledge provide a measure of content knowledge? Given the emphasis on the role of subject-specific knowledge in text comprehension, I would argue that an understanding of the vocabulary that occurs with high frequency within a particular discipline implies a broader understanding of that discipline. If this argument is valid, then to what degree can a measure of vocabulary size be used to estimate academic performance? Although vocabulary forms only one component of language, tests of vocabulary size have been found to relate to language proficiency in general (Diack, 1975; Meara and Jones, 1987; Nation and Waring, 1997; Read, 1997). In the case of second-language speakers, a measurement of vocabulary size reflects the gap between their lexicons and those of mother-tongue speakers and so provides some indication of the status of their interlanguage. What, then, is the average vocabulary size of undergraduate students studying through the medium of a second language? In contrast to this, what vocabulary size is expected of the second-language undergraduate student? These questions refer specifically to the potential gap between the vocabulary found in the prescribed reading material and that in the student's lexicon. The issue addressed is thus the degree to which students' vocabularies correspond with the extent of vocabulary knowledge required to adequately comprehend the course material.

The questions posed above suggest an urgent need for research into the relationship between vocabulary size and academic performance. Furthermore, the nature of the

discrepancy between students' command of lexis and the vocabulary requirements of the study material needs to be clarified if relevant changes in coursework are to be effected. Since vocabulary size presumably affects reading and writing, which in turn would influence academic performance, can vocabulary proficiency be assessed in terms of reading and writing skills in particular, and academic performance in general? The following general questions are appropriate departure points for a study intended to approach these problems:

- How is vocabulary size measured?
- To what extent does vocabulary size correlate with academic performance?
- Is there a relationship between the command of basic, academic and advanced vocabulary, and academic performance?
- To what extent does the students' vocabulary meet the requirements of the prescribed reading material?
- What is the best means to assess overall competence in vocabulary (vocabulary proficiency)?

In addressing these questions in the course of this dissertation I will

- suggest the most appropriate method by which to measure the size of students' vocabulary;
- assess the relationship between vocabulary size and academic performance through a specific study;
- establish the degree to which students' command of vocabulary is at odds with that required by the prescribed reading; and
- examine the effect of vocabulary size on other areas of language competence such as reading comprehension and quality of writing as a means of assessing proficiency.

The primary aim of this study is to establish the extent to which vocabulary size affects academic performance. The goals outlined above are set as the criteria in terms of which this aim may be realised.

This study relies on a body of previous research into the nature and role of vocabulary in second-language learning and within the context of broader academic issues. Aspects of this research which relate most closely to the question of vocabulary size are outlined below.

1.3   Preliminary comments on previous research

Vocabulary is a component of language proficiency that received relatively little attention until the 'vocabulary boom' of the 1980's, as illustrated by the amount of research on vocabulary emanating from this time (Anderson and Freebody, 1981; Carter, 1987 and 1991; Carter and McCarthy, 1988; Laufer, 1986 and 1997b; McCarthy, 1984; Meara, 1987; Nation, 1982; Read, 1988). Laufer (1986:70) speculates that the reason for this neglect results from the specific focus of language acquisition research.

> The linguists have preferred to study grammar and phonology since these are closed systems and therefore lend themselves to much more abstraction and generalization than vocabulary, which is not a closed system but an open set.

This perspective was inflected with exaggerated estimations of the size of the mental lexicon of mother-tongue speakers. Traditional methods of testing vocabulary are based on dictionary sampling, in which lexical items are drawn randomly from the entries in a first-language dictionary. Students are then tested on their knowledge of the selected items. The number of words 'known' by the students is assumed to correspond proportionately to the total number of words in the dictionary. These methods (discussed in detail in §2.3.3) have resulted in figures ranging from approximately 74 500 words (Oldfield, 1963:125) to 157 000 words (Seashore and Eckerson, 1940:33). As Goulden,

Nation and Read (1990:342) observe, these results are particularly significant when seen in terms of the number of words to be acquired by the language learner per day: "Diller (1978), for example, estimates that secondary school children learn 20,000 words a year. This works out to be a rate of around 60 words a day!" However, criticism of the methods used to determine vocabulary size has led to more reasonable estimates, which suggest that, at tertiary level, mother-tongue speakers of English have a vocabulary of approximately 17 000 base words (Goulden *et al.*, 1990; D'Anna, Zechmeister and Hall, 1991).

The discrepancies in earlier vocabulary size estimates have given rise to three central questions concerning vocabulary testing:

- How a word is defined. (What is a word?)
- How one establishes whether a word is considered part of the mental lexicon. (What does it mean to know a word?)
- How dictionaries are selected and what sampling procedures are used. (How should a sample of words be selected?)

One of the core issues to be addressed in this study centres on how many words second-language students have in their vocabularies. In examining those factors which have influenced estimates of vocabulary size, my intention is to identify the central problems surrounding vocabulary testing. Previous studies on the vocabulary of first-language speakers relates closely to research into the vocabulary size of second-language speakers. One advantage to more accurately determining the vocabulary size of English speakers is that this provides the learner with a clearer goal of what is required for lexical proficiency. For instance, it has been established that, of the 17 000 words proposed by Goulden *et al.* (1990) and D'Anna *et al.* (1991), 2000 form a basic, core vocabulary that is used with high frequency across a range of texts (cf. Nation, 1990). This finding allows for the development of a structured vocabulary learning programme for second-language speakers since the objectives of the learning task are clear and teaching methodologies can be goal-driven. A second advantage relates to the relationship between vocabulary size and

reading comprehension (§2.4), and the implications for comprehension where a large proportion of words in a text are unknown. These issues are addressed in Chapter 2.

1.4 Aims and objectives

The general aim of this study is:

- To determine the degree of correlation between vocabulary size and academic performance.

The secondary aims relate directly to the secondary hypotheses, and are:

- To determine the degree of correlation between basic, high frequency vocabulary size and academic performance.
- To determine the degree of correlation between academic vocabulary size and academic performance.
- To determine the degree of correlation between advanced vocabulary size and academic performance.

These aims were addressed by means of a number of specific, short-term objectives: to perform an analysis of the vocabulary in the reading material in terms of frequency levels; to devise a vocabulary test in order to assess the size of students' basic, academic and advanced vocabularies; to compare the results of the vocabulary test with academic performance by correlating the basic, academic and advanced vocabulary scores against each student's final mark; to conduct a multiple regression analysis in order to determine the extent to which vocabulary size contributes to academic performance; and, finally, to establish the relative performance of students in the tests of basic, academic and advanced vocabulary.

These objectives were addressed through deductive research in which hypotheses are formulated and then tested using inferential statistics. The hypotheses on which this study is based are stated below.

1.5   Research hypotheses

This section outlines the assumptions underlying the general and secondary research hypotheses. Each hypothesis is reviewed in detail in Chapter 3 (§3.2).

1.5.1   General hypothesis

There is a significant positive correlation between students' scores on a test measuring vocabulary size and their eventual academic performance.

This hypothesis predicts that academic performance, as measured by the final mark, is related to vocabulary size. It relies on the assumption that one of the factors contributing to academic failure is an inadequate command of the vocabulary required by the discipline. The validity of this hypothesis will be tested by comparing students' scores on a vocabulary test with their final mark using Pearson's product-moment correlation coefficient.

1.5.2   Secondary hypotheses

The three secondary hypotheses which derive from the general hypothesis focus in turn on the relationship between the basic, academic and advanced frequency levels, and academic performance, as stated below:

HA   There is a significant positive correlation between students' scores on a test measuring basic, high frequency vocabulary and their eventual academic performance.

HB   *There is a significant positive correlation between students' scores on a test measuring academic vocabulary and their eventual academic performance.*

HC   *There is a significant positive correlation between students' scores on a test measuring advanced vocabulary and their eventual academic performance.*

These subsidiary hypotheses predict that academic performance will correlate specifically with measures of basic, academic and advanced vocabulary size. The assumption underlying these hypotheses is that the vocabulary at each frequency level serves a particular function within academic discourse and, as such, contributes individually to academic performance. The validity of these hypotheses was again tested using the Pearson product-moment correlation coefficient. In each case the vocabulary test score served as the independent variable, with the final mark as dependent variable.

1.6   Research procedure

A pilot study was conducted prior to the main study in order to determine whether the relationship postulated between vocabulary size and academic performance justified further investigation. The assumption underlying the pilot study was that there is a discrepancy between the vocabulary expected from the undergraduate student and the student's lexical proficiency, particularly in the case of weaker students. In order to test this assumption, the study examined the nature of the relationship between the vocabulary size and academic performance of a small sample of first-year students. A vocabulary test was designed based on the vocabulary in the prescribed reading material, which was analysed and graded into frequency levels to determine the approximate level of difficulty of each lexical item. The test results were then compared with the students' examination results. The results of this study indicated a significant relationship between vocabulary size and academic performance. I inferred from this finding that vocabulary was a reliable indicator of academic performance and, as such, might prove an appropriate and efficient mechanism for identifying students in need of additional or bridging tuition.

In order to achieve the aims of the study, two multiple-choice vocabulary tests were compiled and administered to a sample of first-year students of education. The tests were designed to assess vocabulary size in order to compare it with academic performance, as determined by the final mark. The test items were drawn from the course material which was analysed by means of the Lexical Frequency Profile programme (Laufer and Nation, 1995). This programme classifies words according to frequency of occurrence, so was used to identify the test items as basic, academic or advanced vocabulary (§2.5.1). The results of these vocabulary tests were analysed using Pearson's product-moment correlation coefficient as well as a stepwise multiple linear regression. The correlation analysis was intended to indicate whether the relationship between vocabulary size and academic performance was significant. The regression analysis was used to assess the degree to which the independent variables (the scores for basic, academic and advanced vocabulary) contribute towards the dependent variable (the final mark), and was thus intended to determine the extent to which vocabulary size serves as an indicator of academic performance.

1.7   Structure of the dissertation

This chapter identifies the problems which will be addressed in this study and provides a rationale for the research. In addition, the aims, objectives and hypotheses of the study are outlined and a brief explanation of the research procedures is provided.

Chapter 2 is a review of relevant literature. Studies of the role of vocabulary in reading comprehension and the research surrounding vocabulary testing are examined. Issues related to vocabulary testing, such as the distinction between 'active' and 'passive' vocabulary knowledge and the concept 'word family' are discussed, and the validity of the Lexical Frequency Profile as a measure of vocabulary proficiency is reviewed. This chapter also addresses related areas in the field of vocabulary research. These include learning vocabulary within an academic context, the academic word list, the capacity of

a student to determine the meaning of unknown words from context and the significance of frequency counts.

Chapter 3 provides a rationale for each of the hypotheses and describes the testing procedures used to assess vocabulary size. It then goes on to explain the research design in terms of which the vocabulary size and academic performance of the students were examined, both for the pilot study and the main study. In conclusion it outlines the statistical procedures used to analyse the data.

Chapter 4 presents the results of the study and interprets these results in the light of previous research.

Chapter 5 concludes with a summary of the major findings, an analysis of the limitations of the study and an outline of implications for further research into the relationship between vocabulary size and academic performance. It concludes with a suggestion of possible applications within the context of tertiary education.

Appendix A contains the vocabulary tests administered to the students in this study.

Appendix B outlines the principles underlying hypothesis testing, including the formulation of the null hypothesis, the significance level in terms of which the null hypothesis is evaluated, and the distinction between directional and non-directional hypotheses.

Appendix C contains a discussion of the statistical techniques used in the analysis of the results, and outlines the procedure followed in the multiple regression analysis.

Appendix D tabulates the results of the analyses of each vocabulary test item.

Chapter 2

A measure of vocabulary

2.1   Introduction

The aim of this chapter is to investigate the issues surrounding vocabulary testing, the methods used to test vocabulary size, the concept 'word family', and the validity of the Lexical Frequency Profile as a measure of vocabulary proficiency. These issues provide a framework in terms of which the hypotheses have been formulated, and so serve to justify the problems identified in this study. In order to situate this study within a broader theoretical as well as empirical framework, the role of vocabulary in reading comprehension will be examined. This chapter is divided into three sections:

(a)   A review of the research into the relationship between vocabulary size and academic performance.

(b)   An investigation into the testing of vocabulary size.

(c)   A review of the role of vocabulary in reading comprehension.

Bauer and Nation's (1993) concept of 'word family' is central to the definition of a word, and the criteria in terms of which they include a word in a word family will be critically examined. The methods used to test vocabulary size and some of the problems inherent in these approaches are discussed, with an examination of the difficulties involved in selecting words for testing. Because of the integral role that vocabulary plays in reading comprehension, this chapter includes a discussion of what vocabulary size is required for reading comprehension.

2.2   The relevance of vocabulary within the academic context

> The basic dimension of lexical competence is size. All other things being equal, learners with big
> vocabularies are more proficient in a wide range of language skills than learners with smaller
> vocabularies, and there is some evidence to support the view that vocabulary skills make a
> significant contribution to almost all aspects of L2 proficiency.
>
> (Meara, 1996b:37.)

The view that vocabulary plays an essential role in the academic context is supported by studies such as that of Anderson and Freebody (1981), Corson (1983), Cummins (1981, in Goulden *et al.*, 1990:342), Cunningham and Moore (1993), Meara and Jones (1987), Nation (1990), Saville-Troike (1984) and Sternberg and Powell (1983). It appears that "learners generally agree that many of their difficulties in both receptive and productive use [of language] arise from their inadequate vocabulary knowledge" (Kang, 1995:43).

In one of the more significant studies relating to the question of the role of vocabulary within the academic context, Meara and Jones (1987) developed a placement test based on an assessment of vocabulary size. They found a significant correlation between the vocabulary test scores and a purpose-designed placement test, indicating "a strong link between vocabulary size and other language abilities" (Meara and Jones, 1987:15). Similarly, Anderson and Freebody (1981:77) propose that "[m]easures of vocabulary knowledge are potent predictors of a variety of indices of linguistic ability". In support of this claim, they cite a number of findings which indicate a high correlation between vocabulary test scores, on the one hand, and the results of either intelligence or achievement tests, on the other.

The study by Cunningham and Moore (1993) was designed to investigate students' reading comprehension performance in response to questions which differed only in terms of the type of vocabulary used, i.e. academic versus basic vocabulary. "Differences between means indicated that academic vocabulary in comprehension questions significantly decreased question-answering performance" (Cunningham and Moore, 1993:17). In addition, studies such as those of Astika (1993), Laufer (1994) and Linnarud (1986)

illustrate the importance of vocabulary in writing as a component of academic proficiency. This research illustrates the degree to which poor proficiency in academic vocabulary can result in poor scholastic achievement.

The predominating view, examined below (§2.4), is that lexical competence primarily plays a significant role in reading comprehension. However, on the basis of Grabe's claim that "reading is probably the most important skill for second language learners in academic contexts" (Grabe, 1991:375), it must be inferred that lexical competence is thereby fundamental to academic performance on the whole. Nevertheless, despite evidence in support of this view, it cannot be assumed that a cause-effect relationship between vocabulary and academic performance necessarily exists (Saville-Troike, 1984:200). This study is therefore intended to investigate to what extent vocabulary size is an independent factor in determining students' academic performance.

2.3   Measuring vocabulary size

What is the average vocabulary size of adult English speakers? And why has answering this question proved to be so problematic? Results of studies into vocabulary size, which range from estimates of 3000 words (Fries and Traver, 1960) to 216 000 words (Diller, 1978, in Goulden *et al.*, 1990:342), have shown enormous discrepancies over a relatively short period. Two principal reasons for the widely varying estimates in answer to the question "How many words does a speaker know?" relate to the operational definitions of the concepts 'word' and 'know'. This section examines the problem of how to define a word, and what it means to know a word, within a broader discussion of the procedures used in the design of a vocabulary test, i.e. the nature and size of the dictionary or frequency list used as the target corpus, the sampling technique used to select test items from the dictionary or frequency list, and the method used to test the selected vocabulary items.

The considerable amount of research conducted into the question of vocabulary size stems, to a large degree, from the practical applications which relate to more accurately determining the size and nature of the native speaker's mental lexicon. One of the more direct applications of providing an accurate estimation of vocabulary size, as indicated by Goulden *et al.* (1990:342), is an indication of how many words a second-language learner needs to understand, and at what rate, in order to attain a degree of first-language proficiency within a reasonable period. A second application relates to the significant correlation between vocabulary size and reading comprehension (§2.4). In a paper on the relationship between vocabulary and reading, Nation and Coady (1988:97) assert that "in measures of readability of a text, vocabulary difficult has consistently been found to be the most significant predictor of overall readability". From their review of the literature, it seems clear that if the density of unknown words exceeds a certain level, a decrease in comprehension results. Thus the number of words that readers know will have a direct bearing on their ability to make sense of a text. (This argument is later moderated on the grounds that vocabulary size impedes reading comprehension only in the beginning stages of second-language learning (Brisbois, 1995).)

The conclusion drawn from previous estimates of vocabulary size (§1.3) was that vocabulary must also be learnt from context since teaching vocabulary lists in the classroom can have no significant impact on linguistic competence (Jenkins and Dixon, 1983; Nagy, Herman and Anderson, 1985). However, one implication of the finding that the educated first-language speaker has a vocabulary of approximately 17 000 base words (D'Anna *et al.*, 1991; Goulden *et al.*, 1990; Hazenberg and Hulstijn, 1996), is that vocabulary lists should be considered a valid teaching aid. This claim is based on the grounds that words of high frequency could be taught at an early stage, and would then serve as a framework for the integration into the lexicon of the less frequently occurring words through context.

Given the assumption that vocabulary size to a large extent determines the available vocabulary for the writing or understanding of a text, it may be inferred that quality of writing, level of comprehension when reading, and consequently academic performance

are related to vocabulary size. Previous research (Cunningham and Moore, 1993; Diack, 1975; Engber, 1992; Laufer, 1992; Laufer and Nation, 1995; Linnarud, 1986; Meara and Jones, 1987) has shown some correlation between the extent of a student's vocabulary, measured either in terms of lexical richness, or in terms of vocabulary size in relation to dictionary samples and frequency lists, and academic performance. It is therefore essential to determine whether the correlation between vocabulary size and academic performance is significant, and following from this, to assess the vocabulary size of second-language speakers as opposed to first-language speakers, and to determine the areas in which their vocabularies are significantly different.

Three main issues in vocabulary testing relate to the reasons for the discrepancies in earlier vocabulary size estimates:

- The criteria in terms of which a word is defined. (What is a word?)
- The criteria in terms of which a word is considered to be part of the mental lexicon. (What does it mean to know a word?)
- Inconsistencies in the selection of dictionaries and the sampling procedures used. (How should a sample of words be selected?)

2.3.1 What is a word?

A critical aspect of vocabulary testing is the question of how to define a word. Do inflectional forms such as *eaten, eating* and *ate* constitute three separate words or a single word family? Is it valid to assume that a speaker who knows the word *justify* will understand the derivation *justification*? Carter (1987:4-7) examines several possible definitions, including the orthographic definition in terms of which a word is a sequence of letters bound on either side by a space, the free-form definition in terms of which a word is an autonomous linguistic feature, and the stress-based definition in terms of which a word contains a maximum of one stressed syllable. Carter argues that, while these definitions may appear acceptable, further analysis shows they do not accommodate all word forms in the language. The orthographic definition, for example, does not distinguish

between the multiple meanings of polysemous items such as *head* and *watch*, and would classify inflectional forms such as *wrote* and *writing* as separate word forms, although these are included in a single entry in a dictionary. And, while stress commonly forms one of the criteria in terms of which spoken words may be distinguished in English, unstressed weak forms, such as *the*, also occur. A problem common to all the definitions outlined above is that they do not recognise fixed expressions such as idioms and clichés. A phrase such as *burning the candle at both ends* is intelligible only when seen holistically, since the expression becomes meaningless when each word is interpreted individually.

Carter (1987:6-7) proposes that vocabulary should be defined in terms of a lexeme, an abstract form underlying all grammatical variants. Thus the lexeme CATCH will form the bold-faced dictionary entry in terms of which the inflected forms *catching* and *caught* are defined. The concept 'lexeme' also allows for polysemy in words, where two separate lexemes would represent the related meanings of a word, for instance, *dog*: (a) the domesticated animal, and (b) to follow closely on someone's heels. Read (1988:14) provides the following argument in support of Carter's approach:

> ... one has to decide whether a 'word' is an individual word form or a word family (or lemma) consisting of a base form together with the inflected and derived forms that share the same meaning. Including all such forms as separate words will clearly increase the estimate of vocabulary size, whereas a more conservative approach results in a substantially lower figure. The latter approach seems more realistic, even though it requires a careful definition of criteria for grouping words into families ...

Bauer and Nation (1993:253) describe the concept 'word family' as consisting of

> ... a base word and all its derived and inflected forms that can be understood by a learner without having to learn each form separately. So, *watch, watches, watched*, and *watching* may all be members of the same word family for a learner with a command of the inflectional suffixes of English. As a learner's knowledge of affixation develops, the size of the word family increases.

Thus once the base form of an English word is known, recognition of the inflected and derived forms of the word family is dependent simply on knowledge of affixation. Since

knowledge of a base form should enable a learner to understand the inflected and derived forms in the word family, the meanings of the base and derived forms in a word family must be closely related. The words *art* and *artery*, for example, would therefore not be considered members of the same family, since *artery* is not derived from *art*. Bauer and Nation's notion of a headword as the base of a word family thus corresponds closely to Carter's definition of a lexeme.

Bauer and Nation (1993:255-256) classify affixes such as the inflectional forms *-s, -ed* and *-ing*, and the derivational forms *-ly, -less* and *-tion* into different "levels" according to their phonological and morphological behaviour. Each level is ordered on the basis of certain criteria, such as the frequency of occurrence of the affix, the level of productivity (the extent to which the affix is responsible for the formation of new words), the degree to which the meaning of the affix can be predicted, regularity of form and regularity of function. This definition of 'word family' is therefore dependent on levels of morphological awareness, graded according to the frequency, productivity, predictability and regularity of English affixes. Level 2, for example, consists of all the inflectional suffixes, such as the plural and past tense forms. At Level 3 are the most frequent and regular derivational affixes including *-er, -ish* and *-less*, while Level 7, on the other hand, contains the classical roots and affixes which have to be learnt individually, for example, *embolism*.

2.3.2 'Family planning': Applications of word families

Bauer and Nation's (1993) classification of word families is fundamental to the development of vocabulary tests, since the lack of consistency in defining a word or word family has previously led to the extreme variations in estimates of vocabulary size discussed above. They argue, for instance, that where the vocabulary items in a test have been selected from words listed according to frequency of occurrence, subjects may appear to know an unrealistic number of low frequency words where these are derived from a high frequency base form. The subject is then able to infer the meaning of the test item because of a knowledge of the base form:

Because the learners sitting the test can see the derived-base connection, they succeed on that item, whereas base words of the same low frequency level are not known. If the test was based on a suitable word family description, this problem would be avoided and the expected implicational scaling [between the various frequency levels] would be more likely to occur.

(Bauer and Nation, 1993:265-266.)

Implicational scaling refers to the ratio between frequency level and vocabulary size. It is expected that, as the frequency of a base form decreases, so the number of words known at the lower levels of frequency will reduce. ('Frequency' refers to how commonly a word occurs in any text.) Thus, in terms of implicational scaling, high frequency words that occur regularly across texts, such as *the, and, with, house* and *bread*, will be familiar to the learner from an early stage in the acquisition of a language, while low frequency words, which are more rare, and are often restricted to specialised texts, for example, *acquiesce, endorse, obsequious* and *neologism*, are far less likely to be familiar to the learner. The principle underlying the relationship between word frequency and level of difficulty is that, in order to become known, a word must typically be encountered several times in context (Coomber and Peet, 1993; cf. Meara, 1996b; Nagy, 1997). Thus words which occur frequently are more likely to be assimilated into the learner's lexicon, while low frequency words, being relatively uncommon, are more difficult to assimilate. For example, Hazenberg and Hulstijn (1996:149) found that, in a "42-million-word-token corpus of contemporary written Dutch", only four words had a frequency of over one million occurrences, and covered 13.4% of the text, while the next 54 words had a frequency of over 100 000 occurrences, and covered 45.6% of the corpus. On the other hand, the majority of words (21 053 base forms) had a frequency of less than 5 occurrences, and covered less than one percent of the corpus. Studies by Sutarsyah, Nation and Kennedy (1994) and Nation and Hwang (1995) have shown that the highest frequency words (the first 2000 word families) provide coverage of 80% of most written texts, from which it may be inferred that the estimated additional 56 000 base forms of English, excluding proper nouns, compounds, affixes and homographs with unrelated meanings (Goulden *et al.*, 1990) have a negligible chance of occurring in the remaining 20% of the text. Thus the possibility of acquiring the low frequency words is considerably reduced. This is further exacerbated by Worthington and Nation's (1996:2) claim that the University Word

List "provides around 8.5% coverage of the running words [in academic texts]". Added to the 80% coverage of the 2000 high frequency words, this would provide a coverage of 89%, which suggests that approximately 11% of the running words in academic texts are of low frequency.

The relevance of the concept 'word family' to vocabulary testing may be illustrated, firstly, by briefly examining a number of studies in which the researcher has not defined a word in terms of a word family, and, secondly, by examining the degree to which the concept 'word family' forms the basis for dictionary entries, since the dictionary is conventionally the source of the vocabulary sample to be tested.

Since the focus of this section is on the discrepancies in estimates of vocabulary size, it is necessary to examine different notions of what constitutes a word across a range of studies. Goulden *et al.* (1990:343), for instance, cite a study in which vocabulary size was estimated to be 216 000 words:

> Diller (1978) assumed that the largest Webster's dictionary contained 450,000 entries, which is the figure given in the preface of the dictionary. He selected a sample of 1,000 words to be tested. when he administered the test to high-school seniors, the median score was 480 out of 1,000, or 48 per cent. He thus calculated that their total vocabulary size was 216,000 words: 48 per cent of the 450,000 entries in the dictionary.

Goulden *et al.* allege that Diller's estimation of vocabulary size was based, first of all, on an unquestioning acceptance of the claim that *Webster's New International Dictionary* (1961, in Goulden *et al.*, 1990:358) contained over 450 000 words (cf. Nation, 1993). The second assumption made by Diller was that each entry represented a separate word. Testing a percentage of all bold-faced entries, which included inflected forms as separate words, therefore led to the greatly inflated estimate given above. Similarly, in a study by Seashore and Eckerson (1940:15), a word was defined as "an item listed in a standard unabridged dictionary", in other words, a dictionary entry. On this basis they conclude that "the average college undergraduate in our groups recognised ... a grand total of 155 736 words" (Seashore and Eckerson, 1940:33). In a study by Carroll, Davies and Richman

(1971), a word was regarded as a form, with a change in capitalisation or inflection being viewed as a different word. On this basis a corpus of 86 741 words was developed in the examination of "textual samples from published material to which students are exposed in grades 3 through 9" (Carroll, Davies and Richman, 1971:v). Thorndike and Lorge (1944), in compiling *The teacher's word book of 30 000 words*, regarded inflectional forms as members of a single word family, while derived forms were considered separate words. Goulden *et al.* (1990:155) criticise the Thorndike and Lorge list on the grounds that "of the 30,000 items in the list, only 13, 900 are base words". These studies serve to illustrate that basing the sampling procedure on the assumption that inflected and/or derived forms constitute separate words naturally leads to inflated estimates of vocabulary size. It is obvious that without a standard unit of measurement of vocabulary no fair comparisons of vocabulary size can be made (cf. Meara, 1996a).

Before following the procedure of dictionary sampling in compiling a vocabulary test, it is essential to determine whether the dictionary selected adheres to the principles underlying the concept 'word family' as the criterion for inclusion of a word. Lexicographers, with different aims and target users in mind, may apply different principles in selecting headwords and designing the layout of entries in the dictionary. In addition, dictionary editors are frequently inconsistent in the application of the criteria selected for inclusion of entries, even within a single dictionary, and in the implementation of policies regarding the inclusion of atypical words such as proper names. Thus estimates of vocabulary size depend heavily on the nature of the dictionary chosen for the sample. The importance of accurately determining the amount of words in the source is stressed by Meara (1996b:39).

> Estimating the number of words that make up the vocabulary you are interested in is probably *the* critical variable in constructing a test of vocabulary size. Obviously, if the test you are using suggests that a testee knows fifty per cent of the target vocabulary, it is vitally important to know how big or small this target vocabulary is. Fifty percent of 5,000 words is a lot less that fifty percent of 20,000 words.

With regard to dictionary entries, Bauer and Nation (1993:267) argue that "words formed using inflections and affixes at Levels 2 and 3 are the most likely candidates for non-defined subentries", while those words based on higher levels of affixation, such as *successful* (Level 4), *antibody* (Level 5), *acquisition* (Level 6) and *absolve* (Level 7), should be listed as separate entries. In the course of their discussion, Bauer and Nation (1993:267) also provide recommendations as to which word forms do not need to be included under the headword, whether prefixed derived forms such as *unimportant* should be included in the entry as well as separately, which affixes are indicated by the dictionary makers to be the most frequent and productive, and so worth learning, and, finally, recommendations regarding which of the affixes from Levels 3 to 7 should be listed; this list of affixes is provided as a "word-building device", and is intended to enable the user to generate derivational forms from the base words. For example, the knowledge that the suffix *-ant* changes the grammatical category from verb to noun would enable a reader to derive the Level 5 nominalised form *consultant* from the base word *consult*.

It is important, when examining the entries in a dictionary, to distinguish between full entries, defined sub-entries (where a derivational form of the headword or base form is defined) and non-defined sub-entries (where a derivational form of the headword is listed but is not defined), and to consider which derived forms have not been included under the headword, for example prefixed derivatives such as *insignificant*. In terms of Bauer and Nation's (1993) argument presented above, the dictionary entry for *cook* should include the headword as a full entry, the Level 2 inflectional forms *cooks*, *cooked* and *cooking* as non-defined sub-entries, the Level 3 derivation *cooker* as a non-defined sub-entry (although the British use of *cooker* to mean 'stove' probably falls outside Bauer and Nation's (1993:270) condition of "restricted use", since the *-er* suffix is used to indicate instrument, rather than agent, and so requires a defined sub-entry), and the Level 5 derivation *cookery* as a defined sub-entry.

An analysis of the entries listed under the headword *cook* in the Cambridge International Dictionary of English (1995) suggests that the criteria applied by the dictionary makers for the inclusion of a word correspond to the concept 'word family' (Cooper, 1997). This

dictionary, which is aimed at second-language learners and users of English, thus appears, on initial investigation, to fulfill the criteria stipulated by Nation (1993) in his examination of what is required for dictionary sampling procedures. These procedures are discussed below.

2.3.3 How should a sample of words be selected?

Two primary methods of sampling from a selected text, dictionary or list of words have been employed in the compilation of vocabulary tests. The procedures followed in the implementation of these methods, that is, frequency sampling and dictionary sampling, are outlined below.

2.3.3.1 Frequency sampling

Frequency sampling involves sampling words drawn at random from the higher to lower levels of a word frequency count. The frequency-sampling procedure is based on the assumption that words of high frequency, which occur commonly in a wide range of texts, will be most familiar to learners at all levels of proficiency. The low frequency, less common vocabulary items will be known by learners or native speakers at a higher level of proficiency. Those with a knowledge of low frequency words are likely to be widely read and to have studied in a specialised field at tertiary level (Diack, 1975). However, estimates based on this approach are restricted by the number of words found at a particular frequency level, as well as by the number of words in the overall count. Word frequency lists permit easy classification of words according to level of difficulty, but may obscure the relationship between base and derived word forms. This can inflate the resulting vocabulary estimates, where the meaning of a low-frequency word, such as *fabrication*, is closely related to the high-frequency word from which it is derived, i.e. *fabric*, thereby making an unrealistic number of low-frequency words accessible to the subject (Goulden *et al.*, 1990).

2.3.3.2 Dictionary sampling

Dictionary sampling has been used most extensively in vocabulary testing as the source of the word list (Kirkpatrick, 1891; Lorge and Chall, 1963; Nation, 1993). This method first requires an estimate of the total number of words in the dictionary selected. (With the result that one of the central problems related to dictionary sampling is the exaggeration of the estimated number of words in the dictionary for marketing purposes.) A representative sample of this total is then drawn from the dictionary, and subjects are tested on their knowledge of this sample. The subject's vocabulary size is estimated by multiplying their score by the ratio of the sample to the total number of words, on the assumption that the number of words known by each subject is proportional to the total number of words in the dictionary (D'Anna et al, 1991:110). However, due to the size and nature of the dictionary as well as the method of sampling, this procedure has resulted in widely varying estimates of vocabulary size.

Since the dictionary sampling method is based on the assumption that a selection of words from a dictionary will be representative of vocabulary size, it is essential that the dictionary is appropriate. It should have at least 30 000 base words (Nation, 1993:31), and be sufficiently modern to represent present-day English usage. The dictionary-based test should avoid sampling procedures that result in an unrealistic selection of high-frequency words. Dictionary sampling methods in particular are prone to show a predominance of high-frequency words, which accounts to a large extent for the overestimation of vocabulary size. Nation explains the bias towards high-frequency words in terms of the size of the dictionary:

> Procedures that involved choosing the first word on a page regardless of whether it was the first full entry and whether it was a subsequent homograph were particularly susceptible to this bias. This bias occurred simply because high-frequency words occupied more space per entry and had more entries than low-frequency words. The greater the size of the dictionary, the more space given to high frequency words.
>
> (Nation, 1993:28.)

Formulating a consistent definition of 'word' for the sampling procedure required in the design of any vocabulary test has become more complex in later research. The question of whether proper nouns, abbreviations and compound words should be included as separate vocabulary items or omitted from the item bank is somewhat controversial. Goulden *et al.* (1990) decided against the inclusion of such words on the grounds that they are not consistently represented across the range of dictionaries, and because compounds and abbreviated forms require minimal additional learning if the base words from which they are derived are known. They suggest that including such items in a vocabulary test will, again, inflate the estimated vocabulary size of the subjects. The issue of what constitutes a word, and whether or not to include proper names, technical terms, archaic words, abbreviations, compounds, and so on, in the sample relates closely to the size and nature of the dictionary or, to a lesser extent, the type of frequency list selected as the target corpus from which the sample of test items is drawn.

2.3.4   What does it mean to know a word?

Another area of debate in vocabulary testing is the question of what constitutes knowledge of a word. Crystal (1988) draws a distinction between active knowledge, referring to the ability to use the word appropriately in a sentence, and passive knowledge, where the word is recognised and to some degree understood in context. It has been postulated that passive and active vocabularies form two ends of a continuum (Palmberg, 1987, in Meara, 1996c and 1997). The basis of this model is that words pass along a continuum from a state of passive into active knowledge. Meara (1997:117) criticises this view of vocabulary acquisition on the grounds that it confuses state models with continua models, which "are fundamentally different", and so is inherently contradictory. Meara (1997:117) argues that the "continuum metaphor" is misleading, since it implies a smooth transition from one state to the next. In terms of Palmberg's model, however, words may be acquired as active vocabulary without passing through an initial passive state and are able to move around on the continuum in a non-linear fashion, although they do not appear to revert to a state of passivity from the active state. Meara (1997) recognises that passive and active vocabularies do form two separate states, but suggests that, in the process of acquisition,

"words pass through a number of discrete stages" (Meara, 1996c). Within this framework, Meara (1996c) has developed a multi-state model of vocabulary acquisition which operates on the principle that there are five discrete states of word knowledge, that words move freely between states, and that lexical proficiency or fluency is measured in terms of the speed at which a word is accessed. In this regard, Meara (1978, in Gass, 1989:94) maintains that the question of what it means to know a word is "best described in behavioral terms as the ability to react to a word in ways which are considered appropriate by the speech community". A language user would thus be deemed to understand a word if he or she is able to respond correctly to that word (according to the spoken or written context in which it occurs) within a minimal time frame. In support of this model, Meara (1997:118) proposes that the primary feature of vocabulary acquisition is the integration of a new word into the already existing lexicon:

> The main advantage of our model is that it allows us to think about vocabulary acquisition as a cumulative activity, rather than an all-or-nothing affair. Unknown words are words that have no connections of any kind to the learner's lexicon. Known words are words that are connected, but the number of connections may vary. This gives us a natural mechanism for talking about words which are 'known' to a greater or lesser extent. Poorly-known words are words with few connections, while better-known words are simply words with many connections.

Although Meara (1996c) does not clearly delineate the conditions under which transitions from one state to the next occur, his mapping of the movement of words in the lexicon and the connections between words is an attempt to explain the lexicon as a whole, rather than the knowledge of individual words. In addressing the lexicon rather than the word, Meara (1996c) appears to bring us one step closer to understanding lexical competence:

> A word in any state has a measurable chance of moving to any other state during a given time period. If these probabilities can be assessed for a particular learner, then we can predict long term development in the overall structure of the learner's lexicon.

Meara's model accounts for the distinction between passive and active vocabularies in terms of the nature of the connections between words. While words in the active

vocabulary are interconnected, passive vocabulary items are not related to other words in the lexicon by means of reciprocal connections, and so are activated only by external stimuli (Meara, 1990; 1997). Simply stated, the connections between words in the active vocabulary are multi-directional, while those in the passive vocabulary are uni-directional, occurring exclusively from the particular word to the network of items in the active vocabulary. *Feline*, for example, may be accessed through reading or hearing this word in context, but will not be spontaneously triggered by the words *cat* or *spite*, since it is not directly available via related items in the mental lexicon. A passive vocabulary item will become active when a reciprocal connection to that item from the active network is established. Meara (1997:120) argues that the transfer from passive to active may be transitory, "depending on which other parts of the lexicon are activated". This perspective on the nature of passive vocabulary means "that a new item can become active as a result of a single exposure, but also implies that there is no natural progression from a passive state to an active one" (Meara, 1997:120). This model does not address issues such as which words typically remain passive vocabulary items in the lexicon and the circumstances required to cross the threshold from passive into active. Another question in this regard is the ratio of passive to active vocabulary items which should be examined in relation to the distinction between the lexicons of first- and second-language speakers (cf. Laufer and Paribakht, 1998). This model does, however, reinforce the idea that the clue to understanding lexical competence lies in the organisation of the lexicon rather than in the assessment of words in terms of, for example, Richards' (1976) word knowledge framework.

In contrast to the somewhat simplistic dichotomy between active and passive vocabulary proposed by Crystal (1988), Richards (1976, in Carter and McCarthy, 1988:44) proposes that select criteria form the basis in terms of which a competent speaker has both a theoretical and a practical understanding of a particular lexical item. These criteria correspond closely to the requirements underlying depth of word knowledge outlined below, while Crystal's description of passive knowledge corresponds more closely to the idea of breadth of word knowledge. The distinction between breadth and depth of

vocabulary knowledge therefore relates to the distinction between active and passive (cf. Meara, 1996b:44), or receptive and productive (Nation, 1990:31-32), word knowledge.

The construct 'breadth of knowledge' assumes a limited knowledge of individual words, while the construct 'depth of knowledge' is more stringent, and requires an understanding of the word's spoken and written forms, referential, affective and pragmatic meanings, syntactic behaviour, inflectional and derivational forms, semantic field, including related synonyms, antonyms and hyponyms, as well as expected frequency and typical collocations. In terms of the criteria for depth of knowledge it would seem that knowing a word requires a theoretical and practical understanding of all its features. It is clear, however, that even first-language speakers are unlikely to be familiar with all aspects of many of the words in their mental lexicons. Meara (1996c) challenges the assumption that Richards' (1976) model can be used "as a framework for describing vocabulary knowledge". He argues that Richards' intention in devising the model was to establish "how current thinking in linguistic theory might inform classroom practice" (1996c), rather than to prescribe a set of features delimiting vocabulary knowledge. Meara examines the assumptions underlying Richards' model, and concludes that the focus is primarily on the knowledge of individual words, which he contends is neither feasible nor practical as an approach to the question of lexical knowledge. As indicated above, Meara proposes that a model of vocabulary knowledge should focus on the lexicon as a whole, rather than attempting to delineate knowledge of individual words: "[L]exical organisation might be an important aspect of lexical competence" (Meara, 1996b:48). The assumptions underlying Meara's holistic approach are probably best tested by means of the "word knowledge framework" (cf. Meara, 1996d; Read, 1993; Schmitt, 1995). Those tests devised around the word knowledge framework are examined below, together with various test formats designed to measure both breadth and depth of vocabulary knowledge.

2.3.5 Methods of testing vocabulary

In terms of vocabulary testing, it is necessary to decide which of the word's properties constitute knowledge of that word. In attempting to assess whether or not a word is part

of a speaker's mental lexicon, four test formats have been widely used: constructed answers, yes/no checklists, matching and multiple choice. Constructed answers require the subject to provide a definition, synonym, translation or illustration of the test item, or to use the word in an appropriate sentence to explain the meaning. Yes/no checklists require subjects merely to indicate whether a word is known (in which case the word is simply marked with a tick: ✓). The checklist format is validated either by the inclusion of non-words in the corpus, or by requiring the subjects to define a sample of the known words. Matching involves selecting the test item which most closely corresponds to an appropriate synonym or suitable definition from a number of options. Multiple choice has been widely used in vocabulary testing, since it enables the researcher to test a large number of words in a relatively short time. Compare, for example, the time that would be required to conduct a vocabulary test of 30 items in which the subject has to provide a constructed answer for each test item. With the exception of the constructed answer format, however, each of the above test formats focuses on breadth of knowledge, with very little attention given to depth. The multiple-choice format, for example, usually requires the subject to recognise a word synonymous to the test item, but does not directly assess any other properties of that word. A secondary problem is that this format may, in the choice of distractors, indirectly test words not related to the test item. The selection of distractors is critical to the test, as is illustrated in the following example from Anderson and Freebody (1981:103-104):

> An individual may select the correct synonym for *platitude* from the choices: a) duck-billed mammal, b) praise, c) commonplace remark, d) flatness. He may make the correct selection because he has heard the word used in reference to an utterance and with a negative connotation. This information, however, may not enable him to select correctly from a) commonplace remark, b) nonsense, c) irrelevant question, d) insult. The set of choices constrains the individual's response to different degrees, and different policies for generating distractors will, of course, lead to differences in performance.

The central problem with the multiple-choice format, then, is that the distractors do not necessarily compel the subjects to respond to the test item itself. One means to avoid this problem is the used of the checklist format, which Goulden *et al.* (1990) employ in their

study of receptive vocabulary. (Receptive vocabulary is also referred to as recognition vocabulary, and knowledge of a word in this sense merely requires that the subject is familiar with the word in context.) In the checklist or yes/no format the subject is presented with a list of words and simply asked to tick the ones that are known. There are a number of procedures that may be used to control overrating, i.e. preventing the subject from marking words simply on the basis that they have encountered the word previously, but would not be able to define the word, or have only a vague idea of the meaning. These procedures include requiring the subject to define a number of the known words after completing the test (Diack, 1975); including "nonwords" which may be used to check the reliability of the subjects' responses by ensuring that some of the test items could not be known words (Meara and Buxton, 1987); and using a scale in terms of which the subjects can rate their responses, from "have never experienced the word before" to "know the meaning of the word well enough to give its definition" (D'Anna *et al.*, 1991:117). Despite these controls, the validity of this method of testing is still under debate (Goulden *et al.*, 1990:354-355). One of the key arguments against the checklist format is that, in requiring only a yes/no answer to whether or not a word is known, it simply tests recognition vocabulary. Meara (1996b:44), however, questions this assumption on the grounds that this method allows for a degree of inferencing. He maintains that a knowledge of low frequency words such as *manipulate* must reflect an understanding of a range of higher frequency items. Simply stated, Meara argues that a speaker with a large vocabulary will have regularly encountered the more common words, and so may be presumed to have a good grasp of the meanings of a significant proportion of high frequency vocabulary. This argument is applied to first-language speakers and, to a similar extent, to second-language learners:

> Most people acquire L2 words from exposure to the language, not from learning lists of words
> in the abstract, and it is inevitable that while they are doing this, they also acquire a broader
> knowledge about the words that they already know.

(Meara, 1996b:44-45)

In an attempt to develop a vocabulary test that would measure both breadth and depth of understanding, Read (1993) designed the word association test. This test format is based on the assumption that learners who have a good grasp of the meaning of a particular word will be able to identify related words which represent different aspects of the meaning of the key word, for example:

edit

arithmetic	*film*	*pole*	*publishing*
revise	*risk*	*surface*	*text*

The options provided represent lexical and semantic networks relating to the test item. The relationships between the word and the associates provided are paradigmatic (synonyms), syntagmatic (collocates), and analytic, where associated words are likely to form part of the dictionary definition. This test was designed for university students, and was intended to measure the rate of vocabulary learning. One problem with this format is that random guesses lead to overestimations, since the subjects are able to recognise semantic relationships between test items without reference to the stimulus word. A second problem experienced is that many of the items in Nation's (1984) University Word List do not easily lend themselves to an associated lexical or semantic network. For example, attempts to find synonyms for *electron* or *mathematics* may unnecessarily confuse or obscure the objective of the test.

It should be noted at this stage that, in terms of the definitions of this study, lexical richness is regarded more as a measure of vocabulary proficiency, since it is intended to assess lexical originality, lexical density, lexical sophistication and lexical variation (Laufer and Nation, 1995; Linnarud, 1986), while Diack's (1975) test and others (D'Anna *et al.*, 1991; Goulden *et al.*, 1990; Hazenberg and Hulstijn, 1996; Meara, 1996b; Meara and Buxton, 1987; Meara and Jones, 1987; Nation, 1983) are regarded as measures of vocabulary size. The Lexical Frequency Profile (Laufer and Nation, 1995) falls somewhere between these measures of richness and size. While it measures lexical richness to some degree, since the focus is on breadth of vocabulary at various levels of difficulty, it does not consider

vocabulary depth (§2.3.4), and it correlates significantly with an independent measure of vocabulary size, i.e. Nation's Vocabulary Levels Test (Nation, 1990). Similarly, Corson's (1983) measure of passive vocabulary and Read's (1993) word association test may be regarded as measures of vocabulary proficiency to some extent, rather than vocabulary size. Although they do not examine the full depth of knowledge of any word, they nevertheless require more than passive recognition of vocabulary items, as discussed above.

The degree to which depth of word knowledge is tested thus constitutes another source of discrepancy in the estimation of vocabulary size. In devising a test, a word might be considered 'known' if the subject simply indicates familiarity with one of its several meanings (Diack, 1975; Goulden *et al.*, 1990; Meara, 1996b; Meara and Buxton, 1987; Meara and Jones, 1987). On the other hand, the subject may be required to supply a definition or synonym to demonstrate knowledge of the word (Allen, 1983; Read, 1988). Clearly, the results of tests based on the former assumption of what word knowledge entails are more likely to lead to inflated estimates of vocabulary size. However, the need to test a large sample within a reasonably short period has prompted the development of vocabulary tests which focus on breadth rather than depth of vocabulary, as illustrated by Read (1993:357):

> Generally vocabulary tests focus on breadth of knowledge in that they cover as many words as possible within the time allocated and require only a single response in relation to each word tested.

The theoretical definition of word knowledge which the researcher subscribes to will be manifested, explicitly or implicitly, in the method of testing selected. A test requiring the subject to link the word to an equivalent definition or synonym (either through multiple choice or matching the word to an item listed) assesses breadth of word knowledge, while a test requiring constructed answers corresponds more closely to depth of word knowledge. It should be recognised, however, that, even using the 'constructed answers' method of

testing, the degree to which depth of word knowledge is tested remains relatively superficial due to practical constraints such as time.

2.3.6   In conclusion: A reasonable estimate

Recent studies of vocabulary size, in attempting to answer the question posed earlier of 'how many words a speaker knows', have taken into consideration the issues examined in this section, namely, what constitutes a word, the size and nature of the dictionary used as the target corpus, the sampling technique used to select the representative vocabulary sample, and the method used to test understanding of the representative words, with the result that the present estimations of vocabulary size appear far more reasonable. Goulden *et al.* (1990:341) conclude that "well-educated adult native speakers of English have a vocabulary of around 17,000 base words, [which] represents an acquisition rate of around two to three words per day". D'Anna *et al.* (1991:109) propose an average English vocabulary of approximately the same size:

> By providing a clear rationale for the word source which was sampled, and by using clearly defined operational criteria for what constitutes a word [that is, a base form or lemma], as well as for the procedures used in the estimation task, we found that the average number of different words known by a college student is 16,785.

The need to determine more accurately the size of the gap between the number of lexical items postulated for mother-tongue speakers of English as opposed to second-language learners becomes accentuated when considering the role of vocabulary in reading comprehension, which is unquestionably one of the decisive variables in academic performance.

2.4   The role of vocabulary in reading comprehension

The role that vocabulary plays in reading comprehension has been widely acknowledged in the literature, both from the perspective of studies on the lexicon (Carver, 1994; Coady, 1997; Coady, Magoto, Hubbard, Graney and Mokhtari, 1993; Cunningham and Moore, 1993; Laufer, 1994; Laufer and Sim, 1985; Marshall and Gilmour, 1993; Nation and Coady, 1988; Sutarsyah *et al.*, 1994), and from the perspective of studies into reading comprehension (Brisbois, 1995; Davis, 1972; Grabe, 1991; Pretorius, 1995; Williams and Dallas, 1984).

The issue of the relationship between vocabulary size and academic performance must then be considered within the framework of reading theory to some degree. Since vocabulary forms a significant proportion of the configuration of any text, it should be a significant variable in the comprehension of the prescribed texts. The interactive theory of reading, for example, argues that both micro- and macro-level processing is required to understand a text. According to Cummins' (1981, in Brisbois, 1995) threshold hypothesis, "language transfer is possible only after a threshold level of L2 proficiency has been attained" (Brisbois, 1995:566). Cummins' hypothesis thus addresses the role of linguistic knowledge and bottom-up processes. In line with Cummins, Pretorius (1995:37) argues that these are "variables that cannot be ignored in L2 theories of reading. After all, if L2 readers do not have a sufficient level of proficiency in the L2, they cannot even begin to understand the text." This perspective may be used to support the critical role of vocabulary knowledge in reading, particularly if the focus is on the phrase "sufficient level of proficiency". The lexical threshold hypothesis (Laufer, 1989) proposes that 95% of the running words in a text should be known to the reader for comprehension of the text to occur (the term *running words* refers to the total number of words in any text, and includes base forms as well as all inflectional and derivational forms). This threshold is intended to provide readers with sufficient understanding of the words in a text to enable them to interpret the meaning of most unknown words from context. Below this level of vocabulary proficiency, top-down processing is unlikely to assist the reader in interpreting the text. This argument is corroborated by later research (Laufer, 1992). In this study,

Laufer claims that, for adult second-language learners reading expository texts, "the nature of the language threshold for reading purposes is largely lexical" (Laufer, 1992:126). Following the same argument, Hirsh and Nation (1992) and Nation and Hwang (1995), in an examination of word frequency, coverage and range, propose that the ratio of known to unknown words is a significant factor in the comprehension of academic texts.

In order to determine the discrepancy between the vocabulary size of the students and the level of vocabulary required by the prescribed reading material (§1.4), the students' vocabularies are examined in terms of Laufer's (1989) lexical threshold hypothesis. In other words, the students' vocabulary size evaluated in terms of the level of proficiency required for automatic word recognition that distinguishes poor readers from more competent readers (Grabe, 1991:380).

> It is argued that the automaticity of bottom-up skills is very important for freeing the mind to access top-down processes so that meaningful reading can take place.
>
> (Pretorius, 1995:37)

The students' relative vocabulary size will therefore be examined in relation to the vocabulary required by the reading material to determine whether less than 95% of the running words in the text is likely to be understood.

Frequency counts have been used to produce a number of specialised word lists such as Xue and Nation's (1984) University Word List. These are usually based on a count of words occurring in university textbooks and other academic writing material, and take into account the range of disciplines in which the words are found as well as the number of occurrences. The list is compiled by excluding both high frequency basic vocabulary such as *today*, *remove* and *work*, and low frequency, narrow range vocabulary items which occur only as technical terms in a specialised field, for example, *cataphora*, *phoneme*, *sinusoidal* and *thixotropy*. The academic word lists form an inventory of words which have a high frequency in the field of academic English, and account for a high proportion of the words in any academic text.

Although the precise number of word families required for adequate comprehension of academic texts is still under debate, researchers on the whole agree that a minimum of 95% coverage is required in order for the reader to grasp the contents of the text (Hirsh and Nation, 1992; Laufer, 1989 and 1992; Nation and Hwang, 1995). While a 95% coverage enables the learner to guess any unknown words from the context, a 97-98% coverage allows for far greater comprehension, since the density of unknown words is considerably less. Based on a study of frequency, text coverage and range, Nation and Hwang (1995:40) argue that a general vocabulary of 2000 words provides an 83.4% coverage of the words in non-fiction texts. Sub-technical or academic vocabulary, such as Nation's University Word List, provides an 8.5% coverage of academic texts. The conclusion reached from this study is that the general, high frequency vocabulary together with the list of academic vocabulary provides a 91.9% coverage of the text, which they suggest "approaches" the 95% coverage of academic texts advocated. Laufer (1989:321), on the other hand, argues that a 5000 word basic vocabulary, without the academic word list, is required for a 95% coverage of texts:

> ... lexical coverage of 95% - the understanding of 95% of word-tokens in a text - can ensure reasonable reading comprehension, i.e. a score of 55% and above. Lower lexical coverage is associated with unsatisfactory more often than with satisfactory comprehension. Since the 95% coverage can be achieved by learning 5 000 words, it is suggested that 5 000 words [or lexical items] seems to be the lexical threshold beneath which other facilitating factors in reading comprehension may not be very effective.

These studies by Nation and Hwang (1995) and Laufer (1989) provide a basis for the design of vocabulary tests, the aim of which would be to assess the student's knowledge of the 2000, 5000 and academic word lists in order to establish the percentage of the text covered, and so the degree of comprehension in reading. In her study entitled "How much lexis is necessary for reading comprehension?", Laufer (1992) examines the correlation between vocabulary size and reading comprehension and the lexical threshold distinguishing readers from non-readers. The aim of her study is "to explore the relationship between passive vocabulary size and the comprehension of academic texts" (Laufer, 1992:127). Laufer argues that while vocabulary size is not the only variable to

influence reading comprehension scores, it is the most significant factor, which enables the reader to predict the level of reading proficiency of the subject. The results of this study suggest that the threshold for reading comprehension is 3000 word families, which Laufer converts to 4800 "lexical items or dictionary entries" (Laufer, 1992:130). This indicates that the results of her previous study (Laufer, 1989:321), in which a 'word' was considered to be a 'lexical item' rather than a 'word family', are very similar to the results of the later study on the correlation between vocabulary size and reading comprehension.

In conclusion, then, it appears that there is little discrepancy between the findings of Nation and Hwang (1995), who consider 2000 word families plus the University Word List to form a lexical threshold, and Laufer (1989; 1992), who argues that a knowledge of 3000 word families provides the reader with sufficient vocabulary to grasp the meaning of a text written in a second language.

The issues identified in the course of this literature survey, from which the research hypotheses were derived, are discussed below.

2.5   Deriving the hypotheses

The principal issues which became apparent in the course of the literature review were, firstly, the degree to which vocabulary influences academic performance, particularly in the light of recent developments in reading theory; secondly, precisely what vocabulary size is required by the arts student at first-year level; and, finally, the role played by academic vocabulary within a discipline. A number of subsidiary issues also arose, such as the best means by which to assess vocabulary size, and the question of how to determine students' specific vocabulary needs from a test of general vocabulary. These are addressed below in the discussion of the Lexical Frequency Profile programme and in the description of the University Word List.

2.5.1 The Lexical Frequency Profile Programme

This is a public domain computer programme designed by Laufer and Nation (1995). The Lexical Frequency Profile (LFP) programme compares the lexical items in a text with three frequency-based vocabulary lists in order to identify which words from the text occur in these lists. This programme classifies words according to their frequency of occurrence as a means of determining the approximate level of difficulty of each lexical item (§2.3.2). The LFP also provides an estimation of the level of difficulty of the text through establishing what percentage of words in the text occur in the demarcated lists (§2.4). The vocabulary of written texts is divided by the LFP programme into three broad categories: the first 2000 most frequently occurring words, academic vocabulary, as defined by the University Word List (Xue and Nation, 1984), and other vocabulary items of lower frequency, i.e. words that do not occur in these lists. In each case this programme calculates how often the words listed have occurred in the text. The analysis of frequency levels is based on the concept 'word family', so that in the case of the basic and academic vocabulary, i.e. the 1000 and 2000 word lists as well as the University Word List (UWL), the headwords are listed separately from their related inflectional and derivational forms. Laufer (1992) speculates on this basis that the first 1000 words represent about 3000 word forms.

The first 2000 words are taken from *A general service list of English words* (West, 1953), and include basic vocabulary from all word classes with a low level of difficulty, high frequency and wide range of occurrence, for example, *accept*, *because*, *during* and *various*.

The UWL contains words that occur frequently across a broad range of academic contexts, and which are often not understood by second-language students in particular, such as *correlate*, *deviate*, *equivalent* and *postulate*. Included in this list are subtechnical terms which serve to define concepts in scientific disciplines, but are familiar to language users outside the field of expertise, for example, *colloquial*, *competence* and *homogenous*.

This list was drawn primarily from two other word lists, as explained in Xue and Nation (1984). The first, that of Campion and Elley (1971), "represents vocabulary that students are likely to encounter in their university studies" (Xue and Nation, 1984:215). The second, assembled by Praninskas (1972, in Xue and Nation, 1984), was compiled for second-language students studying through the medium of English, and was drawn from ten books prescribed for first-year reading at university level across a range of courses. In the case of both primary word lists, basic vocabulary had been excluded from the count. The two subsidiary lists from which Xue and Nation's list was compiled were those of Lynn (1973) and Ghadessy (1979). These were drawn up from words which second-language students had annotated in the course of their reading on the assumption that the students had experienced difficulty with these words in particular. Approximately 30% of the words in the subsidiary lists did not overlap with those in the primary lists, and these were added to the larger corpus to form the combined university word list.

Advanced vocabulary items are words with a low frequency and high level of difficulty (cf. §3.2.1.3). Each low frequency item that occurs in the text is not grouped under a headword, however, since these are simply items which have not been entered into the programme under the basic and academic word lists, and are therefore not classified in terms of word families. Thus proper nouns (*Europe, Latin, Plato*), technical terms (*didactic, pedagogic, phenomenological*), compounds (*self-discipline, teacher-pupil, world-wide*), abbreviations (*HOD, Prof., spec.*) and colloquialisms (*shebeen, stayaway*) are listed as low frequency words, while related words such as *enforce, enforcing* and *enforcement* are included as separate items. The inflectional and derivational forms of words in this list are therefore listed individually, rather than being incorporated under a single headword as members of the same word family. (Although the word *force*, which may be regarded as the headword, is listed in the basic word list under the first thousand words, the derivational forms are of a low frequency (Bauer and Nation, 1993), and are not included under the headword, since their meaning is not necessarily transparent despite an understanding of the headword. Similarly, while *large, largely* and *larger* are included in the 1000 word list, the derivational forms *enlarge* and *enlargement* occur in the list of advanced words.) Another considerable problem with the list of low frequency items is

that, as well as the inclusion of related words as separate lexical items, any incorrectly spelt words are included in this list, thereby misrepresenting the percentage of advanced words in the text.

2.6   Conclusion

The aim of this chapter was to provide a review of the research into the relationship between vocabulary size and academic performance, to investigate the testing of vocabulary size, and to review the role of vocabulary in reading comprehension. The principal findings, theories and problem areas which became apparent in the course of this evaluation were discussed in relation to the formulation of the research hypotheses on which this study is based. These hypotheses are presented in Chapter 3 with a description of the research design in terms of which the hypotheses are evaluated.

Chapter 3

Research method

3.1 Introduction

The aims of this chapter are to examine the research hypotheses tested in the study, and to describe the research design and procedures used to test each hypothesis.

3.2 Research hypotheses

In this section the research hypotheses are discussed, special attention being given to the assumptions on which each is based. The general hypothesis, which is based on a global assessment of vocabulary size, is presented initially. The more specific, secondary hypotheses, which address vocabulary size in terms of frequency levels, are then considered.

3.2.1 The general hypothesis

There is a significant positive correlation between students' performance on a test measuring vocabulary size and their academic performance.

This hypothesis predicts that vocabulary size is related to academic performance. Here academic performance refers to the final mark, which comprises an average of the year mark and the examination result (§3.3.2.5). I propose, in other words, that one of the factors contributing to academic failure is an inadequate command of vocabulary. The lexis, as defined in this study, comprises three components: basic, academic and advanced

vocabulary. This hypothesis is therefore broken down into three subsidiary hypotheses, each of which examines one component of the lexis.

3.2.1.1 Hypothesis A

HA   There is a significant positive correlation between students' scores on a test measuring basic, high frequency vocabulary and their eventual academic performance.

Basic vocabulary occurs with high frequency across a broad range of texts (§2.5.1). It is assumed, therefore, that an inadequate command of this vocabulary will affect the student's ability to assimilate and reconstruct the ideas and concepts presented in the literature prescribed at undergraduate level. These ideas and concepts rely on an understanding of the basic vocabulary for their interpretation. The argument that a minimum level of vocabulary is required to comprehend the meaning of a text is derived from the premise that meaningful reading demands a high degree of interaction with the text.

> Skilled readers actively call into play and integrate the knowledge and experiences in their memories with the words on the printed page.
>
> (Durkin, 1981, in Heimlich and Pittleman, 1986:2)

This active participation is only possible if the reader has certain linguistic skills, such as a basic vocabulary which includes many of the cohesive markers, a knowledge of causal relations and anaphora, as well as inferencing skills. Included within these linguistic prerequisites is the demand for a level of understanding of basic vocabulary that will enable the reader to interpret the meaning of unknown words from their context. This supposition is based on a study of vocabulary frequency, text coverage and range conducted by Nation and Hwang (1995:40), who conclude that a general vocabulary of 2000 words provides a 75% coverage of the words in non-fiction texts, while academic vocabulary provides an 8.5% coverage of academic texts. The list of basic vocabulary items, for example, includes logical connectors such as *yet, besides* and *however*. An understanding of these cohesive markers is required in order to understand the way in which concepts within the text are

related, and in order to follow the sequence of assertions and counter-arguments presented in any critical debate. Thus, a reader who does not interpret *although* in a contrastive sense will not realise that it serves to modify another statement or claim in the text. If this hypothesis is valid, it follows that students with a thorough grasp of the basic vocabulary are in a better position to manage the skills required in the academic environment.

3.2.1.2 Hypothesis B

HB   There is a significant positive correlation between students' scores on a test measuring academic vocabulary and their eventual academic performance.

Academic vocabulary is inherent to academic discourse, and to some extent distinguishes this style of writing from others through the use of lexical items such as *assimilate, comprise, intrinsic, propensity, repudiate, stipulate* and *tentative*. These words occur with high frequency across a range of academic texts (§§2.5.1 and 2.5.2). In discussing the composition of the University Word List, Nation (1990:140) maintains that "[s]pecialized word counts which look at a *range* of material within a specialized area naturally end up with lists of words of general usefulness in that area". This claim is illustrated by Lynn (1973):

> Perhaps the most striking feature of the resulting [university word] list is the absence of technical terms. One might expect to find terms like *debenture, blue-collar, inflation* and *debit*, and in fact terms of this type make up the bulk of the vocabulary items in nearly all TEFL texts for commercial students. But instead we find "textbook English" words - non-technical terms from the academic register - presenting the greatest problems to our students[1]. Even such apparently commercial terms as "appraise" and "compensate" were not, in fact, encountered in a commercial context, but in such academic phrases as "appraising the significance of ..." and "factors which compensate for ..."

> (In Nation, 1990:140)

---

[1] This claim is based on the method according to which the word list developed by Lynn (1973) was compiled. This list was assembled from words which second-language students annotated in their university textbooks. It was therefore assumed that these were words with which students had experienced difficulties in the course of their reading.

Academic vocabulary has a specialised function in academic texts which distinguishes it from basic and advanced vocabularies. It serves both a directive role, in that these words frequently require the reader to carry out a particular task, and a semantic role, in that they have a particular meaning in a specific disciplinary context and are related to the remainder of the text by semantic mapping. 'Semantic mapping' refers to the relation between the function of words in a particular text, which creates an associative network among the words. Words are typically associated or mapped through four categories of semantic relations (Aitchison, 1994:84-85):

(a)   Co-ordination. This is a paradigmatic relationship in which words "cluster together on the same level of detail" (Aitchison, 1994:84), for example, *salt* and *pepper*, *butterfly* and *moth*. This category includes colours such as *red, white* and *blue*, and opposites such as *left* and *right*.

(b)   Collocation. This is a syntagmatic relationship in which words commonly occur together, for example, *salt water*, *butterfly net* and *bright red*.

(c)   Superordination. This is a relationship of hyponymy in which the meaning of one word is included in the meaning of another, for example, *butterfly* and *insect, red* and *colour*.

(d)   Synonymy. This refers to the relationship between words that have roughly equivalent meanings within a particular context, for example, *hungry* and *starved*.

The semantic role of academic vocabulary is exemplified with reference to the words *invoke* and *dynamic*. In the sentence *Investigators found that working-class mothers invoke authority figures such as police officers in threatening their children*, the word *invoke* conveys the meaning 'to call on for support', and is related to *authority figures* by means of collocation since both words relate to the notions of control and power. Collocational links range from words which are commonly, although not necessarily, associated (*rude adolescents*) to 'frozen' multi-word units such as *bread and butter*, clichés and idioms (Aitchison, 1994:89).

A second example of semantic mapping is the use of *dynamic* in contrast to *static* in the following context: *Metaphor is a creative process which forges new connections between the meanings of words. This implies that the semantic system is dynamic and not static.* It is essential in this instance that the reader interprets *dynamic* here as meaning 'prone to change', and is cognisant of the antonymous relationship which exists between *dynamic* and *static.* The directive role of academic vocabulary, on the other hand, is clearly illustrated by the words *construct, criticise, debate, evaluate, identify, justify, refute, summarise* and *verify.* When used in the imperative form in the context of an assignment question, for example, each directive demands a particular approach to the question. Thus, while *evaluate* requires that students assess the strengths and weaknesses of an argument, for instance, and conclude in favour of one position, with justification for their conclusion, the directive *refute* requires students to prove that an argument, opinion or statement is incorrect with reference to counter-claims and contrasting findings. The students should conclude, in this case, with seemingly irrefutable evidence to invalidate the original argument.

It is apparent that academic vocabulary has particular relevance within the academic context since these words have a high frequency and wide range of occurrence in academic texts (Nation, 1990; Xue and Nation, 1984). The LFP analysis of the course material on which the vocabulary tests were based (§§2.5.1 and 3.3.2.2) revealed that academic vocabulary comprised 9.7% of the running words in the texts and 20.3% of the total number of lexemes. The relative proportion of academic vocabulary in relation to basic and advanced vocabulary is illustrated in the following table:

Table 3.1

Percentage of text covered by each word list

Frequency list	Percentage of tokens	Percentage of types
1000	77.1	37.0
2000	6.5	13.3
UWL	9.7	20.3
Other	6.7	29.4

The 1000- and 2000-word lists, which together constitute the basic vocabulary, comprise 83.6% of the running words and 50.3% of the lexemes. Viewed in terms of running words, academic vocabulary then comprises the major proportion of non-basic vocabulary items. This degree of coverage appears to support the above argument concerning the semantic and directive roles of academic vocabulary in the academic context. It seems likely then that a considerable number of the directive verbs in academic texts would occur in the UWL, and would thus constitute academic vocabulary (cf. Nation, 1990).

3.2.1.3 Hypothesis C

H1c   There is a significant positive correlation between students' scores on a test measuring advanced vocabulary and their eventual academic performance.

Advanced vocabulary occurs with low frequency across a limited range of texts, and may, to some degree, be considered a measure of lexical sophistication. The words contained in the list of advanced vocabulary items are defined as words not found in any base list (Laufer and Nation, 1995). Advanced vocabulary, by default, includes proper nouns (*English, Saxon, Wernike*), Latinate and Francophone borrrowings, i.e. loan words that have not been anglicised and so have retained the phonological and morphological structure of the language of origin (*ad hoc, inter alia, laissez-faire*), technical terms (*ablative, deixis, metathesis*), the majority of compounds (*baby-talk, male-female, pseudo-science*), abbreviations (*Dr, IQ, café*), acronyms (*LAD: Language Acquisition Device*), colloquialisms (*shebeen, stayaway*), slang forms (*barf, heck, zonked*), nonsense words

(*uglification, gimble*), irregular derivations, where the meaning of the derived form cannot necessarily be inferred from a knowledge of the base form of the word (*default, fortuitous, inviolable, multifaceted, undifferentiated*), as well as a host of infrequently occurring words (*capricious, gubernatorial, heuristic, moribund, rhesus*).

Although the advanced vocabulary comprises only a small percentage of the total number of running words in any text, a knowledge of these words may be a decisive factor in understanding the context in which they occur. For example, in the following sentences, the advanced vocabulary item is basic to the interpretation of the broader text:

> His views have been dismissed far too readily by many linguists, who tend to reject all his claims because a few may be *untenable*.
>
> Language itself is not sexist, just as it is not obscene, but it can *connote* sexist attitudes as well as attitudes about social taboos or racism.
>
> There seem to be specialised areas in the brain and the neurological system that *predispose* human beings to acquire and use language.

Failure to understand the advanced vocabulary item in these instances is likely to result in misconstrual of the overall meaning. A reader who does not know the meaning of *predispose*, for example, would not necessarily realise that an argument is being put forward in favour of the 'innateness hypothesis'. Freebody and Anderson (1983) examined the effect of placing low frequency words in both the important and unimportant parts of a text. Using difficult words as keywords seemed to reduce comprehension in the course of reading the whole text. On the other hand, students were able to summarise the content when low frequency vocabulary occurred in unimportant parts of the text. One explanation of this is that "students did not process many of the unimportant items, lightening the load in terms of length, and helping them to focus on more important items which would be more useful in the formation of summaries" (1983:35). This suggests that students' reactions to unknown words may be simply to ignore them if they do not seem to play a fundamental role in the text (Nation and Coady, 1988:99). The use of a low frequency word in an academic text is often crucial to the meaning of the text, however, and skipping

over these words all too often results in misunderstanding. This claim has support in research conducted by Freebody and Anderson (1983):

> Research indicates that when low frequency words or difficult vocabulary words are found in sentences which contain important concepts, comprehension may be lowered.
>
> (In Gordon, Schumm, Coffland and Doucette, 1992:158).

The use of jargon, which as a rule constitutes advanced vocabulary, obviously plays a significant role in most academic texts and is integral to the understanding of the subject matter. Unlike other instances of advanced vocabulary, jargon, which may include proper nouns (*Chomsky*) and acronyms (*LAD*), is generally recognised by the writers of course material as unfamiliar to the reader (cf. Hubbard, 1996). Each technical term is typically defined, and explained in full, when introduced in the text. However, the majority of advanced words simply occurs in the text without further explanation, and consequently impedes comprehension. It therefore follows that, where students have a poor grasp of advanced vocabulary and not particularly well-developed guessing strategies, this will be reflected in their overall academic performance. Following from Diack's (1975:8) claim that "[t]he number of words in your general vocabulary is an indication of the width and quality of your reading" it seems probable that the student who has access to a number of advanced vocabulary items is a competent reader. This hypothesis claims that a good command of advanced vocabulary will serve as an indication of general academic proficiency.

3.3   Research design

This study is based on a descriptive, quantitative approach, in which vocabulary tests were designed and conducted to assess the validity of the research hypotheses. The results of a pilot study, outlined below, provide incentive for the main study.

3.3.1 Pilot study

A pilot study (cf. Cooper, 1995) was conducted with the aim of determining the relationship between vocabulary size and academic performance in a delimited field. The field in question was the first-year sociolinguistics course offered by the Linguistics Department of the University of South Africa (Unisa). The sample consisted of 38 students. Since the focus of the study was on the lexical proficiency of second-language speakers (§1.2), only three of the students were mother-tongue speakers of English, while the other students spoke a range of first languages. These included Shangaan, Southern Sotho, Northern Sotho, Tswana, Tsonga, Venda and Zulu.

The sociolinguistics section of the Linguistics I study guide formed the corpus for the vocabulary test. This prescribed reading material was analysed using the Lexical Frequency Profile programme (§2.5.1) in order to establish the frequency level of each word in the corpus. A total of 30 words were selected from the text on the basis of this analysis: 6 from the 1000 word list, 6 from the 2000 word list, 11 from the UWL and 7 from the advanced word list. (This uneven distribution of test items was based on the premise that the 1000, 2000 and advanced word lists should be comparably represented, while the majority of test items should be drawn from the UWL. I conjectured, in designing this test, that students experience most difficulty with academic vocabulary. I presumed, for this reason, that items from the UWL would correlate most significantly with the students' examination results and so provide the most worthwhile analysis.) These words formed the basis for a multiple-choice cloze test. The distractors were taken from the word lists tabulated in the analysis of the sociolinguistics corpus, and were selected on the basis of phonological similarity, orthographical similarity or appropriate frequency level, i.e. the test item and distractor were both of the same frequency, and so, by implication, level of difficulty (§2.3.3.1; cf. Nation and Waring, 1997). The examination results of those students who had completed the test were then recorded, and compared with the results of the vocabulary test.

The subjects were divided into three groups according to their examination results: weak students, who had obtained between 10% and 49%, average students (50% to 69%), and more proficient students (70% to 100%). For each of the three groups, the percentage of words known at each frequency level was calculated on the basis of the number of correct answers to each question. A table of lexical frequency profiles was thus obtained on the basis of the examination results (Table 3.2 below). The assumption underlying this profile was that any difference in results would reflect the relationship between the vocabulary levels and academic performance.

The degree of correlation between the vocabulary test results and examination results was plotted by means of the Pearson product-moment correlation coefficient.

Table 3.2

Relationship between percentage of words known at each frequency level and examination results

	2000	UWL	ADVANCED
Group A	83%	69%	51%
Group B	75%	59%	26%
Group C	67%	47%	23%

Table 3.2 provides the mean percentage for each group of students at each frequency level. Group A represents the students who performed well in the examination, Group B represents the average students, while Group C represents the weaker students. The lexical profile of each group is illustrated in the following bar graph:
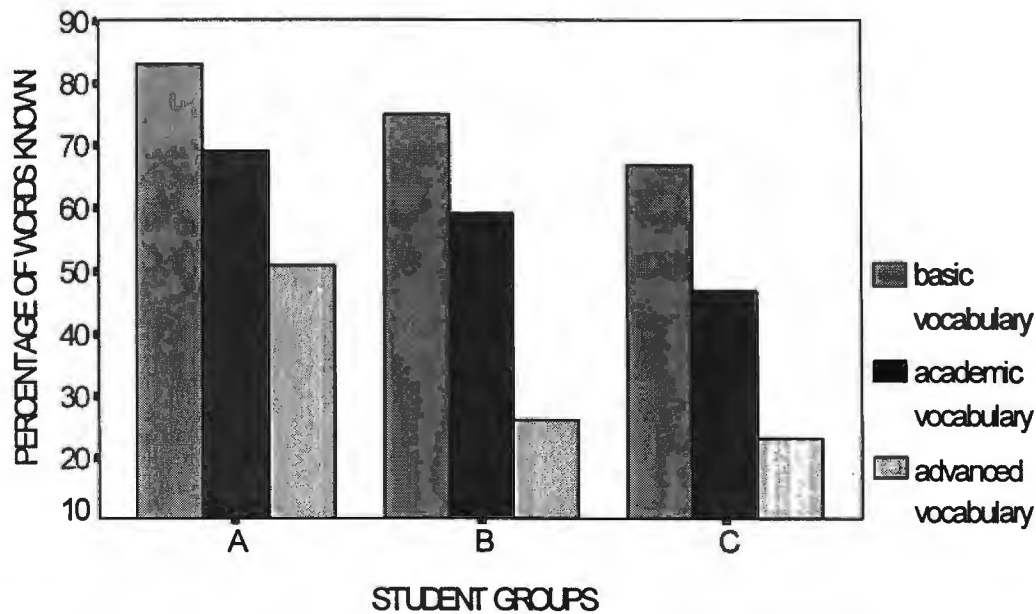
Graph 3.1: Lexical profiles of students graded according to examination results

This graph clearly illustrates the differences between the weak, average and more proficient students in terms of their understanding of basic, academic and advanced vocabulary. While the results indicate a similar intra-group distribution at the levels of basic and academic vocabulary, the most notable difference appears to lie in the students' performance at the level of advanced vocabulary, since the most proficient students have a considerably better grasp of the low frequency lexical items.

The primary research question examined in this pilot study was whether a relationship exists between vocabulary size and academic performance. Analysis of the results indicated a significant correlation between vocabulary size, as determined by the vocabulary test, and examination results. The scores which students obtained for the vocabulary test designed to measure lexical profiles at three levels of word complexity were comparable with their examination results for the macrolinguistics section of the course ($r = .39$, $p<.05$), and with their overall examination results ($r = .41$, $p<.05$). The vocabulary test was therefore shown to be a fairly good indicator of academic performance. The lexical profile of the examination results showed that the students who perform well academically have a greater understanding of words from all three frequency levels, while the weaker students have

significantly fewer lexical items in their receptive vocabularies, particularly in the case of low frequency words (§3.2.1.3).

A key problem in the design of the pilot study was the selection of the test item and distractors from different frequency levels. This may have biased the results, particularly if, in the case of the academic and advanced vocabulary, the distractors were of a higher frequency than the test item, and the subjects were therefore able to select the correct option based on a process of deduction. In other words, the subjects may have been able to determine which of the distractors was inappropriate, and may thus have selected the correct answer without actually knowing the meaning of the test item. In this case, the question would have reflected their understanding of the distractors, but not necessarily an understanding of the test item. Despite this factor, the results appeared to indicate that the method used in the pilot study was appropriate as a basis for further research.

3.3.2   Main study

The main study was based on a single large sample of undergraduate students who completed two vocabulary tests designed to determine the size of their receptive vocabularies. The method of subject selection, the materials used in the development of the vocabulary tests, the design of these tests, the procedures followed in their administration, and the instrumentation used in the analysis of the data generated in the main study are described in this section.

3.3.2.1   Subjects

The students registered for the Education I course at Vista University formed the population on which the main study was based. Of the total number of students originally enrolled for this course (1669), less than one percent (0.6%) were known to be mother-tongue speakers of either English or Afrikaans. The largest single language group was first-language speakers of Xhosa (22.6%), while the remainder spoke Southern Sotho (22.3%), Tswana (18.3%), Northern Sotho (12.8%), Zulu (9.5%), Tsonga (3.9%), Ndebele (2.7%),

Swati (1.4%) and Venda (0.6%) as a first language. (The remaining 5.3% is unaccounted for since the students' mother tongues are unknown.) The distinction between first- and second-language speakers was therefore not considered a variable in the design of the study.

The primary justification for conducting the study through Vista University was the number of students available on campus for vocabulary testing. Vista provided an opportunity to test the research hypotheses using a large sample which increases the statistical validity of the tests.

The two vocabulary tests compiled for this study were administered to 946 subjects in total, of whom only 653 completed both Tests A and B, and obtained a final mark. The correlation analyses of these tests (§§4.4 and 4.5) are therefore based on 653 responses. The analysis of individual test items (§4.8; Appendix D), however, is based on the total number of subjects, that is, 946.

3.3.2.2   Materials

The reading material prescribed for the Education I course at Vista University formed the corpus from which the test items and distractors for the vocabulary tests were drawn. This study material was analysed using the Lexical Frequency Profile programme (LFP) to determine the frequency level of each lexical item in the corpus (§2.5.1). The coursework material was edited to exclude the majority of proper nouns, following the procedure used by Laufer and Nation (1995), as well as to correct any words that had been misspelt in the original text. However, since the LFP programme does not recognise punctuation, any contracted or possessive forms which require an apostrophe, such as *don't* and *children's*, are represented as *dont* and *childrens*, and so included in the advanced word list rather than the basic word list, since they constitute misspellings.

A total of 60 words were selected for the multiple-choice cloze tests: 10 from the 1000 word list, 10 from the 2000 word list, 30 from the UWL and 10 from the advanced word list. As in the pilot study, I conjectured that the most significant results would be obtained

from a test of academic vocabulary (§3.3.1). In order to obtain a representative sample of test items from the corpus, words were selected at regular intervals from each frequency list. The items in each list were themselves ordered according to frequency of occurrence, so that within the basic, academic and advanced word lists, items from a range of frequencies were tested. This was achieved by first counting the total number of words in each list, arranged according to decreasing levels of frequency, then dividing that number by the number of items due to be tested at that frequency level, thereby separating each list into an even number of groups. The test items selected were those words which occurred at the end of each group.

The aim of this method of selection was to attain an even distribution of lexical items from each word list despite varying frequencies of the items within each list. Within the UWL, for example, the word *intellectual* occurs 70 times in the reading material, while the word *perpetuate* occurs only once. Thus, for example, where the 1000 word list contained 802 words in total, and 10 items from the list were due to be tested, every 80th word was selected as the test item. Where the word was regarded as inappropriate because it did not appear in a suitable context in the text (see below), or it was a function word belonging to a closed set (rather than a lexical content word), for example, *be, it, of, she*, or because a more academic word was required, another word in the immediate vicinity of the original item was selected. Thus *we* and *eat*, which were considered inappropriate on the grounds that they would not test the student's grasp of high frequency words typical in an academic context, and so would not be sufficiently representative of the student's academic proficiency, were substituted by the previous items on the list, i.e. *develop* and *receive*, respectively, both of which had the same frequency of occurrence within the 1000 word list.

A number of conjunctions, such as *although*, and adverbials, such as *still*, were deliberately chosen as test items, although they did not necessarily occur at the intervals stipulated in the methodology outlined above. The reason for selecting words from these grammatical classes was to assess students' understanding of function words of this type, since, although these are included as basic vocabulary items, it was postulated that the weaker students

would have a poor grasp of conjunctions and adverbials in particular. It should be noted from this discussion that there were problems and extraneous factors that had to be taken into consideration during the compilation of the tests, and that, as a result, the procedure followed cannot be regarded as absolutely objective.

Once the test items had been selected, appropriate sentences were taken from the text in order to situate each word in context. It was essential that the meaning of the test item, as the required choice, should be represented clearly and unambiguously through the use of collocation, synonymy, antonymy or other semantic relations. The following example serves to illustrate this point, since the test item *conform* is semantically related by means of collocation to the phrases *peer group, group membership*, and *ideas and norms*:

> *Peer groups often assume the form of a club or gang and group members are expected to _____ to the ideas and norms of the group.*

Similarly, the word *develop* was tested within the context of the following sentence, since *immature* and *education* collocate with this test item:

> *The pupil is regarded as first and foremost an individual with a personality of his own, though an immature one. His mind, like his body, is immature and has to be _____ through education.*

A second means of indicating which of the options provided was most appropriate was through the use of near-synonymy, in which the test item and another word in the sentence share similar meanings:

> *The theory of education that forms part of the philosophy of a society also includes guidelines about how schools should function, how teaching should be conducted, what schooling should achieve; these all _____ the nature of schooling in a society.*

The meaning of *comprise* relates to the concept of inclusion. The listing of three guidelines which form part of, or are incorporated in, the theory of education suggest that a synonym for *form part of* or *include* is required.

Antonymy, in which the test item was contrasted against the meaning of another word in the sentence, was also used as a means to highlight the required option:

> *You must learn to argue scientifically, i.e. to debate or discuss a question, and*
> *to use reasoning to support or _____ a theory.*

In this example there is an antonymous relationship between the test item *refute* and the word *support*, which is set in juxtaposition to the test item.

Conjunctions, as logical connectors which clarify the relation between linked clauses, also serve to express semantic relations. Consider the example below in which the test item *although* is an adversative conjunction:

> *One criticism is that _____ Pragmatists claim to place the focus on the*
> *child in the schooling process, the child in fact often becomes secondary to the*
> *experience or the project.*

In this sense the choice of *although* as the most appropriate option is indicated by means of the contrast between theory and practice outlined in the sentence. The adversative nature of this conjunction thus distinguishes it from the distractors *since, because* and *whether* which perform causal, temporal and contrastive functions, respectively. The issue of selection criteria in terms of which the distractors were decided on is addressed in greater detail below.

The distractors were selected from the word lists tabulated in the analysis of the reading material, and were chosen on the basis of phonological and/or orthographical similarity with

the test item as far as possible (cf. Laufer, 1990a and 1997b; Ryan, 1997). The following examples, in which the test item is given in capital letters, serve to illustrate this principle:

Table 3.3

Cases in which the test item and distractors have similar lexical forms

DEPENDS	defends / descends / deserves
DEVISED	derived / deprived / deferred
EMULATION	eradication / evisceration / emancipation
FACILITATE	frustrate / fluctuate / formulate
INTRINSICALLY	intolerably / intensively / insistently
PAUCITY	profanity / periphery / perversity

The selection of similar lexical forms (Laufer, 1990b; Laufer-Dvorkin, 1991) was intended to rigidly constrain the subject's response with the aim of avoiding the problems associated with diverse distractors. Multiple-choice items, for example, may enable the subject to select the appropriate option by elimination or by default without a clear grasp of the meaning of the word being tested (§2.3.5; Anderson and Freebody, 1981). While Nation (1990:82) suggests that "the words in the group [of multiple-choice items] should not be closely related to each other in meaning", the use of distractors which are too dissimilar from the test item restricts the degree of confidence in the findings. Thus, although the distractors should obviously not be closely synonymous with the test item or semantically ambiguous within the context of the test sentence, they should "discriminate between a learner who has a hazy notion [of the word] ... and one who can place the word more precisely within its semantic field" (Meara and Jones, 1987:22). Although Meara and Buxton (1987) argue against the multiple-choice vocabulary test, they do acknowledge that "[this] format can produce items of varying difficulty and refinement through careful manipulation of the distractors" (1987:143).

An exception to these selection criteria was made in testing the conjunctives *still*, *although* and *however*. In each instance the distractors were selected for syntactic similarity with the test item and so were derived from the set of cohesive conjuncts. For each question the test

item and distractors belong to the same grammatical class and have the same frequency, and so, by implication, level of difficulty (§2.3.2). For example the distractors *yet, even* and *besides* and the test item *still*, which are all adverbial, were selected from the 1000-word list.

It should be noted that a few of the distractors used in these tests were selected from frequency lists generated in an analysis of the Linguistics I corpus, since there were insufficient options provided by the word lists drawn up from the education course material to fulfill the criteria of phonological and orthographical similarity required. Words from the linguistics corpus were selected to ensure that both the test item and each distractor were of the same frequency. Once the vocabulary tests had been drawn up they were administered to the subjects according to the procedures outlined below.

3.3.2.3 Procedure

The vocabulary tests were conducted on a voluntary basis at a number of campuses throughout South Africa, namely, Bloemfontein, East Rand, Port Elizabeth, Mamelodi, Sebokeng, Soweto and Welkom. The course leader responsible for Education I at each campus was requested to administer the tests, and was issued with a set of instructions. Before completing the tests, the students were told that the objective was to assess their understanding of the vocabulary used in the study material with the aim of determining which of the words were problematic for students. These problems could then be addressed in the development of study material. This was intended to address the concern that students would otherwise view the tests indifferently and not complete them to the best of their ability.

To ensure that the students understood what was required of them they were given a detailed explanation of how to approach a multiple-choice question paper, and were provided with examples of how to evaluate and answer each question (cf. Appendix A).

In an attempt to ensure that the students did not lose concentration or become distracted, the test items were divided between two 40-minute tests, which were completed in separate sessions to allow the students a short break. Each test contained 5 items from the 1000-word list, 5 from the 2000-word list, 15 from the UWL, and 5 from the advanced word list.

The vocabulary tests were then collected from each campus for marking and analysis at a central point.

3.3.2.4   Data analysis

The initial analysis of the vocabulary in the course material was conducted by means of the Lexical Frequency Profile programme, as described in the section outlining the compilation of the vocabulary tests (§3.3.2.2).

The marking of the multiple-choice answer sheets for the vocabulary tests was done by means of computer, using the Unisa Student System programme. From this process, the results of the tests were available in the form of a binary code, according to whether the subjects had marked the answers correctly (1) or incorrectly (0). The final stage in the analysis of the data, in preparation for the statistical analysis, was the calculation for each subject of the sum of the correct answers at each frequency level. In other words, for each test, the sum of correct answers to the 1000-word list, the 2000-word list, the academic word list and the advanced word list was calculated for each subject. On the basis of an analysis of the frequency levels in each test (§4.3), the sum of correct answers for both the 1000- and 2000-lists was then calculated jointly, since the difference in results at these frequency levels was found to be negligible. Thus the sum of the results for the basic vocabulary test as a whole was determined.

3.3.2.5   Statistical analysis

In order to test the validity of the research hypotheses, the vocabulary test scores for each subject were correlated against the final mark. The final mark was calculated as an average

of the year mark and examination result, with the examination counting 50% of the final mark. The final mark was therefore considered a better reflection of academic performance, since it was based on an adjustment of the examination result according to the student's overall performance in the course of the year.

The degree of correlation between vocabulary size and academic performance was measured by means of a parametric test, namely, the Pearson product-moment correlation coefficient (Appendix C). This is a statistic that indicates the degree to which two variables are related to one another, and is therefore directly applicable to the aims of the study. Where scores on one variable tend to increase as scores on the other variable increase, this indicates a direct, positive relationship between the two variables. A significant positive correlation between students' scores on the vocabulary tests and their final mark would support the research hypothesis: this finding would reflect that students who perform well on the vocabulary tests attain correspondingly high marks during the course of the year, while students who perform poorly on the vocabulary tests attain comparatively poor results overall. Vocabulary size could then be assumed to be related to academic performance.

All correlation analyses were conducted by means of the SAS computer programme, while graphs of the results were drawn using the SPSS statistical programme.

3.4   Conclusion

In the first section of this chapter, the general and secondary hypotheses were set out and discussed, and the rationale behind each hypothesis outlined. The pilot study was then described, and an overview of the research method used in the course of the main study provided. The results of the study are presented in Chapter 4, with a discussion of the statistical analyses.

Chapter 4

Results

4.1   Introduction

The aim of this chapter is to present and describe the findings of the vocabulary tests and the analyses performed on the results. A synopsis of the research method used to obtain the data is provided in order to contextualise the results, which are then discussed together with tabulation and graphs.

4.2   The research method: An overview

The primary objective of this study was to establish the relationship between vocabulary size and academic performance. In order to estimate vocabulary size two multiple-choice cloze tests were conducted. The tests were based on the reading material prescribed for the Education I course at Vista University. This reading material was analysed using the Lexical Frequency Profile programme (§2.5.1) and the test items were then selected on the basis of frequency. Each test comprised 30 items: 5 from the 1000-word list, 5 from the 2000-word list, 15 from the UWL and 5 from the advanced word list. The distractors were taken from the word lists generated by the LFP programme, and were selected on the grounds of phonological and/or orthographical similarity and identical frequency. The students registered for the Education I course completed the tests on a voluntary basis. Students' final marks (comprising an average of the year mark and examination result) were then compared with the results of the vocabulary tests. The findings were analysed by means of the Pearson product-moment correlation coefficient to test the validity of the research hypotheses. Finally, a stepwise multiple linear regression analysis was conducted

to establish to what extent, on the basis of the correlation conjectured between vocabulary size and academic performance, the vocabulary test scores have predictive value.

4.3   Descriptive statistics

The scores for the vocabulary tests are considered as the first stage in the analysis of the data (cf. Appendix D: Tables 1 and 2:135-136). These scores were derived from the computer analysis of the multiple-choice answer sheets, and are based on the sum of correct answers for each test item (§3.3.2.4). This section then presents and examines the results obtained by the subjects on the vocabulary tests.

The total percentage of correct answers at the levels of basic, academic and advanced vocabulary for each test is presented in Table 4.1 below. This analysis of the frequency levels is intended to assess both intra- and inter-group distribution at each level as a measure of reliability. The aim is to establish the degree to which the results correspond, firstly between the matching frequency levels of each test (intra-group distribution), and within the test itself, across frequency levels (inter-group distribution). At each frequency level the average number of correct responses is given together with the corresponding percentage. This percentage is based on the mean total divided by the number of subjects.

Table 4.1

Percentage of correct answers at each frequency level (Tests A and B)

	Basic vocab. (1000-list) 5 items	Basic vocab. (2000-list) 5 items	Basic vocab. (combined) 10 items	Academic vocabulary 15 items	Advanced vocabulary 5 items
Test A (N = 653)	66.4	64.6	65.5	36.4	16.5
Test B (N = 653)	65.5	66.1	65.6	50.0	28.4

This analysis illustrates that the scores for basic, academic and advanced vocabulary decrease relative to frequency level. The finding that the scores decrease according to level of difficulty and that the students performed as expected suggests that the tests are reliable and well motivated.

It should be noted that the scores for academic and advanced vocabulary decrease more rapidly in proportion to basic vocabulary in Test A than in Test B. That is, the drop in mean percentage from basic to academic vocabulary is 29.1% in Test A, but only 15.6% in Test B. The difference in scores between academic and advanced vocabulary is consistent for the two tests, with declines of 19.9% and 21.6% in Test A and Test B respectively. Thus the greatest difference between the scores is the sudden decline in score from basic to academic vocabulary in Test A. Given the similarity between the tests in the mean scores for basic vocabulary, the academic vocabulary section of Test A appears to have been in some sense more difficult. This assumption is explored further in the analysis of individual test items (§4.8). The differences in the percentage of correct answers across the frequency levels is illustrated in the graph below.

Graph 4.1

Percentage of correct answers at each frequency level
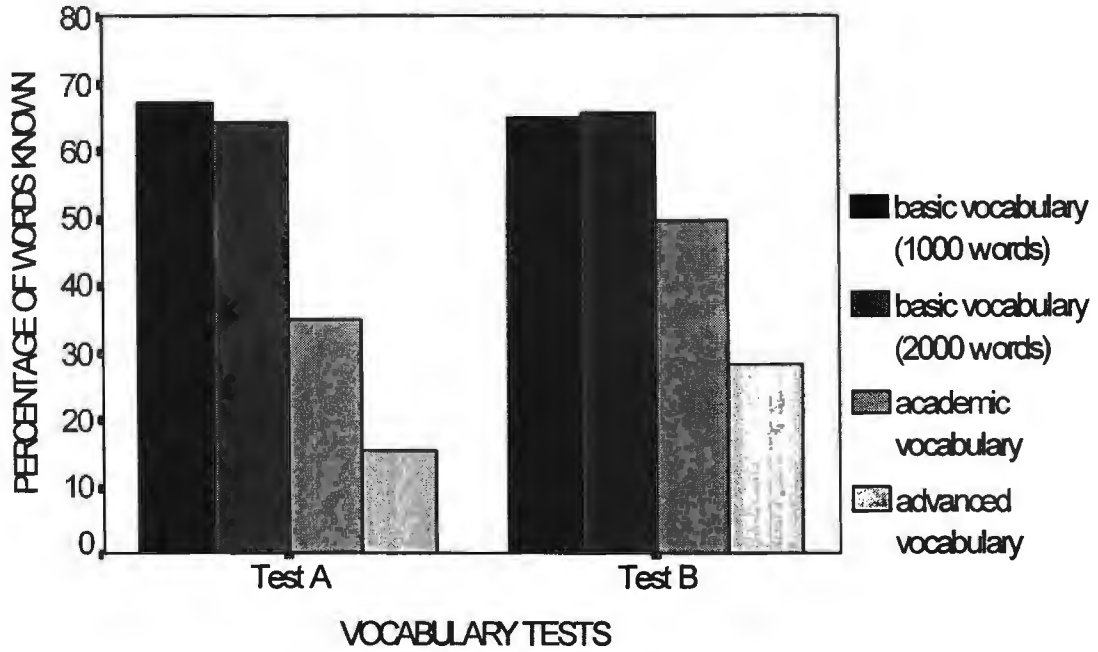
Tests A and B were regarded as separate tests in the analysis of individual items and frequency levels in order to identify any discrepancies between them. However, since the motivation for conducting two separate vocabulary tests was both pragmatic, i.e. to counteract the effects of fatigue in the subjects (§3.3.2.3), and methodological, i.e. to test the reliability of the measure (§4.5.1), the test scores are combined for the purpose of further statistical analysis. The correlation analysis for each research hypothesis is thus based on a larger number of test items at each frequency level. Table 4.2 below presents the results of the analysis of frequency levels for the test as a whole, together with the means and standard deviations.

Table 4.2

**Means, standard deviations and frequency analysis of the combined test**

	Basic vocab. (1000-list) 10 items	Basic vocab. (2000-list) 10 items	Basic vocab. (combined) 20 items	Academic vocabulary 30 items	Advanced vocabulary 10 items
Mean (N = 653)	6.59	6.53	13.13	12.96	2.24
Std. deviation	1.71	1.85	3.05	4.38	1.38
% of correct items	65.9	65.3	65.6	43.2	22.4

These results confirm that the scores for basic, academic and advanced vocabulary decrease in proportion to frequency level, as reflected in the analysis of frequency levels in Table 4.1. Given that the difference between the percentage of correct answers for the test items at the 1000 and 2000 frequency levels is negligible (0.6%), these frequency levels are combined for the purpose of further analysis, forming a single test of basic vocabulary items. The consolidation of these frequency levels appears to be justified in terms of the finding that the 1000- and 2000-word tests were more reliable when considered together as a test of basic vocabulary than when analysed at two separate frequency levels (§4.5.1: Table 4.6). It was assumed, in other words, that the two frequency levels have greater reliability and validity with a larger number of test items. The joint analysis of these frequency levels is further supported by Laufer's (1994:25) study in which she considers lexical proficiency in terms of two "condensed profiles". The first 2000 words form a single group of basic vocabulary items, while all vocabulary not included within the 2000 word families is regarded as non-basic vocabulary.

In order to establish a difference among the frequency levels, the range of scores for each level was calculated and the results interpreted as a percentage of the total score of 60. This range (the confidence interval) illustrates the distribution of 95% of the test scores, having excluded 5% of the extreme scores (the outliers) at either end of the scale. These results are presented in Table 4.3 below.

Table 4.3

Confidence interval for each frequency level of the combined test

Frequency level	No. of items	Confidence interval	Percentage range
Basic vocab. (1000-list)	10	$5.53 \leq \mu \leq 7.65$	$55.3 \leq \mu \leq 76.5$
Basic vocab. (2000-list)	10	$5.38 \leq \mu \leq 7.68$	$53.8 \leq \mu \leq 76.8$
Basic vocab. (combined)	20	$11.80 \leq \mu \leq 14.47$	$59.0 \leq \mu \leq 72.4$
Academic vocabulary	30	$11.39 \leq \mu \leq 14.53$	$38.0 \leq \mu \leq 48.4$
Advanced vocabulary	10	$1.39 \leq \mu \leq 3.10$	$13.8 \leq \mu \leq 31.0$

Table 4.3 illustrates that 95% of the scores range from 59.0 to 72.4% for basic vocabulary, from 38.0% to 48.4% for academic vocabulary and from 13.8% to 31.0% for advanced vocabulary. This indicates that the greatest range in scores occurs at the level of advanced vocabulary, where the difference between the highest and lowest mean scores is 17.2%, while the smallest range occurs at the level of academic vocabulary, where the difference between the highest and lowest mean scores is 10.4% (cf. Best and Kahn, 1993:327). This difference is 13.4% at the level of basic vocabulary. This finding suggests that the most consistent scoring occurred at the level of academic vocabulary. The analysis of confidence intervals indicates a degree of internal consistency at this level and suggests that the academic vocabulary test items were more compatible in level of difficulty than were those items used to test basic and advanced vocabulary. The test of academic vocabulary thus appears most homogenous. This assumption is supported by the analysis of individual test items and the analysis of options (§4.8).

It is clear from this table that each of the frequency levels is discrete since there is no overlap in the range of scores. In other words the lowest mean score for basic vocabulary (59.0%) is notably higher than the highest mean score at the level of academic vocabulary (48.4%). Similarly, the lowest mean score for academic vocabulary (38.0%) is larger than the highest mean score for the advanced vocabulary (31.0%). The evidence of implicational scaling (§2.3.2) across the frequency levels suggests that the test items are appropriate to each frequency level and that the test has reliability and validity (§4.5). This assessment of

the confidence interval suggests that the results of the correlation analysis reasonably reflect the students' vocabulary at each of the frequency levels, as intended.

In conclusion, the analysis of frequency levels was intended to assess the degree to which the results correspond, firstly between the frequency levels across tests (intra-group distribution), and then within each test itself, across frequency levels (inter-group distribution) as a measure of reliability. The finding that the test scores decrease relative to frequency level suggests that the inter-group distribution is reliable (cf. §2.3.2). A superficial exploration of intra-group distribution indicates some discrepancy in scores between Tests A and B at the levels of academic and advanced vocabulary in particular. These discrepancies are examined further in a correlation analysis designed to compare the frequency levels of each test. The results of this analysis are presented later as part of the detailed evaluation of reliability and validity (§4.5).

In the following section the assumptions underlying the study are outlined, and the results for each hypothesis are presented and discussed in turn.

4.4   Testing the hypotheses

The primary aim of this study was to examine the relationship between vocabulary size and academic performance. In order to do so, a single, general research hypothesis was formulated regarding the nature of this relationship. This hypothesis examines the role of vocabulary in general as well as the role of vocabulary at different levels of difficulty. The function of the hypothesis is therefore to test the effect that independent variables such as basic vocabulary, academic vocabulary and advanced vocabulary have on academic performance, as measured by the final mark in a content subject. These independent variables were expressed as more specific, secondary hypotheses, and tested individually. The general hypothesis, which examines the effect of vocabulary as a whole, is tested first; thereafter the three secondary hypotheses are examined together.

4.4.1 The general hypothesis

The general research hypothesis states that the students' academic performance, as measured by their final mark, will be significantly related to the size of their vocabularies (§3.2). For the purpose of this hypothesis vocabulary size comprises basic, academic and advanced vocabularies collectively. The premise on which this hypothesis is based is that vocabulary, as a component of language, underlies skills such as reading and writing. Given that academic performance is contingent on such skills, it may be inferred that vocabulary is basic to academic performance (§§2.2 and 2.4). The null version of this hypothesis was formulated as follows:

H_0 *There is no significant positive correlation between students' scores on a test measuring vocabulary size and their eventual academic performance.*

4.4.1.1   Results

In order to test this null hypothesis, the overall results of the vocabulary test were compared with the results of the final mark using Pearson's product-moment correlation coefficient. Vocabulary size serves as the independent variable with academic performance as the dependent variable. The results are considered significant only at the $p<.05$ level. The results are presented in Table 4.4 below.

Table 4.4

Correlation of combined vocabulary test result with final mark

Correlation (N = 653)	$r = 0.39$
Probability level	$p<.0001$

It is clear from Table 4.4 that the correlation coefficient is both positive and statistically significant. The null hypothesis can therefore be rejected. It must be recognised, however, that the correlation coefficient is low, as illustrated in Graph 4.2 below. This scatterplot

clearly reflects the broad distribution of scores. The issue of statistical versus practical significance is examined further in the discussion of the secondary hypotheses (§4.4.2.2) and in the assessment of reliability and validity (§4.5).

Graph 4.2

Scatterplot of combined vocabulary test score and final mark
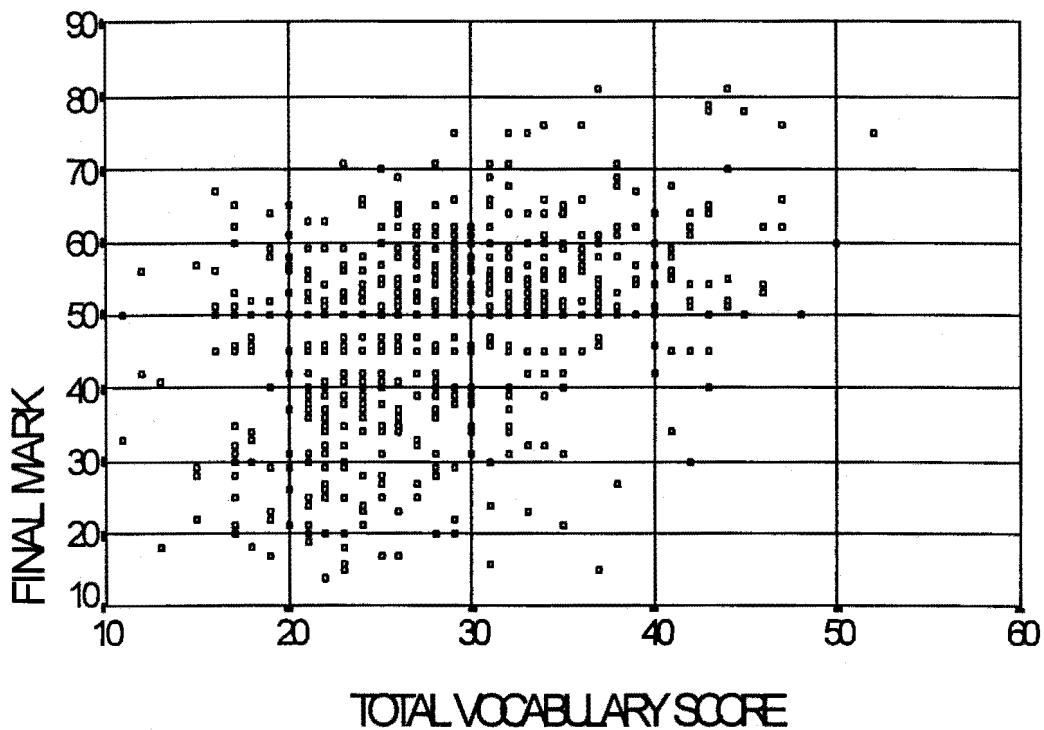
4.4.1.2 Discussion

The finding that the students' scores on the vocabulary test correlate significantly with their final marks indicates that vocabulary size does influence academic performance to some degree. The results for the general hypothesis thus seem to confirm the assumption that vocabulary size plays a significant role within the academic context. It may be inferred from this that the number of words in the mental lexicon directly influences the student's ability to meet the demands of academic study. Thus, while vocabulary is merely one component of language competence, it does seem to underlie many of the cognitive skills required for

academic performance. The finding that the larger a student's vocabulary, the more likely that student is to perform well in tasks such as interpreting and evaluating arguments in text, or writing critically, has significant pedagogical implications, as proposed in Chapter 5.

The corollary of the finding that the extent of a lexicon indicates the degree of competence in broader cognitive skills, is that poor vocabulary proficiency is likely to be associated with an inadequate command of these broader skills. An obvious implication of this is that students whose vocabularies do not satisfy the reading and writing requirements for an academic course are unlikely to perform well. This effect is explained in part by the lexical threshold hypothesis (Laufer, 1989; §2.4) since these students would not have access to the wider vocabulary that enables a reader to understand the meaning of unknown words from context. In support of this explanation, Grabe (1991:390) proposes that "fluent readers need a sound knowledge of language structure and a large recognition vocabulary". In addition, Medo and Ryder (1993:120) argue that "[i]f text specific vocabulary is limited, accurate mental representations and causal connections will not be made and comprehension will be hindered". It is apparent then, that where students' vocabulary is inadequate to meet the basic level of lexical proficiency necessary for the comprehension of texts, academic performance will be affected.

The notion that vocabulary size is indicative of more general linguistic competence is a common theme in research on aspects of vocabulary knowledge (Anderson and Freebody, 1981; Corson, 1983; Diack, 1975; Dupuy, 1974; Gass and Selinker (1994) in Laufer, 1997b; Kang, 1995; Meara and Jones, 1987; Nation, 1990; Saville-Troike, 1984; Sternberg and Powell, 1983). Research on vocabulary proficiency and its relation to measures of verbal intelligence, for example, has yielded correlations of approximately .80 (Jenkins, Stein and Wysocki, 1984). The influence of vocabulary size on general academic proficiency in particular is illustrated through research conducted by Meara and Jones (1987). In a study of the relationship between vocabulary size and language skills as measured by the Eurocentres JET test, they found a significant correlation of .664. In addition, they found that the correlation improved in relation to the degree of homogeneity of the test group. Meara and Jones (1987:6) conclude from this that the vocabulary test

was a better predictor of students' proficiency for some language groups than for others. This finding may to some extent explain the low correlation between variables in this study, given the heterogenous population (§3.3.2.1). Another possible explanation for the low correlation may relate to the finding by Meara and Jones (1987:12) that teachers' assessments of student ability were at variance with the vocabulary test scores.

> This suggests that the vocabulary test is an adequate instrument for placing individuals in broad level categories, but that it might not be accurate enough to detect fine differences in ability within a level.

In a study of the influence of vocabulary size on reading comprehension Laufer (1992:128) found a correlation of .5 for the Vocabulary Levels Test (Nation, 1990) and of .75 for the Eurocentres Vocabulary Test. Laufer concludes from these results that the minimum vocabulary required for reading comprehension is 3000 words (cf. §4.9). Furthermore, vocabulary size has been found to correlate highly with writing proficiency (Engber, 1993; Laufer and Nation, 1995; Linnarud, 1986). In her study of the relationship between vocabulary and the quality of written compositions, Engber (1993) found a correlation of .57 between lexical variation where the lexis did not contain errors and an overall assessment of writing quality. She infers from this that "it is the combination of a varied vocabulary used correctly that is likely to influence a reader to give a higher score" (Engber, 1993:128). Similarly, Laufer and Nation (1995) found a significant correlation between the lexical profiles of students' writing and an independent measure of vocabulary size. In conclusion, Laufer and Nation (1995:319) suggest that learners' productive use of language may be expected to reflect their vocabulary size.

The finding that vocabulary size correlates significantly with academic performance suggests that the results of this study compare with those of previous studies on the role of vocabulary in the academic context (§2.2). This claim is supported by Read (1997:304) in his review of vocabulary testing.

> Multiple-choice vocabulary tests proved to be highly reliable and to correlate very well with tests of reading comprehension as well as psychometric measures of intelligence. Thus, vocabulary

tests were valued ... [for] their apparent validity as indicators of language ability in a broad sense.

The argument that vocabulary size corresponds significantly to other language skills and that vocabulary test scores correlate with the results of both intelligence tests and achievement tests (cf. Anderson and Freebody, 1981; Meara and Jones, 1987) is clearly directly applicable to academic performance in tertiary education. Thus, given that academic performance correlates significantly with vocabulary size, the assumption that vocabulary size is an index of quality of writing, level of comprehension when reading, and consequently academic performance, appears to be valid. From this it may be inferred that an inadequate grasp of vocabulary is symptomatic of poor writing, poor reading comprehension, and thus poor academic performance.

Having established that vocabulary size in general correlates significantly with academic performance, the secondary hypotheses now address the questions of the extent to which the vocabulary at each frequency level affects academic performance, and whether the academic vocabulary does have particular relevance within the academic context, as predicted (§3.2.1.2).

4.4.2   The secondary hypotheses

In order to test these hypotheses, the scores for basic, academic and advanced vocabulary are compared with the final mark using the Pearson product-moment correlation coefficient. The underlying assumptions are outlined; thereafter the results are presented together and discussed.

Hypothesis A tests the prediction that there is a significant positive correlation between basic vocabulary and academic performance. The assumption underlying this hypothesis is that, since basic vocabulary occurs with high frequency across a broad range of texts and thus forms "the essential common core" (Nation and Hwang, 1995:35), an inadequate command of this vocabulary will affect the student's ability to assimilate and reconstruct the

ideas and concepts presented in text at a fundamental level. The null version of this hypothesis is formulated below.

HA_0 *There is no significant positive correlation between students' scores on a test measuring basic, high frequency vocabulary and their eventual academic performance.*

Hypothesis B tests the prediction that there is a significant positive correlation between academic vocabulary and academic performance. This hypothesis is derived from the assumption that an understanding of academic vocabulary is one of the keys to academic discourse since these words have a high frequency and wide range of occurrence in academic texts (Nation and Waring, 1997). The null version of this hypothesis is formulated below.

HB_0 *There is no significant positive correlation between students' scores on a test measuring academic vocabulary and their eventual academic performance.*

Hypothesis C tests the prediction that there is a significant positive correlation between advanced vocabulary and academic performance. Two primary assumptions underlie this hypothesis. The first is that, where advanced vocabulary is basic to the interpretation of the broader text, a knowledge of these items plays a decisive role in understanding the context in which they occur. The second is that a knowledge of advanced vocabulary requires a degree of lexical sophistication and thus serves as an indication of general language proficiency. Students who are proficient in reading and writing should in turn perform well academically. The null version of this hypothesis is formulated below.

HC_0 *There is no significant positive correlation between students' scores on a test measuring advanced vocabulary and their eventual academic performance.*

4.4.2.1   Results

In order to test these null hypotheses, the results of each subject's basic, academic and advanced vocabulary tests were compared with his or her final mark. As in the general hypothesis, the results are considered significant only at the $p<.05$ level.

Table 4.5

Correlation of basic, academic and advanced vocabulary with final mark

	Basic vocabulary	Academic vocabulary	Advanced vocabulary
Correlation (N = 653)	$r = .29$	$r = .38$	$r = .18$
Probability level	$p<.0001$	$p<.0001$	$p<.0001$

These results indicate that, although not very strong, the correlation coefficients are both positive and statistically significant at each frequency level. The null hypotheses formulated with regard to the basic, academic and advanced vocabulary tests can therefore be rejected.

4.4.2.2   Discussion

Despite ranging from very low (between .01 and .19) to low (between .20 and .39) (Mulder, 1986:73), the correlation between vocabulary size and academic performance is significant at the $p<.0001$ level in each case. The strongest correlation occurs at the level of academic vocabulary, while the weakest correlation is for advanced vocabulary.

Although the vocabulary test scores were found to correlate significantly with academic performance at the $p<.0001$ level, it can be assumed that the practical significance of these findings is limited. This assumption is based on the fact that a low correlation can account for only a small variance in the other variable, thereby indicating a weak, but definite, relationship. The following argument by Hatch and Farhady (1982:21) concerns research findings in which the relationship between variables is clearly weak. They propose that

additional variables, extraneous to the test, must then be considered in the interpretation of results.

> Whether you interpret the correlation value using variance overlap or by testing for statistical significance, your interpretation cannot be made simply on the basis of these figures. Your interpretation depends on what variables are being compared and what kinds of decisions must be made on the basis of the discovered relationship.

Similarly, Rosenthal and Rubin (1982:166) argue that the results of research in the social sciences tend towards low correlations because of the multivariate nature of social data. For this reason a number of factors need to be considered in the analysis and interpretation of results. The aim of the following discussion surrounding the correlations for basic, academic and advanced vocabularies is therefore to interpret the results and to consider other possible variables which may account for the low correlations in Table 4.5.

The results for the secondary hypotheses clearly show that, of the three frequency levels tested, academic vocabulary size is the most significant indicator of academic performance. This result confirms the assumption that academic vocabulary comprises the 'tools of the trade' in tertiary education (§3.2.1.2). Academic vocabulary, in particular, is the lexis used to read, write, think and talk about the issues dealt with in content subjects. Students who do not have access to the more formal, abstract and specialised words that comprise academic vocabulary are to some extent denied access to the discourse, and so are unlikely to adequately interact with academic texts.

In a paper entitled "Social dialect, the semantic barrier and access to curricular knowledge", Corson (1983, in Romaine, 1984) proposes the existence of a 'lexical bar' in English, which may account for a proportion of scholastic failure. Corson compared the use of common, high-frequency, largely monosyllabic vocabulary, which is primarily derived from Anglo-Saxon, with vocabulary of Graeco-Latin origin. He found that degree of familiarity with the Graeco-Latin vocabulary of academic discourse varied considerably among scholars of different social backgrounds. The implications of an inadequate

command of this vocabulary become apparent in the light of Corson's claim that Graeco-Latin words predominate in particular semantic fields:

> It was concluded that [an] unequal active access [to this vocabulary] is probably a profound and significant factor in contributing to the speakers' school failure, since words of Graeco-Latin origin populate, almost to the exclusion of other types of word, the specialist knowledge categories of the school curriculum.

(Corson, 1983:5.)

Corson's argument centres on types of vocabulary, or frequency levels. He estimates that between 65 and 100 per cent of specialist vocabularies in academic disciplines are derived from Graeco-Latin words, the use of which increases progressively in school texts. This argument obviously applies to tertiary education, where students who have "good active and passive control over this part of the English lexicon have easier access to the semantic fields and genres in which they predominate" (Romaine, 1984:215). Romaine examines the issue of active vocabulary in terms of the demands of narrative as opposed to expository text. She concludes that expository texts require greater evidence of language ability as these involve explanation rather than description. Thus, in terms of Corson's (1983:5) description of education as a process of "language on display", Romaine illustrates the importance of increasing proficiency in Graeco-Latin vocabulary (1984:216).

The 'Graeco-Latin' vocabulary discussed above clearly incorporates both the academic and advanced frequency levels. This argument is based on the assumption that vocabulary relating to "specialist knowledge categories" (Corson, 1983:5) would include jargon in addition to academic vocabulary not specific to a particular domain of study. The extent to which advanced vocabulary is derived from Graeco-Latin origin is substantiated by Goulden, et al. (1990:356), who state that "around 66 per cent of the low frequency words of English come from French, Latin, or Greek". In view of Corson's assertion of the impact of these frequency levels on academic success, the finding that students seem to be familiar with only 43.2% of the academic vocabulary and 22.4% of the advanced vocabulary on average (Table 4.2), suggests an urgent need to address these areas of the students' vocabularies, with a focus on explicit teaching of the academic vocabulary. In support of this argument, Beck, McKeown and Omanson (1987:155) propose that high frequency

academic vocabulary items are candidates for instruction as a knowledge of these words "can have a significant impact on verbal functioning". The question of vocabulary instruction is addressed in Chapter 5.

The impact of academic vocabulary on academic performance has been examined by Cunningham and Moore (1993). In a study which explored the relationship between the vocabulary of written comprehension questions and students' reading comprehension performance, they found that the inclusion of academic vocabulary in questions resulted in a significant decrease in comprehension. On the basis of their findings, Cunningham and Moore (1993:178-179) propose that the role of academic vocabulary within the classroom deserves attention.

> Knowing the terms *cause, effect, topic, antagonist*, etc. seems to us to enable a degree of precision and economy when communicating about a passage that every-day vocabulary cannot ordinarily provide. Moreover, it is possible that knowing such terms enables or, at least, facilitates the ability to cognitively manipulate abstract rhetorical entities the way explicitly knowing the terms *word* and *sound* helps emerging readers learn to identify words.

Similarly, Marshall and Gilmour (1993:69), in investigating students' knowledge of 'subtechnical' or academic vocabulary, argue that insufficient lexical knowledge affects the activation of content schemata when reading, which in turn affects the comprehension of texts. As a means of addressing comprehension problems, they recommend both extensive reading and explicit teaching of academic vocabulary.

The aim of this discussion was to illustrate the fundamental role of academic vocabulary in education while concurrently addressing the subsidiary role of advanced vocabulary. In conclusion I refer to the study by Cummins (1981, in Goulden *et al.*, 1990). Cummins examined the effect of the gap between the vocabulary size of second-language speakers and that of native English speakers. He argues that the difference in vocabulary size between first- and second-language speakers of the same age has significant implications for their performance on verbal tests of intelligence. Cummins links his findings to the distinction between basic interpersonal communication skills (BICS) and cognitive-

academic language proficiency (CALP). He proposes that these models of language proficiency relate to the distinction between basic and academic vocabulary. Second-language speakers who acquire fluency in BICS are likely to give the impression of general competence in the language. This communicative ability, however, is restricted to a command of high-frequency, basic vocabulary and does not reflect proficiency at the levels of academic and advanced vocabulary.

> [T]ests of vocabulary and reading comprehension reveal that they lack the underlying proficiency (or CALP) that is required to cope with academic study through the medium of their second language. Therefore, in this context, measures of vocabulary size - particularly the size of academic vocabulary - are important indicators of the ability of second language learners to achieve academic success.
>
> (Goulden *et al.*, 1990:342.)

Finally, the role of basic vocabulary is best illustrated with reference to a study by Coady *et al.* (1993) which examined the relationship between high-frequency vocabulary and reading proficiency. The results were significant and indicate that "increased proficiency in the high-frequency vocabulary of English [through computer-assisted learning] will lead to an increase in reading proficiency" (Coady *et al.*, 1993:217; cf. Grabe, 1991). The significant contribution of basic vocabulary to reading comprehension is logical if we consider that basic vocabulary provides 83.4% coverage of non-fiction texts (Nation and Hwang, 1995; §2.4). A second argument for the role of basic vocabulary relates to Laufer's (1994:30) finding that "there is no correlation between the progress in vocabulary size and the progress in lexical variation". Laufer argues that lexical variation, that is, writers' ability to express themselves by effectively varying existing vocabulary, impacts on the quality of writing. In terms of this argument, where the second-language student has a relatively small vocabulary, the effective use of basic vocabulary alone should contribute towards academic performance.

The roles of basic vocabulary in reading and writing outlined above seem to corroborate the findings in this study. While academic vocabulary obviously plays a key role in academic discourse, basic vocabulary provides a framework in terms of which the student gains

access to both academic and advanced vocabulary. Thus, while academic vocabulary relates directly to the concepts presented within academic domains, and is therefore the most significant indicator of academic performance, basic vocabulary provides entrance to the higher frequency levels. Evidence of implicational scaling across frequency levels (§§2.3.2 and 4.3) clearly indicates that the largest proportion of the students' vocabulary comprises basic vocabulary, with considerably fewer items comprising the academic vocabulary. A command of basic vocabulary then provides initial access to academic discourse, as reflected in the significant, positive correlation between the basic vocabulary scores and the final marks. Of even more significance, however, is the students' command of academic vocabulary which relates more closely to broader cognitive skills. The academic and, to a lesser extent, the advanced vocabulary thus enable students to both assimilate concepts and to demonstrate their language ability.

> It is important to recognize that the core vocabulary argument, that the 2,000 most frequent vocabulary items account for 80% of all words in texts, may be useful for basic reading instruction (e.g., Nation, 1990); however, it falls far short of the needs of academically oriented ESL students who, in fact, need to know many of the less frequent words.
>
> (Grabe, 1991:392.)

The issue of practical significance in this study is addressed again in the regression analysis and in the final evaluation. The degree to which these results may be considered reliable and valid is addressed below.

4.5   Estimating reliability and validity

A measure of reliability was considered to determine the degree to which the results are consistent and dependable. Leary (1991:57) states that "reliability reflects the proportion of the total variance in a set of scores that is systematic, true-score variance". A finding of low reliability would therefore suggest that the tests contain variables which have not been accounted for. In order to show reliability the results of Tests A and B should correlate.

Validity, on the other hand, was measured in terms of content validity and construct validity. For the purposes of this study, content validity concerns the degree to which the tests measure knowledge of vocabulary, while construct validity concerns the degree to which the classification of the vocabulary items as basic, academic and advanced is justified. Both validity measures are evaluated on the basis of prior research.

4.5.1 Reliability

Given that both vocabulary tests were intended as equivalent measures of basic, academic and advanced vocabulary size, an alternate-form reliability test was used to establish their compatibility. This assessment of reliability is based on a correlation analysis of the two tests. The analysis indicates the degree of similarity between the tests and provides a measure of reliability in estimating the amount of variance accounted for. The vocabulary scores for each frequency level of Test A were thus compared with the corresponding scores for Test B by means of the Pearson product-moment correlation coefficient. The degree of correlation between Tests A and B is illustrated in Table 4.6 below.

Table 4.6

Alternate form reliability of Tests A and B

		TEST A				
		Basic vocab. (1000-list)	Basic vocab. (2000-list)	Basic vocab. (combined)	Academic vocabulary	Advanced vocabulary
TEST B	Basic vocab. (1000-list)	$r = .24$ $p<.0001$				
	Basic vocab. (2000-list)		$r = .33$ $p<.0001$			
	Basic vocab. (combined)			$r = .46$ $p<.0001$		
	Academic vocabulary				$r = .45$ $p<.0001$	
	Advanced vocabulary					$r = .14$ $p<.0001$

N = 653

The correlations between the frequency levels of each test are statistically significant at the $p<.0001$ level, thereby indicating that the tests are reliable. It should be noted, however, that the most important correlations are moderate for basic and academic vocabulary and very low for advanced vocabulary (Mulder, 1986:73) and account for a relatively small proportion of the variance in the results. More specifically, it may be inferred that the relationship between the variables at the level of advanced vocabulary is weak since the coefficient is only .14. Despite the relatively low correlation coefficient, this finding is nevertheless significant since significance is a function of sample size. It must be acknowledged, however, that the relationship between the variables appears to be negligible, and so has limited practical significance. The low correlation accounts for only a small variance in the dependent variable, thereby indicating a weak, but definite, relationship. It is possible that this very low correlation is due to guessing on the part of the students. As second-language speakers, the students are likely to have a poor grasp of

advanced vocabulary in particular. If a number of students selected answers at random this would to some extent explain the small amount of variance accounted for, and hence moderate the poor reliability estimate at the level of advanced vocabulary (cf. §4.4.2.2).

A second consideration is the number of test items at each frequency level. The small number of items at the levels of basic and advanced vocabulary may account further for the low correlations. It is apparent that, in the case of the 1000- and 2000-word lists, where the lists were combined, and the calculation based on 10 rather than 5 items, the reliability estimates improved. Thus the weak reliability of this test may be due to the number of items tested rather than the design of the test itself.

It is clear from the above discussion that, while the correlations have statistical significance on the basis of the p-values and are therefore reliable, they are of limited practical significance since they account for a low level of variance. I return to this issue in the evaluation of this study.

4.5.2   Validity

The validity of this study was evaluated in terms of content and construct validity, as outlined above. Content validity is inferred on the basis that the use of the multiple-choice cloze format to test vocabulary knowledge is well precedented. In support of this assumption Read (1997:304) claims that multiple-choice vocabulary tests "were valued both for their technical qualities and their apparent validity as indicators of language ability in a broad sense".

The tests are assumed to have construct validity on the basis of the results in Tables 4.1 and 4.2 (§4.3). These results illustrate a substantial degree of implicational scaling across the three frequency levels (cf. Read, 1988 and 1997). They suggest that the classification of test items according to frequency level using the LFP is a valid assessment of level of difficulty of basic, academic and advanced vocabulary. The results are therefore compatible with Laufer and Nation's (1995:319) conclusion that the LFP is valid and reliable (cf. Read,

1997:314). This assumption of validity was derived from the finding that the LFP correlated well with Nation's Vocabulary Levels Test (Nation, 1990; Laufer and Nation, 1995).

The following section presents and discusses the results of a multiple regression analysis of the data. This analysis was intended as a further test of the extent to which vocabulary is "one of the best predictors of educability" (Diack, 1975:27).

4.6   Analysis of the findings using multiple regression

The purpose of the multiple regression analysis was to elaborate on the results of the correlation analysis by examining the contribution of each independent variable in more depth. This further investigation of the results is motivated by the apparent lack of clarity regarding the effect of vocabulary knowledge on reading comprehension and, by extension, on academic performance.

> [A]ttempts to establish a direct causal link that supports this strong statistical relationship [between vocabulary knowledge and comprehension] have been equivocal and inconclusive. That is, we cannot say that vocabulary acquisition in itself increases comprehension; from the evidence available, the most we can say with assurance is that sometimes it does and sometimes it doesn't.

> (Ruddell, 1994:414.)

While correlation merely expresses the relationship between independent and dependent variables, regression provides a more precise assessment of the extent to which each independent variable affects the dependent variable. The regression analysis was thus intended to investigate how much each frequency level contributed to the variance in the final mark (cf. Astika, 1993; Brisbois, 1995).

In line with Diack's (1975) suggestion that vocabulary is a significant indicator of academic performance, one of the aims in conducting the regression analysis was to determine the

extent to which the scores for basic, academic and advanced vocabulary could be used to provide an accurate estimation of academic performance. Because regression yields a model that can be used for prediction, a forward stepwise multiple linear regression analysis was conducted, using the SAS programme to analyse the data (Appendix C). The students' scores for basic, academic and advanced vocabulary were taken as the independent variables, with the final mark as the dependent variable.

The linearity of the relationships between the dependent and independent variables was examined by plotting the data of the dependent variable against each independent variable. (The procedure followed in this analysis is outlined in Appendix C.) The results show that the relationships between the dependent variable (final mark) and the independent variables (vocabulary scores) varied considerably, with academic vocabulary accounting for the largest amount of variance. These results are presented in Table 4.7 below.

Table 4.7

Results of the stepwise multiple linear regression analysis

Independent variable	Partial R^2	Cumulative R^2	F-ratio	Probability level	N
Academic vocab.	0.1460	0.1460	111.29	$p<.0001$	653
Basic vocabulary	0.0066	0.1526	58.52	$p<.0001$	653

The results indicate that basic and academic vocabulary scores together predict 15.26% of the students' academic performance, as measured by the final mark. The most significant predictor of academic performance is the academic vocabulary score, which alone accounts for 14.60% of the variance in the final mark. The second most significant independent variable, basic vocabulary, accounts for 0.66% of the variance in the final mark. These results meet the $p<.0001$ level of significance, as determined by the F-ratio which tests the effect of each variable. The advanced vocabulary score did not meet the .10 level required for inclusion in the model, which accounts for the omission of the third independent variable

from the model. The academic vocabulary thus contributes most significantly to the total score variance. The effect of the basic vocabulary is statistically significant, but shows minimal real contribution to the variance in the final mark.

The limited nature of the regression analysis is illustrated in a calculation of the range of marks which the model is able to predict. (The equation is given in Appendix C.) In terms of this model, the highest possible mark estimated for academic performance, where the academic vocabulary score is 30 and the basic vocabulary score is 20, is 66.5%. The lowest possible mark, in the case where the scores for both academic and basic vocabulary are 0, is 32.2%. This suggests that the greatest proportion final marks clustered between approximately 32% and 67%. The actual spread of the students' final marks is illustrated in Graph 4.3 below.

Graph 4.3

Std. Dev = 12.45
Mean = 48.1
N = 653.00

FINAL YEAR MARK
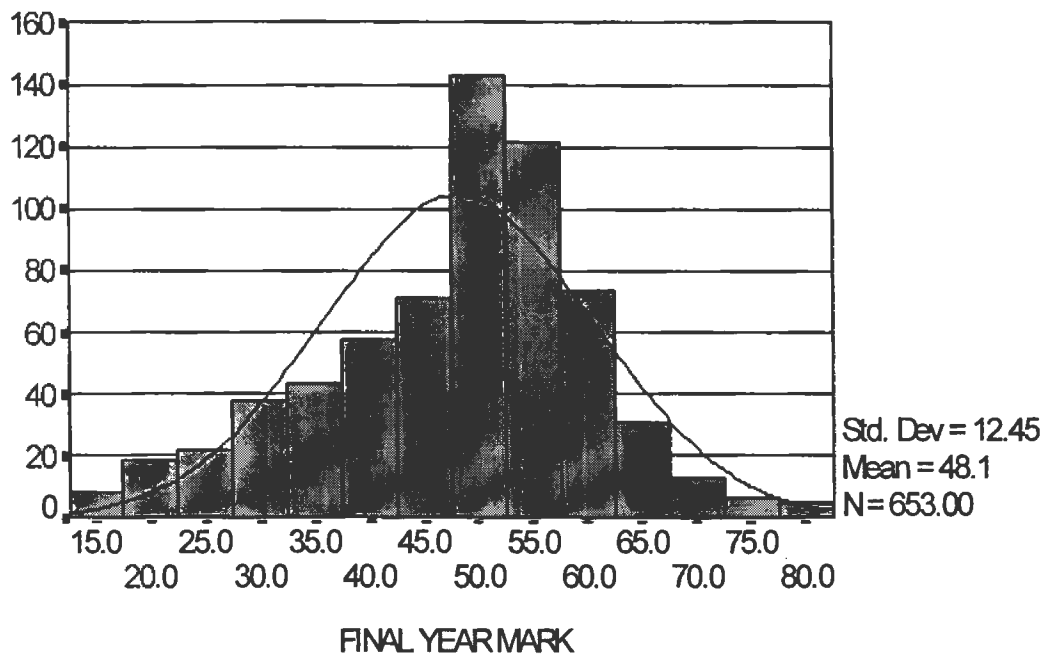
Distribution of final marks

The distribution of final marks could then account for the low correlation since there can be little interaction between vocabulary scores and final marks outside this range. The very low correlation between the final marks and the test scores for advanced vocabulary in

particular provides support for this argument since the results for advanced vocabulary cluster at the lower end of the scale (Table 4.3), between 13.9% and 31.0%. Thus, given that the largest proportion of scores for the advanced vocabulary test fall below 31%, while a similar proportion of final marks occur above 32%, there will obviously be minimal interaction between the variables.

The results of this regression analysis can be compared with those of Astika (1993) and Brisbois (1995). Astika's study showed that, of the five components which contribute to the assessment of writing (content, organisation, vocabulary, language use and mechanics), vocabulary accounted for 83.75% of the variance in the total scores. Vocabulary was thus found to be the best predictor of writing proficiency scores. One possible explanation for the large discrepancy between the findings in my study and that of Astika is the nature of the dependent variables. Astika (1993) focused specifically on writing proficiency, with the result that the dependent variable was more strictly controlled (cf. Engber, 1993; Laufer and Nation, 1995; Linnarud, 1986).

Brisbois's study examined the contribution of first-language reading, second-language vocabulary and second-language grammatical skill to reading scores in the second language. She found that in all cases the vocabulary scores accounted for more of the variance in reading comprehension than did grammar. For beginners, vocabulary scores were the primary contributor, accounting for 10.11% of the variance in reading scores, while for upper level subjects it accounted for only 6.26% of the variance. Brisbois's study suggests that the relationship between reading comprehension and vocabulary proficiency changes as reading ability increases. Thus vocabulary accounted for most of the variance in beginner learners, while first-language reading ability contributed more in the case of upper level subjects. She concludes that "as L2 vocabulary knowledge gradually increases, and readers gain in automaticity, various reading strategies and skills can be used more and more" (Brisbois, 1995:578). It must be noted that where the dependent variable is reading comprehension rather than writing proficiency, the degree of variance accounted for by the vocabulary scores is far lower, and thus more in line with my own findings. This could be accounted for both in terms of productive and receptive uses of language, and in terms of

the broader cognitive and linguistic skills required for writing as opposed to reading. Thus, while writing proficiency tests the students' active vocabulary, reading comprehension may rely to a greater extent on both active and passive vocabularies. This suggests that the type of vocabulary measured in each case then varies considerably.

Despite the low correlations for the regression analysis, it may be argued that the academic vocabulary scores in particular generally predict final marks in the correct direction. Given this argument, the academic vocabulary scores may be used as a rough indicator of academic performance. That is, where students pass the academic vocabulary test they are likely to pass at the end of the year, and where students fail the academic vocabulary test their final mark is predicted to be less than 50%. This argument is supported by the correlation between academic vocabulary and the final mark as illustrated in Table 4.8 below.

Table 4.8

Cross-tabulation of academic vocabulary score and pass rate

| | | ACADEMIC VOCABULARY | |
		Fail	Pass
FINAL MARK	Fail	45.2%	16.7%
	Pass	54.8%	83.3%

N = 653

This table indicates that 83.3% of the students who passed the academic vocabulary test passed at the end of the year, while 45.2% of those who failed the academic vocabulary test failed the year ($r = .35, p < .05$). These results then suggest that academic vocabulary alone could be used to determine trends in academic performance.

4.7   Summary of results

The results indicate that there is a positive correlation at each frequency level between students' basic, academic and advanced vocabulary and their academic performance. In addition the results indicate that, of the three frequency levels, academic vocabulary relates most significantly to academic performance. In contrast to this, advanced vocabulary was shown to be the least significant variable, having the weakest correlation with academic performance.

However, based on the results of the regression analysis, vocabulary as a whole appears to be a weak predictor of academic performance, since 84.74% of the variance in the final mark is *not* accounted for by the vocabulary test scores. Although this finding does not negate the results of the correlation analysis, it does bring into question previous assumptions made with regard to the importance of the role of vocabulary. The statement that "vocabulary difficulty has consistently been found to be the most significant predictor of overall readability" (Chall, 1958; Klare, 1974 in Nation and Coady, 1988: 97) is certainly disputable. It may be argued that the notion of 'vocabulary' is not sufficiently nuanced, with the result that a general, sweeping statement about vocabulary fails to take into account the complexity of the relationship between vocabulary size and academic performance. In fact, careful consideration of the results of the regression analysis serve to reveal more precisely the role that academic vocabulary in particular has in the academic context. In addition, these results suggest that more sensitive instruments are required to measure vocabulary and its degree of influence, since the measures used in this study merely reflect trends in academic performance. The following argument by Ruddell (1994:41) reinforces this conclusion.

> [A]ny number of variables impinge on our understanding of the links between vocabulary knowledge and comprehension and serve both to obscure that understanding and to emphasize the complexity of the relationship itself.

The above conclusions were drawn from the results of the correlation and regression analyses at each frequency level. In contrast, the discussion which follows focuses on the

analyses of individual test items and distractors, and is based on a more qualitative assessment of the vocabulary test items. The aim of this analysis was to identify anomalous items, as explained below. The results of the two item analyses referred to in this section are presented in Appendix D.

4.8   Analyses of the test questions

The purpose of these analyses was to assess the responses to each of the multiple-choice questions in the two vocabulary tests. In the analysis of individual test items, the total number of correct responses to each item was examined in relation to other scores within the same frequency level. The aim was to establish which of the test items did not correspond with others at the same level of frequency, and thus to provide a qualitative assessment of the students' performance at each frequency level. The second analysis explores the degree to which each question discriminated between weak and good students. In order to do so, the correct option is examined in relation to the distractors. For each question, if the majority of students who obtained higher overall scores for the test selected the correct option, it may be inferred that the option discriminates in the right direction. However, where those students who obtained higher scores overall selected a distractor, it is apparent that the test item discriminates in the wrong direction, and is thus problematic (cf. Appendix D). This analysis focuses on those items which did not discriminate in the right direction, and possible reasons for this failure to discriminate among students of different proficiency levels are then proposed. The implications of this analysis for researchers in the field of vocabulary testing and for educators are discussed in Chapter 5.

4.8.1 Analysis of individual test items

The first item analysis considers the number of correct answers for each test item (Appendix D: Tables 1 and 2:135-136), and is intended to provide an indication of the students' overall performance at each frequency level. This analysis indicates that, with the exception of items such as A2, A3, A9, A16, A30, B16, B24 and B28, the scores generally correspond with the level of difficulty on which the frequency lists were based. In other words, there were more correct responses to questions testing basic vocabulary than to those testing academic vocabulary. Similarly, students performed better at the level of academic vocabulary than at the level of advanced vocabulary. Those instances in which students performed extremely poorly in the test of basic vocabulary could be attributed to a confusion of similar lexical forms (§3.3.2.2). The incorrect responses to the test items *attempted* (A3) and *preferred* (A9), for example, appear to justify the assumptions regarding synforms on which the selection of distractors was based. The results of the item analysis for *attempted* show that, while 45% of the students selected the correct option, 37% chose the distractor *accepted.* In terms of the broad interpretation of synform proposed earlier in this study (§3.3.2.2), the orthographical similarity between this distractor and the test item may account for the poor result. Similarly, the item analysis for *preferred* shows that the majority of students (51%) selected the distractor *referred* in preference to the test item. (*Prefer* and *refer* constitute Category 4 synforms in terms of Laufer-Dvorkin's (1991:207-208) classification of similar lexical forms.) In addition, the analysis of test scores confirms the assumption that students have a poor grasp of conjunctives and adverbials (§3.3.2.2), as reflected in the results for the items *although* (A2: 37.6%) and *however* (B1: 49.6%).

The other extreme, in which students performed exceptionally well on an item in the advanced vocabulary test (B28: 60.8%) is less easily accounted for. It is possible in the case of *integral* that students associated the test item with the related word *integrate*, which occurs in the UWL, and is therefore more likely to be familiar. If this conjecture is valid, then the classification of *integral* as a separate item rather than a member of the related

word family is problematic. Since there are only a few test items at the level of advanced vocabulary, the high score for this word may have skewed the overall result.

4.8.2 Analysis of each test item and related distractors

The second item analysis explores the degree to which each question discriminates among students of differing proficiency levels (Appendix D: Tables 3 and 4:137-138). A test item discriminates in the right direction when students who perform best overall select the correct option. On the other hand, a test item discriminates in the wrong direction if those students who perform best overall select a distractor rather than the correct option. In examining the results of both analyses I hoped to identify problematic items which either did not compare with other results at the same frequency level or did not discriminate in the right direction.

The results of this analysis illustrate that only 5 of the 60 test items (8.3%) did not discriminate in the correct direction, and so did not have the expected discrimination value. In two instances (A1 and A4) a small sample of students (4% and 1%, respectively) who obtained the highest overall score selected the incorrect option. However, since the majority of students selected the correct response, these items were nevertheless regarded as reliable. The remaining items which did not adequately discriminate between weak and good students occurred at the level of advanced vocabulary (A28, A29 and B30). The frequency level of these items suggests that the failure to discriminate may be due to either guessing by the students or students' confusion of similar lexical forms (cf. §2.3.2). These items were nevertheless retained in the analysis of results since the overall score at the level of advanced vocabulary is low, thereby suggesting that these results are a true reflection of the students' weak grasp of low frequency items in general.

This analysis of options seems to support the argument that the selection of lexically similar distractors results in a more discriminating multiple-choice test. As predicted, potential synform errors possibly account for the apparent difficulty of test items A11, A30, B14 and B30, as predicted (§3.3.2.2). This may explain, for example, the results of A11, where the

majority of students selected the distractor *formulate* in favour of the test item *facilitate*, and A30 where the majority selected the distractor *emancipation* rather than the test item *emulation*. Besides items which are orthographically similar, the analysis suggests that students also confuse semantically similar items. This is illustrated in A27 where the majority of students chose the distractor *predictable* rather than the test item *preconceived*. These words are arguably similar in meaning since both relate to forming an opinion regarding an event or idea in advance and so share the underlying semantic feature 'to anticipate'. Further examples of possible semantic confusion are A21, in which the majority of students selected the option *self-evident* instead of the test item *superficial* (where the similarity is 'apparent') and A24, in which students chose *imply* rather than *infer* (where there is a confusion of agency). Since the students who selected the incorrect options in these instances did not perform best overall, these items may be considered good discriminators of student proficiency.

If the argument regarding lexically similar items is valid, it may be inferred that test items such as A6 (*depends*) and B18 (*devised*), in which the majority of students selected the correct option despite similar distractors, reflect a greater degree of understanding of the test item (cf. Meara and Jones, 1987; §3.3.2.2). The test may then be assumed to be a reliable indicator of vocabulary knowledge. Those questions which did not adequately discriminate between students will be examined further in Chapter 5 when the practical implications of these results are considered.

4.9   Conclusion

In examining the relationship between vocabulary size and academic performance, what emerged from this study is that these variables may be causally connected. In other words, an increase in a student's vocabulary size could result in an overall improvement in academic performance (cf. Brisbois, 1995; Coady, *et al.*, 1993; Laufer, 1994 and 1997a; Medo and Ryder, 1993; Nagy and Herman, 1987). This inference is drawn from the results of the regression analysis which were significant at the $p<.0001$ level for both basic and

academic vocabulary. Thus, although basic and academic vocabulary appear to be weak predictors of the final mark, there is a definite relationship, particularly in the case of academic vocabulary (cf. Cunningham and Moore, 1993; Marshall and Gilmour, 1993). Nevertheless, it cannot be stated explicitly that an increase in vocabulary size alone will inevitably effect reading comprehension (Nation, 1990).

While vocabulary size influences academic performance, it clearly cannot determine academic performance. This conclusion is based on the fact that vocabulary is merely one component of language, and so must be considered in relation to other variables. Thus, although vocabulary size correlates significantly with academic performance, the correlation is weak. This finding clearly corresponds with Laufer's (1992:130) claim regarding the vocabulary requirements for reading comprehension.

> [T]he turning point of vocabulary size for reading comprehension is 3,000 word families. In other words, good L1 readers can be expected to transfer their reading strategies to L2 when they have reached this level. Until then, reading in L2 will be hampered by an insufficient knowledge of vocabulary.

However, given that Laufer's focus is on *good* first-language readers, this claim must be interpreted with caution. The assumption is that these readers already have the necessary reading skills and are able to transfer them to the second language, provided they have the required vocabulary. However, this raises a question regarding students whose reading skills and lexical proficiency are poor. Would increasing the vocabulary size of these students necessarily enable them to comprehend texts in the second language? It seems clear that, where readers lack other cognitive tools required for reading comprehension, it is very unlikely that the exclusive development of vocabulary will affect the comprehension of texts. Obviously, in these instances, any investment in improving vocabulary proficiency must be considered in relation to other skills. This condition on the role of vocabulary is reinforced by Nation and Waring.

[I]t should not be assumed that if a learner has sufficient vocabulary then all else is easy. Vocabulary knowledge is only one component of language skills such as reading and speaking. It should also not be assumed that substantial vocabulary knowledge is always a prerequisite to the performance of language skills.

(Nation and Waring, 1997:6.)

Nevertheless, if one considers the possible number of variables which contribute towards academic performance, for instance a knowledge of logical relations, coherence and text structure (Pretorius, 1995:39), then the finding that academic vocabulary alone may account for 14.6% of the variance has potential practical significance. The implications of these results are examined in Chapter 5.

Chapter 5

Conclusion

5.1   Introduction

This chapter reviews the contribution of this study to research on vocabulary size and examines the limitations and practical implications of the study with a view to further research in the field.

5.2   Review

The general aim of the study was to contribute to the debate surrounding the extent to which vocabulary size affects academic performance. This section considers the relevance of the findings with respect to the aims and research problems outlined in Chapter 1.

5.2.1   Summary of major findings

The results of this study clearly illustrate that, of the three levels of vocabulary investigated, academic vocabulary contributes most significantly to academic performance. This finding then supports the arguments presented in Chapter 3 regarding the relevance of academic vocabulary to academic discourse (§3.2.1.2). The consequences of this finding for students studying through the medium of a second language are explored below (§5.2.2).

The students involved in this study show an understanding of less than 70% of the basic, high-frequency vocabulary of English. Again, the implications of this finding, bearing in

mind that these words provide coverage of 80% of most written texts (Sutarsyah *et al.*, 1994; Nation and Hwang, 1995), are extensive. The results of advanced vocabulary testing indicate that these items correlate most weakly with the measure of academic performance used. It is perhaps surprising that the correlation between advanced vocabulary and academic performance was particularly low, with the result that the advanced vocabulary contributed very little to the variance in the final mark (§4.6). A common-sense perception would be that the students' proficiency in higher levels of vocabulary would be most directly an index of academic potential. In other words, one would expect that students who perform well academically have a particularly good knowledge of advanced vocabulary (cf. Diack, 1975). However, the low frequency of the advanced vocabulary items in the reading material, and the nature of the words that comprise this vocabulary (§3.2.1.3), accounts for the fact that academic vocabulary correlates far more strongly with academic performance.

5.2.2   Contribution of the study

Research on reading comprehension and writing quality has emphasised the relevance of vocabulary in both productive and receptive use of language. One finding in this regard suggests that skilful reading requires a good knowledge of at least 5000 words (Coady, 1997a:287). In reviewing the literature on aspects of vocabulary size, vocabulary testing and vocabulary proficiency, this study has presented a range of views on the extent to which vocabulary influences broader cognitive and linguistic skills. While the majority of sources concur that vocabulary is intrinsic to linguistic competence, there is little agreement on the nature of the relationship between vocabulary proficiency and general language ability.

Tests of vocabulary size in this study have shown that the second-language students' overall grasp of basic, academic and advanced vocabularies does not meet the lexical demands of the prescribed reading material on which their studies are based. It is evident from these findings that the students require explicit instruction in basic vocabulary in general and academic vocabulary in particular if they are to attain the minimum vocabulary required for comprehension. This conclusion is based on the assumption that learners' recognition

vocabulary must pass a particular threshold before the learners are able to transfer their first-language reading skills to the second language (Brisbois, 1995).

In light of the findings of this study, my own contributions to the field of vocabulary testing derive from this investigation into the relationship between vocabulary size and broader cognitive skills such as reading and writing, as reflected in students' academic performance. A significant contribution of the study was to establish more clearly the effect of academic vocabulary on academic performance through the implementation of a test of vocabulary size which took account of frequency level. The results of this study suggest that tests of academic vocabulary may be used to obtain a broad, general indication of a student's potential performance. This finding points to my second contribution to the field. This study raises a number of questions regarding the common assumption that vocabulary correlates significantly with other measures of verbal aptitude, and thus that the contribution of vocabulary size to academic performance is significant (§2.2).

Taking into account the number of extraneous variables that may have contributed to the low correlations in the results, my findings suggest that the matter of the relation of vocabulary size to academic performance remains a point of debate. This claim is made in view of the fact that the findings were shown to be highly significant statistically ($p<.0001$), despite the low correlations. Thus, the results of this study indicate that vocabulary size in general contributes very little to academic performance on the whole since it accounted for very little variance in the final mark.

5.2.3   Limitations of the study

A central concern in the analysis of results in the present study is that the assessment of vocabulary size at each frequency level was not equally weighted. The unequal distribution of test items at the levels of basic, academic and advanced vocabulary could account for the differences between the scores. The small number of test items at the level of advanced vocabulary might account for the low correlation coefficient at this level. This is confirmed, in part, by the finding that the 1000- and 2000-word tests correlated more highly when

combined in the test of reliability (§4.5.1: Table 4.6). The effect of the number of test items at each frequency level should be taken into account in any further tests of this nature.

A second concern, related to that outlined above, involves the limited sample of items used to test vocabulary size. Due to time constraints surrounding the administration of the vocabulary tests, and the effects of fatigue on the subjects' performance, it was not feasible to conduct a multiple-choice cloze test of more than 60 items. A small sample of test items inevitably increases the chances of misrepresentation of vocabulary size since the opportunities afforded the subjects to illustrate their vocabulary knowledge is clearly reduced (cf. Meara, 1987; §4.5.1).

Although every effort was made to control extraneous variables in the design and implementation of the tests, this study was subject to the methodological limitations inherent in studies of this sort. Thus, the fact that the tests were completed on a voluntary basis may have influenced the findings. The conjecture in the analysis of test questions (§4.8.2) that students were more inclined to guess at the level of advanced vocabulary, could equally be accounted for by the possibility of students' growing disinterest in the test due to increasing difficulty. Further research could possibly account to a greater extent for any independent variables not considered in this study, such as motivation, the increase in difficulty of the test items in relation to frequency level and the effect of the sentential context of the test item on the students' answers.

The discrepancy in distribution of vocabulary test scores in relation to the distribution of final marks is another limitation of this study (cf. §4.6). Given that academic performance was measured in terms of this mark, which comprises an average of the year mark and examination result, the results were constrained by this assessment of student performance. The finding that the largest proportion of the final marks ranged between approximately 32% and 67% (§4.6: Graph 4.3) while the range of vocabulary scores was considerably lower accounts to a large extent for the low correlation between the vocabulary test scores and the final marks. On the one hand, of course, this correlation could simply be a valid reflection of the degree to which vocabulary influences academic performance. On the

other hand, however, the measures of student ability used to compile the final mark, such as the assignments and the examination, are not optimal assessments of linguistic and cognitive competence. Given this assumption, the vocabulary test, which is arguably an implicit assessment of reading and writing skills (§§2.2 and 2.4), may be a more sophisticated measure of student performance. The results of this study do appear to bring into question the level of competence required of these students over the course of the year.

Another limitation of the study was uncovered in the analysis of the test items and related distractors (§4.8.2). The few items (8.3%) that did not discriminate in the right direction may have affected the results of the study, particularly at the level of advanced vocabulary. The decision to retain these test items should possibly be re-evaluated in further studies. In addition, given the nature of the multiple-choice cloze test, it is apparent that this test was a measure of breadth rather than depth of vocabulary knowledge (§2.3.4). Future studies of vocabulary knowledge, based on Read's (1993) word association test, could possibly investigate the relationship between vocabulary size and academic performance in more depth.

Finally, with regard to the students themselves, the issue of the problems surrounding the testing of a heterogenous group are addressed in the discussion of the general hypothesis (§4.4.1.2). The finding by Meara and Jones (1987) that the correlation between vocabulary size and an assessment of language skills improves in relation to the degree of homogeneity of the test group is clearly a variable which may have affected the results of this study. Another consideration relating to the student group is the fact that this study does not account for possible growth in vocabulary size in the course of the year. While the vocabulary tests were conducted in the middle of the academic year, 50% of the final mark comprises the examination result, which provides some indication of the students' level of competence at the end of the year. The lag between the time at which the test was conducted and the examination period could account for the discrepancy between vocabulary test scores and the final marks. It is possible, of course, that the students' vocabulary knowledge improved substantially in the six months between the test and the

examination. In future studies of a similar nature, such maturational variables should be more carefully controlled.

The final limitation to be addressed is the issue of the administration of the tests. Although explicit instructions were issued to students regarding requirements for the completion of the tests, it cannot be assumed that the students' approach to the tests was consistent or in line with the instructions issued.

In conclusion, this section outlines a number of factors which could possibly have affected the results of this study. There is a growing need for more fine-tuned research into the effect of vocabulary on linguistic competence in general and aspects of academic performance in particular. Despite the fact that interest in vocabulary has increased over the last two decades (§1.3), we still seem to be at the starting gate in many respects (cf. Meara, 1996a). The following section poses some of the questions which have arisen out of this study.

5.3   Implications for further research

A significant implication of this study relates to the degree of coverage of academic vocabulary (cf. §2.3.2). Given the assertion by Worthington and Nation (1996:2) that the UWL provides coverage of approximately 8.5% of academic texts, and that, together with the 2000-word list, it "brings the total coverage to 91.9%" (Nation and Hwang, 1995:40), it is apparent that both the basic and academic vocabulary are required to meet the 95% coverage stipulated by Laufer (1989) for adequate comprehension. In support of this claim, Hirsh and Nation (1992) propose that a vocabulary size of around 5000 word families is required for 97-98% coverage of unsimplified texts.

The findings of this study suggest that these students do not have the level of either basic or academic vocabulary required to meet the demands of academic discourse. Considering that the items in the academic word list occur at the 4000 to 6000 frequency level (Laufer

and Paribakht, 1998:377), it may be inferred that these students, who seem to know less than 70% of the 2000-word list on average (§4.3: Table 4.3), are certainly unlikely to have an adequate grasp of words at the lower frequency levels. This is clear from their performance at the levels of academic and advanced vocabulary in particular. The vocabulary of these students, therefore, does not seem to approach the minimal requirement for adequate reading comprehension.

Laufer (1989:320) refers to studies of the vocabulary size of university students from a number of different countries. She reports that the English vocabulary of the second-language students tested in these studies ranges from 500 to 3000 word families.

> A gap of 2000-4000 words between the amount of words they know and they should know turns reading into 'mission impossible'. It is through increasing vocabulary to 5000 and ensuring a lexical coverage of 95% that reading has a good chance of becoming 'mission surmountable'.
>
> (Laufer, 1989:320.)

What, then, are the pedagogical implications of these findings? To what degree does vocabulary affect the academic performance of second-language students whose cognitive and linguistic skills do not address the requirements of the academic context, and whose first-language CALP skills are particularly poor? Would explicit teaching of vocabulary adequately contribute to academic performance when students lack primary reading skills in their mother tongue? Finally, what should tertiary institutions be doing to address these problems?

In my conclusion to the discussion on the contribution of this study (§5.2.2), I referred to Brisbois's (1995) finding that students must achieve a minimum degree of competence in vocabulary before they are able to draw on their first-language reading skills for the comprehension of texts in the second language. This finding, however, rests strongly on the assumption that students have, in fact, attained first-language reading skills. Clearly, where students lack adequate CALP skills in their mother-tongue, improving lexical proficiency alone does not address the question of how students then cross the reading threshold without attention being given to reading skills. I would argue on the basis of the

relatively low correlations attained in this study that other variables must be drawn into the equation when assessing factors which influence student performance.

Given these conditions surrounding explicit vocabulary teaching, I would nevertheless strongly recommend that developmental programmes take both basic and academic vocabulary into account in their design. Considering the degree of coverage by the high-frequency and academic word lists in relation to the students' inadequate command of these frequency levels, it is apparent, in terms of the lexical threshold hypothesis, that the students have relatively few lexical tools at their disposal. This suggests that their reading comprehension and writing ability is severely hampered at present by insufficient linguistic knowledge. These students obviously require a minimum level of vocabulary, without which they will be unable to develop the higher order skills required to cross the threshold referred to by Brisbois (1995) above (§5.2.2).

In conclusion, the idea that tests designed simply to measure vocabulary size may not be of much benefit to second-language students relates directly to Read's assessment of the present role of vocabulary tests. Read (1997:320) has proposed that vocabulary tests must be designed to include other aspects of linguistic competence:

> While there is still a role for tests that assess how well learners know words that occur frequently in the language or ones that are useful for the learners' own communicative purposes, these tests need to be located within a broader framework of communicative lexical ability. ... In this kind of framework, other lexical measures will be needed in addition to the relatively 'pure' tests of learner knowledge of individual words that currently dominate our thinking about what a vocabulary test is. The future trend in vocabulary testing is likely to be towards the design of integrative test formats that have a strong lexical focus but in which vocabulary ability is one of several factors that contribute to test-taker performance.

5.4   Conclusion

> South Africa's transition from apartheid and minority rule to democracy requires that all existing
> practices, institutions and values are viewed anew and rethought in terms of their fitness for the
> new era. Higher education plays a central role in the social, cultural and economic development
> of modern societies. In South Africa today, the challenge is to redress past inequalities and to
> transform the higher education system to serve a new social order, to meet pressing national
> needs, and to respond to new realities and opportunities. It must lay the foundations for the
> development of a learning society which can stimulate, direct and mobilise the creative and
> intellectual energies of all the people towards meeting the challenge of reconstruction and
> development.
>
> (Department of Education, 1997: White Paper 3)

The introductory statement to the White Paper on Higher Education in South Africa clearly illustrates the urgency of the need to redress inequalities in higher education in this country. The legacy of discrimination against black scholars outlined in Chapter 1 (§1.2) remains evident in the lack of appropriate teaching material, inadequate teaching facilities, inappropriate teaching methodologies and the inadequate qualifications of teachers in addition to their lack of English proficiency.

If the imbalance in educational standards between white and black scholars in this country is to be redressed, a primary consideration must be language teaching. For this reason, determining more efficient methods to facilitate the process of language learning by second-language students should play a fundamental role in primary and secondary as well as tertiary education. In light of this, the development of courses which focus on the second-language student in particular and are aimed at providing academic support has gained impetus at undergraduate level. These courses are therefore typically developed with the aim of improving the students' linguistic competence and, thereby, their academic performance.

Given these considerations, and in view of the results of this study, it is clear that attention to basic vocabulary and, more specifically, to the development of academic vocabulary

knowledge, would contribute to improved academic performance. However, without due consideration for more general academic skills, developmental courses which aim simply to increase the size of students' vocabularies will have very little effect on overall competence. Knowledge of vocabulary, while integral to reading comprehension and the writing of expository texts, and thus basic to any academic discipline, is nevertheless a single component of language ability, and must be considered in relation to the number of other variables which contribute to academic performance.

References

Aitchison, J. 1994 *Words in the mind: An introduction to the mental lexicon.* Oxford: Blackwell.

Alderson, J.C. and Urquhart, A.H. (Eds) 1984 *Reading in a foreign language.* London: Longman.

Allen, V.F. 1983 *Techniques in teaching vocabulary.* Oxford: Oxford University Press: 104-133.

Anderman, G. and Rogers, M. (Eds) 1996 *Words, words, words: The Translator and the Language Learner.* Clevedon: Multilingual Matters.

Anderson, R.C. and Freebody, P. 1981 "Vocabulary knowledge." In Guthrie, J.T. (Ed.) 1981:77-117.

Arnaud, P.J.L. and Béjoint, H. (Eds) 1992 *Vocabulary and Applied Linguistics.* London: Macmillan.

Astika, G.G. 1993 "Analytical assessments of foreign students' writing." *RELC Journal* 24(1):61-72.

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

Beck, I.L., McKeown, M.G. and Omanson, R.C. 1987 "The effects and uses of diverse vocabulary instructional techniques." In McKeown, M.G. and Curtis, M.E. (Eds) 1987:147-163.

Best, J.W. and Kahn, J.V. 1993 *Research in Education.* Boston: Allyn and Bacon.

Blacquiére, A. 1989 "Reading for survival: text and the second language student." *SAJHE/SATHO* 3(1):73-82.

Brisbois, J.E. 1995 "Connections between first- and second-language reading." *Journal of Reading Behavior* 27(4):565-584.

Brown, G., Malmkjaer, K. and Williams, J. (Eds) 1996 *Performance and competence in second language acquisition.* Cambridge: Cambridge University Press.

Carroll, J.B., Davies, P. and Richman, B. 1971 *The American Heritage Word Frequency Book.* New York: Houghton Mifflin.

Carter, R. 1987 *Vocabulary: Applied linguistic perspectives.* London: Routledge.

Carter, R. and McCarthy, M. 1988 *Vocabulary and language teaching.* London: Longman.

Carver, R.P. 1994 "Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction." *Journal of Reading Behavior* 26(2):413-437.

Chall, J.S. 1987 "Two vocabularies for reading: Recognition and meaning." In McKeown, M.G. and Curtis, M.E. (Eds) 1987:7-17.

Cleary, L.M. and Linn, M.D. (Eds) 1993 *Linguistics for teachers.* New York: McGraw-Hill.

Coady, J. 1997 "L2 vocabulary acquisition through extensive reading." In Coady, J. and Huckin, T. (Eds) 1997:225-237.

Coady, J. and Huckin, T. (Eds) 1997 *Second Language Vocabulary Acquisition.* Cambridge: Cambridge University Press.

Coady, J., Magoto, J., Hubbard, P., Graney, J. and Mokhtari, K. 1993 "High frequency vocabulary and reading proficiency in ESL readers." In Huckin, T., Haynes, M. and Coady, J. (Eds) 1993:217-228.

Coomber, J.E. and Peet, H.D. 1993 "Vocabulary teaching: Active involvement in contexts." In Cleary, L.M. and Linn, M.D. (Eds) 1993:577-589.

Cooper, P.A. 1995 "In search of sufficient vocabulary: Testing the vocabulary levels of undergraduate students." *South African Journal of Linguistics* 26:25-37.

Corson, D.J. 1983 "The Corson measure of passive vocabulary." *Language and Speech* 26(1):3-20.

Crystal, D. 1988 *The English Language.* London: Penguin.

Cunningham, J.W. and Moore, D.W. 1993 "The contribution of understanding academic vocabulary to answering comprehension questions." *Journal of Reading Behavior* 25(2):171-180.

D'Anna, C.A., Zechmeister, E.B. and Hall, J.W. 1991 "Toward a meaningful definition of vocabulary size." *Journal of Reading Behavior* 23(1):109-122.

Davis, F. 1972 "Psychometric research on comprehension in reading." *Reading Research Quarterly* 7:628-678.

Department of Education 1997 *White Paper 3: A Programme for the Transformation of Higher Education.* Pretoria. [http://www.polity.org.za/govdocs/white_papers/highed.html]. 24 July 1997.

Diack, H. 1975 *Wordpower: Your vocabulary and its measurement.* Herts: Paladin.

Dupuy, H.J. 1974 *The rationale, development and standardization of a basic word vocabulary test.* Washington, DC: US Government Printing Office, (DHEW Publication No. HRA 74-1334).

Engber, C.A. 1992 *A study of lexis and the relationship to quality in written texts of second language learners of English.* Michigan: UMI Dissertation services.

Freebody, P. and Anderson, R.C. 1983 "Effects on text comprehension of differing propositions and locations of difficult vocabulary." *Journal of Reading Behavior* 15(3):19-39.

Fries, C.C. and Traver, A.A. 1960 *English Word Lists.* Ann Arbor: George Wahr.

Gass, S.M. 1988 "Second-language vocabulary acquisition." *Annual Review of Applied Linguistics* 9:92-106.

Gordon, J., Schumm, J.S., Coffland, C. and Doucette, M. 1992 "Effects of inconsiderate versus considerate text on elementary students' vocabulary learning." *Reading Psychology: An International Quarterly* 13:157-169.

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

Grabe, W. 1991 "Current developments in second language reading research." *TESOL Quarterly* 25(3):375-406.

Guthrie, J.T. (Ed.) 1981 *Comprehension and teaching: Research reviews.* Newark, DE: International Reading Association.

Hair, J.F. (Jr), Anderson, R.E., Tatham, R.L. and Black, W.C. 1995 *Multivariate Data Analysis.* Englewood Cliffs: Prentice-Hall.

Hatch, E. and Farhady, H. 1982 *Research design and statistics for applied linguistics.* Rowley: Newbury House.

Hazenberg, S. and Hulstijn, J.H. 1996 "Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation." *Applied Linguistics* 17(2):145-163.

Heimlich, J.E. and Pittleman, S.D. 1986 *Semantic mapping: Classroom activities.* International reading association. Newark: Delaware.

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

Hubbard, E.H. 1996 "Profiling contextual support for technical and general academic terms in writing for students." *South African Journal of Linguistics* 32:95-106.

Huckin, T., Haynes, M. and Coady, J. (Eds) 1993 *Second language reading and vocabulary learning.* New Jersey: Ablex.

Jenkins, J.R. and Dixon, R. 1983 "Vocabulary learning." *Contemporary Educational Psychology* 8:237-260.

Jenkins, J.R., Stein, M.L. and Wysocki, K. 1984 "Learning vocabulary through reading." *American Educational Research Journal* 21(4):767-787.

Kang, S.-H. 1995 "The effect of a context-embedded approach to second-language vocabulary learning." *System* 23(1):43-55.

Kirkpatrick, E.A. 1891 "Number of words in an ordinary vocabulary." *Science* 18(446):107-108.

Laufer, B. 1986 "Possible changes in attitude towards vocabulary acquisition research." *IRAL* 24(1):69-75.

Laufer, B. 1989 "What percentage of text-lexis is essential for comprehension?" In Laurén, C. and Nordman, M. (Eds) 1989:316-323.

Laufer, B. 1990a "Ease and difficulty in vocabulary learning: Some teaching implications." *Foreign Language Annals* 23(2):147-155.

Laufer, B. 1990b " 'Sequence' and 'order' in the development of L2 lexis: Some evidence from lexical confusions." *Applied Linguistics* 11(3):281-296.

Laufer, B. 1992 "How much lexis is necessary for reading comprehension?" In Arnaud, P.J.L. and Béjoint, H. (Eds) 1992:126-132.

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

Laufer, B. 1997a "The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess." In Coady, J. and Huckin, T. (Eds) 1997:20-34.

Laufer, B. 1997b "What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words." In Schmitt, N. and McCarthy, M. (Eds) 1997:140-155.

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

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

Laufer, B. and Sim, D.D. 1985 "Taking the easy way out: non-use and misuse of contextual clues in EFL reading comprehension." *English Teaching Forum* 23(2):7-10.

Laufer-Dvorkin, B. 1991 *Similar lexical forms in interlanguage.* Tubingen: Gunter Narr Verlag.

Laurén, C. and Nordman, M. (Eds) 1989 *Special language: From humans thinking to thinking machines.* Clevedon: Multilingual Matters.

Leary, M.R. 1991 *Introduction to behavioral research methods.* California: Wadsworth.

Linnarud, M. 1986 *Lexis in composition: A performance analysis of Swedish learners' written English.* Lund: CWK Gleerup.

Lorge, I. and Chall, J. 1963 "Estimating the size of vocabularies of children and adults: An analysis of methodological issues." *The Journal of Experimental Education* 32(2):147-157.

Macdonald, C.A. 1990 *Crossing the Threshold into Standard Three in Black Education: The Consolidated Main Report of the Threshold Project.* Pretoria: Human Sciences Research Council.

Marshall, S. and Gilmour, M. 1993 "Lexical knowledge and reading comprehension in Papua New Guinea." *English for Specific Purposes* 12:69-81.

McCarthy, M.J. 1984 "A new look at vocabulary in EFL." *Applied Linguistics* 5(1):12-22.

McKeown, M.G. and Curtis, M.E. (Eds) 1987 *The nature of vocabulary acquisition.* New Jersey: Lawrence Erlbaum Associates.

Meara, P. 1987 *Specialised bibliography 4: Vocabulary in a second language: Volume 2.* London: Centre for Information on Language Teaching and Research.

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

Meara, P. 1996a "The classical research in L2 vocabulary acquisition." In Anderman, G. and Rogers, M. (Eds) 1996:27-40.

Meara, P. 1996b "The dimensions of lexical competence." In Brown, G. *et al.* (Eds) 1996:35-53.

Meara, Paul. 1996c "The vocabulary knowledge framework." Vocabulary Acquisition Research Group. University of Wales: Swansea. [http://www.swan.ac.uk/cals/vlibrary/pm96d.html]. 19 November 1996.

Meara, Paul. 1996d "Word associations in a foreign language." Vocabulary Acquisition Research Group. University of Wales: Swansea. [http://www.swan.ac.uk/cals/vlibrary/pm83a.html]. 9 December 1996.

Meara, P. 1997 "Towards a new approach to modelling vocabulary acquisition." In Schmitt, N. and McCarthy, M. (Eds) 1997:109-121.

Meara, P. and Buxton, B. 1987 "An alternative to multiple choice vocabulary tests." *Language Testing* 4:142-151.

Meara, P. and Jones, G. 1987 "Tests of vocabulary size in English as a foreign language." *Polyglot* 8:1-40.

Medo, M.A. and Ryder, R.J. 1993 "The effects of vocabulary instruction on readers' ability to make causal connections." *Reading Research and Instruction* 33(2):119-134.

Mulder, J.C. 1986 *Statistical techniques in Education.* Pretoria: Kagiso.

Nagy, W.E. 1997 "On the role of context in first- and second-language vocabulary learning." In Schmitt, N. and McCarthy, M. (Eds) 1997:64-83.

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

Nation, I.S.P. 1982 "Beginning to learn a foreign vocabulary: A review of the research." *RELC Journal* 13(1):14-36.

Nation, I.S.P. 1990 *Teaching and learning vocabulary.* New York: Newbury House.

Nation, P. 1993 "Using dictionaries to estimate vocabulary size: Essential, but rarely followed, procedures." *Language Testing* 10(1):27-40.

Nation, P. and Coady, J. 1988 "Vocabulary and reading." In Carter, R. and McCarthy, M. 1988:97-110.

Nation, P. and Hwang, K. 1995 "Where would general service vocabulary stop and special purposes vocabulary begin?" *System* 23(1):35-41.

Nation, P. and Waring, R. 1997 "Vocabulary size, text coverage and word lists." In Schmitt, N. and McCarthy, M. (Eds) 1997:6-19.

Oldfield, R.C. 1963 "Individual vocabulary and semantic currency: A preliminary study." *British Journal of Social and Clinical Psychology* 2:122-130.

Perkins, D.M. 1991 "Improvement of reading and vocabulary skills at the University of Transkei." *South African Journal of Education* 11(4):231-235.

Pretorius, E.J. 1995 "Reading as an interactive process: Implications for studying through the medium of a second language." *Communicatio* 21(2):33-43.

Proctor, P. 1995 *Cambridge International Dictionary of English.* Cambridge: Cambridge University Press.

✓ 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. 1997 "Vocabulary and testing." In Schmitt, N. and McCarthy, M. (Eds) 1997:303-320.

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

Romaine, S. 1984 *The language of children and adolescents: The acquisition of communicative competence.* Oxford: Basil Blackwell.

Rosenthal, R and Rubin, D.B. 1982 "A simple, general purpose display of magnitude of experimental effect." *Journal of Educational Psychology* 74(2):166-169.

Ruddell, M.R. 1994 "Vocabulary knowledge and comprehension: A comprehension-process view of complex literacy relationships." In Ruddell, R.B., Ruddell, M.R. and Singer, H. (Eds) 1994:414-447.

Ruddell, R.B., Ruddell, M.R. and Singer, H. (Eds) 1994 *Theoretical models and processes of reading.* Newark: International Reading Association.

Ryan, A. 1997 "Learning the orthographical form of L2 vocabulary - a receptive and a productive process." In Schmitt, N. and McCarthy, M. (Eds) 1997:181-198.

Santos, T. 1988 "Professors' reactions to the academic writing of nonnative-speaking students." *TESOL Quarterly* 22(1):69-90.

Saville-Troike, M. 1984 "What really matters in second language learning for academic achievement?" *TESOL Quarterly* 18(2):199-219.

Savini, M. 1992 "Vocabualry errors in ESL writing: Some classroom suggestions." *ELTIC Reporter* 16(2&3):50-63.

Schmitt, N. 1995 "A fresh approach to vocabulary: Using a word knowledge framework." *RELC Journal* 26(1):86-94.

Schmitt, N. and McCarthy, M. (Eds) 1997 *Vocabulary: Description, Acquistion and Pedagogy*. Cambridge: Cambridge University Press.

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

Sternberg, R.J. and Powell, J.S. 1983 "Comprehending verbal comprehension." *American Psychologist* 38(8):878-893.

Sutarsyah, C., Nation, P. and Kennedy, G. 1994 "How useful is EAP vocabulary for ESP? A corpus based study." *RELC Journal* 25(2):34-50.

Thorndike, E.L. and Lorge, I. 1944 *The teacher's word book of 30 000 words*. New York: Columbia University Teachers College.

Vorster, J. and Piper, S. 1995 "Divide and rule: On the rationalisation of vocabulary teaching." *South African Journal of Linguistics* Supplement 27.

Vorster, J. and Reagan, T. 1990 "On the lexical development of L1 and L2 speakers." *South African Journal of Linguistics* 9(3):80-84.

Wenborn, N. 1981 *Wordpower: A test yourself guide*. London: Kogan Page.

West, M. 1953 *A general service list of English words*. London: Longmans, Geen and Co.

Williams, R. and Dallas, D. 1984 "Aspects of vocabulary in the readability of content area L2 educational textbooks: a case study." In Alderson, J.C. and Urquhart, A.H. (Eds) 1984:199-212.

Worthington, D. and Nation, P. 1996 "Using texts to sequence the introduction of new vocabulary in an EAP course." *RELC Journal* 27(2):1-11.

Xue, G. and Nation, P. 1984 "A University Word List." *Language Learning and Communication* 3(2):215-228.

Appendix A

Vocabulary Tests A and B

NAME: _____

STUDENT NUMBER: _____

MOTHER TONGUE: _____

Instructions

In each of the sentences below, one word has been left out. Select the word which best fits the sentence from the four choices provided. Circle the corresponding letter in each case, for example:

Pre-primary education aims to develop the young child to a state of readiness for _____ schooling.

(a)   firm
(b)   fierce
Ⓒ   formal
(d)   frequent

Many children come from _____ slums on the periphery of large cities, which, with their overcrowded living conditions, offer a limited range of stimulation to the child.

ⓐ   squalid
(b)   affluent
(c)   appropriate
(d)   prestigious

TEST A

1.   Although schooling is now practised in almost every western country, there are _____ millions of children in underdeveloped countries who do not receive any schooling.

(a)   yet
(b)   still
(c)   even
(d)   besides

2. One criticism is that _____ Pragmatists claim to place the focus on the child in the schooling process, the child in fact often becomes secondary to the experience or the project.

    (a) since
    (b) because
    (c) whether
    (d) although

3. The process, or the way in which one achieves goals, is more important than the goal itself. This means that in a classroom situation, the fact that a pupil does not produce a perfect piece of work is less important than the fact that the work was actually _____ .

    (a) allowed
    (b) adopted
    (c) accepted
    (d) attempted

4. It is expected that the individual who _____ schooling will pursue some occupation which will contribute to the general development of society.

    (a) contains
    (b) concerns
    (c) completes
    (d) commands

5. The child's evaluation of himself in relation to his self-identity, his behaviour and the emotions he experiences as a result of the feedback he _____ from others leads to the formation of the self-concept.

    (a) regards
    (b) respects
    (c) receives
    (d) represents

6. Whether this influence should be considered as positive or negative largely _____ on one's perception of the information available.

    (a) defends
    (b) depends
    (c) descends
    (d) deserves

7. Since Humanists believe that the focal point of education is the human being, they _____ education and schooling which ignore human problems or which give human problems minor importance.

    (a) annoy
    (b) offend
    (c) oppose
    (d) approve

8. People react in certain ways to the impressions they have of others, and it is _____ this reaction which influences the development of someone's personality.

    (a) purely
    (b) probably
    (c) peculiarly
    (d) profitably

9. Meaningful learning is _____ to rote learning because it is more easily learned and is transferable from the school context to the real world.

    (a) referred
    (b) reflected
    (c) preferred
    (d) prejudiced

10. Because we think in language, knowing the correct vocabulary is _____ in the formation of concepts.

    (a) exact
    (b) efficient
    (c) essential
    (d) excessive

11. Many students use underlining or highlighting as a means to _____ studying.

    (a) frustrate
    (b) fluctuate
    (c) facilitate
    (d) formulate

12. You must learn to argue scientifically, i.e. to debate or discuss a question, and to use reasoning to support or _____ a theory.

    (a) react
    (b) refute
    (c) retain
    (d) reinforce

13. Peer groups often assume the form of a club or gang and group members are expected to _____ to the ideas and norms of the group.

    (a) confer
    (b) confine
    (c) contend
    (d) conform

14. Conflict exists in relationships when parties believe that their aspirations cannot be achieved simultaneously, or perceive a divergence in their values, needs or interests and purposefully employ their power in an effort to defeat, neutralise or _____ each other to protect or further their interest in the interaction.

    (a) evoke
    (b) evaluate
    (c) elaborate
    (d) eliminate

15. There are other statements that theorists make that can't be listed as facts since they are open to _____ .

    (a) dispute
    (b) decline
    (c) diversity
    (d) deficiency

16. A model of schooling includes the aim, content and methods of the schooling provided by society to _____ pupils to be educated.

    (a) evolve
    (b) enable
    (c) ensure
    (d) exploit

17. The theory of education that forms part of the philosophy of a society also includes guidelines about how schools should function, how teaching should be conducted, what schooling should achieve; these all _____ the nature of schooling in a society.

    (a) comply
    (b) comprise
    (c) compound
    (d) compensate

18. One of the most important ways in which a country's language policy _____ itself is in the kind of provision it makes for the education of children.

    (a) modifies
    (b) manifests
    (c) manipulates
    (d) misinterprets

19. In many areas of children's learning, a goal is set and activities are provided to assure that mastery of the concept is _____ .

    (a) adapted
    (b) attained
    (c) adjusted
    (d) ascribed

20. One reason for the speculative and disputative nature of this debate is due to the fact that cognition is not observable and hence not open to _____ testing in any direct form.

   (a) eventual
   (b) empirical
   (c) emotional
   (d) elemental

21. The theorists give a rather _____ overview of the child's development during this period without placing it in the broader context of the child's overall development.

   (a) subtle
   (b) stratified
   (c) superficial
   (d) self-evident

22. While Idealist education has emphasised the intellectual side of man, it has tended toward intellectualism to the _____ of the affective and physical side.

   (a) detriment
   (b) deviation
   (c) derivation
   (d) deprivation

23. It is the duty of the learner to _____ or absorb the truths presented to him.

   (a) affect
   (b) assert
   (c) assume
   (d) assimilate

24. The restricted code of language leaves statements unfinished, compelling the listener to _____ or guess what is meant.

   (a) infer
   (b) imply
   (c) impose
   (d) impress

25. The homelands adopted the use of one language medium, generally from standard three upwards, and further _____ that this language should be English.

   (a) signified
   (b) stipulated
   (c) stimulated
   (d) speculated

26. Children are particularly _____ to the influence of the media because of their uncritical acceptance of the messages imparted.

   (a) subjective
   (b) supportive
   (c) submissive
   (d) susceptible

27. An important point in this regard is that whenever we investigate a phenomenon, we already have certain ideas about that phenomenon. These _____ ideas are called basic assumptions.

   (a) predictable
   (b) perceptible
   (c) predisposed
   (d) preconceived

28. His views have been dismissed far too readily by many critics, who tend to reject all his claims because a few may be _____.

   (a) untenable
   (b) unanimous
   (c) unequivocal
   (d) uncontroversial

29. The exercises at the end of every segment of the study manual are designed to _____ the academic skills that are practised in the course of the year.

   (a) convey
   (b) counteract
   (c) consolidate
   (d) compromise

30. The teacher who does the presentation must be many things to the child. He must be an example of a being who is nearer the perfect ideal of personality, and therefore worthy of _____.

   (a) emulation
   (b) eradication
   (c) evisceration
   (d) emancipation

TEST B

1.   Throughout the nine months before birth, the environment in which the child is growing plays an important role. The child will benefit if the expectant mother lives a healthy life-style and eats well. If the expectant mother drinks, smokes or takes drugs, or if she is subjected to starvation, illness or abuse, _____ , it will have a detrimental effect on the child and can influence the development of the child after birth.

   (a)   but
   (b)   instead
   (c)   however
   (d)   moreover

2.   The pupil is regarded as first and foremost an individual with a personality of his own, though an immature one. His mind, like his body, is immature and therefore has to be _____ through education.

   (a)   decided
   (b)   declared
   (c)   described
   (d)   developed

3.   The home environment in which children live _____ the rate and quality of their language acquisition.

   (a)   intends
   (b)   includes
   (c)   increases
   (d)   influences

4.   The inadequate financing of schools by the state does not _____ the quality of schooling expected by parents, teachers and pupils.

   (a)   prevent
   (b)   present
   (c)   produce
   (d)   provide

5.   At school the dominant social group accepts the information taught because it is part of its own cultural capital. This means that the information taught _____ the attitudes and values of this group.

   (a)   relates
   (b)   refuses
   (c)   reduces
   (d)   represents

6. The educative process should be conducted in such a way as to _____ infringing upon the rights of each person as an individual.

   (a) avoid
   (b) advise
   (c) attend
   (d) approve

7. Not all teachers accept the explanation given above and other explanations have been put _____ to account for the same observation, i.e. different interpretations of this observation have been proposed.

   (a) firmly
   (b) further
   (c) forward
   (d) formally

8. To synthesize means to put together or to _____ into a coherent whole.

   (a) combine
   (b) compete
   (c) compose
   (d) complain

9. Children become increasingly independent as they manage to _____ more complex physical tasks which enable them to enter school and cope with its demands.

   (a) practise
   (b) perform
   (c) review
   (d) reserve

10. The pupil's self-learning is the core of education; the ultimate _____ for learning must rest with them, not with the teacher.

    (a) rivalry
    (b) reflection
    (c) responsibility
    (d) recommendation

11. Some of this criticism against the pragmatic approach has merit because curriculum reform based on pragmatic ideas has usually been implemented too hastily and without _____ preparation of the teaching staff.

    (a) adjacent
    (b) acquired
    (c) adequate
    (d) available

12. The neo-Marxist _____ on social conflict includes a description of power, disputed interaction and alienation that occurs with this type of conflict.

    (a) perspective
    (b) pertinence
    (c) persistence
    (d) perpetration

13. Recent developments in education have _____ the role of the community marginal, so that during the teachers' strikes, parents found themselves without a voice.

    (a) revived
    (b) restored
    (c) rendered
    (d) restricted

14. People who are _____ motivated behave in a certain way because they derive some personal satisfaction from that behaviour, or because it satisfies a need for achievement or competence, rather than being motivated by external approval.

    (a) insistently
    (b) intolerably
    (c) intensively
    (d) intrinsically

15. This developmental stage _____ between roughly the ages of twelve and eighteen years, and is the last phase of childhood.

    (a) prevails
    (b) pertains
    (c) provokes
    (d) propagates

16. There were considerable increases in the total budgets _____ to black education.

    (a) asserted
    (b) allocated
    (c) assumed
    (d) assimilated

17. The poor child will generally fall behind because of his dependence on the school only for learning. Schooling therefore _____ social class distinction.

    (a) persists
    (b) perpetrates
    (c) perpetuates
    (d) participates

18. The classroom is an artificially _____ environment in which the outside world is softened and introduced in doses small enough for the individual child to absorb.

    (a) derived
    (b) devised
    (c) deprived
    (d) deferred

19. The Idealist influence on education has been of great magnitude, but there are serious reservations about whether that influence has always been _____ or advantageous.

   (a) beneficial
   (b) conventional
   (c) controversial
   (d) contradictory

20. The most vocal critics attack Realism because it has _____ the idea of a fixed, intelligible universe, capable of being perceived objectively by the observing intellect.

   (a) alleged
   (b) attained
   (c) adhered
   (d) advocated

21. Writers who describe schools in terms of a Consensus Model assume that the pupils who come to schools are a _____ group and that they all agree to behave as required.

   (a) indigenous
   (b) homogeneous
   (c) hierarchical
   (d) hypothetical

22. The programme of general education is most fundamental in the Realist's model; vocational education, which is the education for living or survival, is _____, although also necessary.

   (a) sufficient
   (b) subsequent
   (c) spontaneous
   (d) supplementary

23. Humanistic psychology accepts that learning should be free, open and meaningful. This is implemented in schools where pupils are encouraged to challenge the _____ or truth of what is taught.

   (a) validity
   (b) variability
   (c) verification
   (d) stratification

24. In a school that functions as a Consensus Model of Socialization children wear uniforms; varieties of uniform are not _____, for example, hair must be short, skirts must cover the knees, and only school shoes may be worn.

   (a) sustained
   (b) suppressed
   (c) tolerated
   (d) transmitted

25. When private schools were established in South Africa they aimed to _____ many British private school traditions, and so practices such as corporal punishment for boys were assimilated into the school policy.

(a) induce
(b) indicate
(c) implicate
(d) incorporate

26. We are thus faced with the dilemma that a loose interpretation is too broad, while a strict interpretation is too narrow. We can resolve this dilemma if we propose a compromise between the two _____ views.

(a) extremist
(b) expedient
(c) explicable
(d) extenuating

27. The educational interests of the pupils were not regarded as paramount, but were _____ to ideological and political factors that were concerned with white interests.

(a) subsumed
(b) substantial
(c) subservient
(d) submerged

28. One other objective of the school, and hence of education, is to help the individual to see his knowledge as an _____ part of a much larger whole and to interpret events in the context of the whole system instead of from his personal point of view alone.

(a) integral
(b) inimical
(c) ineffable
(d) intermittent

29. It is important not to isolate the first problem from the rest, but to work within a _____ approach in which the selection of an official language is seen as part of a total social and economic reform plan.

(a) holistic
(b) heuristic
(c) hippocratic
(d) holophrastic

30. The kinds of literature available are extremely limited, there being a _____ of everyday reading materials that might appeal to the young, for example, comic books, mystery stories, and collections of folk tales.

(a) paucity
(b) profanity
(c) periphery
(d) perversity

Appendix B

Hypothesis testing

This section outlines the principles underlying the testing of hypotheses. The formulation of the null hypothesis, the significance level in terms of which the null hypothesis is evaluated, and the distinction between directional and non-directional hypotheses are explained.

A hypothesis is an assumption made about the nature of the relationship between variables, and so is also "a prediction regarding the outcome of a study" (Leary, 1991: 338). The formulation of a hypothesis is based on an observation which is considered to be a problematic phenomenon. This observation is then construed as a problem statement, from which the hypothesis is derived. A critical aspect of any hypothesis is that it must be falsifiable, in other words, the hypothesis must be stated in such a way that the underlying assumption can be rejected or supported by empirical findings. Kerlinger (1986: 17) argues that, in terms of the above criteria, the hypothesis statement should "contain two or more variables that are measurable or potentially measurable and [should] specify how the variables are related".

In order to determine the validity of the research (or alternative) hypothesis, a contradictory hypothesis, known as the 'null hypothesis' is formulated and tested. "The null hypothesis for an experiment states that the independent variable *did not* have an effect on the dependent variable." (Leary, 1991: 160) The two hypotheses are therefore "mutually exclusive" and "exhaustive" (McCall, 1990: 160). The objective of empirical research is to present findings that enable the researcher to reject the null hypothesis. The criterion in terms of which the null hypothesis is evaluated is based on the probability that similar results would be obtained if the test were to be repeated under the same conditions. This degree of probability is referred to as the significance level. "The significance level (or critical level), symbolized by α (alpha), is the probability value that forms the boundary between rejecting and not rejecting the null hypothesis" (McCall, 1990: 194). The value of α is usually taken as $p<.05$ in behavioural research, where p represents the probability that the finding occurred *by chance*. Thus the researcher is not able to reject the null hypothesis, and thereby confirm the research hypothesis, "if analyses show a high probability that the difference between the group means reflects nothing more than the influence of error variance" (Leary, 1991: 160).

A research hypothesis can be expressed as either directional or non-directional. A directional hypothesis predicts that one sample mean will be bigger, while a non-directional hypothesis simply predicts that there will be a difference between the means, without stating the direction of the outcome.

> A directional hypothesis states explicitly which of the two condition means is expected to be larger. That is, the researcher predicts the specific direction of the anticipated effect. A nondirectional hypothesis merely states that the two means are expected to differ, but no prediction is ventured regarding which mean will be larger.
>
> Leary (1991: 169)

The validity of a directional hypothesis is assessed in terms of a one-tailed test. Since the direction of the findings has been predicted, the analysis of results focuses only on the relevant area of distribution, that is, on one side of the mean. The validity of a non-directional hypothesis, on the other hand, is determined by means of a two-tailed test. Since no specific prediction has been made with regard to the outcome, it is necessary to evaluate both tails derived from the distribution of results, that is, the researcher looks for "a significant difference between the mean in either direction" (Beer, 1982: 69).

This study is based on a directional hypothesis. The general research hypothesis posits a significant positive correlation between vocabulary size and academic performance. The assumption of directionality is supported by previous research, as argued in Chapter 2.

Appendix C

Statistical techniques

This section provides further explanation of the two statistical methods used in the analysis of results, i.e. the Pearson product-moment correlation coefficient and multiple linear regression. The rationale behind each of these statistical techniques is outlined, with a brief explanation of the calculations on which they are based. In addition, the procedure used to obtain the results of the regression is outlined, and applications of the model are presented.

The Pearson product-moment correlation coefficient

A correlation coefficient is a statistical indicator that determines the degree to which two variables are related to one another. The sign of a correlation coefficient (+ or -) indicates the direction of the relationship between the two variables. Variables may be either positively or negatively correlated. A positive correlation indicates a direct, positive relationship between the two variables. If the correlation is positive, an increase in one variable is accompanied by an increase in the other. A negative correlation indicates an inverse, negative relationship between the variables. As the value of one variable increases, the value of the other variable decreases. Thus the Pearson product-moment correlation coefficient, symbolised by r, reflects the extent of a linear relationship which ranges from -1.00 to +1.00.

> 'Correlation coefficient' is a measure of the interdependence, the varying together, the simultaneous increase or decrease of two sets of numerical values. The *direction* of a relation is whether it is positive or negative. The *magnitude* of a relation is the extent to which two sets of measures vary together (covary) positively or negatively.
>
> Kerlinger (1979: 52-53)

The magnitude of the correlation, that is, its numerical value, expresses the strength of the relationship between the variables. A correlation of zero ($r = .00$) indicates that the variables are not at all related. As the numerical value of the coefficient increases, so does the strength of the relationship. Thus, a correlation of +.78 indicates that the variables are more strongly related than does a correlation of +.30. The relationship between any two variables can be portrayed graphically on x-y axes by plotting a point that represents a

combination of scores in the two variables. The resulting representation of the data is known as a 'scatter plot'.

Where there is high positive correlation subjects scoring above the mean on one test will score above the mean on the other test, since the product of their deviations from the mean is positive. Similarly, subjects scoring below the mean on one test will score below the mean on the other test as the product of these two negative deviations is also positive. However, where there is high negative correlation the reverse is true. Subjects scoring above the mean on one test will score below the mean on the other. The products of these two deviations (one positive, one negative) will be negative. Where there is little or no correlation there will be a mixture of positive and negative deviation products and the mean of these will be small - somewhere near zero. (Beer, 1982: 117-118.)

The use of the Pearson product-moment correlation coefficient in this study serves to illustrate the degree of statistical significance between variables. The correlation coefficient indicates both the direction and the magnitude of the relationship between two variables, and was therefore considered an appropriate measure of the relationship between vocabulary size and academic performance. The hypothesis assumes a strong positive correlation between the variables, which should be reflected by the results of the correlation coefficient. Thus, because of the large number of subjects in this study, a result ranging from less than .16 to +1.00 (Kerlinger, 1986: 188; Tuckman, 1988: 477) would suggest that the hypothesis is valid at the .05 level, where the strength of the interdependence between variables is reflected by the proximity of r to +1.00. (To be statistically significant, the value of r is, of course, dependent on the sample size, according to the law of large numbers. Kerlinger (1986: 188) proposes that "with 100 pairs of measures the problem [of statistical significance] is less acute; to carry the .05 day, an r of .16 is sufficient".) In the case that the hypothesis is *not* valid, this would be reflected either by a negative correlation, or by a negligible result, where r is close to .00.

Multiple Linear Regression

> The goal of regression analysis is to develop a **regression equation** from which we can predict one score on the basis of one or more other variables.
>
> (Leary, 1991:115.)

The linear regression equation assumes an a priori mathematical relationship between the independent and dependent variables. The technique uses the principle of 'least squares' to evaluate the best fitting straight line through the data points on a scatter plot. This line portrays the nature of the relationship between test scores and performance ratings. In multiple regression equations, the relationship is still linear although it

does not have a clear graphic representation. Since this study has three independent variables, the linear regression equation requires solving for the constants b_0, b_1, b_2 and b_3, according to the formula:

$$y = b_1 x_1 + b_2 x_2 + b_3 x_3,$$

where x_1, x_2 and x_3 are independent variables. (These represent the scores in the three subsections: basic vocabulary, academic vocabulary and advanced vocabulary.) In addition to multiple regression, this study employs stepwise regression to investigate whether any of the independent variables is significantly responsible for the variation of the dependent variable.

As indicated in Chapter 4 (§4.6), the linear regression equation was applied to the results of this study with the aim of determining the degree to which vocabulary size may be used to predict academic performance. Following from the assumption that vocabulary size is directly related to academic performance, it was assumed that the independent variable, x, could be used to estimate the value of the dependent variable, y, to some degree. The significance of the relationship between the set of independent (or predicting) variables and the dependent variable (y) is indicated by the coefficient of determination, R^2, which ranges from .00 to 1.00. The greater the value of R^2, the more likely the degree of prediction, since an R value of 1.00 indicates that 100% of the variation in y is dependent on the variation of the independent variable(s). The practical implications of the general research hypothesis would therefore be reflected by the degree to which x approximates y on the straight line graph. Thus, if R^2 approaches 1.00, it may be assumed that a test of vocabulary size will serve to predict academic performance to some degree, as determined by the value of R^2. The lower the value of R^2, on the other hand, the less significant the degree of correlation. The actual value of R^2 at which the hypothesis is considered invalid is dependent on the sample size.

The stepwise multiple linear regression model indicates which of the independent variables (basic vocabulary: x_1, academic vocabulary: x_2, or advanced vocabulary: x_3) accounts for the most significant variance in the dependent variable, and ranks the variables x_1 to x_3 in order of significance. This model thus indicates which of the independent variables serves as the best predictor of academic performance by adding the independent variables to the model according to their degree of influence on the dependent variable. This process, in which the independent variable that serves as the highest predictor is selected initially, followed by the next highest predictor, and so on, until the addition of further variables does not have any significant influence on the dependent variable, is aimed at calculating the best regression model, "i.e. the model which leaves the smallest residual variance after it is fitted" (Woods, Fletcher and Hughes, 1986:243-244). In the model, only those variables found to be significant at the 0.10 level were included, as determined by SPSS. (It should be noted that this 0.10 significance level is for the entry of an additional variable into the model, and is not the level of significance of the model. The significance of the coefficient of determination (R^2) is calculated at the standard probability level, i.e. 0.05.)

The procedure followed in the course of the analysis is outlined below, in the presentation of results. Each of the steps in the model is provided in turn, with a brief explanation as to how the results should be interpreted. The final results of the stepwise multiple linear regression analysis are illustrated in Table 4.8. As academic vocabulary was found to be the most significant indicator of academic performance in the regression analysis, the first step in the model is $y = b_0 + b_2 x_2$, as illustrated below.

(a)   Step 1: Stepwise regression for the model $y = b_0 + b_2 x_2$ (where x_2 is the second independent variable, i.e. academic vocabulary, and y is the dependent variable, i.e. academic performance).

R^2	df	F-ratio	Probability	b_0	b_2	N
0.1460	1	111.29	$p<.0001$	34.06	1.09	653

In this model, R represents the multiple correlation coefficient, and describes the degree of relationship between the dependent and independent variables. The square of this correlation coefficient, R^2, reflects the proportion of the variance in the dependent variable that can be accounted for by the independent variable(s). In other words, the independent variable, in this case academic vocabulary, accounts for 14.6% of the variance in the dependent variable, academic performance.

The F-ratio is the interaction of between-groups variance and within-groups variance, and is used to determine the degree of significance of the R-squared value, which in this case is significant even at the 0.1% level. The intercept constant or the point at which the regression line crosses the y-axis, is represented by b_0, while x_2 is the score of an individual on the second predictor or independent variable, i.e. academic vocabulary.

(b)   Step 2: Stepwise regression for the model $y = b_0 + b_2 x_2 + b_1 x$ where x is the second independent variable, i.e. academic vocabulary, x_1 is the first independent variable, i.e. basic vocabulary, and y is the dependent variable, i.e. academic performance).

R^2	df	F-ratio	Probability	b_0	b_1	b_2	N
0.1526	2	58.52	$p<.0001$	30.85	0.40	0.92	653

The most significant distinctions between Steps 1 and 2 is the increase in the value of R^2 from 0.1460 to 0.1526. This indicates the degree of variance accounted for by both independent variables. The final result of the regression analysis is illustrated below.

Results of the stepwise multiple linear regression analysis

Variable	Partial R^2	Cumulative R^2	F-ratio	Probability level
Academic vocab.	0.1460	0.1460	111.29	$p<.0001$
Basic vocabulary	0.0066	0.1526	58.52	$p<.0001$

These results indicate that basic and academic vocabulary scores together predict 15.26% of the students' academic performance, as measured by the final mark. The most significant predictor of academic performance is the academic vocabulary score, which alone accounts for 14.6% of the variance in the final mark. This is indicated by the incremental coefficient of determination (partial R^2), which illustrates the extent to which a single independent variable accounts for variance in the dependent variable. Thus, while academic vocabulary accounts for 14.6% of the variance, as calculated in Step 1, the second most significant independent variable, basic vocabulary, alone accounts for 0.7% of the variance in the final mark, as indicated in Step 2. The final coefficient of determination ($R^2 = 15.26\%$) is derived from the sum of both partial coefficients. The advanced vocabulary score did not meet the 0.10 significance level required for inclusion in the model, which accounts for the omission of the third independent variable from the model.

Predicting from the results of the regression analysis

The statistical significance of these findings can be illustrated by means of a comparison between students' predicted results in terms of the formula for the regression model and their actual results. The results generated by the regression analysis of the test scores may be applied to the formula $y = b_0 + b_2 x_2 + b_1 x_1$ as follows:

academic performance $=$ 30.85 (intercept) + 0.92 × academic vocabulary score + 0.40 × basic vocabulary score.

Thus, where a student obtained a total score of 25 for the academic vocabulary and 15 for the basic vocabulary, the mark predicted by the regression analysis would be 59.9%. In this instance the predicted mark differs by a margin of only 2.1% from the final mark of 62%. However, the potential for errors in prediction is exemplified in an application of the model to extreme ends of the range of final marks obtained by students. A student who obtained 14% for the final mark, for example, obtained a score of 8 on the academic vocabulary test and 11 on the basic vocabulary test, in terms of which the predicted mark is 42.6%. This indicates a discrepancy of approximately 28% between predicted and actual scores. At the other end of the scale, the predicted mark for a student who obtained a final mark of 81% was only 57.3%, where the score for both academic and basic vocabulary was 20. In the cases outlined above, the model has either

overestimated or underestimated the students' academic performance. These examples are intended to sketch the limited nature of this regression model, as discussed in Chapter 4.

Appendix D

Item analyses

This section tabulates the results of the analyses of each vocabulary test item. The conclusions drawn from each analysis are presented and discussed in Chapter 4 (§4.8).

Tables 1 and 2 illustrate the total number of correct answers for each question. In every case the proportion of students who answered the question correctly is given as a percentage.

Tables 3 and 4 present the results of an analysis of the options in each question. These tables contain two sets of results for each of the four options provided. The first figure (tabulated under the heading *prop.*) denotes the proportion of students who selected that option and is given as a fraction of 1.00. The second figure (*score*) provides an indication of the overall achievement of the students who chose that particular option and is given as a mark out of 60. These figures are based on the total number of subjects who completed Tests A and B (n = 946), and include cases with missing variables. It must be noted that, because approximately one third of the students only completed Test A, the overall scores are on average lower than those in the analysis of results, owing to the missing variables. (Note, too, that where the percentages for a question do not total 100 a number of test items were omitted by the students). The mean score at each frequency level thus differs from those in Table 4.1 (§4.3). The results for the correct option (the test item) are given in bold in each case.

Table 1 : Percentage of correct answers for each item in Test A

	Test item	Total (n = 653)	Percentage
1000- LIST	1 (*still*)	555	85.0
	2 (*although*)	246	37.7
	3 (*attempted*)	280	42.9
	4 (*completes*)	550	84.2
	5 (*receives*)	536	82.1
2000-LIST	6 (*depends*)	584	89.4
	7 (*oppose*)	407	62.3
	8 (*probably*)	492	75.3
	9 (*preferred*)	167	25.6
	10 (*essential*)	460	70.4
UNIVERSITY WORD LIST	11 (*facilitate*)	261	40.0
	12 (*refute*)	85	13.0
	13 (*conform*)	359	55.0
	14 (*eliminate*)	239	36.6
	15 (*dispute*)	276	42.3
	16 (*enable*)	515	78.9
	17 (*comprise*)	238	36.4
	18 (*manifests*)	168	25.7
	19 (*attained*)	260	39.8
	20 (*empirical*)	134	20.5
	21 (*superficial*)	161	24.7
	22 (*detriment*)	119	18.2
	23 (*assimilate*)	392	60.0
	24 (*infer*)	122	18.7
	25 (*stipulated*)	235	36.0
ADVANCED WORD LIST	26 (*susceptible*)	67	10.3
	27 (*preconceived*)	119	18.2
	28 (*untenable*)	78	11.9
	29 (*consolidate*)	84	12.9
	30 (*emulation*)	190	29.1

Table 2: Percentage of correct answers for each item in Test B

	Test item	Total (n = 653)	Percentage
1000- LIST	1 (*however*)	324	49.6
	2 (*developed*)	589	90.2
	3 (*influences*)	482	73.8
	4 (*provide*)	409	62.6
	5 (*represents*)	333	51.0
2000-LIST	6 (*avoid*)	382	58.5
	7 (*forward*)	306	46.9
	8 (*combine*)	568	87.0
	9 (*perform*)	382	58.5
	10 (*responsibility*)	519	79.5
UNIVERSITY WORD LIST	11 (*adequate*)	388	59.4
	12 (*perspective*)	487	74.6
	13 (*rendered*)	183	28.0
	14 (*intrinsically*)	221	33.8
	15 (*prevails*)	379	58.0
	16 (*allocated*)	528	80.9
	17 (*perpetuates*)	311	47.6
	18 (*devised*)	246	37.7
	19 (*beneficial*)	370	56.6
	20 (*advocated*)	155	23.7
	21 (*homogeneous*)	400	61.3
	22 (*supplementary*)	109	16.7
	23 (*validity*)	368	56.4
	24 (*tolerated*)	482	73.8
	25 (*incorporate*)	274	42.0
ADVANCED WORD LIST	26 (*extremist*)	222	34.0
	27 (*subservient*)	54	8.3
	28 (*integral*)	397	60.8
	29 (*holistic*)	201	30.8
	30 (*paucity*)	53	8.1

137

Table 3 : Analysis of items in Test A

	(a)		(b)		(c)		(d)	
	Prop.	Score	Prop.	Score	Prop.	Score	Prop.	Score
Q.1	0.04	26.56	**0.86**	**24.38**	0.07	22.87	0.01	18.40
Q.2	0.46	23.79	0.06	23.34	0.07	21.24	**0.38**	**25.69**
Q.3	0.02	22.66	0.12	23.95	0.37	22.87	**0.45**	**25.77**
Q.4	0.01	26.00	0.07	20.35	**0.84**	**24.86**	0.05	20.93
Q.5	0.05	21.32	0.05	18.80	**0.82**	**25.17**	0.06	21.14
Q.6	0.02	17.81	**0.90**	**24.89**	0.02	23.38	0.04	17.65
Q.7	0.09	23.11	0.05	19.25	**0.61**	**26.35**	0.23	20.66
Q.8	0.11	21.77	**0.77**	**24.87**	0.06	23.55	0.04	22.36
Q.9	0.51	23.97	0.13	23.44	**0.26**	**26.14**	0.07	21.71
Q.10	0.11	20.94	0.14	22.18	**0.67**	**25.69**	0.05	20.85
Q.11	0.02	22.59	0.09	21.15	**0.36**	**28.36**	0.51	22.12
Q.12	0.16	21.37	**0.12**	**28.60**	0.24	22.70	0.45	25.12
Q.13	0.13	20.32	0.23	23.11	0.08	22.95	**0.53**	**26.09**
Q.14	0.18	24.78	0.30	23.11	0.15	21.46	**0.35**	**26.39**
Q.15	**0.40**	**25.81**	0.18	22.14	0.28	24.36	0.12	22.38
Q.16	0.02	20.26	**0.78**	**25.15**	0.16	21.42	0.02	21.68
Q.17	0.35	23.35	**0.35**	**26.84**	0.09	22.65	0.19	22.21
Q.18	0.40	23.60	**0.24**	**28.47**	0.31	22.40	0.03	21.37
Q.19	0.41	23.08	**0.38**	**27.40**	0.14	20.89	0.04	20.46
Q.20	0.17	24.46	**0.18**	**27.21**	0.48	22.95	0.14	24.79
Q.21	0.09	26.08	0.18	23.76	**0.24**	**26.93**	0.47	22.82
Q.22	**0.17**	**27.83**	0.33	23.15	0.13	22.75	0.34	24.41
Q.23	0.04	17.31	0.10	22.50	0.27	21.44	**0.57**	**26.46**
Q.24	**0.17**	**28.92**	0.45	24.27	0.20	22.27	0.16	22.12
Q.25	0.28	22.38	**0.34**	**27.52**	0.19	21.48	0.17	24.12
Q.26	0.49	24.11	0.31	23.25	0.10	23.83	**0.08**	**30.21**
Q.27	0.60	23.55	0.18	23.03	0.04	24.36	**0.16**	**28.80**
Q.28	**0.11**	**25.22**	0.19	23.54	0.08	27.01	0.59	24.01
Q.29	0.45	25.65	0.23	22.84	**0.13**	**25.32**	0.16	22.02
Q.30	**0.27**	**26.56**	0.25	22.40	0.06	20.49	0.38	24.73

Table 4 : Analysis of items in Test B

	(a)		(b)		(c)		(d)	
	Value	Percent	Value	Percent	Value	Percent	Value	Percent
Q.1	0.05	23.88	0.16	27.34	**0.38**	**28.95**	0.15	28.05
Q.2	0.02	22.72	0.02	21.10	0.01	23.22	**0.68**	**28.71**
Q.3	0.00	19.28	0.06	24.41	0.12	24.90	**0.55**	**29.36**
Q.4	0.07	25.77	0.06	25.29	0.14	25.95	**0.46**	**29.47**
Q.5	0.26	26.75	0.01	22.50	0.09	23.29	**0.37**	**30.46**
Q.6	**0.43**	**30.62**	0.14	23.08	0.04	24.82	0.12	26.30
Q.7	0.13	26.99	0.16	25.82	**0.35**	**30.39**	0.09	25.00
Q.8	**0.65**	**28.78**	0.04	21.90	0.03	25.90	0.01	21.61
Q.9	0.25	26.29	**0.44**	**29.92**	0.03	21.70	0.01	20.53
Q.10	0.02	24.78	0.05	24.54	**0.59**	**29.14**	0.07	23.88
Q.11	0.04	23.97	0.17	24.67	**0.43**	**30.68**	0.09	24.63
Q.12	**0.56**	**29.57**	0.02	23.13	0.09	24.37	0.07	23.16
Q.13	0.07	28.91	0183	26.60	**0.21**	**28.95**	0.27	28.32
Q.14	0.08	25.98	0.11	26.66	0.29	26.58	**0.25**	**31.36**
Q.15	**0.43**	**30.31**	0.09	26.67	0.15	24.45	0.06	24.67
Q.16	0.04	25.31	**0.60**	**29.41**	0.04	23.07	0.05	20.20
Q.17	0.15	26.27	0.09	27.81	**0.35**	**30.76**	0.14	23.91
Q.18	0.22	26.82	**0.28**	**31.30**	0.18	25.54	0.05	26.83
Q.19	**0.41**	**29.47**	0.07	23.45	0.17	27.17	0.08	27.83
Q.20	0.10	27.70	0.29	26.89	0.15	25.87	**0.18**	**32.62**
Q.21	0.12	27.22	**0.46**	**29.40**	0.09	24.96	0.06	25.63
Q.22	0.33	27.73	0.13	28.49	0.14	28.20	**0.12**	**29.22**
Q.23	**0.42**	**30.58**	0.09	25.47	0.09	26.18	0.13	23.78
Q.24	0.07	25.94	0.06	23.36	**0.55**	**29.40**	0.03	24.18
Q.25	0.11	27.14	0.11	24.90	0.18	27.49	**0.31**	**30.29**
Q.26	**0.25**	**29.54**	0.12	25.76	0.29	28.03	0.05	28.77
Q.27	0.07	23.58	0.38	28.31	**0.06**	**30.19**	0.21	28.81
Q.28	**0.46**	**30.29**	0.05	23.43	0.06	24.29	0.15	25.06
Q.29	**0.23**	**30.13**	0.09	27.04	0.29	26.52	0.10	29.20
Q.30	**0.06**	**26.75**	0.18	27.38	0.26	29.47	0.22	27.72