

# *Measuring the contribution of vocabulary knowledge to proficiency in the four skills*

James Milton  
Swansea University

This chapter examines the way vocabulary knowledge relates to the ability to perform communicatively in a foreign language and in particular the ability to perform in the four language skills of reading, writing, listening and speaking. It reviews recent research designed to investigate the way vocabulary knowledge and performance inter-relate. There is a tradition of research which demonstrates that measures of vocabulary knowledge are particularly good predictors of performance in the four skills, and recent research suggests that when measures of different dimensions of vocabulary knowledge are combined this predictive-ness can be enhanced. Large vocabularies, and speed and depth of vocabulary knowledge, appear indispensable to the development of good performance in any language skill and it is now possible to enumerate the scale of vocabulary that is needed for the CEFR levels of communicative performance.

## **1. Lexical knowledge and language learning**

A feature of the English language literature on language learning and language teaching methodology over the last 60 years or so is the way vocabulary as a subject for teaching has been side-lined. It receives little attention in much of the literature on second language acquisition as a general process (e.g. Mitchell & Myles, 2004; Lightbown & Spada, 2006). It is almost entirely absent from major books on the syllabus and theory of language teaching (O'Dell, 1997, p. 258). Wilkins (1972, p. 109) suggests this may have been a product of the development of structural approaches to linguistics after the Second World War and the way that, in these approaches, vocabulary could be reduced to the minimum needed to illustrate the structural content. However, the absence of vocabulary is notable even after structural approaches to language teaching became unfashionable and were replaced by communicative and other approaches. Definitive works in these areas either omit to mention the topic entirely, as in Littlewood (1983), or dismiss the subject as one which is unsystematic and incidental at best to language learning, as in Brumfit (1984). It is true that at an academic level there is much renewed interest in the subject but,

as Schmitt (2008) notes, the insights gained have failed to make their way into the mainstream literature on language pedagogy. An example of the prevailing attitude to vocabulary in pedagogy can be seen in the comment by Harris and Snow that “few words are retained from those which are ‘learned’ or ‘taught’ by direct instruction ... [and learners] extend their vocabulary through sub-conscious acquisition” (Harris & Snow, 2004, pp. 55-61). With this attitude, the explicit teaching of vocabulary, and the systematic organisation of vocabulary in the curriculum, is not a priority.

In academic circles, the place of vocabulary in language learning has been significantly revised over the last decade and current academic thinking is very much at odds with much classroom and textbook practice. Far from being an element which is merely incidental to language learning, current thinking advocates that vocabulary may be crucial to the development of language performance overall. In a recent version of generative grammar, the Minimalist Program (Chomsky, 1995), the differences between languages are seen to be mainly lexical in nature and this leads Cook (1998) to suggest that the Minimalist Program is lexically-driven. The properties of the lexical items shape the sentence rather than lexical items being slotted into pre-existent structures. The task the language learner faces, therefore, is principally one of learning the vocabulary of the foreign language. The acquisition of vocabulary items in sufficient quantity triggers the setting of universal grammatical parameters. This approach is reflected in the Lexical Learning Hypothesis (Ellis, 1997) according to which vocabulary knowledge is indispensable to the acquisition of grammar.

One of the outcomes of the recent academic interest in vocabulary has been the development of ways for describing and testing vocabulary knowledge, which are both principled and systematic. Recently developed methods allow normalised data to be produced so the growth of a foreign language lexicon over the course of learning can be modelled. With this information it becomes possible to measure the contribution of vocabulary knowledge to language development and confirm whether the close relationship between vocabulary growth and language level exists in practice.

## **2. Dimensions of vocabulary knowledge**

A feature of our understanding of vocabulary, whether in a first or foreign language, is that knowledge of this aspect of language is multi-faceted. Potentially, there is a lot involved in knowing a word. The ancient Greeks, for example, clearly identified three elements of word knowledge: knowledge of aural and written forms and knowledge of the meaning of the word. For at least a century, too, a distinction has been made between receptive knowledge, and produc-

tive word knowledge. Some words, it seems, exist in the minds of language speakers primed for use and can be called to mind in speech or in writing easily and quickly. Other words are not used in this way but can, nonetheless, be called to mind for comprehension if they occur in the speech or writing of others. Each of these facets of knowledge can contribute to language performance in its own different way. A language user with extensive knowledge of words in their phonological form but no knowledge of the written form of words, for example, has the potential at least to speak and understand speech but no capacity for reading or writing. There is no definitive list of what comprises word knowledge and even native speakers will not know every facet of every word in their lexicon. In measuring vocabulary knowledge in order to assess how it impacts on overall language performance, therefore, decisions have to be made as to exactly what it is that is being measured.

The nearest thing we have to a definitive list of what it means to know a word is Nation's (2001) table shown in table 1. This table usefully encapsulates knowledge of the various forms of a word, the various aspects of meaning a word can carry with it, and the elements of use which are also part of word knowledge. Knowledge of form includes not just knowledge of the written and sound forms of a word but also knowledge of affixation, knowledge of the way extra parts can be added, or the ways in which a word can change, to reflect changes in its grammatical function or to add to its meaning. Knowledge of meaning includes not just knowledge of a core meaning, perhaps a link with a direct foreign language counterpart, but also the concepts, referents and associations, which a word may carry with it. Words in different languages often carry differences in nuances of meaning, which, if a learner is to perform fluently, may need to be known. And knowledge of use includes knowledge of the grammar of a word but also the way words like to behave in relation to each other. Some words like to occur in combination with other words, in particular idioms for example, and some words, like swear words, may be restricted in the occasions where they can be used appropriately, and this knowledge will also be needed if the language is to be used fluently and skilfully. Each facet of knowledge is sub-divided into receptive and productive knowledge.

This is a very useful and insightful list, and makes apparent just how much is involved in fully knowing a word. It is also clear that designing a test that can capture knowledge in all this diversity is scarcely practical. A single test could not possibly hope to encompass every aspect of knowledge described in this table. There is a further difficulty inherent in this table in that the various forms of knowledge are characterised but not precisely defined. In assessing knowledge of word parts, for example, it is unclear at what point the additions and changes to a word will form a new word rather than a derived form of an existing one. Nor is it clear, for example, how frequently a word must co-occur with another

word for a collocation to be created. But if vocabulary knowledge is to be tested and understood, then these are constructs which must be precisely defined.

**Table 1.** Description of “what is involved in knowing a word”, from Nation (2001: 27).

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

In order to reduce this complexity to manageable proportions, therefore, it has become common to think of vocabulary knowledge in terms of dimensions rather than a lengthy list of discrete elements. Each dimension can encompass a range of the separate elements in Nation’s list, which are linked in some way to form a single, larger entity. A common distinction, instigated by Anderson and Freebody (1981), is that between vocabulary breadth, that is, the number of words a learner knows regardless of the form they are known in or how well they are known, and vocabulary depth, which is how well or how completely these words are known. Neither of these terms is completely unambiguous. Vocabulary breadth, sometimes called vocabulary size, may be used to reflect a learner’s recognition vocabulary only: their ability to recognise the form of a word as a real word in the foreign language, and distinguish it from an artificially created non-word. The term may also be used to reflect a learner’s ability to recognise a word and link it to meaning or to a translation in the first language. Defining a vocabulary item like this entails a higher order of knowledge than defining it in terms of sheer recognition and it might be expected that

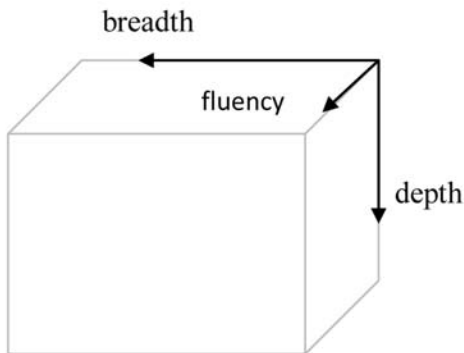
measurements of knowledge made using a higher order of knowledge criterion would be smaller than measurements made using a recognition requirement only. Notwithstanding these differences, vocabulary breadth has become sufficiently well acknowledged to be included in several well-established tests. Meara and Milton's (2003) X-Lex measures recognition knowledge of the most frequent 5000 lemmatised vocabulary items in a number of languages. Nation's (2001) Vocabulary Levels Test tests the ability of learners to recognise vocabulary items and link them to a definition among a selection of items drawn from a range of frequency bands and lists. Usefully, there is also a productive version of this test (Laufer & Nation, 1999).

Vocabulary depth is less well defined. It can be characterised in terms of knowledge of any of the several facets which Nation lists and which might involve knowledge about a word rather than just recognising it: associational knowledge, collocational knowledge, inflectional and derivational knowledge, knowledge of concepts and referents, and knowledge of constraints on use (Read, 2000). It has also been characterised in terms of movement along a continuum from partial to precise knowledge of a word (Henriksen, 1999). These approaches have been criticised because it is difficult to find a concept that holds together the variety of elements, which might fall into this category (Milton, 2009). However, Meara and Wolter (2004) can make a case for doing this by defining depth in terms of the number of links between words and the networks words can create. A word that is recognised as a word in a language, but where nothing more is known about it, has no links and is not networked. Once a meaning is attached to that form and some idea is gained as to how the word can be used, then it develops links with other words and begins to network and it does not matter whether these are grammatical or associational or collocational links. Words, which have this network of links around them can be said to be known more deeply than those, which are not known in this way. Depth of knowledge could be measured by counting the number of links in a word's network. While this is an interesting approach, there is an argument too that breadth and depth are, in effect, the same thing. As Vermeer (2001) points out, a condition of having a large network of links to a word is knowing a large number of words to make those links. A condition for developing depth in vocabulary knowledge will be to develop vocabulary breadth. The two are interdependent. As these points demonstrate, there is something of a contradiction in the way vocabulary specialists handle the distinction between breadth and depth. On the one hand, it can be assumed that breadth and depth will be closely related so that scores on tests of breadth can be used to validate newly created tests of depth (e.g. Gyllstad, 2007; this volume). On the other hand, it is still common to talk of vocabulary knowledge in terms of breadth and depth as two separate and contrasting dimensions which should not be closely connected.

Perhaps because the term is ill-defined, there is an absence of well-established and standardised tests in this field. Attempts, such as Wesche and Paribakht's (1996) Vocabulary Knowledge Scale (VKS), to fill this void are not without their difficulties and in us, as Wolter (2005) points out, they function as breadth rather than depth measures.

Meara (1996) adds a third dimension to breadth and depth by also characterising vocabulary knowledge in terms of the automaticity with which the words a person knows can be recognised and processed, or accessed for use in language. Daller et al. (2007) call this fluency and the presence of this dimension, and with it the attempt to characterise and measure the ability to activate what would otherwise be receptive knowledge, retains the productive and receptive distinction, which has proved so useful in measuring vocabulary knowledge. The presence of this third dimension allows Daller et al. to suggest a hypothetical, three-dimensional lexical space (Figure 1).

**Figure 1.** The lexical space: dimensions of word knowledge and ability (based on Daller et al., 2007: 8)



This hypothetical space allows learners with different types of vocabulary knowledge to be positioned differently in this space and systematically distinguished from each other. As Meara and Wolter (2004) explain, these distinctions might be used to explain how learners can behave differently from each other in their ability to perform in their foreign language. A learner with high vocabulary breadth but low fluency and low depth might be usefully distinguished from a learner with the same vocabulary breadth but higher fluency and depth. Although the number of words they know might be the same, the latter learner might be expected to be more skilful and able in using the foreign language in communicative situations. There is an increasing body of research evidence to support the idea that vocabulary knowledge and performance in a for-

eign language are linked (e.g. Alderson, 1984; Laufer, 1992; Laufer & Nation, 1999; Qian, 1999; Zimmerman, 2004) and it is the nature and extent of this link that this chapter intends to make more clear.

### **3. Performance in the four skills in a foreign language and vocabulary knowledge**

The goal for any foreign language learner is to use the language in some way. This may be for speech and casual conversation, or for translation of texts, or for study through the medium of the foreign language. It has become a commonplace in the assessment of language to consider language in terms of four separate skills: the receptive skills of reading and listening, and the productive skills of speaking and writing. In reality, of course, these distinctions are not so clear and the ability to read and listen fluently requires the learner to actively anticipate the language that is likely to occur and then monitor input to check that the possibilities which have been created are occurring. Nonetheless, the distinction is enshrined in formal and assessment schemes. The University of Cambridge Local Examinations Syndicate (UCLES) exams, such as the International English Language Testing System (IELTS) test, administer separate papers for each of these skills and devise separate grading schedules for them. The Council of Europe's (2001) Common European Framework of Reference for Languages (CEFR) hierarchy uses both global descriptors of language performance as a whole (p. 24), and descriptors separated into the four skills (pp. 26-27). These descriptors are couched in terms of performance of language rather than in terms of the language knowledge, which is likely to underlie performance. The example below of the CEFR's global descriptor for performance at C2 level illustrates this (Council of Europe, 2001, p. 24).

Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.

There is an assumption that language knowledge, such as vocabulary knowledge, will develop in relation to language performance and this is reflected both in the wording of the descriptors for the four skills but also in the presence of levels criteria specifically for vocabulary range and vocabulary control in the CEFR document (p. 112). These two terms are not explicitly defined but range appears broadly to refer to the vocabulary size available to the learner, and con-

trol appears to be closer to vocabulary depth in that it refers to the accuracy and appropriateness of vocabulary selection and use. Table 2 presents the descriptors for vocabulary range.

**Table 2.** CEFR vocabulary range descriptors

<b>level</b>	<b>descriptor</b>
<b>C2</b>	Has a good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning.
<b>C1</b>	Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions; little obvious searching for expressions or avoidance strategies. Good command of idiomatic expressions and colloquialisms.
<b>B2</b>	Has a good range of vocabulary for matters connected to his/her field and most general topics. Can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution.
<b>B1</b>	Has a sufficient vocabulary to express him/herself with some circumlocutions on most topics pertinent to his/her everyday life such as family, hobbies and interests, work, travel, and current events. Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics.
<b>A2</b>	Has a sufficient vocabulary for the expression of basic communicative needs. Has a sufficient vocabulary for coping with simple survival needs.
<b>A1</b>	Has a basic vocabulary repertoire of isolated words and phrases related to particular concrete situations.

In terms of the dimensions of vocabulary knowledge described in the first section, it is clear that what is anticipated is that learners will grow an increasingly large lexicon as they progress through the levels. The ‘basic vocabulary’ at A1 Level becomes ‘a very broad lexical repertoire’ at C2 level. It is implied that only at the most advanced levels will developments in vocabulary depth be relevant. At C1 level ‘Good command of idiomatic expressions and colloquialisms’ is expected, and at C2 level ‘awareness of connotative levels of meaning’ is added. The use of expressions like command in the descriptors also implies that learners have these items available for use and that vocabulary knowledge has progressed along the fluency dimension as well as the breadth and depth dimensions. Implicit within this framework, therefore, is the understanding that a requirement of making progress in communicating through foreign language is acquiring greater volumes of vocabulary and acquiring ever greater sophistication and control in the use of this vocabulary. The Framework document, there-



fore, also suggests that it might be possible and useful for vocabulary size and depth measurements to be attached to the different levels.

There is some empirical evidence that links vocabulary breadth measures with the CEFR language levels. Milton (2010), shown in Table 3, provides EFL vocabulary sizes (out of the most frequent 5,000 lemmatised words in English) gained from over 10,000 learners in Greece taking both recognition tests of their vocabulary size and also formal UCLES exams at levels within the CEFR framework.

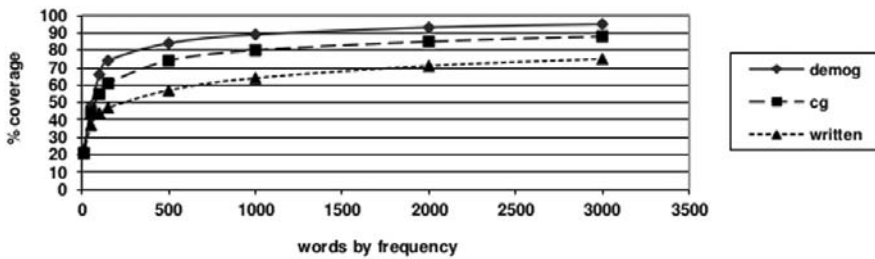
**Table 3.** Vocabulary size estimates, CEFR levels and formal exams (Milton, 2010, p. 224)

CEFR Levels	Cambridge exams	XLex (5000 max)
A1	Starters, Movers and Flyers	<1,500
A2	Kernel English Test	1,500 – 2,500
B1	Preliminary English Test	2,500 – 3,250
B2	First Certificate in English	3,250 – 3,750
C1	Cambridge Advanced English	3,750 – 4,500
C2	Cambridge Proficiency in English	4,500 – 5,000

While there is some individual variation around these ranges, Milton is able to conclude that “the assumption made in the CEFR literature, that as learners progress through the CEFR levels their foreign language lexicons will increase in size and complexity, is broadly true” (2010, p. 224). Variation may be explained by the way vocabulary knowledge and language performance are imperfectly linked. Learners with the same or similar vocabulary sizes – and remember these are based on knowledge of the 5,000 most frequent lemmatised words in English and so are not absolute vocabulary size estimates – may make different use of this knowledge to communicate more or less successfully. Milton and Alexiou (2009) report similar vocabulary size measurements for CEFR levels in French and Greek as foreign languages.

If vocabulary breadth predicts overall language performance well, then it might be expected that vocabulary breadth will link well also with the four separate skills. However, there are reasons for thinking that the oral skills, speaking and listening, will have a different relationship with vocabulary knowledge from the written skills, writing and reading. Figures for coverage (the proportion of a corpus provided by words in the corpus arranged in frequency order) in spoken and written corpora suggest that written text is typically lexically more sophisticated than spoken text. A comparison (Figure 2) of coverage taken from written and spoken sub-corpora of the 100 million word British National Corpus illustrates this (Milton, 2009, p. 58).

**Figure 2.** Coverage from written and spoken corpora in the BNC



The relationship between text coverage and comprehension (and by extension communication more generally) in a foreign language is now quite well understood. In this example the most frequent 2,000 lemmatised words in English provide about 70 % coverage of the *written* sub-corpus. This suggests that a learner with a vocabulary size at about this level would struggle to communicate at anything but the most basic level through the medium of writing. The same vocabulary size provides much more coverage in both the spoken sub-corpora, around 90 %. The *demographic* (demog in figure 2) sub-corpus is taken from samples of general conversation and the *context governed* (cg in figure 2) sub-corpus from examples of rather more formal and organised spoken language: lectures, court room exchanges and sermons. A figure of 95% coverage is often associated with full comprehension and even though this is probably an over-simplification Figure 2 does nonetheless suggest that a learner with a vocabulary size of about 2,000 lemmatised words might be quite communicative in speech and listening. It has been suggested (Milton et al., 2010) that the relationship between vocabulary size and performance in the written skills might, for all practical purposes, be linear and that up to the highest levels of performance greater vocabulary breadth is associated with better language performance. A learner's vocabulary would have to be well beyond the 5,000 words measured in this paper before the additional benefit of extra words ceased to add to comprehension. However, the relationship in the oral skills may not be linear and Milton et al. suggest that beyond a comparatively low level, 2,000 or 3,000 lemmatised words, there may be little benefit to performance in learning more vocabulary. Beyond this level, therefore, learners will experience diminishing returns in their oral performance for the time invested in learning vocabulary.

## 4. Examining the relationship between vocabulary knowledge and language proficiency

It has been acknowledged for some time that vocabulary knowledge is a good predictor of general proficiency in a foreign language. However, most research on the relationship has been conducted with measures of vocabulary size only, and within the realm of reading skill only (Stæhr, 2008). Generally, such studies have found strong correlations between receptive vocabulary size tests and reading comprehension tests, ranging from 0.50 to 0.85, with learners from different proficiency levels (e.g. Laufer, 1992; Qian, 1999; Albrechtsen, Haastrup & Henriksen, 2008).

A feature of recent work in vocabulary studies has been to try to investigate more fully the links between lexical knowledge and learner performance, and investigate the scale of the contribution which vocabulary, in all three of its dimensions, can make to a variety of communicative skills in foreign language performance. By extension, such research also tests the credibility of theories such as the Lexical Learning Hypothesis (Ellis, 1997), and contributes firmer evidence to the place that vocabulary should have in the structure of the foreign language learning curriculum, since in this view of learning it is vocabulary knowledge which drives learning in other aspects of language. However, the considerations above have suggested that the relationship between vocabulary knowledge and overall language skill may potentially be difficult to model and to measure. Different dimensions of vocabulary knowledge might need to be measured separately and their effects combined if the full nature of the relationship with language skill is to be seen. Further, it might be that the relationship will vary according to the level of the learner and the skills the learner needs. The following sections will examine particular pieces of research in this area, which illustrate the state of our knowledge and from which broader conclusions can be drawn.

### 4.1. *Stæhr (2008)*

In this paper Stæhr attempts to investigate the relationship between vocabulary size and the skills of reading, writing and listening among EFL learners in lower secondary education in Denmark. He further investigates the role played by the most frequent 2,000 word families in achieving pass grades in the tests of these skills.

Stæhr measures vocabulary size using the improved version of the Vocabulary Levels Test (VLT) made by Schmitt, Schmitt and Clapham (2001). The test consists of five separate sections, which represent four levels of word frequency and a similarly structured test of knowledge of Coxhead's Academic Word List

(Coxhead, 2000). However, the academic word level was excluded from Stæhr's study as not relevant for low-level learners. The test assesses learners' receptive knowledge of word meaning at the 2,000, the 3,000, the 5,000 and the 10,000 level, and the test results can thus give an indication whether learners master the first 2,000, 3,000, 5,000 or 10,000 word families in English. Although the VLT was originally designed as a diagnostic test intended for pedagogical purposes, researchers (e.g. Read, 2000; Schmitt et al., 2001) acknowledge its use as a means of giving a good guide to overall vocabulary size. Tests of language skills were assessed as part of the national school leaving examination. Reading and listening skill abilities were measured using pencil-and-paper multiple-choice tests. Writing ability was measured using the scores awarded for an essay task where the participants had to write a letter to a job agency applying for a job.

Stæhr's results indicate a correlation between vocabulary size and reading, which is comparable with the findings of other research mentioned above and suggests a strong and statistically significant relationship between the amount of vocabulary a learner knows in the foreign language and their ability to handle questions on a text designed to test their ability to fully comprehend the text. His analysis, using binary logistic regression, shows that as much as 72% of the variance in the ability to obtain an average score or above in the reading test is explained by vocabulary size (Nagelkerke  $R^2 = 0.722$ ). The results also illuminate the relationship with other language skills. The correlation between vocabulary size and both writing and listening ability is also statistically significant and reasonably strong. Stæhr suggests that 52% of the variance in the ability to obtain an average or above-average writing score is accounted for by vocabulary size (Nagelkerke  $R^2 = 0.524$ ), and that 39 % of the variance in the listening scores, in terms of the ability to score above the mean, is accounted for by the variance in the vocabulary scores (Nagelkerke  $R^2 = 0.388$ ). His interpretation of this is that this amount of variance is substantial. Even the contribution towards listening, the smallest in this study, is considerable, given the fact that it is explained by one single factor. This confirms the importance of receptive vocabulary size for learners in all three skills investigated.

Stæhr's findings further indicate the importance of knowing the most frequent 2,000 word families in English in particular and he suggests that knowledge of this vocabulary represents an important threshold for the learners of his study. Knowledge of this vocabulary is likely to lead to a performance above average in the listening, reading and writing tests of the national school leaving exam. The results seem to emphasize that the 2,000 vocabulary level is a crucial learning goal for low-level EFL learners and suggest that the single dimension of vocabulary size is a crucial determiner of the ability to perform in the three foreign language skills tested. The more vocabulary learners know, the better they are likely to perform through the medium of the foreign language.

#### 4.2. *Milton, Wade & Hopkins (2010)*

Even if it appears that a single dimension, vocabulary size, contributes hugely to a learner's ability to perform through a foreign language, this is still far from a complete explanation of the nature of the relationship. There is a gap, it appears, in vocabulary's explanatory power. Studies investigating the relationship between vocabulary and language proficiency almost never offer an insight into the relationship between vocabulary and speaking ability. Zimmerman (2004) is possibly the only study, which demonstrates that such a relationship exists. The correlations between vocabulary size scores and aural skills, such as listening, are generally weaker than the correlations with the written skills of reading and writing. This may be due to the way oral language is lexically less sophisticated than written language, but may also be due to the nature of the tests to measure vocabulary size, which are invariably delivered through the medium of writing. Milton et al. (2010) also point out that it is quite conceivable that word knowledge may be held in the lexicon in aural form only, and not necessarily in written form and, particularly if vocabulary learning comes about predominantly through oral input as Ellis (Ellis, 1994, p. 24) suggests. With regard to aural skills, therefore, the vocabulary tests may be tapping into vocabulary knowledge in the wrong form for the relationship to be made fully clear.

Milton et al. (2010) therefore conducted a study which investigated whether vocabulary knowledge can be held in different formats, written and/or aural, and whether the measurements of vocabulary size made using two different formats can better explain performance in all the four skills. To measure written vocabulary size they used X-Lex (Meara & Milton, 2003), where the words tested only ever appear in written form, and to measure vocabulary size in aural format they use A-Lex (Milton & Hopkins, 2006), which is designed as an equivalent of X-Lex but where words are just heard. To measure performance in the four skills they used IELTS sub-scores. Data was collected from 30 students at intermediate and advanced level, from a variety of national and language backgrounds, attending a pre-session course in UK. Milton et al. hypothesised that scores from the written vocabulary size test would correlate best with IELTS scores for reading and writing, that scores from the aural vocabulary size test would correlate best with the IELTS scores for speaking, and that a combination of the two vocabulary tests would best explain scores from the IELTS listening test where the test format involved words in both written and aural form.

Milton et al.'s results, like Stæhr's, show statistically significant correlations between vocabulary size and language performance measures, although the strength of the relationships is weaker. They also very strongly suggest that vocabulary knowledge need not be held in both written and aural form combined, and that words can be held in one form only. The correlation between the

two vocabulary size tests is moderate to poor at 0.41, even if the relationship is still statistically significant. Interestingly, it appears that elementary level learners have knowledge predominantly in aural form, while the more advanced learners tend increasingly to grow lexicons where words appear to be known through written form only (see also Milton & Hopkins, 2006; Milton & Riordan 2006). It seems that vocabulary size can predict oral skills comparably with written skills provided that vocabulary size is measured appropriately. The correlation between A-Lex and speaking scores (0.71) is very similar to the correlations observed between X-Lex and reading and writing scores (0.70 and 0.76).

Regression analysis suggests that vocabulary size can explain broadly similar amounts of variance in all the four skills. If the relationship is assumed to be linear, and one should bear in mind that for oral skills in particular this need not be the case, then between 40 % and 60 % of variance in sub-skills scores can be explained through the single variable of vocabulary size. Variance in the listening sub-test, which involves both reading questions and listening for answers, is best explained through a combination of the written and aural sub-scores. Analysis using binary logistic regression, used because the relationship may not be linear, produces comparable results explaining between 41% and 62% of variance in the ability to score grade 5 or above on the IELTS sub-tests. The fact that binary logistic regression explains more variance in the speaking scores (Nagelkerke  $R^2 = 0.61$ , Cox & Snell  $R^2 = 0.45$ ) than the linear regression (Adjusted  $R^2 = 0.40$ ) is tentatively suggested by Milton et al. as evidence that the relationship between vocabulary size and performance in tests of speaking skill is non-linear, although differences in the way these scores are calculated make this a highly subjective interpretation.

The significance of these results is to confirm the importance of the vocabulary size dimension in all aspects of foreign language performance. Vocabulary size, calculated appropriately, appears consistently to explain about 50% of variance in the scores awarded to learners for their performance in the sub-skills of language, including speaking skills where hitherto the relationship has been assumed to be less strong. The fact that, as in explaining listening sub-scores, measurements for different aspects of vocabulary knowledge can be aggregated to enhance the explanatory power of vocabulary in the four skills suggests that continuing to investigate the various dimensions of vocabulary knowledge may yield useful insights.

#### *4.3. Schoonen (2010)*

Recent work by Schoonen and his colleagues has investigated the influence of the dimensions of vocabulary knowledge additional to size on the development of language skills. They have tried to use a combination of size and depth and flu-

ency measures to better explain variation in performance in the language skills. These have been usefully summarised in a conference paper (Schoonen, 2010).

Data on the influence of lexical variables on reading comprehension and writing proficiency is drawn from a number of studies (including Schoonen et al., 2011), which includes vocabulary size and automaticity (or fluency) with other predictor variables. Vocabulary size was measured using the Vocabulary Levels Test (VLT) using the Schmitt et al. (2001) version. Speed was measured using two computer-delivered tests where speed of word recognition and speed of word retrieval could be measured. The other variables were collected using a metacognitive knowledge questionnaire and grammatical knowledge tests. It is acknowledged that one of the problems associated with this approach and the analysis of the data it produces is multicollinearity. Because the lexical variables and other factors such as grammatical knowledge may all be influenced by the frequency of their occurrence in natural language and the degree of exposure a learner has had to the foreign language, these variables may correlate closely with each other. Separating out the impact of individual variables from each other may be difficult.

The results show that vocabulary size and the vocabulary speed measures produce statistically significant and positive correlations with scores from the reading and writing tests. It is concluded that the predictiveness of vocabulary size can be enhanced when combined with their measures of speed or fluency. Lexical variables in this study explain in the range of 30% of the variance of reading and writing scores, slightly smaller than obtained in either of the Stæhr or Milton et al. studies. In this analysis, too, vocabulary is a good predictor of performance and the proportion of variance it explains is substantial.

#### *4.4. Vocabulary knowledge and the four skills*

Perhaps the most important conclusion that emerges from the research is the importance of vocabulary knowledge in being able to understand and communicate in a foreign language. The studies reported above, among others, demonstrate this clearly, showing a moderate to strong relationship between vocabulary measures and the ability to read, write, listen, and it seems also speak, in the foreign language. Generally speaking, the more words a learner knows, the more they are likely to know about them, and the better they are likely to perform whatever the skill. The single factor of vocabulary can explain up to 50% of the variance in performance in scores gained from tests of the four skills. This is a large figure, given that variation might be expected in learners' ability to apply their lexical knowledge - some are likely to be more skilled in using what they know than others. Nonetheless, because this close connection between vocabulary knowledge and skill exists, it is not perhaps surprising that

vocabulary sizes can be linked to language levels as those presented in the CEFR and that vocabulary size can be used as a reliable placement measure. The expectation that oral skills would not be so closely linked to vocabulary size has not emerged in these studies possibly because the measures of skill used relate to measures such as IELTS scores, which are rather academic and might favour a more linear relationship than would be the case if the skills were measured in a non-academic context. Unusually in the spoken register, the skills rewarded in the IELTS speaking sub-test may benefit from the more extensive use of infrequent vocabulary. This conclusion has emerged despite the clear evidence that in successful language performers words are held predominantly in the written form and have presumably been learned by reading rather than through oral interaction.

Stæhr (2008) has remarked that the explanatory power of vocabulary size in explaining variance in scores on language skills suggests that vocabulary size may be *the* determinant factor, pre-eminent among the other factors which may be at work in performing in and through a foreign language. Schoonen's findings, however, suggest that this may be an exaggeration, since size and other factors appear so closely linked and the importance of other variables exceeds vocabulary in his study. Nonetheless, vocabulary knowledge, and vocabulary size in particular, are clearly a very major contributor to success in language performance. It has emerged that knowledge of the most frequent 2,000 words, in particular, is an important feature in successful communication through a foreign language. There is a caveat here, in that the findings suggest that in oral skills the importance of vocabulary knowledge diminishes with increasing size rather faster than it does in skills that involve the written word. The reason for this is worth consideration and the best explanation available is that this is connected with coverage and differences in the way we handle written and spoken language. Corpora suggest that, in English language for example, the most frequent words in a language are even more frequent in spoken language than in written language. Adolphs and Schmitt's (2003) analysis of spoken data in CANCODE indicates that important coverage thresholds such as the 95% coverage figure for general comprehension might be reached with between 2,000 and 3,000 words; perhaps half the figure needed to reach the same threshold in written discourse.

The studies by Stæhr (2008), Milton et al. (2010) and Schoonen (2010) discussed above suggest that, because the dimensions of vocabulary knowledge are so closely linked, a single measure of vocabulary knowledge is likely, by itself, to be a good indicator of skill and level in a foreign language. Because vocabulary breadth in English is now easily measurable using reliable tests for which we have normalised scores, perhaps it is not surprising if vocabulary size or breadth has become particularly closely associated with performance in the four



skills. It seems from the studies above, however, that other dimensions also contribute to performance, perhaps as much as size, and that a combination of scores for size and depth, or size and speed, for example, can add up to 10% to the explanatory power of vocabulary knowledge in skills performance. Very crudely, the more sophisticated the measures of vocabulary knowledge, the more they are likely to explain variance in performance in the four skills, up to the level of around 50%. Beyond that point other factors will be needed to improve the explanatory power of any model. These could be knowledge factors, such as grammatical knowledge, or skill factors in the ability that users have in applying their knowledge when listening, reading, speaking or writing. This is clearly an avenue for further research.

The studies discussed above also allow us to reconsider the concept of lexical space explained at the outset of the chapter: the idea that learners can be characterised differently according to the type of knowledge they have of the words they know in their foreign language, and this can explain how they vary in performance. One interpretation why the depth and size dimensions correlate so well is that they are essentially the same dimension, at least until learners become very knowledgeable and competent and sufficient words are known for subtlety in choice or combination to become possible (see Gyllstadt, this volume). The convenient rectangular shape in Figure 1 is transformed into something much narrower at the outset of learning where lexical size is paramount, and becomes wider at the most competent levels where increased depth becomes a possibility and a potential asset. Co-linearity is noted by Schoonen who suggests another possibility (Schoonen, personal correspondence), that there will be an 'equal' development in all three dimensions, and all three will be strongly correlated, but this is probably a spurious correlation due to language exposure as common cause. Theoretically, it remains possible to have uneven profiles, including differences in breadth and depth, but to evaluate this experimental studies would be required where one dimension only is trained, for example speed, as in Snellings, Van Gelderen & De Glopper (2004).

#### *4.5. Vocabulary knowledge, theories of language learning, and implications for pedagogy*

At the outset of this chapter I suggested that there was a contradiction between much pedagogical theory and practice and recent SLA theories, as regards the importance and relevance of vocabulary knowledge to the process of acquiring proficiency in a foreign language. Current methods and approaches to language teaching fail to consider how vocabulary should be systematically built into the curriculum or suggest that this would not be appropriate assuming that the acquisition of vocabulary is merely incidental to the process of lan-

guage learning as a whole. Learners will not need a systematically constructed language lexicon of any particular size or content in order to achieve success in their second language (see for example Häcker 2008; Milton 2006). Contrary to this, recent theories, for example Ellis's Lexical Learning Hypothesis (Ellis, 1997), suggest that learning a large vocabulary is central to successful language learning and drives the learning of other aspects of language. I suggested that some of the recent research into vocabulary knowledge and performance in the four skills of language might illuminate this debate and provide better guidance for best practice in teaching and learning foreign languages.

If vocabulary were really marginal to the process of developing communicative ability in a foreign language, it might be expected that learners with large vocabularies and extensive lexical knowledge would, presumably, perform comparably with learners with much smaller vocabularies. If, however, the volume of vocabulary a learner knows is driving the acquisition of the other aspects of language and overall proficiency, then a much closer association might be expected. Learners with small or poorly developed vocabularies could not be as proficient nor as fluent in performing through the foreign language. It was suggested that because of the complexity in describing vocabulary knowledge completely and in measuring the variety of aspects that can be involved, the relationship between vocabulary and language learning might be difficult to capture and to measure. The research in this area suggests that the relationship between vocabulary and the development of skills in performing in the foreign language can be modelled and measured and several key features emerge.

Part of this modelling process can now make important distinctions in our understanding of the structure of the mental lexicon and the nature of the vocabulary needed to achieve communicative goals. It appears that word knowledge may be in phonological form or orthographic form and that important communicative goals are likely to be achieved with fewer words in speech than in writing. Surprisingly, it appears that a substantial volume of a highly fluent foreign language user's knowledge may reside in the realm of orthographic knowledge only. Speakers of languages using alphabetic systems of writing, where the spelling clearly relates to the pronunciation, still manage, it seems, to avoid storing this information or storing it correctly and so fail to recognise by sound words they can recognise in writing. Such a possibility has been discussed by Suarez and Meara (1989) and Segalowitz and Hulstijn (2005), who suggest that advanced learners develop a direct route to meaning from the written form, cutting out any intermediate phonological form, but there has been little to demonstrate that this can occur before. It implies that high fluency is linked to high literacy and the ability to access large amounts of particularly written text, to access the necessary infrequent words, and recognise them by shape or by morphological structure providing a route to meaning, which does not rely on

phonological coding. Learners without this high literacy and who are tied to phonological decoding may develop more balanced lexicons with orthographic and phonological word knowledge more equal in size as suggested in Milton and Hopkins (2006) and Milton and Riordan (2006). However, the price to be paid for this, perhaps through the slowness of the reading process and the extra burden on memory, is that the lexicon tends to grow more slowly, limiting communicativeness in the written domain.

The research summarised above appears to support theories such as Ellis's Lexical Learning Hypothesis. Vocabulary development, however measured, appears to mesh very closely with other features of language such as grammatical development, and also with overall language ability. Developing learners' vocabulary knowledge appears to be an integral feature of developing their language performance generally. The link has not been established in a strongly causal sense and while it is not yet clear that the vocabulary knowledge is driving the other aspects of language development, vocabulary certainly appears to develop in size and depth alongside every other aspect of language. This very strongly supports the idea, as in the lexical approach (Lewis & Hill, 1997), that vocabulary should be built more explicitly into the development of any good language curriculum. This could be in the form of indicating particular words to be learned, as in the most frequent words in any language, but it might imply the introduction of size as a metric into curricula as a means of setting appropriate targets and monitoring progress without dictating the content of learning directly.

Even though this may seem quite commonsensical, we have evidence from the UK that details of vocabulary can be systematically downplayed from formal curricula in line with methodological approaches such as the Communicative Approach. Curriculum descriptions for B1 level foreign language exams in UK (e.g. Edexcel, 2003, for French) routinely contain only minimal core vocabularies of around 1,000 words, levels of vocabulary which are incompatible with performance attainment at B1 level observed elsewhere in Europe (Milton & Alexiou, 2009). We also have evidence that the teaching of foreign language vocabulary following these curricula rarely extends beyond 1,000 words at B1 level (Milton, 2006; 2008; David 2008). In other countries (as indicated in Milton & Alexiou, 2009) CEFR levels have an expectation of rather greater vocabulary knowledge than in the UK and since it is highly unlikely that learners can be as communicative with 1,000 words at B1 level as with the 2,000 or more words required for this level elsewhere in Europe, there is a clear mismatch in the applications of the CEFR level which vocabulary size estimates can demonstrate.

## References

- Adolphs, S. & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24(4), 425-438.
- Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Basingstoke: Palgrave Macmillan.
- Alderson, J. C. (1984). Reading in a foreign language: A reading or a language problem? In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 1-24). London: Longman.
- Anderson, R. C. & Freebody, P. (1981). Vocabulary Knowledge. In J. T. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp. 77-117). Newmark: International Reading Association.
- Brumfit, C. (1984). *Communicative Methodology in Language Teaching*. Cambridge: Cambridge University Press.
- Chomsky, N. (1995). *The Minimalist Program*. (Current Studies in Linguistics, 28.) Cambridge, MA: MIT Press.
- Cook, V. (1998). Review of Skehan, P. (1998) *A Cognitive Approach to Learning Language*. Oxford: Oxford University Press. Accessed at <http://homepage.ntlworld.com/vivian.c/Writings/Reviews/SkehanRev.htm> on 16 Feb 07.
- Council of Europe (2001). *Common Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Daller, H., Milton, J., & Treffers-Daller, J. (2007). Editors' introduction: conventions, terminology and an overview of the book. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge* (pp. 1-32). Cambridge: Cambridge University Press.
- David, A. (2008). Vocabulary breadth in French L2 learners. *Language Learning Journal*, 36(2), 167-180.
- Edexcel (2003). Edexcel GCSE in French (1226) First examination 2003 June 2003. Accessed at <http://www.edexcel.com/quals/gcse/gcse-leg/lang/french/Pages/default.aspx> on 08.03.2011.
- Ellis, N. (1997). Vocabulary acquisition: Word structure, collocation, word-class, and meaning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description Acquisition and Pedagogy* (pp. 122-139). Cambridge: Cambridge University Press.
- Ellis, R. (1994). Factors in the incidental acquisition of vocabulary from oral input: a review essay. *Applied Language Learning*, 5(1), 1-32.
- Gyllstad, H. (2007). *Testing English Collocations - Developing Receptive Tests for Use with Advanced Swedish Learners*. Lund University: Media-Tryck.
- Häcker, M. (2008). Eleven pets and twenty ways to express one's opinion: the vocabulary learners of German acquire at English secondary schools. *Language Learning Journal*, 36(2), 215-226.

- Harris, V. & Snow, D. (2004). *Classic Pathfinder: Doing it for themselves: focus on learning strategies and vocabulary building*. London: CILT.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(2), 303-317.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. & Nation, P. (1999). A productive-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Lewis, M. & Hill, J. (1997). *The Lexical Approach; the state of ELT and the way forward*. Boston, Mass: Heinle.
- Lightbown, P. & Spada, N. (2006). *How Languages are Learned (3<sup>rd</sup> Ed)*. Oxford: Oxford University Press.
- Littlewood, W. (1983). *Communicative Language Teaching*. Cambridge: Cambridge University Press.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meara, P. & Milton, J. (2003). *X\_Lex, The Swansea Levels Test*. Newbury: Express.
- Meara, P. & Wolter, B. (2004). V\_Links, beyond vocabulary depth. *Angles on the English Speaking World*, 4, 85-96.
- Milton, J. (2006). Language Lite: Learning French vocabulary in school. *Journal of French Language Studies* 16(2), 187-205.
- Milton, J. (2008). French vocabulary breadth among learners in the British school and university system: comparing knowledge over time. *Journal of French Language Studies*, 18(3), 333-348.
- Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.
- Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Vedder, I. Bartning, & M. Martin (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (pp. 211-232). Second Language Acquisition and Testing in Europe Monograph Series 1.
- Milton, J. & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review*, 63(1), 127-147.
- Milton, J. & Riordan, O. (2006). Level and script effects in the phonological and orthographic vocabulary size of Arabic and Farsi speakers. In P. Davidson, C. Coombe, D. Lloyd, & D. Palfreyman (Eds.), *Teaching and Learning Vocabulary in Another Language* (pp. 122-133). UAE: TESOL Arabia.
- Milton, J. & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference for Languages. In B. Richards, H.M. Daller, D.D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary Studies in First and Second Language Acquisition* (pp. 194-211). Basingstoke: Palgrave Macmillan.

- Milton J., Wade, J. & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torrealblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.
- Mitchell, R. & Myles, F. (2004). *Second Language Learning Theories*. London: Hodder Arnold.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- O'Dell, F. (1997). Incorporating vocabulary into the syllabus. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 258-278). Cambridge: Cambridge University Press.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, 56(2), 282-307.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Review article: instructed second language vocabulary learning. *Language Teaching Research* 12(3), 329-363.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18 (1), 55-88.
- Schoonen, R. (2010). The development of lexical proficiency knowledge and skill. Paper presented at the Copenhagen Symposium on *Approaches to the Lexicon*, Copenhagen Business School on 8-10 December 2010. Accessed at <https://conference.cbs.dk/index.php/lexicon/lexicon/schedConf/presentations> on 03.03.2011.
- Schoonen, R., Van Gelderen, A., Stoel, R., Hulstijn, J., & De Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary-school students. *Language Learning*, 61(1), 31-79.
- Segalowitz, N. & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 371-388). Oxford: Oxford University Press.
- Snellings, P., Van Gelderen, A., & De Glopper, K. (2004). The effect of enhanced lexical retrieval on L2 writing. *Applied Psycholinguistics*, 25(2), 175-200.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139-152.
- Suárez, A. & Meara, P. (1989). The effects of irregular orthography on the processing of words in a foreign language. *Reading in a Foreign Language*, 6(1), 349-356.
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics* 22(2), 217-234.
- Wesche, M. & Paribakht, T. A. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *The Canadian Modern Language Review*, 53(1), 13-40.
- Wilkins, D. A. (1972). *Linguistics in Language Teaching*. London: Arnold.
- Wolter, B. (2005). *V\_Links: A New Approach to Assessing Depth of Word Knowledge*. PhD Dissertation, University of Wales Swansea.
- Zimmerman, K. J. (2004). *The role of Vocabulary Size in Assessing Second Language Proficiency*. MA dissertation, Brigham Young University.