USING A WORD KNOWLEDGE FRAMEWORK

TO RESEARCH VOCABULARY

by Norbert Schmitt, MEd

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy, May 1997

TABLE OF CONTENTS

ABSTRACT

The study of vocabulary acquisition is not exactly a new area, but previous research and hypothesizing has failed to produce a coherent overall theory which adequately describes it. This is partly because of the complexity of the subject. One method of reducing the complexity is to work with the individual components of vocabulary knowledge, in an attempt to understand the whole by first better understanding the parts. The word knowledge listing proposed by Nation (1990) is adopted in this thesis as a framework from which to study vocabulary.

Chapter 1 introduces the word knowledge framework. Chapter 2 provides a literature review which summarizes the research concerning each of the eight types of word knowledge. Chapter 3 reports on a study which attempts to quantify native and nonnative intuitions of word frequency. Chapter 4 describes how a procedure for weighting word association responses was developed. Chapter 5 does the same for a measure of collocational knowledge. Chapter 6 applies the word knowledge research paradigm to the evaluatation of the vocabulary items on the TOEFL test. Chapter 7 reports on a longitudinal study of four nonnative subjects which tracked their incremental acquisition of spelling, association, collocation, grammar, and meaning knowledge for eleven words over one year. Chapter 8 examines the data from the longitudinal study to see if the various kinds of word knowledge are learned in a developmental sequence. Chapter 9 concludes the thesis by giving the author's opinions about the strengths and weaknesses of the reported course of research.

DEDICATION

For my father

May I someday be half the man he was

ACKNOWLEDGEMENTS

In the course of three years, numerous `good shepherds' helped me to move this project along to fruition. Chronologically, the first which need thanking are the faculty of Temple University, Japan for initially opening my eyes to the wonders of Applied Linguistics, particularly Kenneth Schaefer, Michael Rost, Steven Gaies, Gladys Valcourt, and Rod Ellis. Next, I am lucky to have had the support and encouragement of some of the leading figures in vocabulary studies, namely Ron Carter, Michael McCarthy, Paul Meara, and Paul Nation. It also gives me a great deal of pleasure to call them my friends. Research requires a great deal of cooperation from teachers, students, and institutions. I would like to give a blanket thanks here; specific assistance from individuals will be acknowledged in each study. The Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom (CVCP) provided partial financial support for my research through an Overseas Research Students Award. Special thanks are due to my supervisor Michael McCarthy for his unending encouragement and guidance, and particularly for the freedom with which I was allowed to pursue my ideas. Finally, to Diane, without whom I would have folded up long ago.

Norbert Schmitt

Nottingham, April 1997

CHAPTER 1 INTRODUCTION

People have been interested in improving their mother-tongue vocabulary since the earliest orators in ancient Greece and before. In this century, the growing number of people attempting to learn a second language has focused attention on the problems of mastering L2 vocabulary as well. The vocabulary research done in the last 100 years has run the gamut from early association studies (eg. Galton, 1879-1880; Kent and Rosanoff, 1910) to statistical analyses of English lexis (such as West, 1953) to more recent studies into mnemonics (eg. Pressley, Levin, and Miller, 1982). Vocabulary research throws up many interesting questions and problems, not the least being the definition of the term *word* (Carter, 1987, Chapter 1). If the issue is vocabulary acquisition, the question 'What does it mean to know a word?' becomes critical. Read (in press) suggests that there are two main approaches to addressing this question. The first involves a *developmental* approach in which scales are used in an attempt to quantify the degree of word mastery. The second is a *component* approach, in which the various types of knowledge which make up the total knowledge of a word are described. Let us look at each of these approaches in turn.

Developmental Scales

Various vocabulary studies have utilized a variety of scales ranging from fairly simple to rather more complex. A commonly cited one, designed for L1 students, is the four-stage scale devised by Dale (1965, p. 898):

Stage 1: I never saw it before

Stage 2: I've heard of it, but I don't know what it means

Stage 3: I recognize it in context - it has something to do with

____

Stage 4: I know it

('it' refers to the target word)

Dale also mentions knowledge which could be a fifth level - being able to distinguish the word from others which are closely related.

Another L1 scale is that by Drum (1983, also Drum and Konopak, 1987). In this scale, learners are asked to give definitions of the target words and then those definitions are placed into one of the following categories. The assumption is that the categories represent an increasingly complex understanding of the word.

Stage A: Perceptual - physically similar words

1. substituting a look- or sound-alike

   "horse" for *house*; "gorilla" for *guerrilla*

2. defining a look- or sound-alike

   "buddy" for *pall* (pal)

Stage B: Syntactic - internal structure or grammatical function of words

3. defining a morpheme

   impropriety as "not a proprietor"

4. using the word in a phrase or sentence

   "a person's *reflection*"; "follow a *schedule*"

Stage C: Semantic - general meaning dimensions of a word

5. giving a general semantic attribute

serendipity as "a feeling"

6. giving a more precise attribute

aberration as "bending the rules"

Stage D: Correct - a specific definition

7. giving a part of a correct meaning

icon as "a holy picture"

8. giving a complete correct meaning

finesse as "elegance or smoothness of manner"

(Drum and Konopak, 1987, p. 79-80)

Schmitt and Meara (1997) used a scale which was designed to be suitable for the more conservative Japanese judgements of L2 English word knowledge. The English translation of the scale is:

Stage 1: I don't know this word

Stage 2: I think I might have some sense of what this word means

Stage 3: I think I know this word's meaning, but I am not sure

Stage 4: I know this word

Meara (personal communication) often uses a very similar scale in his L2 research:

Stage 1: I don't know this word

Stage 2: I'm not sure if I know this word

Stage 3: I think I know this word

Stage 4: I'm sure I know this word

The scales used by Schmitt & Meara and by Meara are comparable to Dale's, but a third scale, designed for L2 reading research in Canada, is built along slightly different lines. The *Vocabulary Knowledge Scale* [VKS] (Paribakht and Wesche, 1993) combines self-report with a productive demonstration of vocabulary at the higher stages.

Stage 1: The word is not familiar at all

Stage 2: The word is familiar but the meaning is not known

Stage 3: A correct synonym or translation is given

Stage 4: The word is used with semantic appropriateness in a sentence

Stage 5: The word is used with semantic appropriateness and grammatical accuracy in a sentence

Perhaps the main advantage of scales is that they promote an incremental notion of vocabulary acquisition, rather than a dichotomous knows/doesn't know view. But there are serious problems as well. Let us take the VKS as an example and examine it in more detail in order to illustrate what these might be. The first difficulty is that scales attempt to measure stages of knowledge in vocabulary acquisition. This can be rather problematic since vocabulary knowledge is likely to be learned incrementally on some form of continuum. In order to have useable stages, the stage boundaries must first be defined. This leads to the question of whether any naturally occurring discrete stages exist. Without a solid theoretical foundation or much empirical evidence on which to base the description of such stages, any current stage boundaries are likely to be somewhat subjective. Scales also tend to suffer from uneven intervals between the scale categories, with many of the gaps being too large between the categories. There needs to be more research to better define the stages of vocabulary acquisition before we can develop a scale we can confidently use in vocabulary research.

Notice how receptive knowledge gets rather short shrift in the VKS scale. Level 2 does tap orthographical receptive knowledge, but after this, if a learner cannot use the word productively, then the system assumes the meaning is not known. There is also a big jump from Level 2 where no meaning is assumed to Level 3, where the student is able to produce a synonym or translation. Learners who can recognize and use a word receptively will be shortchanged by this rating system. Perhaps the scale could be widened to include more categories, some of which address receptive knowledge. However, there must always be a compromise between accuracy and practicality. It seems possible that a scale with many more than five categories would be too daunting and confusing for learners and so researchers could run into the problem of diminishing returns.

A third possible problem is that the VKS may favor advanced students since they presumably would be better able to complete a sentence to illustrate their vocabulary knowledge. Requiring learners to write sentences does give information about their knowledge of a word's grammatical properties, its meaning, and perhaps also something of their collocative and associative knowledge, but we must be aware that the VKS is measuring more than just vocabulary knowledge. It also measures syntactical knowledge, and to a certain degree, the learner's writing ability. Beginning learners may have a reasonably good grasp of a basic word, but might not be linguistically advanced enough to be able to prove it by producing a correct sentence.

Another weakness of this scaling system (and most other forms of vocabulary measurement) is that it does not measure how fluent the learner is with a word. Just because a learner can write a synonym or a sentence on a test does not mean they can use the word in a conversation, or even write it correctly on a test if they were under time pressure. Paribakht and Wesche do not mention how much time learners should be given to complete the VKS, but the default may be 'all the time they need'. This would be especially true if it is considered a 'power' test to examine all the vocabulary knowledge learners have for the target words. Some sort of 'speed' element would need to be injected into the VKS in order to get an indication of the automaticity with which the words could be used.

The reader will notice that, at least at the lower levels, the accuracy of the measurement relies on learners' self-evaluation. Unfortunately, these judgements may be less than precise. Schmitt and Meara (1997) found that Japanese subjects who rated verbs as unknown could usually at least attach inflectional suffixes to them, while students who rated verbs as known showed an inability to produce native-like associations for them. Thus learners may have only a very general idea of how well they know a word. So it would seem that every scale utilizing self-assessment data needs some form of verification to see if the learner assessments are accurate. The VKS uses produced synonyms, translations, and sentences to do this, but only for the higher levels. This, however, introduces an element of subjectivity into the 'grading' of the scale. Who or what is the authority which decides if the evidence produced is sufficient? In the likely event that it is teachers or others proficient in the target language, what are the judgement criteria? This is an example of the problem of scoring such a scale objectively. It is made even more difficult because different learners approach scales with different ideas of what it means to know a word.

Word Knowledge

The above discussion shows that, although scales do have promise in measuring depth of vocabulary knowledge, their use is still currently problematic. The second approach mentioned by Read (in press) involves attempting to measure total knowledge of a word by measuring its component types of knowledge. The construct of overall vocabulary knowledge may well be too complex to capture in any single measure. The advantage of dealing with the component types of word knowledge (hereafter *word knowledge*) is that less complex, more manageable units can be manipulated, while still acknowledging the diversity of knowledge necessary to master a word.

An early discussion of behaviors involved in understanding a word was given by Cronbach (1942). In the context of L1 vocabulary testing, he mentions five types: *generalization* (defining a word), *application* (using it appropriately), *breadth of meaning* (knowing its different meanings), *precision of meaning* (being able to use it ~~in~~ correctly in different situations), and *availability* (productive use). However, the current discussion on word knowledge can be traced back to a seminal article by Jack Richards in 1976. He was the first to make explicit the idea that there are several different kinds of word knowledge necessary for the mastery of a word. He presents them as eight assumptions concerning the nature of lexical competence:

1. Native-speakers continue to increase their vocabulary into adulthood, but

their knowledge of syntax is nearly complete by puberty.

2. Native-speakers know how frequently a word usually appears in speech or print. They also know the frequent collocations for many words.

3. Native-speakers understand the register constraints imposed upon a word by variations in situation and function, including temporal variation, geographical variation, social variation, social role, field of discourse, and mode of discourse.

4. Native-speakers know the syntactic behavior of a word.

5. Native-speakers know the underlying root of a word and the derivations that can be made from it.

6. Native-speakers have knowledge about the network of associations between a certain word and the others in a language.

7. Native-speakers know the semantic value of a word.

8. Native-speakers will know many of the meanings of a polysemous word.

Meara (discussion paper) suggests that Richards was attempting to give an account of contemporary research and its implications for vocabulary teaching rather than formulate a systematic account of word knowledge. This can explain some of the gaps, such as knowledge of the form aspects of word. But intentionally or not, Richards planted the seed for later thought in this area. Some later authors discussed word knowledge indirectly in order to make related points about language, while others explicitly tried to give a comprehensive listing of word knowledge types. In either case, it is natural that they built upon Richards' lead. Alexander (1982), in discussing the relationships between individual words in a lexicon, went a fair way towards specifying word knowledge:

1. Phonological links
2. Morphological links
3. Syntactic links
4. Paradigmatic sense relationships
5. Collocational patterns (syntagmatic relationships)
6. Style and register factors
7. Knowledge of fixed expressions and idiomatic phrases
8. Connotative meaning
9. Allusional meaning

Blum-Kulka (1981) suggested that learning a word entails mastery of four aspects:

1. Semantic mapping (linking words and their referents)
2. Morpho-semantic restrictions (accurate application of morphology and grammar in the use of vocabulary)
3. Collocational restrictions (collocational appropriacy)
4. Communicative functions (stylistic appropriacy)

Reviewing previous vocabulary research, Laufer (1990, in press) isolated a number of 'intralexical factors' which affect the learning of words: phonological factors, grammatical characteristics, semantic features, register restrictions, and multiple meanings. She finds that these factors can facilitate vocabulary learning, inhibit vocabulary learning, or have no consistent effect, depending on similarities or differences of the second language vocabulary to the first language vocabulary and how regular (not being exceptions) the target words are. Table 1 summarizes the intralexical factors. As with Alexander

Table 1 Intralexical factors which affect vocabulary learning

Facilitating factors	Difficulty-inducing factors	Factors with no clear effect
familiar phonemes	presence of foreign phonemes	
phonotactic regularity	phonotactic irregularity	
fixed stress	variable stress and vowel change	
consistency of sound-script relationship	incongruency in sound-script relationship	
		word length
inflexional regularity	inflexional complexity	
derivational regularity	derivational complexity	
morphological transparency	deceptive morphological transparency	
	synformy	
		part of speech
		concreteness/abstractness
generality	specificity	
register neutrality	register restrictions	
	idiomaticity	
one form for one meaning	one form with several meanings	

(Laufer, in press)

above, although Laufer focuses on learning burden rather than word knowledge specifically, her discussion is valuable in that it indirectly points out some of the types of word knowledge that are necessary for knowing a word.

Laufer's paper particularly hints at the complexity of placing the various kinds of word knowledge into neat categories. Take meaning for instance. She shows that it makes a difference in learning whether a word has a single meaning or several. Although it is not obvious that learning the first of several polysemous meanings is any more difficult than learning the single meaning of a monosemous word, learning several polysemous meanings and the proper usage of each must clearly involve more learning effort. Even when working with a single semantic meaning, the factors of abstractness, specificity, and idiomaticity tend to cloud the picture. Taken together, this shows how complex the idea of *knowing a meaning* is (is the meaning known productively or only receptively, how quickly and automatically can it be used, can it be used with the proper stylistic and collocational restraints, is it only one of several polysemous meanings, etc.). Other word knowledge categories surely hide a similar variety of complexities.

This brings us to perhaps the most complete and explicit description of word knowledge to date. In Chapter 3 of his wide-ranging book, Nation (1990) looks at what is involved in knowing a word. His list nicely captures the key elements presented in the above lists, especially if one considers morphological knowledge as part of grammatical knowledge.

1. The *spoken form* of a word
2. The *written form* of a word
3. The *grammatical behavior* of the word
4. The *collocational behavior* of the word
5. How *frequent* the word is
6. The *stylistic register* constraints of a word
7. The *conceptual meaning* of a word
8. The *associations* a word has with other related words

Since Nation's listing of word knowledge is the best and most elegant to date, it will be used as the basis for the rest of this thesis. Therefore it is worth briefly reviewing what he says about each type in greater detail. In the chapter he mainly explains how these word knowledge aspects correspond to learning a word, with particular emphasis on the effect of the mother tongue on L2 vocabulary acquisition. Reviewing the literature, he concludes that "The more predictable and regular the features of a word, the lighter the learning burden" (p. 35).

Form: English words which use the same sounds and sound arrangements as the learner's L1 will present little problem in pronunciation. Likewise, if the scripts of the two languages are the same, learning will be easier. Nation suggests it makes sense to teach words with spellings that 'follow the rules' in an attempt to teach the spelling regularities before dealing with exceptions.

Grammatical patterns: Similar to the behavior of word form above, grammatical patterns which mirror L1 patterns will be easier to use. Patterns congruent with the normal patterns for English will also be easier to learn.

Collocation: If a word's collocations can be guessed from the L1 translation equivalent, the meaning of a word, or the form of a word, then learning the collocations will be easier.

Frequency: Learners can get clues about a word's frequency from how often it appears in classroom English lessons, the frequency of its translation equivalent, and from its form. Since the most frequent words are monosyllabic, learners choosing between a long and short word should usually choose the shorter.

Stylistic Register or *Appropriateness*: Since many words have stylistic constraints, teachers should make students aware of these constraints when teaching words which might be misused.

Meaning: Learning will be easier if the meaning of a word can be predicted from its form or from the meaning of the mother tongue word, and when the word's various meanings relate to the same underlying concept.

Associations: Learners have L1 associations but there is little research to say whether they are carried over and formed into L2 associations. Teaching words together which are closely associated is unwise, as it may cause 'cross-association', in which the learner becomes confused about which meaning goes with which L2 word.

The Word Knowledge Framework: Potential Uses

Nation has provided us with a list of word knowledge types. This leads to the bottom-line question, "So what? What can we do with it?" As Schmitt and Meara (1997) conclude, listings like Nation's are purely descriptive and have no explanatory power. None of the above authors have tried to fit their listings into any kind of theory, framework, or order. Perhaps this is not surprising given the current state of knowledge about vocabulary acquisition. But Schmitt and Meara also suggest that such descriptive summaries can be used as frameworks for research which can be explanatory. Using a word knowledge framework as a basis for research is one possible use for a word knowledge framework. Two other possibilities are suggested by Schmitt (1995a). One which will be explored in this thesis is as a framework to examine what vocabulary tests are actually measuring. The other is as a framework to examine what aspects various vocabulary learning activities address. Let us examine the applications of research/theory building and testing each in turn.

Research and Theory-Building

As both Meara (discussion paper) and Read (in press) point out, attempting to measure every kind of word knowledge for a word is a daunting task. It is likely to be very time-consuming and so a word knowledge approach must be wildly impractical for everyday applications like testing or teaching activities, especially if large numbers of words need to be addressed. Even for a very small number of words, the effort required would be considerable. However, for research purposes, where a great deal of preparation and effort is the norm in pursuit of new insights, this unwieldiness need not disqualify the approach.

At the moment there is no satisfying theory of vocabulary acquisition, and part of the reason for this must be because the process is incredibly complex (or at least so it seems to the unknowing - us). Of course other research paradigms must continue to be explored as well, but if the word knowledge framework can be shown to be informative, there is no reason why it should not be pursued. One advantage already mentioned is the simplifying effect of dealing with more manageable components rather than a sometimes impenetrable whole. If it can be discovered how each individual word knowledge is acquired, that must take the field closer to an understanding of how words are acquired globally. Learning more about how the various types of word knowledge are interrelated would surely also be useful. Intuition strongly suggests that such relationships exist (eg. between frequency of occurrence and formality register; between word class and derivational suffixes) and Schmitt and Meara (1997) have recently demonstrated some of these interrelationships correlationally. In addition, better awareness of word knowledge may help explain the movement of vocabulary from receptive to productive control. Where it has been normally assumed that a word is either receptively or receptively and productively known, the actual situation is likely to be that each of the different types of word knowledge is known to different receptive and productive degrees. Research into how the underlying word knowledge states of receptivity/productivity affect the overall ability to use words in a receptive vs. productive manner could prove to be quite exciting indeed.

Another intriguing possibility is that the types of word knowledge are hierarchical, that is, learned in some type of developmental order. If this could be demonstrated, it would be a breakthrough in the way we understand vocabulary acquisition. It would also have the effect of instantly transporting the word knowledge framework into practical applications. Because the types of word knowledge would fall into an implicational order, only a small number would have to be measured in order to get the larger picture. It would seem improbable that word knowledge is not at least partially hierarchical. One might suppose that all that is usually picked up about a word on the initial exposure is some basic kernel meaning and perhaps some orthographical or phonological impression of the word's form. After more exposures (or giving conscious attention to the word), a learner would gradually learn the other kinds of word knowledge, with perhaps collocational and stylistic knowledge being the last to be mastered. It doesn't seem reasonable that a learner would have a rich associative and collocational network built up without a knowledge of the word's form. This would suggest that some kinds of word knowledge are acquired before others. It remains to be seen whether the cognitive mechanisms of the mind work in a way which enforces an order developmentally, or whether the acquisition order is more probabilistic, depending more on the type of word, learner, and learning context.

This section has argued that the word knowledge framework is worth pursuing because it may well prove a productive avenue of research. This should not leave the reader with the impression that it is an end in itself however. The author agrees with Meara (discussion paper) that a desirable goal for vocabulary researchers is to eventually arrive at explanations and measurement procedures which deal with the lexicon as a whole, rather than at the level of individual words. I would argue that the word knowledge framework may enable research which will bring this final goal closer. As such, the word knowledge framework should be considered transitional, because once we better understand the lexicon, it will no longer be needed.

Until that day arrives, perhaps the best initial research using the word knowledge framework would be a longitudinal study which follows the acquisition of individual words over time to see how each type of word knowledge develops. Such a study will be the centerpiece of this thesis. It is hoped that a great deal of useful insights will come out of it. But in addition to such insights, the viability of the word knowledge framework will also be tested in that study, for if such a study proves to be uninformative, then the framework itself may be incapable of producing any substantial results.

Testing

Schmitt (1995a) states that the word knowledge framework can be useful in analyzing vocabulary tests to gain better insights into what they are actually measuring. This evaluative use should not be too controversial, but when it comes to applying the framework to construct an everyday vocabulary test, the story is different. As mentioned before, such a time-consuming test is unlikely to be practical. If however, word knowledge turns out to be hierarchical, then the situation changes, as only a limited number of word knowledge aspects would have to be measured in order to quantify the depth of knowledge. Even if word knowledge proves not to be hierarchical, the same end result might be achieved by giving subjects a battery of word knowledge tests. If one or several of the component tests correlated highly enough with the total scores from all the tests together, these few component tests could be used as a measure of overall depth of vocabulary knowledge.

Another possibility is a hybrid vocabulary test, combining breadth and depth measures. A computer program like the EVST (Meara and Jones, 1990) could be modified to do this. What would result is a checklist test where a relatively small number of words indicated as known would be selected by the program, and learners questioned on the different kinds of word knowledge. In addition to a vocabulary size estimate, this kind of program could give estimate scores for various word knowledge as well. The test could measure receptive knowledge by giving multiple choice questions and productive knowledge by having a store of information gathered from native speakers, such as associations and collocations, in its memory which the program could match against the answers the learner types in. Teachers and administrators would receive a much more comprehensive picture of their students' vocabulary knowledge if a program like this could be developed.

A final use of the word knowledge framework in testing is in validation. One way of checking to see if a vocabulary item is really measuring vocabulary knowledge is to go in after the item is answered and find out what the testee knows about the target word. A word knowledge framework can inform the interviewer about what types of knowledge to probe for and what kind of questions to ask. The framework will be used in just this way to explore what the vocabulary items found on the TOEFL test (1995) are measuring and how well.

Conclusion

This introduction has briefly explored the background to a word knowledge framework and has suggested some possible ways in which it might prove useful to the field of vocabulary studies. In this thesis, I hope to use a word knowledge framework to provide insights into some individual types of word knowledge, and to explore whether the framework is of use in researching vocabulary tests and vocabulary acquisition. While the focus will be on the information and insights coming out of the individual studies, a question always in the background will be whether the word knowledge framework is proving viable as a means of informing and underpinning such research. By the end of the course of my research, I would hope to have at least preliminary answers to the following questions about the framework itself:

Is the word knowledge framework feasible for use in vocabulary research? Can reasonable tests for the various types of word knowledge be developed? Finally, is the word knowledge framework informative in vocabulary research?

Before we launch into the studies proper, it is first important to give a more comprehensive backgrounding on word knowledge. The next chapter provides a short literature review for each of the eight types of word knowledge Nation isolated. The sections on frequency, associations, and collocations are slightly more comprehensive than the others, since these are the three kinds of word knowledge which will involve individual studies to develop new measurement procedures. In some cases, especially the meaning section, lack of space lead to glaring gaps in what could be covered, but it is still hoped that the reader will come away from the literature review with a reasonably good understanding of the type and scope of research which has been done in each of the word knowledge areas.

CHAPTER 2 LITERATURE REVIEW

INTUITIONS OF WORD FREQUENCY

Subjective Frequency Estimates of Words Occurring in Language

It is widely accepted that how often a word occurs in language affects how we use and process that word. In fact, the effects of word frequency are so pervasive in language processing, that word frequency is one of the main factors that needs to be controlled for in linguistic experiments. Moreover, there is a widespread assumption that native speakers have intuitions about how frequently individual words occur in their language. This assumption was made explicit by Richards (1976) and Nation (1990) when they included knowledge of word frequency in their lists of what must be known to have full mastery of a word. It is not difficult to understand why this assumption is held, since the relationship between frequency and some other kinds of word knowledge is obvious. Knowledge of frequency must facilitate register decisions, since, for example, words in a spoken register tend to be more frequent than words in a written register, formal words tend to be less frequent than informal words, and words are archaic simply because they have become so infrequent. Mental word associations have also been shown to relate to frequency (Howes, 1957)

A number of research studies beginning in the late 1960s moved past the assumption that native-speakers have frequency intuitions and attempted to measure how accurate those intuitions are. These studies compared the subjective estimates of frequency (SFE) from subjects with objective frequency count data (OF) for the same words and generally found quite high correlations: .74-.78 (Tryk, 1968); .92-.97 (Shapiro, 1969); .92-.97 (Carroll, 1971); .69 (Thrasher, 1973); .57 (Richards, 1974); .91-.94 (Backman, 1976); .67-.90 (Ringeling, 1984); and .64-.79 (Arnaud, 1989, cited in Arnaud, 1990)[1]. However, the relationship between SFEs and OF is not a simple linear correspondence; indeed, we would not expect that people could precisely register every time they had been exposed to a word. Rather, we might expect that the first few times one was exposed to a new word, those exposures would be relatively salient, while after numerous exposures, each individual exposure would become less distinct and important. This is exactly what the studies showed. The typical relationship between SFEs and OF forms a negatively-accelerated curve.

Researchers developed two basic ways of describing this relationship. The first was with the use of logarithms. Since a logarithmic scale becomes more and more compressed, the result is that the relationship can be expressed as a linear one on a logarithmic scale. Thus many frequency studies express their results in log terms, typically using Carroll's (1970) standardized method of expressing frequency, the Subjective Frequency Index (SFI), which uses the following formula:

$$\text{SFI} = 10(\log_{10}\text{frequency of word in corpus}+10).$$

An alternative method of expressing the relationship is with the power law:

$$\text{judged frequency} = (a)(\text{frequency of word in corpus})^m$$

where *a* depends on the scale the measurements are made in and *m* expresses how rapidly subjective frequency increases as a function of objective frequency (Carroll, 1971).

Studies have attempted to elicit SFEs in two basic ways. The first method asks subjects to give *absolute* frequency estimates of separate words, with answers like *very frequently used, seldom used, used once a week,* or *used once a year.* This method draws upon work done by Stevens (1956, 1958), who developed a way to measure estimates of sensory magnitudes (such as loudness) which are not easily quantified. Thus, the method is also called *magnitude estimation* (ME). The second method requires subjects to rank words on a list according to frequency. In this *relative* method, randomized words may need to be reordered according to frequency, or each word given a frequency figure relative to the others (Arnaud, 1990). It is also called *multiple rank orders* (MRO). In MRO tasks which require frequency figures, a starting benchmark figure can be given or not. When benchmark figures were not given, Carroll (1971) found that it did not matter if the first word (in relation to which subjects tend to rate the other words) was of high or low frequency.

Studies examining SFEs have mainly concentrated on native-speakers, and have produced a number of interesting results. The most important one has already been mentioned - that SFEs tend to correlate strongly with their corresponding objective frequency counts. Another is that subjects are able to judge not only the frequency of words in their own personal situation and context, but also the general frequency of words in society. Tryk (1968) had 50 American university students rate 100 words, taken from a logarithmic sampling of the Thorndike and Lorge word count (1944), for both their perceptions of public use of the words and for their own personal use of the words. The resulting estimates for public and private use were essentially the same when correlated to the Thorndike-Lorge OF data (.74-.78). It seems even advanced nonnative speakers can give accurate SFEs for the general use of words, although their estimates of their personal use of words do not match OF data as well. Ringeling (1984) found that five advanced Dutch English-speakers and five native English speakers (all members of staff at an English department at a Dutch university) were able to give accurate SFEs when asked to do so for general use in the language, but when asked to do so for personal use, the correlations with the objective frequency count were much lower for the Dutch subjects (.66,.66,.69,.78,.61) than the native-speakers (.87,.85,.79,.75,.67). Ringeling suggests that differing instructions regarding rating for personal use or general language occurrence could be behind the different correlation strengths found in frequency studies. Another seemingly obvious possibility presents itself. While corpus OF is probably a reasonable baseline for general language use in society, it may not accurately portray the way any *individual* uses the language. It is thus possible that the subjects were better at accurately estimating their frequency of use than the OF data.

Native-speaker SFEs seem to be reliable as well. Tryk's subjects were retested three weeks later and the test-retest correlations were very high (.96 and .98). Arnaud (1989), cited in Arnaud (1990), examined the SFEs of 51 French university students. They were asked to rank two lists of 30 French words. The students were tested five weeks later on one list and the test-retest reliability was .80. This shows that SFEs can be accurate and reliable for languages other than English. In addition, although he found large individual differences in SFE performance, he also found an interesting pattern where students who provided the most accurate SFEs also had the most stable intuitions, as shown by higher individual test-retest correlations.

Native-speakers as a group tend to give similar SFEs. Thasher (1973), cited in Upshur (1975),[2] found that correlations of SFEs were high among 5 third-graders (.90) and among 5 educated adults (.88), indicating that native-speakers tend to give SFEs which are consistent between individuals (although note the small number of subjects). Similarly, Carroll (1971) found correlation figures of between .97 and .99 for native speakers. Shapiro (1969) studied sixth-graders, ninth-graders, college sophomores, industrial chemists, elementary school teachers, and newspaper reporters and found no difference in their SFEs.

In contrast with these results, the SFEs of nonnative-speakers are less consistent among respondents. Thasher compared the above native speakers to four groups of English learners consisting of: 1) 4 advanced Japanese 2) 4 beginning Japanese 3) 4 mixed-nationality advanced students and 4) 4 mixed-nationality beginning students. They all judged 60 high frequency verbs and gave SFE judgements gathered from a ME task. The nonnative-speakers only had intercorrelations of .40 to .64.

When Thasher (1973) compared the SFEs with the frequency data in *The Word Frequency Book*, the correlations were .42 (L1 third-graders), .69 (L1 adults), and .63-.75 (L2 learners). In addition, the advanced learners produced SFEs which were closer to those of the native adults than those of the beginning students. These results suggest that intuitions about frequency develop with our language ability. That the SFEs of the L2 subjects improved with their language ability is not surprising, but the fact that the L1 children had such a low correlation figure indicates that frequency intuitions are also developed over time in an L1. Note the discrepancy with Shapiro's (1969) finding of no difference between L1 sixth-graders, ninth-graders, and adult SFEs. Perhaps it is in the period between the third and sixth grade with exposure to more reading material that L1 students consolidate their intuitions about frequency. The nonnative subjects in Ringeling's (1984) study were very advanced speakers of English, and their high proficiency allowed them to give accurate SFEs of words in general language. So nonnative weaknesses in frequency judgements are likely to be caused by a general lack of L2 language proficiency, and can possibly improve to a native-speaker level. Carroll (1971) showed metacognitive knowledge about language can also sharpen frequency intuitions. He asked 15 lexicographers and 13 people with no specialist linguistic knowledge to give absolute SFEs for 60 words. As expected, the lexicographers produced better correlations with the objective word count data (Thorndike & Lorge, Kucera & Francis) (.97) than the nonspecialists (.92), although the correlation figure for the nonspecialist group shows just how accurate the SFEs of average people can be. On balance, it seems that SFEs improve with increased proficiency in a language.

A study by Richards (1974) indicates that the average SFE is more closely related to written language than spoken language. He extracted 4495 concrete nouns from *The Advanced Learner's Dictionary* and *The New Merriam-Webster Pocket Dictionary*, which he then had 1000 Canadian college students judge in a ME task. The student results correlated poorly with two oral word counts: *A Study of the Oral Vocabulary of Adults* (Schonell, Meddleton, and

Shaw, 1956) at .37, and *A Word Count of Spoken English* (Howes, 1966) at .39. The student results correlated much better (.57) with the main written word count used, the *Computational Analysis of Present Day American English* (Kučera & Francis, 1967), although this figure is lower than those arrived at in other frequency studies. Earlier, Shapiro (1969) found that the spoken/written distinction made no difference when subjects tried to distinguish between them; instructions to rate in terms of spoken language produced results no different from instructions to rate in terms of written language. To put it another way, in Shapiro's study, subjects were unable or disinclined to differentiate between spoken and written frequencies. Yet there is a difference between frequency counts based on spoken and written data. Given this difference, it is not clear why the SFEs should have corresponded better with written than spoken counts in Richard's study. Possible explanations are a) that subjects gain frequency intuitions mainly from written sources, b) that they reply to frequency tasks on a basis of what they believe the written frequency to be, c) that the tests themselves were in the written mode, or d) the objective frequency lists themselves were faulty. In sum, there is evidence that people do not differentiate between spoken and written frequencies, but that, in fact, their intuitions correspond more closely with written frequency.

So far, we have discussed native-speaker intuitions about the single words only, but there is some evidence that people have accurate frequency intuitions for multiword units as well. Backman (1978) had 15 Swedish university students rank the frequency of 18 different word combinations, like *kanske är det* (it may be) and *i stort sätt* (on the whole) against a benchmark phrase (the Swedish equivalent of "at first hand"). The correlation between the SFEs and objective data was .56. He concluded that the collocational recurrences also have parallel cognitive counterparts, in other words, it seems people do have some sense of the frequency of phrases in language.

There have been far fewer studies that have looked at the SFE ability of nonnative-speakers. In addition to Thrasher (1973) and Ringeling (1884) mentioned above, Arnaud (1990) studied 126 French first-year and 87 American second-year university students. He extracted 30 French nouns and adjectives from the Juilland, Brodin, and Davidovitch (1970) list and 30 nouns and adjectives from the *Word Frequency Book* (Carroll, Davies, and Richman, 1971). On the English list, the native-speakers outperformed the nonnative-speakers, but on the French list, the American and the French results were not reliably different statistically. Arnaud believes the American students were of a generally higher caliber, echoing the idea that language proficiency can affect the accuracy of SFEs. In his discussion, Arnaud proposes an interesting explanation of *how* SFEs are arrived at. He suggests that, in addition to directly rating the L2 words, subjects may also consider the frequency of the L1 translation equivalents in their mother tongue. To show that this could be possible, he compared the frequency rankings of the translation equivalents in French and English. They correlated at .84 to .89. This is in harmony with the figure (.84) reported by Kirsner *et al.* (1984). So subjects relying on their L1 equivalents alone could get reasonable SFE scores in their L2. Considering the large degree of influence of the L1 in acquiring an L2, this suggestion seems reasonable, although this strategy may only be effective when the two languages involved are similar.

From the above studies, a number of factors are seen as affecting SFEs. By far the most important one is how frequently a word occurs in language, as indicated by objective word counts. This includes multi-word units as well as single lexical items. Age may or may not make a difference, but if it does, it is probably tied to language proficiency. But even the effects of language proficiency is unclear, as Shapiro found no difference between elementary school students and adults. SFEs seem to match written word counts more closely than spoken word counts, but people do not seem to be able to explicitly indicate the difference between written vs. spoken frequency. Native and nonnative-speakers are able to judge reasonably accurately the frequency of words in general language, but judgements of their personal use are more difficult to assess. Speakers of an L2 have frequency intuitions about L2 words, but unless they are very advanced, the intuitions are not as accurate as native-speakers. There may be other factors as well; Backaman (1976) found that an ANOVA analysis did not separate subjective frequency from other the attributes he studied: familiarity, pronounceability, and comprehensibility.

In sum, the literature indicates that native, and to a lesser extent, nonnative-speakers, *as a group*, to able to give reasonably accurate and reliable subjective estimates of frequency for words in general societal use.

Subjective Frequency Estimates of Words Occurring in Experimental Situations

In contrast to the above studies, which attempted to discover how well subjects could judge the frequency of words in the 'real-life' world, quite a number of studies have followed a different research direction. In this paradigm, researchers attempt to quantify SFEs not from natural exposure, but for words presented in limited lists in a controlled environment. There is the question of how much the results derived from such experiments can be generalized to natural language use and intuitions, but they provide a body of evidence which should not be ignored. Discovering the factors which affect the SFEs of limited-exposure input may inform us of what happens in language at large.

The usual situational frequency study presents one or more long lists of words to subjects, in which a number of target words are embedded. These target words are repeated a varying number of times in the lists, typically at lower repetition rates (0[3],1,2,3,4) and sometimes also at higher rates (6,10). The target words are separated by a varying number of filler words chosen from the same criteria as the target words (typically 0,1,2,4,8, and 16 intervening items). Thus the subjects do not know the target words from the filler words. After exposure to these lists, the subjects are tested on their SFEs, recall, or recognition of words on the lists.

Gregg, Montgomery, and Castaño (1980) summarize some of the most commonly found effects, "Free recall of lists containing only high-frequency words ... is generally superior to recall of lists composed of only low-frequency words... In contrast to the results obtained with pure lists (containing only high- *or* low-frequency words), low-frequency words are more likely to be recalled than high-frequency words from mixed lists (containing equal numbers of high- and low-frequency words)" (p. 240). Another recurring finding is that "subjects tend to overestimate items presented relatively few times and underestimate words presented relatively often in the list" (Rose and Rowe, 1976, p. 142).

One of the most interesting things to come out of situational frequency experiments is a list of factors which do and do not affect frequency judgements. Let us first look at factors with little or no effect. Hasher and Chromiak (1977) found that age did not affect situational frequency judgements. Children in the primary school (second, fourth, and sixth grades) gave frequency estimations which were just as accurate as those of college students. Thus, the ability to give accurate situational SFEs is either not a skill which develops over time, or is already developed by the second grade.

It does not seem to matter whether the method of exposure is visual or auditory. In two experiments, Hintzman, Block, and Summers (1973) asked 47 and 191 American adults to judge the frequency of occurrence of 45 three-letter high-frequency nouns. The words were presented either visually or orally (0, 1, or 2 times in the first experiment; 0, 1, 5, or 15 times in the second), but the level of judged frequency was not affected much by the modality of the exposure.

Evidently, spatial tags do not reinforce frequency information either. Howell (1973a) had 50 college students either sort word cards into category columns or pronounce the word and put all cards into one slot. Although the subjects could remember where they had placed the cards (spatial recall), Howell did not find any differences in frequency judgements for items presented in the same or varied spatial locations.

One would expect that orthographical form may influence the recognition of words, which in turn might facilitate accurate estimates of how often a word was seen in situational frequency experiments. Hintzman and Summers (1973) showed that when a word appearing as a test prompt had an identical orthographical form to that in which it was presented (same size and type font), subjects were able to recognize that word more quickly. Likewise, Kirsner (1973) found a small performance improvement in the same-print condition. In addition, her subjects could reliably report the orthographical form of words they could recall for at least 90 seconds after presentation. When the target words were scrambled into nonsense letter strings, the same-print advantage was increased.

In spite of this evidence showing the importance of orthographical form and visual memory, orthographical variation had no effect upon the frequency judgements of 28 Canadian undergraduates in an experiment by Rowe (1974). Considering this non-effect together with the lack of effect of modality of presentation (Hintzman, Block, and Summers, 1973) and spatial location (Howell, 1973a), Rowe hypothesized that nonsemantic attributes in general do not affect situational frequency estimates. To study this, he set a follow-on experiment which studied semantic vs. nonsemantic variables. 12 subjects watched slides of words and were told to write down the number of consonants and syllables for each word (nonsemantic manipulation), while the other 12 were asked to make a connotative strength rating (semantic manipulation) on a 7 point scale (*horse* had high connotations for strength, while *whisper* had weak connotations). He found that the semantic encoders gave much higher estimates of situation frequency than the form encoders. They were more accurate for words which were presented 4 and 6 times, but less accurate for words which were presented 1 and 2 times. Using a *depth* *of processing* framework, he suggests that 'deeper' semantic encoding is the key factor in developing frequency estimates.

Two years later, Rose and Rowe (1976) reinforced this conclusion. They conducted a study with 90 frequent words appearing 2,3 or 5 times under spacing of 0,1,2,4,8 and 16 items judged by 144 university undergraduates split into 3 groups. Each group received different instructions. The first group were told a frequency task would follow (intentional), the second group were told a nonspecific memory task would follow (nonspecific), and the third group were told to rate each word on its connotative strength with no mention of the following test (incidental). The different instructions had an effect on the frequency estimates and the effects of item spacing: the incidental condition lead to the highest estimates, intentional next, and non-specific last (although non-specific was most accurate for 2 and 3 exposures). These results led the experimenters to do a second study investigating incidental instructions. Similar to the 1974 study, they found that tasks which required semantic or acoustic processing enhanced both frequency and spacing effects compared to a nonsemantic task (counting vowels), but mainly at more frequent levels of exposure. It seems that deeper processing does contribute to the formation of accurate frequency intuitions (at more frequent exposures), at least in situational conditions.

Other factors besides semantic encoding affect frequency judgments. The spacing between the repetition of target words seems to affect the resulting SFEs. Hintzman (1969) found that subjects (118 and 163 in two experiments) judged words as more frequent as the spacing between the two presentations of the same word in a list increased from 0 through 1, 2, and 4 repetitions.

The SFEs at spacing levels of 8 and 16 were nearly the same as at 4 intervening items. Hintzman, Block, and Summers (1973) found a similar pattern: SFEs rose progressively through spacings of 0, 1, and 5 intervening items, but then fell slightly at a spacing of 15 items.

There is mixed evidence as to whether the duration of exposure of the target words affects frequency judgments. Underwood (1972) showed that *recognition errors* decreased by 2% for each additional second of exposure duration (2, 4, or 6 seconds). Since he also found that frequency judgements are tightly related to recognition judgments, he suggested that exposure duration also affects frequency judgements. However, Rowe (1973a, 1973b) found that frequency judgments and recognition of target words are affected differently by the type of homonym prompt, calling this close relationship into doubt. Hintzman (1970) directly tested the effects of duration of exposure on situational SFEs. He exposed words for 2, 3, 4, 5, and 6 seconds and then had subjects give SFEs using absolute rating tasks. He found that the exposure duration did not affect the subjective frequency judgements.

It is clearer that the type of homonym presented makes a difference. Rowe (1973a) had 80 Canadian university students judge homonyms they saw in four conditions: repeated single-word homonym, identical repeated phrase which included the homonym, different phrases which included the same homonym meaning (chocolate bar, candy bar), and different phrases which included different homonym meanings (sand bar, bar service). The equally good best conditions were repeated homonym word and identical repeated phrase. They were better than different phrases which presented the same-meaning homonyms, which was better than phrases containing homonyms of different meanings. In a near-replication of the above study, Rowe (1973b) found the same results: for exposure frequencies of 3 to 5 (but not 1 and 2), the most accurate SFEs stemmed from exact repeated phrases, phrases containing homonyms with the same meaning, and phrases containing homonyms with different meanings, in that order. Interestingly, Rowe reports that the three different presentation conditions had no systematic effect on the subjects' recognition memory. In contrast to the studies showing no effect of different orthographies of single words, when it comes to phrases, it seems that repeating similar form and meaning can improve frequency performance.

Many studies point to the conclusion that frequency intuitions are implicitly and automatically acquired. Evidence for this comes from studies in which subjects are either made aware of the frequency purpose of the study or not. Howell (1973b) found that estimations of how often words were presented in lists (0, 1, 2, 4, 6, or 10 times) did not depend on whether the subjects (192 college students) were warned beforehand that they would have to make frequency estimations or whether they had to simply recall the words. On the other hand, instructions to remember the words definitely aided recall of the words. Subjects were able to estimate frequency to a reasonable degree, although they exhibited the common tendency to underestimate the higher frequencies of exposure (6 and 10 exposures) and overestimate lower frequencies (1, 2, and 4). Hasher and Chromiak (1977) found that giving specific instructions beforehand to focus on frequency did not improve the accuracy of the SFE scores for lower exposure frequencies, although it did for higher exposure frequencies.

The frequency instruction/noninstruction effect seems to apply for nonwords as well as words. Flexser and Bower (1975) tested 20 Americans on 117 nonsense syllables (CVC) which occurred in a list from 1 to 6 times. There was no difference in the SFEs between the subjects who were given specific frequency instructions and those who were given general instructions.

Other evidence points to automatic acquisition of frequency intuitions as well. The SFEs produced in Howell's (1973b) study were not affected by the length of the word list the subjects were exposed to (from 10 to 80 words), showing that words in longer lists can be judged as well as words in shorter lists. Since it would be much harder to consciously process the longer list (with a strategy like counting), the alternative, implicit learning, is more likely. Hasher and Chromiak (1977) found that second-, fourth-, and sixth-grade primary school students had essentially the same frequency judgements as university students. Also, when university students were given feedback on their frequency estimates, it failed to facilitate improvements in their scores. Information does not have to be particularly meaningful to create frequency impressions. All this suggests that people gain some representation of word frequency in situational conditions regardless of a) whether their attention is focused on frequency or not, b) the level of memory load as determined by the number of total words exposed to, c) age, and d) feedback. In other words, frequency counting or tagging seems to be "an automatic or essential aspect of information consumption at least as early as grade school" (Hasher and Chromiak, 1977, p. 182). To add evidence from the naturalistic side, Underwood (1971) found that people have a sense of how frequently bigrams (two-consonant clusters like *GP* and *GB*) occur in words, even though they do not carry meaning in themselves and subjects would certainly not have consciously noticed their frequency of occurrence in everyday situations.

To summarize the above findings, in situational contexts people have reasonably accurate SFE intuitions regardless of age. The frequency intuitions seem to stem mostly from semantic processing activity and are more or less automatically acquired. From this, a logical next question would be *how* are frequency intuitions formed? Several studies have tried to build towards an answer to this question. In his 1973(c) survey of possible theoretical explanations of situational frequency results, Howell concluded that a multiple-trace hypothesis was most likely. This means that each time one is exposed to a word, it leaves a separate trace or representation. These traces coexist and the magnitude of frequency judgement relies on the *number* of traces. There is also a possibility that the strength of traces could play a part, but at the time no stronger conclusions were possible.

Evidence for this conclusion comes from Hintzman and Block (1971), who found that subjects could remember the two different positions a word was presented in on a single list and that subjects retained separate frequency information about a word presented at different frequencies in two different successive lists. Thus the new traces must be established rather than a single one being strengthened. Also, Jacoby (1972) found that frequency judgements of a sentence were independent of the frequency of the component words of that sentence. This means that the occurrence of a word in a sentence did not increases the SFE for the individual word. Moreover, presenting synonyms of a word did not affect the frequency of that word.

Tversky and Kahneman (1973) suggest that frequency estimates come from judgements of the availability of the item, in other words, the ease in which it can be retrieved, while Begg (1974) found that subjects could maintain a fairly stable impression of the average frequency of words in a list. Rose and Rowe (1976) combine these two ideas to derive a two-stage estimation procedure: a person makes an initial judgement of frequency based on how easy it is to recall the word, and this rough estimate is refined by reference to the average frequency of all the words in the context.

It is possible for the automatic frequency acquisition process to be overridden by conscious strategy use, however. Begg and Rowe (1972) found almost perfect correlations between running estimates of frequency (where ratings were asked for after each word in a list) and actual presentation frequency. Begg (1974) did a follow-up study which showed this effect was a result of the online frequency estimation, and not the time lag between the exposure to words on the list and the post-list test. Flexser and Bower (1975) review these results and suggest that subjects doing running frequency judgements are adopting a counting strategy, which could account for the extraordinary correlations. If so, people may derive their frequency information from a mental count if the task is small enough and the importance of retaining frequency information seems great enough.

A final question is what practical use can be made of SFEs? It has been suggested they could be used as a possible measure of bilingualism. Cooper and Greenfield (1969) found that the absolute SFEs given by Spanish-English bilinguals correlated at the magnitude of .37 to .69 with other more usual measures of bilingualism: self-ratings of which language they used more at home, self ratings of their speaking facility for the two languages, word fluency as measured by a 60-second word naming task, an analysis of their accentedness, and a rating of how well they used various English speech styles. On the other hand, Upshur (1975), in his examination of whether SFEs could be used as an indirect test of overall language proficiency, came up with mostly nonsignificant results, indicating that SFEs, at least as elicited by his form of test, could not be used for testing purposes. A safe conclusion seems to be that we need to know more about SFEs before we are able to incorporate them into any worthwhile test.

WORD ASSOCIATIONS

Early History

The use of associations in the description and study of language has a very long history, with some scholars pointing all the way back to Aristotle's "Laws of Association". The three 'laws' of Similarity, Contrast, and Contiguity were one of the earliest attempts to describe the systematicity in language by specifying the sequence of ideas in a train of thought. According to Woodworth (1954), British philosophers in the eighteenth and nineteenth centuries expanded upon these ideas, with Locke (1700, p. 43) first introducing the term *association*. This was all 'armchair psychology', however, and it was left to Galton (1879-1880) to carry out the first experiment on word associations. Using himself as a subject, he gave two associations for each of 75 stimulus words (mostly nouns), on four trials about a month apart. He found that of the 505 total responses, about one-quarter recurred at least twice over the successive trials. He was also able to divide his responses into three categories: visual and other images of past scenes or events 32.5%, 'Histrionic representations' acting out an event or an attitude 22.5%, and strictly linguistic connections like names, phrases, and quotations 45.0%. This first experiment set the tone for future association research, with an emphasis on three elements: reaction time, how often a particular response is repeated, and the classification of associations (Woodworth, 1954).

The next step was to gather associations from more than a single subject. Cattell and Bryant (1889) carried out the first large scale association study, collecting association responses from about 500 subjects. They classified the responses in a more detailed way than Galton, using 10 categories, some of them anticipating those in use today. They discovered that different stimulus words typically elicit only certain kinds of responses, ie. oppositeness or whole-part.

The new century saw a considerable amount of interest in association research, with one of the most important studies aiming to use associations as a measurement tool for mentally-ill people. Kent and Rosanoff (1910) wanted to establish what 'normal' associations were like, in order to be able to determine abnormal ones. To accomplish this, they asked 1,000 mostly adult men and women to give one association each for 100 common nouns and adjectives. They tallied the results and fixed them on norming tables for each word. These early norming tables illustrated many of the features we now consider typical of group associative behavior: a relatively limited number of responses were given by many of respondents, the majority of responses were given by only a few respondents, and there were many responses which were given by only one respondent. Table 1 shows the results for the stimulus word *needle*.

Kent and Rosanoff succeeded not only in establishing the idea of collected norms as a criterion reference, but their tables were probably the most influential ones for over 40 years. In addition, their stimulus list became the standard for later association studies, which had the benefit of comparability, but had the disadvantage of including only frequent nouns and adjectives (see Meara, 1983a, for a discussion of the weaknesses of the Kent-Rosanoff stimuli for nonnative-speakers).

Table 1 Kent-Rosanoff Association Frequency Table for *NEEDLE*

160	thread
158	pin(s)
152	sharp
135	sew(s)
107	sewing
53	steel
40	point
26	instrument
17	eye
15	thimble
12	useful
11	prick(s)
9	pointed
7	cotton
6	work
5	implement
5	tool
4	cloth
4	darning
4	knitting
4	sharpness
3	article
3	fine
3	metal
2	books, button(s), clothes, coat, dressmaker, hurt, hypodermic, industry, pricking, small, sting, thick, thin
1	blood, broken, camel, crocheting, cut, diligence, embroidery, handy, help, hole, home, housewife, labor, long, magnetic, material, mending, nail, ornament, patching, pincushion, shiny, slippers, stitching, surgeon, tailor, use, using, weapon, wire, woman
1000	

(from Woodworth, 1938)

A number of other norming lists soon followed Kent and Rosanoff, and focused on other subject groups. Woodrow and Lowell (1916) compiled an association frequency table for the responses from 1000 9-12 year-old children. They used 100 stimulus words, 90 from the Kent-Rosanoff list. O'Conner (1928) used 1000 blue-collar males to create his association norms, once again using the Kent-Rosanoff stimuli. Soon after, Schellenberg (1930) compiled norms based on 925 University of Minnesota freshmen. In 1954, Russell and Jenkins noticed that the Schellenberg norms no longer seemed accurate for their University of Minnesota students, so they studied 1008 mostly sophomore psychology students and compiled the set of norms which finally superseded the Kent-Rosanoff ones. One of the more recent compilations is Postman and Keppel (1970), who put together a number of norms in one volume.

Methodology

It would seem that association tasks are simplicity itself, all that is required being an experimenter to give stimulus words and subjects to give back responses. To a certain extent this is true, which is one reason why association studies have been so popular. It would be very difficult to reach the kind of subject population numbers connected with association studies (often 1,000 or more) when using other research paradigms. There are several variations of the basic elicitation technique however. One can ask for the first response which comes to the subject's mind, it which case the response would be a free-association (although almost all individuals under the instructions to 'give the first word that comes to mind' impose the additional constraint that the response must be meaningful (Moran and Swartz, 1970)). On the other hand, one can restrict the response in some way, for example asking for the first *noun* to be thought of, which would be a controlled-association. Most often single responses are called for, but it is possible to ask for multiple responses, or ask for continuous responses within a certain time period (write as many associations to the word *X* as you can in 30 seconds). The mode of both stimulus and response can be either written or oral. Still, all-in-all, the elicitation of word associations is relatively straight-forward.

The more difficult part is the scoring and analysis of the responses once they are elicited. Woodworth (1938) suggested that individuals could be scored in three different ways:

1) count their idiosyncratic responses
2) count the number of high frequency responses
3) check the median frequency value of their responses.

Note that all these methods require reference to norming information and so are measures of the 'communality' of the individual's responses with those of the norming group. This is the first major type of association analysis.

A second type involves looking at the nature of the associations themselves. This has typically taken into account the association's word class. Responses which have a sequential relationship to the stimulus word are called *syntagmatic*, and usually, but not always, have differing word classes. Examples from Table 1 would be adjective-noun pairs like *sharp-needle* and *steel-needle* or noun-verb pairs like *needle-sews* and *needle-pricks*. Responses of the same word class as the stimulus are labeled *paradigmatic*. Examples are noun-noun pairs like *needle-thread, needle-pin* and *needle-article*. These terms are relatively recent to association studies, first being used by Jenkins (1954). Before that, similar ideas were represented by the terms *heterogenous* and *homogenous*. McNeill (1963) points out that this dichotomy may be inadequate, since there is potential confusion between whether the terms refer simply to words of the same or different word classes, or whether they refer to substitutability versus sequentialness. He suggests a tripartite distinction would be more accurate: *homogenous* - same word class, *heterogenous* - different classes, and *syntagmatic* - responses of either type which are sequential in nature, although this suggestion does not seem to have been taken up by subsequent researchers.

While syntagmatic relationships involve the contiguity of words in language, paradigmatic relationships are more semantic in nature. Sometimes paradigmatic pairs are roughly synonymous (*blossom-flower*) and sometimes they have other kinds of relationship (*black-white, table-furniture, reflect-affect*). This diversity has prompted attempts to categorize the possible associative relationships. We have already seen that Galton (1879-1880) and Cattell and Bryant (1889) began this classification process, and by 1938 Woodword was able to discuss the issue in some depth. He presents a rather complicated 18-class system first used by Wells (1911):

	Class	Example
1.	Failure to respond	
2.	Egocentric	success-I must
3.	Egocentric predicate	lonesome-never
4.	Evaluation	rose-beautiful
5.	Matter-of-fact predicate	spinach-green
6.	Subject-verb	dog-bite
7.	Object-verb	deer-shoot
8.	Cause-effect	joke-laughter
9.	Coordination	cow-horse
10.	Subordination	fruit-apple
11.	Supraordination	table-furniture
12.	Contrast	black-white
13.	Coexistence	Sunday-church
14.	Identity	blossom-flower
15.	Phrase completing	forward-march

16.	Word completing or compounding	black-board
17.	Assonance	pack-tack
18.	Syntactic change	deep-depth

This system proved rather cumbersome, and Wells (1927) himself suggested condensing the above classifications down into five: 1) Egocentric or subjective, 2) Supraordinate, 3) Contrast, 4) Miscellaneous, and 5) Speech habit. However, as late as 1958, Flavell, Draguns, Feinberg, and Budin were using a system containing 16 categories. They were separated into three sections which were supposed to reflect the cognitive maturity of the associative relationship:

Immature categories

1. Completion	needle-haystack	glow-worm
2. Distant	love-Mary	book-darkness
3. Perseveration	related to or repetition of previous stimulus or response	
4. Perseveration-meaningful	as above but also closely related to its stimulus word	
5. Clang	cold-gold	glow-fast
6. Emotional	love-good	snake-ugh!
7. Repetition	expect-expect	deep-depth
8. Multi-word-discrete	jewel-"ruby" then "diamond"	

Mature categories

9. Synonym	liberty-freedom
10. Supraordinate	egg-food
11. Subordinate	fruit-apple

Indeterminate categories

12. Attribute	deep-ocean	
13. Verb	snore-father	cheese-eat
14. Contrast-coordinate	dark-light	wall-ceiling
15. Multi-word	citizen-a person who pays taxes	
16. Blocking	No response or slower than 10 seconds	

Most researchers opted for more streamlined systems (see Marshall and Cofer, 1963, for a summary of methods to compute association scores). Woodworth (1938; also Woodworth and Schlosberg, 1954) suggested four categories: Definition, including synonyms and supraordinates, Completion and prediction, Coordinates and contrasts, and Valuations and personal associations. Goodglass and Baker (1976) used six: Superordinate, Attribute, Contrast coordinate, Function associate, Functional context, and Clang. Deese (1965) believes that non-syntagmatic associations can largely be explained by only contrast or grouping.

The categories in these systems were derived solely by linguistic analysis of subjects' responses. Moran (1966) and Moran and Murakawa (1968) applied the statistical procedure of factor analysis to try to extract categories in a more empirical way. When analyzing their subjects' responses, they found that four factors emerged. They labeled these as Perceptual referent (adjective-noun, noun-adjective), Object referent ('functional' ie. *foot-shoe*), Concept referent (synonym or superordinate) and Dimension referent (contrast or coordinate). Although most subjects usually give responses in all these categories, they discovered that about half had a definite tendency to give one category of associate.

Goodglass and Baker (1976) used speed and accuracy of response to isolate an ordering of associative dimensions or categories. They propose that an 'inner circle' of associations within a semantic field consists of "the [concept] Identity label, the Superordinate association, the most common descriptive adjective (Attribute), and terms related by situational contiguity (Functional Context). Further towards the fringe of the field, in terms of readiness of acceptance as associates, are objects of the same category (Contrast Coordinates) and verbs denoting characteristic actions carried out on or by the target stimulus (Functional Associates)" (p. 371).

Deese (1962) suggests that the distribution of responses to any free-association stimulus forms the associative meaning of the stimulus word and thus the associative concept named by that word. A corollary of this is that associative meaning should predict the words which will occur in the verbal environment of any particular word, but since it stems from a distribution, can not predict the tendency of any two words to elicit one another. In other words, one needs to look at the patterns, not at the responses themselves to get at the underlying meaning.

A third major way to analyze associations is according to the time it takes to give the results. The underlying assumption is that responses given more quickly are somehow stronger. This method was more common in earlier studies but has faded away in modern times, in favor of communality and descriptive approaches.

Instead of describing the communality or attributes of single associations, some scholars have looked for ways to describe the overall association system in a subject's mental lexicon. The method suggested by Kiss (1968) and Meara (1992) is to adapt Graph Theory to the problem. The graphs do not describe individual associations, but attempt to model the complexity of the overall association network by taking into account the total number of associations, the number of connections (valency) between various associations, and distance between any two associations. It must be said that this application of Graph Theory has not captured the attention of the field, perhaps because of the difficulty in establishing workable figures for any of the three parameters mentioned.

What We Know About Word Associations

Communality

Word associations have been empirically studied for far longer than many other aspects of language. Fox (1970) claims to have found approximately 75 original collections of word association data from the period of 1900 to 1970. Much of this was relatively recent at the time, with 48 coming from the period 1960-1970. With this amount of research, one would expect that much has been learned in the century or so since Galton and Cattell and Bryant. Although it is unlikely that associations will ever be as explainable as other 'rule-based' aspects of language, we do have a fair degree of understanding into their behavior. Perhaps the best starting point is the statement that associations exhibit a great deal of systematicity. This is well illustrated by the responses to *needle* in Table 1. The results are not random, otherwise one would expect nearly 1000 different responses from 1000 subjects. Rather we find that over 70% of the responses (712/1000) consisted of one of only five words (*thread, pin(s), sharp, sew(s), sewing*). Clearly there is a great deal of agreement among the members of the norming group. This tendency toward common responses is what makes association frequency norming tables feasible. Of course, human beings are far too creative to be totally uniform, and the remaining 288 responses are spread over 63 different words. This pattern describes very well the distribution of responses for almost any stimulus word for almost any norming group: a small number of responses being relatively frequent, with a larger number of responses being relatively infrequent.

The pattern of communality has been demonstrated across numerous studies. For example, for Lambert and Moore's (1966) English-speaking subjects, the primary response covered about 1/3 of the total responses and the primary, secondary, and tertiary responses together accounted for between 50-60%. This is congruent with the 57% figure reported by Johnston (1974) when she studied the first three most popular responses of 10-11 year-olds.

Typically, there are also a number of responses which are unique to a single respondent. Woodworth and Schlosberg (1954, p. 50-51) state that "Normal persons with only common-school education have given an average of 5.2 individual reactions for the list of 100 stimulus words, while college-educated subjects, perhaps because of their larger vocabulary, have averaged somewhat more, 9.3 individual reactions." Mintz (1969) found that about 10% of responses from hospital nurses were idiosyncratic, but this was with only 39 subjects. Idiosyncratic responses accounted for about 25% of the total given by 84 5th-grade students (Johnston, 1974), with verbs producing the most and adjectives the least from the four major word classes. In the same study, schizophrenic students of the same age typically gave almost 50% idiosyncratic responses. In Table 1, although there are more words in the '1 each' frequency category than any other, they still comprise only about 3% of the total (31/1000).

Word Class

Stimulus words of a certain word class tend to elicit responses of the same class. Woodworth (1938) looked at several studies and came up with the following generalizations:

	Response (%)		
Stimulus	noun	adjective	verb
noun	70	20	8
adjective	45	50	3
verb	50	5	35

From this we can see that nouns are the most common responses and that, for verbs at least, seem to overcome the paradigmatic verb-verb tendency.

Brown and Berko (1960) looked at word class in more detail and found that adults (N=20) consistently gave responses with the same part of speech. Out of a maximum of 6 stimulus words per word class, the means were 5.10 for count nouns, adjectives 5.00, adverbs 4.95, intransitive verbs 4.8, and transitive verbs 4.45. The only exception was mass noun stimuli, which only elicited an average of 2.35 mass noun responses. Rosenweig (1964) studied French and American workers and students and found that when all responses to the first 10 stimuli of the Kent-Rosanoff list were tallied, between 75% and 79% of them were of the same word class, with the exception of the French workers with 50%.

Nouns and adjectives in particular seem to elicit their own word class in responses. Deese (1965) states that irrespective of frequency, the associations of nouns are nouns. Palermo (1971) studied 4th-grade students and found that nouns elicited nouns 81% of the time with verbal prompts and 74% of the time with written prompts. Adjectives were the next most consistent (67% verbal, 62% written), followed by verbs (58% verbal, 25% written) and adverbs (44% verbal, 21% written). As will be commented upon later, in every case the percentages are higher for verbal than written elicitation.

Paradigmatic Shift

Probably the most famous finding in association studies is that responses tend to shift from being predominately syntagmatic to being predominately paradigmatic as a person's language matures. Early studies, such as Woodrow and Lowell (1916), showed that children have different associations from adults, although the differences were normally thought of in semantic terms (Brown and Berko, 1960). Ervin (1961) is usually credited with being the first to develop the principle of using the maintenance of word class to differentiate associations. She gave free and two-choice association tasks to 23 kindergarten, 10 1st-grade, 52 3rd-grade, and 99 6th-grade students and found that as the students' age increased, their proportion of paradigmatic responses also increased. Conversely, there was a decrease with age in clang associations. One year earlier, Brown and Berko (1960) similarly established that the tendency to associate words within a word class increases with age (although they credit an earlier conference presentation by Ervin (1957) for the classification concept).

The syntagmatic→paradigmatic shift (S→P) has been well documented (e.g. Entwisle, Forsyth, and Muuss, 1964; Entwisle, 1966; Sharp and Cole, 1972), but this is not to say that it is entirely uniform. For one thing, the timing of the change seems to vary. Entwisle, Forsyth, and Muuss, (1964) gathered 96 associations each from 500 children aged from 5 to 11 years. They concluded that there was a definite shift from syntagmatic responses to paradigmatic responses from ages 5 to 10, but that most change occurred between ages 6 and 8 (1st to 3rd grade). This was particularly true for high-frequency adjectives and verbs. For high-frequency nouns, there was not as much change, so they speculate whether the S→P shift had already occurred before kindergarten or 1st-grade for those nouns. By age 8, adjectival paradigmatic responses appear to reach asymptote, but verbs are just starting to yield markedly increased paradigmatic responses. On the other hand, syntagmatic responses drop to adult levels during the 8-10 year-old period. Entwisle published her complete study of 1,360 children in 1966 and reached similar conclusions. She suggests that "only a small portion of the vocabulary may evolve through this pattern [S→P] at any particular age" (p. 120) and that word class is a major factor. To illustrate this she presents the following figures:

Percent of Paradigmatic Responses

	K	1	3	5I	5G	UNI
Adjective stimulus	16.8	31.7	70.6	78.5	55.6	65.8
Verb stimulus	16.6	20.4	47.5	59.6	56.1	60.0

K=kindergarten
1=1st grade
3=3rd grade

5I=5th grade, individual administration
5G=5th grade, group administration, high IQ
UNI=University

(Adapted from Entwisle, 1966, p. 59)

From this comparison of adjective and verb results, it is clear that the paradigmatic shift begins later for verbs and is more gradual. For adult university students, the final percentage of paradigmatic responses for verbs is also lower than for adjectives. Preece (1977) found that the number of associations to technical physics terms increased steadily into adolescence, so that associative fluency may continue to improve into young adulthood for these types of technical words. Once the associative connections are developed, it seems that they do not easily fade away. Howard, McAndrews, and Lasaga (1981) found that even elderly subjects (70+ years old) benefit from associative organization. In a lexical decision task, their responses were faster when the words were associated than when they were not. This suggests that automatic processes, such as association, do not deteriorate with age.

Sharp and Cole (1972) studied subjects who spoke Kpelle, an African language structurally different from most European languages, and also found the S→P shift. It occurred at different ages for different word classes and different word classes differed in the proportion of paradigmatic responses they elicited regardless of age. Nouns elicited most, then adjectives, then verbs. In fact, 72% of all the responses were nouns, but the task instructions may have encouraged this to some extent.

Goodglass and Baker (1976) studied aphasics and concluded their responses were more concrete than normals, with their responses becoming more similar to those given by young children.

As we have seen, word class affects the timing of the paradigmatic shift, but it also affects the percentage of paradigmatic responses given. Deese (1962, p.81) studied 100 university undergraduates and found that the "generalization that adult associations are largely paradigmatic is unconditionally true only for nouns. Adjectives and verbs are about equally syntagmatic and paradigmatic; adverbs yield largely syntagmatic associations." In terms of figures, nouns yielded syntagmatic associations about 21% of the time, verbs and adjectives 48% and 50%, and adverbs 73%. This suggests the possibility that a child could give more paradigmatic responses than an adult if the child responded to mainly noun stimuli and adults adverb stimuli. In other words, the word class of stimuli needs to be controlled if valid comparisons are to be made. Entwisle's (1966) 200 university students produced similar results for nouns (77% paradigmatic), but gave a higher percentage for adjectives (66% paradigmatic), verbs (60% paradigmatic), and especially adverbs (79% paradigmatic). So different word classes do seem to elicit different rates of syntagmatic vs paradigmatic responses, but the exact figures are still less than clearly defined.

Ingersoll (1974) expands upon this discussion by pointing out that the predominance of paradigmatic responding in adults has been overgeneralized. He gives a number of arguments to back his assertion. First, adults provide fewer paradigmatic responses to low frequency stimuli than to high frequency stimuli. Second, adults and 6th graders do not differ overall in responses to high frequency stimuli (when it comes to paradigmatic responses), but 6th graders have more paradigmatic responses to adjectives, verbs, and mass nouns. Finally, adults given low-frequency stimuli are similar to 2nd graders given high frequency stimuli, but slightly more paradigmatic. He suggests that more frequent words are better known for word class, so that information can be more easily taken into consideration when giving associations (leading to more paradigmatic responses). Thus the frequency of stimulus words may be as important as age in describing developmental trends, and also needs to be controlled for in studies.

Deese (1962) also looked at how the frequency of occurrence of the stimulus words affected their responses. Frequency did not seem to make a difference for nouns, verbs or adverbs, but high-frequency adjectives elicited fewer syntagmatic associates than low-frequency adjectives. Deese checked the adjective stimulus-response pairs and concluded that high-frequency adjectives are more likely to elicit contrasts (*black-white*) than low-frequency adjectives, which tend to elicit sequential relationships like *absurd-situation*, accounting for the difference in paradigmatic responses. Entwisle (1966) found only minor effects for frequency, but this is not surprising because she had to use relatively frequent words exclusively in view of her young subjects. As for the relationship between frequency and association responses, a study by Howes (1957) indicated that the probability of a word being given as a response on a word association task is closely related to its frequency of occurrence in language.

Palermo (1971) synthesizes his study with previous research (Palermo, 1963; Palermo and Jenkins, 1963) and summarizes the major trends as follows (p. 122):

> [The] frequency of popular responses increases from first grade through to the college level, superordinate responses rise from first to sixth grades and decline thereafter, contrast responses rise steadily from first grade through the college level, and paradigmatic responses rise steadily from

> first grade through college (although the percentage of such responses varies considerably from one grammatical class to another). Of particular interest ... are the rather dramatic differences in the characteristics of the responses from first grade to the second grade. With the possible exception of superordinate responses, where the shift is rather gradual, the changes from Grade 1 to Grade 2 are very large relative to other grade differences.

Idiodynamic Sets

Although most respondents will use most or all of the above categories of association in the process of responding to a list of stimulus words, it is now clear that a substantial proportion of subjects will have a definite tendency towards a particular category. This tendency has been labeled *idiodynamic set.* Building on earlier work (Moran, Mefferd, and Kimble, 1964), Moran (1966) ran a series of studies featuring factor analysis and concluded there are four types of sets. The first is a Perceptual-referent set, in which subjects tend to give associates which are of the object-attribute type (*hammer-steel, sour-pickle, crow-black*). Subjects with an Object-reference set tend to give associates which are concrete objects functionally and commonly related in the real world, such as *foot-shoe* or *boat-dock.* Concept-reference set subjects will tend to give superordinates or synonyms as responses (*banana-fruit, small-little*). Coordinates (*apple-orange*) and contrast (*soft-hard*) would be common responses for subjects in the Dimension-referent set. The factor analyses actually indicated three factors with the perceptual-referent and dimension-referent sets being the opposites poles of a single factor. Moran believes that the above empirically derived sets can be used as a parsimonious categorization system for associations. He also suggests it is a hierarchy of increasing linguistic sophistication.

There is some evidence that these idiodynamic sets are universal. The sex of subjects did not seem to have any effect on the sets derived from the responses. Subjects from other cultures also exhibit the sets. Moran (1966) found that Mexican subjects doing the association task in Spanish also exhibited these sets, and Moran and Murakawa (1968) found similar results with Japanese subjects. In fact, Moran and Murakawa reviewed Moran and his colleagues' work and concluded that older L1 men (age 35), acutely psychotic schizophrenic men, U.S. university students, Mexican university students, and Japanese university students all display these sets (although retardates did not have the dimension-referent set).

Although factor analyses indicated the same factors for all the above groups, of course there are some set preferences among the groups. For example, US university students gave more associations belonging to the dimension-referent and concept-referent association categories, while Japanese university women gave more in the perceptual-referent category. When American and Japanese subjects were checked for individual idiodynamic set tendencies, the distribution in Table 2 was discovered:

Table 2 Idiodynamic Set Membership, US and Japanese University Students Both N=258

Subject	Dimension Referent	Concept Referent	Object Referent	Perceptual Referent	No Set
Japanese	20	25	42	54	117
US	58	29	29	23	119

From these figures, we can see that almost exactly the same percentage of Japanese and American students exhibit idiodynamic set behavior (US 53.9%, Japanese 54.6%). Moran and Swartz (1970) found a similar figure for 15 year-old students (50%), but younger students had somewhat lower percentages (12 year-olds 36%, 9 year-olds 48%). Thus, for adults, we can expect about half of the respondents on a single-free-response association task to show a clear preference for a certain category of response.

Set membership seems to be consistent; if a subject fell into a certain set on a list of 40 stimulus words, they were very likely to exhibit the same set behavior on a second list of 40 different words. Moran and Swartz review a number of studies which indicate that the sets are also stable over varying periods of time: over 4 days with a different list each day (Moran, Mefferd, and Kimble, 1964), 90 days with the same list (Moran, 1966), and 90 days with different lists and different elicitation procedures (Moran, 1967a & b). They went on to study 85 9.7 year-olds, 88 12.7 year-olds, and 107 15.7 year-olds over a period of two years and also found the sets stable over that longer time frame. So the sets seem stable, at least after 9 years of age.

Among the other points of interest to come out of this study is that although the sets themselves are reliable over time, the stimulus response pairs themselves are not. Overall, about 75% of the word pairs were different, even though part of the same set. This indicates that the idiodynamic sets are enduring, but not the word-word pairs. The suggestion is that the field could profit from research into ways to judge subjects on set membership rather than individual word pairings. Another point is the relationship of the four idiodynamic sets to the syntagmatic/paradigmatic dichotomy. It seems that the dimension set promotes paradigmatic responses, the perceptual set promotes syntagmatic responses, and the concept set is not consistently either paradigmatic or syntagmatic. In fact, dimension-referent subjects tend to be paradigmatic and perception-referent subjects tend to be syntagmatic, even when giving associations that are unrelated to their semantic (idiodynamic) set.

So Moran and his colleagues seemed to have established four underlying categories of association relationship. But if these four set categories are real, how can we account for the commonality so obvious in group association tables? Moran and Murakawa suggest that subjects have two association hierarchies: one which is revealed by idiodynamic set responses and also a 'common' hierarchy which people can use if mutual communication is necessary. They cite findings by Jenkins (1959) and Horton, Marlowe, and Crowne (1963) which indicate that subjects can produce 'popular' responses if required. Time pressures in free association tasks tends to move responses away from a particular idiodynamic set and towards more commonality, so the 'common' hierarchy may be the more easily accessed of the two. Thus reference to a 'common' hierarchy can explain the high degree of agreement of association responses. This dual-hierarchy model is interesting, but at the moment we simply do not know enough about the relative strengths or interaction between the two types to assess how accurate or useful the model is.

Variation of Responses over Time

Moran and Swartz's (1970) results indicate that responses to any stimulus vary over time. There is evidence for this from other studies as well. Dalrymple-Alford and Aamiry (1970) asked subjects to twice give responses to the same stimulus word within the same class session (the stimulus words were mixed with buffer words on the T1 and T2 lists). Half of the stimulus words had a primary response which accounted for 50% or more of the responses from a norming group, while other half had a primary responses which accounted for less than 20% of the norm responses. Only 71% of subjects' responses were the same from T1 to T2 for the 'high primary' stimuli, while the figure was 48% for the 'low primary' stimuli. Fox (1970) analyzed responses given 60 days apart by undergraduates at the University of Minnesota and found that, of 1365 total responses, about 52% were different. He also cites Gekoski and Riegel's (1966) figure of 57% different responses as showing a similar amount of variation.

Simpson and Voss (1967) took association elicitation to extremes and asked for 32 responses to each of 6 stimulus words and then repeated the process 10 times. Only 1.7-2.5% of responses were repeated in all 10 trials, while 57-62% were given only once in 10 trials. This shows a great deal of variation in the responses, although this is obviously an unnatural task, and by the later associations, the subjects were probably struggling for any response imaginable to fill their quota. The initial responses are probably the most informative, as it was found that the more frequently a response is given by a group, the more likely it is to be among the first few responses given in a series. If a response was given in all 10 trials, then it was likely to be given within the first 10 responses out of 32, giving support to the intuition that the 'better' responses are given at the beginning of a series.

Despite this individual variation, Fox confirmed that group norms tend to be very stable over time. The higher the frequency of the primary response in a norm group, the greater the likelihood that any individual subject will repeat that primary from T1 to T2. If the percentage of the primary is only about 20% of the respondents, then it is about equally likely that the T2 response will be the primary or another word. But if the primary is given by 60% or more of the norming group, then there is about an 80% chance of repetition. In addition, even if the subject did not give the primary on the T1, it is less likely that he would again fail to give the primary on the T2 if it has a high degree of communality.

What Engenders Associations?

The question of what causes associations or how they are created is a very difficult one to answer, because as Woodworth and Schlosberg (1954) point out, associations are already present when we measure them, and to get at association formation, we would have to cause *new* associations to be formed, not merely study already existing ones. There is also the confounding between connections of lexical label name and the objects themselves; are the referents associated, the word labels, or a mixture? Nevertheless, these problems have not stopped scholars from hypothesizing about the mechanisms that produce associations.

Early scholars working with associations believed that they were created mainly by contiguity, that is, the repeated exposure of words which often co-occur in close proximity in language. Although this explanation seems satisfactory for syntagmatic responses, it is not very convincing for paradigmatic responses like *deep-shallow*, where the two words are unlikely to occur in the same bit of discourse (water will be either deep or shallow, not both at the same time). An early explanation of the S→P shift suggested by Ervin hypothesized that paradigmatic responses are related to forward contiguity, but later research showed that the two were not related (McNeill, 1966). So an explanation for the creation of paradigmatic associations which did not involve contiguity was required.

McNeill (1963) ran an experiment in which subjects were exposed to nonsense triads (*MIP, KOJ*) in sentence contexts. The more frequently triads introduced as nouns in these sentences were presented, the more frequently they were associated together, leading him to hypothesize that the ability of words to intersubstitute for each other (an attribute of having the same word class) leads to association. This linguistic origin could explain the vast majority of noun-noun associations, like *man-woman*. A smaller number of noun-noun associations could be explained by coexistence in the real world, like *table-chair*. This extralinguistic influence could also partially explain the major association category of contrast; in the real-world, an item can have only one attribute of any kind (as in the *deep-shallow* example above), reinforcing the linguistic representations of these opposed qualities. Still, despite this extralinguistic influence, he suggests that paradigmatic associations have mainly a linguistic source.

He goes on to propose that associations are generated by manipulating the semantic features of the stimulus word. Johnston (1974, p. 664) paraphrases his argument quite nicely:

> Words are assigned lists of features, some of which are syntactic and some of which are semantic. For example, *man* would have the following list of features: noun, animate, human, male, adult, etc. By the age of five, the child has acquired the syntactic features of words, and is sophisticated in using grammatical rules, but he knows little about semantics. It is with the child's acquisition of knowledge about the semantic properties of words that the paradigmatic shift occurs. McNeill hypothesized that the associative response is that which can be made with the least conceptual effort, and is thus most often the word which shares the maximum number of features with the stimulus word. As children acquire semantic features, the set of words that minimally contrast to the stimulus word will more often include words of the same grammatical class, so that a paradigmatic response becomes more likely.

Contrasts, which are a major category of association, make a good example of this. Normally all that is required to generate a contrastive word is to change one feature: *man* [+male] - *woman* [-male]. As expanded upon by Clark (1970) below, the process of generating an association may involve changing the least number of features possible, starting at the 'lowest' level, until one arrives at a reasonable association. This makes contrasts a very common type of association. Carrol, Kjeldergaard, and Carton (1962) suggest that the tendency to give opposite responses to opposite-eliciting stimuli accounts for a great deal of the communality on lists like the Kent-Rosanoff norms.

Knowledge of word class seems to be a key component of this linguistic source of associations. Entwisle (1966) interprets her data as showing that by age 8, children have come to understand the substitution privilege of word classes which leads to paradigmatic responses. Brown and Berko (1960) also discovered that the linguistic skills necessary to deal with word class correlated with the frequency of paradigmatic responses. They found that both the number of paradigmatic associations and awareness of word class increased with age and were closely related, causing them to wonder whether the syntagmatic→paradigmatic shift is a manifestation of a developing appreciation of syntax.

Deese (1962) also showed that the paradigmatic/syntagmatic response depends to a large degree on the word class of the stimulus word. He suggests that the responses for adverbs, verbs, and low-frequency adjectives are probably determined by exposure to these sequential pieces of language, ie *amazingly-bright.* On the other hand, subject responses to nouns and high-frequency adjectives are mainly schematic (contrastive for these adjectives), ie. *difficulty-hard* and *black-white.*

Looking at young children (ages 4-5, kindergarten, and 1st-grade) Entwisle, Forsyth, and Muuss (1964) found that the most frequent response to stimuli of any word class were nouns. They suggest that since nouns in general are the first words to be learned and the easiest to conceptually delineate, they are naturally the most available as a class to form associations. But by the 3rdgrade (when presumably the children's vocabulary and cognitive maturity has developed to some degree), their responses are overwhelmingly of the same word class as the stimulus. Very young children (ages 4-5) seem to have trouble reliably associating words according to a semantic basis, resulting in frequent 'clang' responses. Kahana and Sterneck (1972) found that 5 year-olds often made 'illogical' responses to stimuli, such as clang responses, unrelated responses, and 'I don't know' or no responses. By the ages of 7-9 however, clang responses in particular become 'practically nonexistent'.

Petrey (1977) believes that focus on word class and the syntagmatic→paradigmatic shift cannot totally explain children's word associations, arguing that it only captures syntactic information (since it is based solely on the notion of word class). She thinks that in the beginning children give associations based on episodic memories, so that initially situations induce associations. These situational associations can be especially strong if the linguistic environment is congruent and supports them. For example, *salt-pepper* always go together in the real world, but they have a linguistic coordinate relationship as well, making them strong associates. She analyzes Entwisle's (1966) data and concludes that episodic experiences can account for a great deal of the diversity in the kindergarten subjects' responses. Using the stimulus *examine* as an example, she starts by stating that their primary and secondary responses only add up to 11.5% of the total, indicating a great deal of diversity. The adult figure, on the other hand, is 41%. Citing responses like *doctor, x-rays, stethoscope, pill, needle, shot,* and *bed*, she concludes that they all are part of children's experiences with doctors and medical checkups. Together she claims that they make up between one-third and one-half of the different responses.

Thus while pairs like *salt-pepper* and *table-chair* will always be defined as paradigmatic, the actual source of the association may vary with age. Young children are likely to form this pair from their own autobiographical experiences, while more mature subjects will form the pair on a semantic basis. This can explain the cases where young children apparently already form adult-like paradigmatic associations, which then drop in favor of syntagmatic responses before eventually becoming more frequent again. Petrey ultimately suggests that the S➔P shift should be modified to a episodic➔syntagmatic➔paradigmatic shift, in which young children associate primarily to the stimulus's perceived contexts and older subjects to its abstract semantic content.

Preece (1977) postulates a 2-stage learning process, with associates first being learned indiscriminately, and then becoming limited to more direct relationships. He tested grammar school and university physics students for associations having to do with the technical vocabulary of physics and mechanics. This vocabulary was divided into words with a direct relationship, meaning words which occur together in technical formulas, eg. *speed* = *distance* ÷ *time*, and mechanical words not co-occurring in a formula (indirect relationship). He found there was a sharp rise between the 12 year-old subjects and the 15 year-old subjects in both direct and indirect associations, indicating a greater availability of mechanics terms, and therefore learning of those terms. But later falls in indirect, but not direct, associations suggests there is a second stage of learning where these associations become better integrated as a result of better physics knowledge. This kind of reorganization from relatively loose, indiscriminate ties to tighter and more defined relationships mirrors that of the clang➔syntagmatic➔paradigmatic shift.

Lambert and Moore (1966) see associations as being the manifestation of a connotative meaning network. When one word is brought to cognitive attention, it 'activates' other words which it is associated with. This view suggests how associations operate, but does not really give insights into how they originate in the first place.

The traditional approach to the study of semantic field starts with the concept label as stimulus and explores the network of associations which spread outwards from this core. But associations need not be only the *responses* to a verbal label, they may lead to a label as well. A flood of semantic associates may come to mind and be verbalized in unsuccessful attempts to name a concept (such as in tip-of-the-tongue experiments). So we must assume that the semantic field exists with or without the availability of the name for a concept, and may be concurrent with, precede, or even facilitate production of a concept name. In this sense, the ability to name may depend on the integrity of the semantic field, reflecting the convergence of associations on the target.

Goodglass and Baker (1976) tested the above proposition with 32 aphasics. They found that high comprehension (Broca's) aphasics gave normal responses, but more slowly; low comprehension (Wernicke's) aphasics gave eccentric responses, suggesting disruption in the associative structure of aphasics who are most impaired both in comprehension and word finding.

Clark (1970) has taken what is known about associations and has attempted to condense it into a number of 'rules' which he believes can explain the behavior of associations. In doing so, he relies heavily on the idea of semantic features.

Paradigmatic Rules:

The Minimal-Contrast Rule - If a stimulus has a common 'opposite', it will always elicit that opposite more often than anything else. These responses are the most common in word association tasks. It appears that for adjectives and possibly nouns, the most frequent paradigmatic response tends to be a word with the maximum number of features in common with the stimulus. For minimal contrast, only one feature needs to be changed ie MAN-WOMAN [+/- male] Clark suggests it is the change of last feature which is easiest, but this assumes being able to order features in a principled way.

The Marking Rule - This is the tendency to change a feature from, rather than to, its *marked* value in word association data. So marked *dogs* should elicit *dog* more than vice versa. But *man* is unmarked and produces marked *woman* more often than vice versa, so it is not a completely general rule. A marked and restricted word (*shallow*) should elicit an unmarked but polysemous word (*deep*) more often than vice-versa because the unmarked word has more meanings to associate with.

Feature Deletion and Addition Rules - Deletion should have precedence over addition of features, since there are many possible features that might be added, but those to be deleted are already specified. Deletion generally produces superordinates, like *fruit* from *apple*, while addition of features produces subordinates, like *apple* from *fruit*. Subjects generally offer superordinates more quickly than subordinates.

Category Preservation Rule - Word class is preserved because it is a 'high' feature.

All of these rules dealing with paradigmatic responses could be generalized as a *Simplicity of Production* rule which might be stated as "Perform the least change on the lowest feature, with the restriction that the result must correspond to an English word" (Clark, 1970, p. 280). Increasing difficulty of change can be summarized as 1) changing the sign of a feature, 2) deleting a feature, and 3) adding a feature.

Syntagmatic Rules:

The Selectional Feature Realization Rule - A word often has features that restrict the other words it can be placed in context with, ie. *young* has [+animate] so naturally its associates may have these features. In other words, the associates of *young* will naturally be animate entities like *boy* or *people*. This rule can be more specifically stated as "Take the features specified by a selectional feature, adding as many features as necessary for a surface realization; in addition, restrict yourself to the "significant" part of the selectional feature, the portion specifying a lexical word" (Clark, 1970, p. 281). He uses this rule to explain why the different word classes elicit different rates of syntagmatic responses, suggesting the reason that adjectives, adverbs, and verbs elicit more syntagmatic responses than nouns is because they have selectional features, while nouns do not. He thinks this explains why nouns have much fewer syntagmatic responses, as shown in data from

Deese (1962):

-nouns produce 21% syntagmatic responses

-verbs produce 48% syntagmatic responses

-adjectives produce 50% syntagmatic responses

-adverbs produce 73% syntagmatic responses

The selectional features for adjectives specify the nouns they modify, so adjectives elicit nouns most often with 80% of syntagmatic responses.

<u>The Idiom Completion Rule</u> - Find an idiom of which the stimulus is part and produce the next main word ie *whistle stop; justice peace*. The rule only works like language, left-to-right, ie. *cheese cottage* is not a typical associate.

Clark's explanation is the most comprehensive yet articulated, and is certainly an advance on the older formulations which rely solely on word class. Still, there is the sense that Clark's rules are more descriptive than explanative in nature. Clearly a number of factors affect the way associations develop. In the end, it may prove most profitable to view the maturation of associations as a developmental process moving from relatively fewer organizational constraints and a great deal of diversity towards more and more obvious choices based on mainly semantic criteria, resulting in a high degree of commonality of responses within a mature group. At the moment we can only describe associations with any degree of confidence, explaining how they are created still seems to be beyond our grasp.

The Changing Of Group Norms Over Time

We have seen that the associations an individual gives changes over time, especially during the period of primary school. But what of the associations given by a large group of subjects? Jenkins and Russell (1960) found that they change as well. When comparing the 1927 and 1952 University of Minnesota norms, the primary responses increased in frequency. The first three responses accounted for 59% of all responses in 1952, but only 49% in 1927. The primary responses themselves were largely the same (71% identical), although one would probably have not expected even this much change. The primaries which did change were largely superordinate responses with only a small frequency advantage in the first place. They were commonly replaced by coordinate and syntagmatic associations. The primary responses were more stable than secondary or tertiary responses, which showed a greater degree of change. Expanding their horizon and analyzing three other major norms (Kent-Rosanoff, 1910; O'Connor, 1928; Keene, 1951), they found their conclusions based on the two Minnesota norms were supported. There was some indication that there is a tendency for the frequency of popular responses to increase over time. Also, the responses themselves changed over time, with the most common being the most stable, while superordinates decreased in popularity. Entwisle, Forsyth, and Muuss (1964) compared their results to Woodrow and Lowell's 1916 norms and concluded that the syntagmatic→paradigmatic shift occurred 4 or 5 years earlier than in 1916, perhaps because of more and earlier exposure to language, including radio, TV, and general urbanization. In sum, there seemed to be a change towards conformity within a group rather than individuality.

Associations Among Different Languages

Up until now, only the associations of a single L1 group have been discussed. This leads to the question of whether association behavior is the same across language groups. Rosenweig (1961) compared the primary responses to Kent-Rosanoff words and their translations in English, French, German, and Italian and made a number of discoveries. The primary responses to the same stimulus word (as translated) tended to be equivalent in meaning across the languages. Overall, about half of the primaries were translation equivalents for the different languages. In each language, the greater the frequency with which a particular primary response was given, the greater the probability it agreed in meaning with the corresponding primary responses of the other languages. In addition, when both stimulus and primary response words were adjectives or when the response was opposite in meaning to the stimulus word, there was an even higher agreement among the languages. The two categories 'opposite' and adjective-adjective have about 7/8ths of their responses in common across languages. Another finding was that the more the frequency of the primary response for an item exceeds the frequency of the secondary response in one sample, the more likely that response will be primary in another sample. The frequency of the primary response correlates with the difference between primary and secondary responses (.97 & .94) and the more frequent a primary response is in one sample, the more stable it is from sample to sample. So in general, the stronger an associate was for any language, the stronger it was in other languages as well.

Although the stronger primaries were similar across languages, this does not hide that fact that there were major differences, including the overall communality of the different groups. American subjects had a much higher frequency of primary responses (greater communality) than French, German, or Italian subjects, a result confirmed in a later studies using American and French subjects (Rosenweig, 1964) and American, Canadian, and French subjects (Lambert and Moore, 1966). Both studies also showed that the French give fewer superordinate responses than Americans, although French Canadians gave more superordinates than French students, suggesting that lower superordination is not necessarily built into the French language. Rosenweig (1964) raises the possibility that subjects of certain nationalities may have more in common with certain language groups than others. He suggests that a study by K.M. Miller (no reference given) indicates that Australian and English norms are similar to American ones and that German (Russell and Messeck, 1959) and Italian (Levi, 1949) norms are more similar to those of French. Lambert and Moore (1966) showed that American and Canadian university students had very similar association behavior, with a .69 coefficient of overlap and with 78% of the primaries being the same.

These results are from monolinguals from different language groups, but when association tests are given to bilinguals, the percentage of common responses is less. Kolers (1963) tested the associations of 10 Germans, 10 Spanish and 14 Thai students studying at Harvard University (indicating fluency in English) in the following directions:

English stimulus→English response

English stimulus→L1 response

L1 stimulus→English response

L1 stimulus→L1 response.

He found about 1/3 of the subjects' associations translated to the other language regardless of stimulus or response language, indicating only this number were semantically equivalent. Thus a subject tends to give different responses to an L1 word than he does to that word translated into English, and vice-versa. If association is done interlingually (English-L1 or L1-English), little similarity (about 20%) is seen. About half or little more of all responses in the interlingual tests were unique and about 15% of the responses to the L1-L1 and English-English sections were unique to that task. There was an indication that stimuli referring to concrete, manipulable objects elicit more similar associations across languages than those referring to abstract states or emotion, however. Kolers concludes that the experiences and memories leading to associations are not in held in a common store, but are tagged in one language or the other.

Other results from association research indicate that words are organized according to languages, and not a unitary store. Macnamara (1967) asked bilingual Irish-English subjects to say as many words as possible to the stimulus words, but instead of analyzing the relationship between stimulus and response, he counted the number of associations (spew test/word naming task). The bilinguals could give more words in either of their languages separately than when they had to alternate between the two languages when giving associations or when they had to give both L1a and L1b words for a concept. So although *table* and *chair* are closely associated, *table* and *cathaoir* (=chair) are not.

Taylor (1971) carried out a similar experiment with English speakers skilled in French (although they were unlikely to be truly bilingual). Congruent with Macnamara's results, when subjects switched language every association, or every 3rd association, they produced fewer associations than if they switch every 5th word, were free to switch, or worked monolingually. But the really interesting findings came from the free-to-choose-and-switch-languages condition included in this study. When free to choose language of associations, 65% (English-English) to 73% (French-French) initially chose the language of the stimulus word. After the initial association, subjects switched between languages quite a lot (out of 180 cases they switched 166 times, even though in 30 cases they were instructed not to). On average, they switched about 2.7 times per stimulus word. Still, once in a language, the statistical probabilities were that they would stay in the language. At no point in the course of giving continuous associations did the chance of switching languages reach .5 (even probability of staying in language or switching). The average probability of switching out of English was .27 and out of French .32. Taylor concluded that words from two languages tend to be organized according to languages, most likely in unilingual clusters. In a free-switch association task, bilinguals may make transitions between such unilingual clusters by switching languages.

Using a free-recall task in which subjects attempted to remember words presented on a list, Dalrymple-Alford and Aamiry (1969) also found that staying in a single language is facilitative. But they additionally found that when categories of words on the list (apparel, animals, parts of the body) were held constant, this also aided recall. Although the two research methods are not directly comparable, the results are suggestive that bilingual recall involves both language and category aspects, and when both are consistent, the associations are stronger than they would be if they sprung only from an intralingual or from an intracategory source. (See Dalrymple-Alford and Aamiry, 1970, for more in the same vein.)

Rhyming can be used as an organizational framework when no other structure is available in free-recall tasks, but association according to semantic criteria is usually preferred if both are available (Dolinsky, 1972). There is some indication that recall is best when words can be associated along both rhyme and semantic dimensions. This, taken with the above language-category results, suggests that words with multiple types of associative connection may be the easiest to retrieve.

How Subject Attributes Affect Association Behavior

Rosenzweig (1964) suggested that some language groups have similar association behavior, but this does not mean that any particular language group is essentially homogenous however. The main thrust of his 1964 study was to discover whether social class and education make a difference. He concluded that "adult members of the same 'language community' may have verbal habits that differ systematically according to social groupings within the community" (p. 68). For example, the primary responses of French blue-collar workers were the same as those of French high school and university students in only 39 out of 98 cases. In comparison, the responses of American workers and students were much more uniform, being the same in 68/100 cases. In fact, the primary responses of the French students agreed (in translation) with those of the American students (48/99) more than with the French workers' responses. The French students tended to give paradigmatic responses while the French workers often gave syntagmatic responses. Overall, the students (both countries) gave about half the number of superordinate responses as their respective working compatriots.

Rosenzweig showed that there are clear differences between workers and students, but was not able to isolate the causative factors. The American workers were much older than the American students (median age 43), but the French workers and students were matched for age, so this should not have caused the differences in association behavior. It seems most likely that the differences were due to education or social class. It has been difficult to study the effect of education as a discrete factor, because of its almost universal connection with age. Sharp and Cole (1972) were able to find a group of subjects in which the two factors were not necessarily related. They studied 16 Kpelle 12-14 year-olds in Liberia, 8 of which were nonliterate and had never been to school, and another 8 who were in junior high school. The subjects were asked to give 8 responses to each of 15 stimulus words. Both age and education independently increased the proportion of paradigmatic responses. It seems that education helps depending on age: at the 8-9 year level there is little effect, at 12-14 the effect is larger, and at 18-21 is far the largest.

The stress felt by subjects has also been demonstrated to have an effect on associations. By chance, Mintz (1969) was carrying out an association study on a group of 15 nurses in America on the day that John F. Kennedy was assassinated. They were obviously very upset, and when their responses were compared to those of similar nurses a year later, the effects of this stress could be isolated. He checked their responses against those listed in the Palermo-

Jenkins college female norms (1964) according to a 4-category classification: Most Common - response given by 20% or more of norm group, Common - between 2% and 20%, Unusual - between one individual and 2%, and Unique - response not in norms. The 'stressed' subjects gave fewer Most Common associates, and made up for this deficiency by giving more responses in the Unusual category. Both groups gave equal numbers of Common and Unique responses. Methodologically, it is interesting to note that use of this scale reveals that lower communality scores is not necessarily caused by a higher number of unique responses; it can also be caused by fewer most common and common responses and more unusual responses.

The mode of elicitation affects association responses as well (Palermo, 1971). For children in Grade 4, oral administration resulted in a higher percentage of primaries and secondaries being given than with written techniques. The category of association is also affected, as verbal elicitation led to twice as many contrast associations and about 20% fewer superordinate associations. This doubling of contrastive responses under an oral procedure was also noted by Woodrow and Lowell (1916).

Gender has not generally been shown to make much difference in association behavior. A number of studies have failed to show male vs female as a significant factor in the associations produced (Flavell, Draguns, Feinberg, and Budin, 1958; Rosenweig, 1961; Entwisle, 1966; Moran, 1966; Moran and Swartz, 1970). But at least two studies have shown some gender effect. Palermo (1971) reported that a higher percentage of girls than boys in Grades 1-4 gave the primary and secondary response from a norming list. He did not run a statistical analysis to discover if the differences were statistically reliable however. Rierdan (1980) found that socially-isolated sophomore high school and junior high school students emitted a greater proportion (.32) of idiosyncratic responses vs non isolates (.19), but this was significant for only females subjects and not males.

Timing Association Responses

Early association studies often timed the speed at which responses were given, on the assumption that quicker responses indicated stronger associations (Woodworth and Schlosberg, 1954). A very early study by Cattell and Bryant (1889) cites a figure of about 1/2 second to see a stimulus and name a response. Flavell, Draguns, Feinberg, and Budin (1958) measured 84 U.S. university students under different timing conditions. Those asked to take their time when giving their responses averaged 2.27 seconds, while pressed students took 1.60 secs, and pressed students who were also distracted took 1.55 secs. When 18 of the subjects from the 'take your time' condition were retested and paid for response times quicker than 1.20 seconds, their responses were significantly faster, from a mean of 2.31 seconds to 1.44 seconds. However, this was not significantly different from prior subjects in the pressured or pressured and distracted conditions. The nature of the associations also changed with more clang and repetition associations and fewer subordinate responses. Moran (1966) found that time pressure also affected idiodynamic sets: there were fewer set-representative associates and more associates of other set types, as well as an increase in communality.

Association Studies Of L2 Learners

As one can see from this review, association studies have mainly attempted to find answers to questions relating to L1 language development and the organization of the mental lexicon. Additionally, they have also been explored as a way to measure bilingualism (Macnamara, 1967). Association research has looked at non-meaning-based connections as well, with Forrester and Tambs' (1976) association norms for vowel sounds. There have also been a number of applications for association tasks outside the field of linguistics. For example, they have been used to measure attitudes in an indirect way (Szalay, Windle, and Lysne, 1970) and social isolation (Rierdan, 1980). The area of association research has focused on monolingual speakers, leaving the potential for use in L2 situations unexplored to a large degree. Let us now turn our attention to association studies focusing on L2 learners.

Research into L2 association behavior did not really get going until the 1960s and has not attracted anything like the attention given to L1 associations. There has always been a limited number of people working in the area and this continues to be true today. Probably the most influential current commentator on L2 association behavior is Paul Meara. When at Birkbeck College he and his students studied associations in a long-term research program called the Birkbeck Vocabulary Project. This work has continued with his postgraduate vocabulary network based in Swansea. In an interim report of the Birkbeck project's findings (Meara, 1983a) and a survey of the state of vocabulary acquisition (1980), he lists some of the major traits of L2 associations. Although L2 learners typically have smaller vocabularies than native speakers, their association responses are much less regular and often not of the type which would be given by native speakers. This is because L2 responses often include clang associations. Also, L2 subjects frequently misunderstand the stimulus words, leading to totally unrelated associations. Non-native speakers, like L1 children, tend to produce more syntagmatic responses, while native-speaking adults tend towards paradigmatic responses. In addition, L2 responses are relatively unstable. However, with increasing proficiency in the language, L2 responses seem to become more like those of native speakers. Meara (1983a) cites a short study which showed a gradual progression in the learning of the associations of words which were taught in a class, suggesting that association tests may be valuable in capturing the incremental nature of vocabulary learning.

Methodologically, Meara (1983a) suggests that using lists of high frequency words, like the ubiquitous Kent Rosanoff list, may not be appropriate for nonnative-speakers for several reasons. First, high frequency words tend to produce similar responses in the target and native language, making it difficult to know whether the answer is a direct L2 response to the stimulus word, or a translation from the subject's L1. Second, high frequency words are likely to be acquired at the beginning of a subject's L2 learning experience. Subsequent, more advanced learning may be accomplished in radically different ways, possibly leading to different patterns of association response. Third, since basic vocabulary is probably well-established in the learner's mental lexicon, more interesting research might be focused on newly-learned words that are only partially acquired.

These summaries by Meara capture the major findings of L2 associative research to date. The following are descriptions of studies which have either lead to Meara's conclusions, or have built upon them.

Riegel, Ramsey, and Riegel (1967) asked 24 American subjects who were L2 learners of Spanish and 24 Spanish subjects who were learners of English to give associations to 35 stimulus words, once in their L1 and once in their L2. The responses were restricted to seven categories: superordinates, coordinates, synonyms, contrasts, functions (*fork-eat*), attributes, and parts (*fork-handle*). In general, the subjects as L2 learners left more blanks than did the native speakers. The overall trend was for there to be more variability for Spanish responses than for English responses, as Spanish L1 subjects had a higher number of different responses than L1 English subjects for all categories. As for L2 subjects, the Spanish learners (Americans) had a higher number (except for 'parts'), but the difference was much smaller than for the L1 comparison. L2 learners had a higher degree of repetition of responses among the 7 tasks than did the L1 speakers; in other words, the L1 speakers repeated responses less, presumably because they have a larger repertoire of words to use. Finally, the authors suggest that associations may differ according to whether an L2 is acquired naturally mainly through exposure and use, or whether it is learned through formal training.

In an experiment studying intralingual vs interlingual associations, Riegel and Zivian (1972) looked at 24 American undergraduate learners of German. One would assume that since these subjects knew less German than English, that their response variation would be smaller for German as well. In fact they found that second language association responses were more varied than first language responses in a free-response task. In a restricted response task (similar to the one in Riegel, Ramsey, and Riegel (1967) above), there was little difference. There was less variability in interlingual association tasks than in intralingual tasks, because subjects tended to translate the stimuli or give associations with formal similarities. The variation was greatest for L2-L2 (German-German) associations and for L2-L1 associations in which the subjects had no knowledge of the L2 (French-English). In sum, the above two studies by Riegel and colleagues provide good evidence for the conclusion that even though L2 learners have smaller vocabularies, their associations are more varied than native speakers, presumably because the organization of their mental lexicon is less advanced.

Sökmen (1993) looked at the associations of 92 beginning, 59 intermediate, and 47 advanced learners of English. The majority of the responses (4,284/9049) fell into a category she termed 'Affective' which is not a usual classification in the literature, making it difficult to interpret the results. The next most common categories were Collocations, Contrasts, and Coordinates. Her results mirrored the normal finding that nouns tend to elicit nouns in a paradigmatic fashion (68%), while adjectives (61%) and verbs (59%) also tend to elicit nouns in a syntagmatic relationship. Sökmen compares the L2 associations with native norms, but much of the similarity is likely due to her use of Kent-Rosanoff stimuli, which tends to produce stereotyped responses in any case. As she is using native norms to judge the value of the nonnatives' associations, her paper would have been more illuminating if she had done this in more detail.

Söderman (1993) looked at Scandinavian second language learners to see if they would have a similar syntagmatic to paradigmatic shift as L1 children. She found that the number of clang and syntagmatic responses decreased and the number of paradigmatic responses increased as subjects had increasingly more exposure to the L2 through the various education levels: 7th form in comprehensive school, 2nd form at a gymnasium, and 1st year at university. Similar changes were also found in relation to proficiency level. Söderman believes that the shifts were caused more by an increased knowledge of the stimulus words themselves than by any increase in the level of general second language proficiency. The implication of this is that each word passes through the S→P shift at a timing independent of others in the lexicon, and that every person (including native speakers) has words at different stages of the shift depending on how well they were known.

Van Ginkel and van der Linden (1996) compared the associations of Dutch learners of French with native French speakers. Using a continuous word association task, they found that native French high school students gave more associations than Dutch high school students, Dutch university students gave more associations than Dutch high school students, and French university students gave more associations than French high school students. They conclude from this that there is "a correlation between the proficiency of the subjects and the number of association responses that they produce" (p. 31). There was also marginally more overlap between the associations of the most proficient L2 learners and the native speakers. A further conclusion was that there was a positive relationship between the variability of the association responses and proficiency. Unfortunately, weaknesses in their methodology severely limit what can be taken from this study. The researchers only used 5 stimulus words, the number of subjects used were far less than Deese's suggestion of 50, and the analysis of the number of associations given showed only three significant differences out of many possibilities. Also, there is no statistical analysis of the variability and overlap figures, which renders then unusable, since the differences are too small to assume any difference by eye.

Schmitt and Meara (1997) studied 95 Japanese learners of English and twice measured their associative behavior and knowledge of inflective and derivational suffixes for 20 verbs, once near the beginning of their academic year and once near the end. The subjects did not show very good control of the verbs' word associations, when compared to a native-speaking norm group. Even for verbs rated as known, the students as a group were able to produce only about 50% of the possible word associations. Word association knowledge was shown to correlate with suffix knowledge and with total vocabulary size. This was the first time that association 'nativelikeness' has been empirically shown to relate to other kinds of vocabulary knowledge. If it could be shown to relate closely with a number of aspects of lexical knowledge, then the argument for using association tasks as vocabulary tests would be strengthened.

Word association tasks have been considered as measures of language proficiency or lexical knowledge, but so far the results have been inconclusive. Kruse, Pankhurst, and Sharwood Smith (1987) elicited associations from 15 advanced Dutch learners of English via a computer program and compared them to the responses from 7 native speakers. The findings were that the association measure did not discriminate very well between the Dutch and native speaker results, did not predict general language proficiency at any great degree (best correlation was .576), and was not very reliable (r=.759, .658, .554). So the association test did not prove very successful, but this is hardly surprising considering the native norming group consisted of only 7 subjects and that they neglected to compare the association scores with the more obvious kind of linguistic knowledge - lexical knowledge.

Another attempt to explore the viability of association tests has been carried out by Read (1993, 1994). He chose a receptive form of association test as a means to limit the variability of associations which would have to be dealt with. In its initial version (1993), a stimulus word (which could a noun, adjective, or verb) was followed by eight words, four of which were associates and four which were not. The associates could be synonyms, collocates, or have an analytical component relationship (*electron-tiny*). The following is an example using the stimulus word *diffuse*:

diffuse

circulate	government	holiday	light
optional	scatter	tolerate	vague

(Read, 1993)

When piloted on 100 learners of English in New Zealand, the results showed that the format was reasonably reliable, and gave an indication of the overall knowledge of words on the University Word List, from which the stimulus words were drawn. However, guessing proved a problem. Taking into account the weaknesses highlighted by the 1993 study, Read revised the test to include only adjectives as prompt words, and divided the options into two sets: one paradigmatic and one syntagmatic. The format now looked like this:

sudden

beautiful quick surprising thirsty	change doctor noise school

(Read, 1994)

The new version was tried out on 38 learners and the associates score correlated with a traditional vocabulary matching test at .86, showing concurrent validity. In addition, 15 subjects were interviewed in depth and these scores correlated with the Word Associates scores at .76, although this was far lower than the interview-matching correlation of .92. It seemed easier to reach a criterion level for knowing a word on the Word Associates test than in the interview, which may indicate that the problems with guessing have not been solved in the new format. In the final analysis, the Word Associates Test probably needs more refinement before it can be recommended for general use, but it remains a fascinating attempt to measure the depth or complexity of vocabulary knowledge as well as the number of words known.

COLLOCATION

In the forty years since Firth (1957, p. 194) brought the idea of collocation into prominence, there has been an increased awareness of its importance as a part of lexical knowledge. However, most work on collocations have concentrated on finding and describing the collocational relationships between words; few studies have been carried out on how collocational knowledge is acquired. This has left the field in a state where we can describe collocational relationships to a fair degree, but are still at an embryonic stage when it comes to characterizing how they are learned or should be taught.

Defining Collocation

Perhaps the best way to start a discussion on the definition of collocation is with Sinclair's (1991) distinction between the *open-choice principle* and the *idiom principle*. The open-choice principle tries to cover the idea that language is creative, and in most instances there is a wide variety of possible words which could be put into any 'slot'. This is the traditional way of viewing language, and Sinclair states that "virtually all grammars are constructed on the open-choice principle" (p. 110). However, complimentary to this freedom of choice, he notes that language also has a systematicity which constrains vocabulary choice in discourse; constraints which the open-choice principle does not capture. To some extent this sytematicity merely reflects real-world phenomena: *fishing* is often done in close physical proximity to a *lake*, so the words expressing these concepts will naturally co-occur as well. But much of the systematicity is strictly linguistic: there is no reason why we do not say **to put something on fire*, but fluent members of the English-speaking speech community know that the appropriate phrase is *to set/start something on fire*. The idiom principle highlights the fact that there are regularities in which words co-occur with each other.

A good definition of collocation will describe those regularities. Most of the attempts have proved to be quite broad and general. One example is the definition proposed by Van Roey (1990, p. 46, quoted in Granger, forthcoming): "[collocation is] the linguistic phenomenon whereby a given vocabulary item prefers the company of another item rather than its 'synonyms' because of constraints which are not on the level of syntax or conceptual meaning but on that of usage." Cowie (1978) presents a somewhat more detailed description incorporating three kinds of variables he considers important in defining collation. The first concerns how institutionalized a collocation is, ie. whether it is generally regarded by native speakers to be acceptable. The second is whether a particular collocation can be explained by more general principles of co-occurrence (*flag down a bus/taxi/tram* [all are vehicles]) or whether it is unexplainable and idiosyncratic (*keep a weather eye on someone*). The third is how common the collocation is, in other words, how preferred it is over other possible choices. A later definition (Cowie and Howarth, 1995) has been retooled into four aspects, with a psycholinguistic element added. The first aspect remains the same: that collocations are familiar and are institutionalized among a speech community. Second, collocations are mentally stored and memorized. Third, collocations have only limited variability. If they had unlimited variability, then they would be totally unpredictable in future instances. Lastly, they may be either semantically transparent or opaque. Collocations that are semantically opaque may have no other reason for their 'connectiveness' than the chance way a language develops, and so can be unmotivated. (Note however that collocations which are now opaque may have once been transparent.) Cowie's efforts cover key issues, like 'standardness' and frequency of occupance, and map out the general idea of what collocation is. Unfortunately, they also lack the clear (numerical) parameters necessary for a truly operational and applicable definition. This seems to be true of other collocation definitions as well, which leaves the field with definitions useful for linguists and teachers, but inadequate for researchers wishing to carry out empirical studies. This deficiency will be illustrated in Chapter 5 where the lack of precision proved to be a problem in the development of a test of collocational knowledge.

Describing Collocation

Most authors agree that there are two basic kinds of collocations: grammatical/syntactic collocations and semantic/lexical collocations (Benson, 1985; Biskup, 1992; Bahns, 1993). Grammatical collocations are the type in which a dominant word 'fits together' with a grammatical word, typically a noun, verb, or adjective followed by a preposition. Examples are *abide by*, *access to*, and *acquainted with*. Lexical collocations, on the other hand, normally consist of combinations of two basically 'equal' words such as noun+verb (*ball bounces*), verb+noun (*spend money*), and adjective+noun (*cheerful expression*), in which both words contribute to the meaning (Benson, 1985). In addition to these two basic collocational categories, Allerton (1984) proposes a third, consisting of collocations which are not based on grammatical or semantic patterning. The relatively arbitrary prepositions attached to time fit in this category, since there does not seem to be any clear logical reason why we should say *at six o'clock*, but *on Monday*.

Cowie and Howarth (1995) suggest that collocations can be placed on a 4-level scale of complexity. It is illustrated in Figure 1.

Figure 1 Levels of Collocational Complexity

LEAST COMPLEXITY AND VARIATION

1. IDIOM

   fill the bill, shoot the breeze

2. INVARIABLE COLLOCATION

   break a journey, foot the bill

3. COLLOCATION WITH LIMITED CHOICE AT ONE POINT

   take/have/be given precedence [over noun phrase]
   give/allow/permit access to [noun phrase]
   have/feel/experience a need [for noun phrase]

4. COLLOCATION WITH LIMITED CHOICE AT TWO POINTS

   find/experience trouble/difficulty in [doing noun phrase]
   get/have/receive a lesson/tuition/instruction [in noun phrase]

MOST COMPLEXITY AND VARIATION

At Level 1, idioms are multi-word lexemes that have frozen collocation. If any variation is inserted into the idiom, it ceases to exist as a unit. Thus idioms are the least complex because they allow no variation. Moving to Level 2, the string is made up of individual words all contributing to the overall meaning, but the collocation is still fixed. This level is somewhat more complex because the meaning has to be composed from several lexemes, rather than coming from a single one as in Level 1. Level 3 has a slot which can be filled from a limited list of words, most of which are similar in meaning. This choice introduces both variation and complexity. Level 4 has two slots instead of one, adding increased variation and complexity. Thus each level contains increasingly more variation, and therefore complexity. Although they do not discuss the possibility, there could presumably be even longer strings of discourse with three or more slots. However, as the variability increases, the balance must shift from stretches of language with collocational ties to language which is increasingly creative.

At this point, it is appropriate to introduce computers and the large-scale corpora they have allowed to be compiled. These developments have driven collocation inquiry for the last 25 or so years, and have slanted it in a distinctly descriptive direction. Where before it was very laborious to manually count and tally the co-occurrences of target words, modern computer concordancing programs can quickly extract co-occurrences from a corpus and present the instances on a computer screen in such a way that collocation patterning is much easier to notice. This type of corpus research has given us insights into collocations, showing that words have greater collocational connections than ever expected before. Also these collocational connections stretch further afield, being part of the cohesion which tie sentences and larger pieces of discourse together. The fact that collocation is not always a local phenomenon has lead to a debate about how far away collocations can extend. Nattinger and DeCarrico (1992, p. 22), for instance, believe that we must look more than five words away to find every collocational relationship.

Another phenomenon brought to light by corpus research is that "words may habitually collocate with other words from a definable semantic set" (Stubbs, 1995). The words in these semantic sets may carry either positive or negative connotations. For example, Stubbs reports how *cause* typically collocates with unpleasant things like *problems, trouble, damage, death, pain,* and *disease. Provide,* on the other hand, collocates mainly with positive things like *facilities, information, services, aid, assistance,* and *money.* Using the collocate *work* with both of these words further illustrates the difference: *cause work* is usually considered a bad thing, while *provide work* is usually looked upon favorably. The term *collocational prosody* is used to describe this phenomenon. Stubbs also suggests that this aspect of lexical knowledge can be described quantitatively from the information extracted from very large corpora. Based on other estimates, he calculates that an average person will be exposed to about 1 million words per month.[4] Thus a large corpus, like the COBUILD corpus with 320 million running words (as of late 1996), would equal the number of words that a person might be exposed to in 26 years. Such a massive amount of linguistic data must exhibit the kinds of regularities (such as lexical prosody) found in language at large.[5]

A final point in the description of collocation is that it does not only occur between individual words. Nattinger and DeCarrico (1992) have been particularly influential in raising awareness that lexical *phrases* are an important part of language. One reason that collocation within multi-word phrases did not receive adequate attention until relatively recently has to do with computing constraints. The computer is an extremely useful tool, but it can only search where and for what it is told. This makes researching multi-word phrases more difficult than single words, because they may be too discontinuous to be noticed. With better concordancing techniques and more powerful software, we may well find that various idioms, lexical phrases, phrases, and other strings of language collocate with each other in a way similar to how individual words collocate (Ramesh Krishnamurthy, Jane Willis, personal communication).

Usage of Collocation

One might assume that nonnative speakers are less proficient than natives in collocation and so use less collocational connections in their discourse. Research into native and learner corpora by Granger (forthcoming) generally supports this, but she also found that it depends on the specific collocation. The French learners of English used some collocational combinations far more than native English speakers. Granger look at amplifiers like *very, completely, highly, dazzlingly* and *strongly*, and found that the nonnatives tended to overuse the ones which collocated most freely with a wide variety of words. She suggests this liberal use of 'all-rounders', especially *very*, is a 'safe bet' strategy designed to minimize error. She also found that most of the English collocations used by the French learners were congruent with collocations in their L1, and so might have transferred. This is exactly what Bahns (1993) suggests happens.

Granger went on to do an experiment examining the how salient various collocations were. She asked 56 French learners of English and 56 native speakers to complete a receptive collocation judgement task. For each of eleven amplifiers (*highly, seriously, readily, blissfully, vitally, fully, perfectly, heavily, bitterly, absolutely,* and *utterly*), a list of 15 adjectives were given and the subjects asked to circle collocationally-appropriate ones, while putting an asterisk next to a collocate if it was perceived as the most frequent in the group. The format was as follows:

readily	significant reliable ill different essential aware miserable available clear happy difficult ignorant impossible cold important
bitterly	significant reliable ill different essential aware miserable available clear happy difficult ignorant impossible cold important

The native speakers marked 384 collocates with an asterisk, while the nonnatives only marked 280, indicating the natives were much more confident in identifying the most common collocates. In addition, the French subjects marked a greater variety of adjectives than the natives, suggesting less discrimination in collocational judgement. Granger concludes that the most appropriate collocational combinations are often not very salient for nonnatives. Biskup (1992) cites three Polish references (Grucza and Jaruzelska, 1978; Marton, 1978; Arabski, 1979) in stating that "collocational errors constitute a high percentage of all errors committed by L2 learners" (p. 87). This lack of saliency may be one reason why nonnatives have so much trouble in this area.

Psycholinguistic Research into Collocations

In contrast to the considerable amount of corpus-based descriptive research being done (largely motivated by lexicographical concerns), research into acquisitional aspects of collocation is thin on the ground. A select bibliography from the International Symposium on Phraseology at the University of Leeds in April 1994 contains over 175 entries, but only a handful have titles suggesting a focus on the acquisition or learning of collocations. From the small amount of work in this vein, it seems that collocational links are "powerful and long-lasting" links between words in the mind (Aitchison, 1987, p. 79). In association tasks (described in more detail in the Word Association section), collocations are the second most common response type after coordinates (Jenkins, 1970). Aphasics retain collocational information fairly well, even though they have lost much of their other use of language (Goodglass and Baker, 1976). These collocational links do not seem to fade as people grow older (Howard, McAndrews, and Lasaga, 1981). So it does seem that collocations are a psycholinguistic reality and that the mind does organize words according to their collocational links to some extent.

Collocation knowledge seems to be relatively difficult to acquire however. Channell (1981) briefly refers to a study which suggested that even relatively well known words are underutilized collocationally. The advanced subjects in the study could recognize incorrect collocations for words they knew, but they failed to identify large numbers of acceptable collocations for those same words. One of the few other studies directly studying collocations was carried out by Bahns and Eldaw (1993) It showed that collocational knowledge lagged behind general vocabulary knowledge.

As useful as the above studies are, none really give much insight into *how* collocation knowledge is acquired. Cowie and Howarth (1995) do go as far as proposing an acquisition mechanism. They suggest that collocations are learned by extending from a known collocation to a new one, ie. extension by analogy (*acquire a language→acquire proficiency*, *adopt methods→adopt techniques*), but this is not always successful (*adopt an approach→*adopt* *problems, carry out research→*carry out principles*). In fact, assuming synonyms have comparable collocational behavior is inadvisable, because differing collocations is one of the main things which differentiate synonyms.

Cruttenden (1981) advanced the idea that language learning takes place in two stages: item learning and system learning. This means that items are often learned as wholes and then segmentalized into the components. During this segmentalizing, the systematic ties holding the sequence together become salient. Nattinger and DeCarrico (1992) use this approach in their discussion of lexical phrases (which are more idiomatic), but it can be used for collocations as well. (See Peters (1983) and Lewis (1993) for more on this process.) When nonidiomatic strings of language are decomposed, the individual words become apparent, but if the string is frequent enough, the learner will also sense that there are linguistic ties among the words. This emerging sense of linguistic systematicity among words in strings is another way to explain the acquisition of collocation.

Even if we do not know how collocation knowledge is acquired, it seems that this knowledge is something that separates native speakers from nonnative. In a study that has yet to be published, Levenston found that when native speakers responded to a completion task, they relied on collocational criteria to a large extent, while even advanced L2 learners were much less likely to respond with collocationally-based responses. On the receptive side of the coin, Levenston believes that while native speakers rely greatly on collocational information when recognizing words, L2 learners tend to use word analysis to guess meanings. He suggests that, for native speakers, "collocational links are probably so strong that they [can] override semantic analysis" (p. 18), making native recognition patterns (and therefore possibly lexicon organization) different from nonnative speaker ones. To the extent Levenston is correct, collocation is a very important factor in how words are processed (and perhaps stored).

Classroom Teaching of Collocation

This paucity of research into collocation acquisition leaves us in the unenviable position of addressing the problems associated with teaching collocations without the benefit of a firm theoretical foundation. The problems themselves are formidable. One of the main hurdles is the sheer number of collocational possibilities to deal with. It is probably impossible to reliably estimate the number of collocations for even a small part of the total English lexicon, say the 10,000 most frequent words. Still, the figure must be enormous. Bahns (1993) proposes one principled way to at least minimize this huge number. He suggests that learners are able to transfer directly-translatable collocations from their native language to the target language, making it necessary to address only the collocations with no direct translational equivalence. There is some evidence supporting this position. Farghal and Obiedat (1995) found that about 10% of the collocations given by their 57 Jordanian learners of English were transfers from Arabic. In addition, the subjects were much better at producing acceptable English collocations when they were predicable from Arabic, then when they were not predictable. Still, even if we disregard these transferable collocations, the remaining number must be unmanageable.

Another problem is that there seems to be no reason behind many of the collocational combinations, and so they need to be learned individually in a piecemeal fashion. The noticeable exception is when target lexemes collocate with members of a lexical set, as in the '*flag down* a *bus/taxi/tram*' example above. Cowie (1978) suggests using exercises that exploit the kind of item which have collocates clustering in some recognizable group or pattern. For example, students can be made aware that the object of *flag down* is usually a vehicle. Note also how this exercise not only attempts to make the collocational regularity salient, but also tries to make the learner aware of other, sometimes unusual, possibilities.

> i. Look at some of the nouns which can combine with *flag down* (=stop by waving with a flag or the hand):
>
> *flag down* ... a car, a bus, a taxi
>
> What do the nouns have in common? All are *powered vehicles, controlled* by a driver, travelling on *roads*, equipped to carry *passengers*. Now say which of the following could equally well combine with *flag down*:
>
> coach, hitch-hiker, tram, swimmer
>
> Say why you accept some words and reject others.
>
> ii. Try to explain why the following nouns, which are rather unusual choices for *flag down*, might still be used:
>
> lorry, steam-roller, water-bus, horse and carriage, elephant
>
> (Cowie, 1978, p. 42)

While it is very plausible that exercises like this may help students with collocational pairs that have a degree of regularity, it is difficult to imagine how they could work with the large number of combinations which are not so transparent, consistent, or predictable (*seal off a door* but **seal off a jar*). To my knowledge there is no estimate of what percentage of collocations are somewhat predictable like the ones above, but even if it is relatively high, there must remain thousands of unpredictable ones to bedevil language learners. Thus the problem remains: the type of unpredictable collocations which learners are likely to have the most trouble with are precisely those which present no obvious approach to teachers trying to help their students.

The majority of collocation exercises developed so far have been variations of a basic matching task requiring students to fit together collocates from various word groups, check or be given the correct answers, and then presumably memorize them for future use (Brown, 1974; Nation, 1994). If the students do not know the collocations, they have no option but to guess. These exercises are fine for *post hoc* learning, but it would be useful if some guidelines could be discovered which could address the dilemma stated in the prior paragraph. Such exercises can be particularly useful for advanced learners who already have partial knowledge of the words. Rudzka et al. (1981, 1985) and Channell (1981) used collocational grids to show common collocations for words in a semantic field, which may well help learners to differentiate between the near-synonyms on the grid. A problem is that grids large enough to cover the main possibilities may well be too large to be 'learnable'. However, it must again be noted that there is very little research evidence available as to whether even intuitively-appealing exercises like the one above actually work in the classroom. In fact, the key question remains unanswered of whether we as teachers should spend time on collocations at all, or whether we should be content to let them be learned implicitly through language exposure.

Another aspect that is difficult to teach is the 'strength' of a collocation. Not all combinations have the same level of exclusiveness; they can vary between being rather weak (*nice* can collocate with almost any noun to give a sense of pleasantness) to being very strong (*blonde* collocates almost exclusively with *hair*). The strength of collocations have normally been addressed by giving a brief explanation, but it has not been shown that this is the best way to handle the issue. The advent of downsized concordancing programs that will run on everyday computers offer another approach. It may prove most useful for students to explore concordances for themselves to discover which words collocate and to what degree. There has been some discussion about the merits of such a hands-on approach (eg. Willis, 1997), but hard evidence on this matter has yet to move through the pipeline into the published realm.

Previously, we mentioned Bahns' (1993) belief that L1 collocations can transfer into the L2. As a consequence of this belief, Bahns also suggests that collocation workbooks and other materials need to be written for specific L1s. If this is indeed necessary, it would make this kind of project far less attractive for publishers, who prefer to publish materials saleable in a number of countries. So do learners from different countries and L1s haver different collocation behavior? Biskup (1992) studied Polish- and German-learners of English to see if there were any differences in the way they handled English collocations. She found that there were. The Polish students had more 'correct' collocations, but also left more missing answers. Biskup concluded that the Polish students relied more heavily on their L1, while the German students looked for more creative strategies to form collocations, such as giving definitions ([target] *winding a watch* - [answer] *to make a clock working*). This suggests that learners from different countries and school systems may not only have different L1s, but also different attitudes and strategies towards language transfer. This supports Bahns's call for discrete materials. Unfortunately, it also means there is a tension between theoretical best practice and publishing reality.

Reference Materials and Collocation

If a learner wishes to obtain some information on a lexical item (in this case collocation information) and nobody is around to help, one important strategy is to appeal to reference materials. Of the various lexical reference possibilities, the dictionary must surely be the most popular (Schmitt, in press; Scholfield, in press). However, this does not mean that it is automatically the best choice. For one thing, it is not easy to provide good collocation information in a dictionary. (See Cowie (1981) for a discussion of the difficulties of inserting collocation information into dictionaries.) For another, learners may not be very adept at extracting the information that is there. Béjoint (1981) found that some learners tend to use dictionaries mainly for decoding information and largely ignore the encoding information, including that on collocation, which is already included in the dictionaries. This leads Cowie (1981) to believe that dedicated collocation dictionaries may be a better resource for learners. There are several of these available. *The BBI Combinatory Dictionary of English* (Benson, Benson, and Ilson, 1986) is mostly of use for grammatical collocations, since it does not include a great number of lexical collocations. Among the collocation dictionaries with a more lexical bent are the *Collins COBUILD Dictionary of Phrasal Verbs* (Sinclair, *et al.*, 1989), the *Longman Dictionary of Phrasal Verbs* (Courtney, 1983), the *Longman Dictionary of English Idioms* (Long and Summers, 1979), and the *Oxford Dictionary of Current Idiomatic English, Volumes 1 & 2* (Cowie and Mackin, 1975; Cowie, Mackin, and McCaig, 1983).

WORD MEANING

Core Meaning and Encyclopedic Knowledge

It seems fair to speak of word meaning as the primary kind of word knowledge, perhaps along with word form. Nevertheless, over the years, philosophers and psychologists have been perplexed in their attempts to understand and explain its nature. A number of commentators have made the distinction between some type of basic, fundamental meaning of a word and all of the other personal and cultural world knowledge relating to that word which a person might know. This distinction has been formulated with various terminology: extension meaning and intension meaning (Lyons, 1977), denotation and connotation (Hammerly, 1979), definitional information and contextual information (Stahl, 1983), and basic domain and abstract domain (McCarthy, 1990). Predating these, Katz and Fodor (1963) made the distinction with the terms core meaning and encyclopedic knowledge, which I shall use.

To evaluate this distinction, we must first examine the question of whether a word inherently holds any of its meaning within itself, independent of context. Previously, it was commonly believed (and often still is) that <u>all</u> of a word's meaning information was inherently held in this way, in a kind of 'container'[6] view of word meaning. This led to efforts to explicitly describe the presumed intrinsic meaning. One of the best-known methods involved 'semantic features', in which grids of distinguishing attributes of a word's meaning could be listed. Semantic feature grids proved effective in distinguishing the differences between similar words within a grouping (Rudzka *et al.*, 1981, 1985), (indirectly supporting the idea that words can only truly be defined in contrast to other words). Still, the ability to completely define words eluded a semantic features approach. Part of the problem lies in deciding which features are necessary and sufficient to define a word. This is perhaps not such a problem with narrowly-used and precisely defined technical words, but for the vast majority of words the boundaries of their word meanings are 'fuzzy' (Aitchison, 1987). For example, a word with a seemingly transparent meaning like *bachelor* could be defined as +human, +male, +adult and -married. A divorced middled-aged man with several children meets these criteria, but can he be considered a bachelor? For most people, the answer to that question would probably depend on other factors, such as his lifestyle. In reality, cases like this on the boundaries of word meanings are probably not that uncommon, and their categorization is at least partially determined by the context. This realization has led to the view that words have fuzzy meanings, uncharacterizable by a finite number of semantic features, which has lead away from the view that words have all their meaning held 'in the container'.

Still, words must have some meaning information inherently bound to them. This is proven by the simple fact that people are able to draw up some information about a word's meaning when it is presented in isolation. This information may well include a great deal of encyclopedic knowledge, but will almost certainly entail aspects of the basic, underlying core meaning, without which it would be impossible to connect with the represented concept. This core meaning is what is known in common by people using the word. There is empirical evidence that some form of core meaning is a psychological reality, and that it is learned before encyclopedic knowledge. Children learning their native language typically first learn words of low complexity, like proper names, which have a one-to-one referent, advancing later to more complex words. Words are usually learned first at a medium level of generality, then proceed toward both more specific and to more general class names. In the course of this learning, children start using words before they have full adult meanings (deVilliers and deVilliers, 1978). This developmental preference for words with an easily-discernable semantic core suggests the importance of core meaning in learning, making it likely that core meaning will be one of the first things learned when acquiring any word.

Second language studies also suggest that learners acquire the core meaning of a word before more figurative aspects. Ijaz (1986) found that the basic core meanings of polysemous prepositions were better learned by advanced nonnative-speakers than non-typical meanings. In addition, literal meaning senses were more attempted by L2 learners (see also Kellerman, 1978). Levenston (unpublished manuscript) found that Hebrew learners of English tended to know the literal meanings of compound words and transitive verbs which are used both literally and metaphorically, even if native-speakers used the metaphorical meanings almost exclusively. If learners didn't know the meaning of a compound word, they usually guessed a possible literal meaning. From this, he postulates that when learning polysemous words, literal meanings are learned first, and then gradually metaphorical meanings are mastered. Similarly, Lovell (1941) found evidence that knowledge of multiple meanings of a word is closely related to understanding the most common meaning of a word. In addition, core meanings are generally universal, as

Hammerly (1979) notes that they are likely to have equivalents across languages (English *milk* = white liquid produced by mammals to feed their young = French *lait*). Also, when learners do not know a word, they often resort to lexical simplification, simplifying down to core meanings (Blum and Levenston, 1978). All of these pieces of evidence suggest words have an intrinsic core meaning which is relatively universal and is likely to be acquired before other more metaphorical meanings.

Core meaning is only a fragment of the total meaning information people can know about words however. There is also the potentially vast store of meaning knowledge which derives from personal experience, knowledge of the world, and native culture. This encyclopedic knowledge may not be essential for a core definition of a word, but is necessary for an understanding of a word's full meaning. This encyclopedic knowledge must exist because words represent different things in different contexts. "To know a live word is to be able to use it or to understand it in situations in which the person has not experienced it before" (Lado, 1964, p. 118). This flexibility of meaning indicates that there must be some fluid meaning information in addition to core meaning. In the case of *bachelor*, for example, we may know not only the core attributes mentioned above, but also a host of other details: bachelors often have more disposable income than a married man, may like to go out on the town a lot, and may choose to live carefree lives. In fact, someone might have to go on at length to cover everything they knew about bachelors. The exact nature of the encyclopedic knowledge a person has about bachelors depends on their personal experience of bachelors and the culture lived in. Thus, contrary to core meaning, the encyclopedic knowledge a person has about a word is idiosyncratic, although people in a culture tend to have a great deal of encyclopedic knowledge in common.

From their review of a variety of research, Anderson and Nagy (1989) argue that word meanings cannot usually be contained by either a definition or a series of semantic features. Although these may give some sense of the word's meaning (basically core meaning), context plays a large part in 'filling in' the other information necessary to make use of the words. Thus, definitions must be too general to cover all the meanings of a word. They illustrate this with sentences that are uninterpretable without context, like "The haystack was important when the cloth ripped" , showing that words in a sentence cannot always be decoded from strictly intrinsic meaning properties. With the context "parachutes", the reader is able to find a context which is congruent with the sentence to be deciphered. Thus, the core meanings of a word are often not enough; the listener or reader needs to be able to use his/her available encyclopedic knowledge. Context can allow this to happen.

This is related to the idea that meanings act as the bridge between words and concepts. Hirtle (1994) points out that the referents we use for words are not the actual world reality, but rather, our experiences of that reality. These abstract concepts exist in our minds, and need to be labeled linguistically with words. In a similar vein, Drum & Konopak (1987, p. 73) note that there is no inherent connection between a word and its referent; the relationship is arbitrary until formalized by a culture. The spotted animal with a very long neck in Africa could have been called a *golf*, a *glisten*, or a *glabnab*; only consensus in the English-speaking discourse community that the word label for this animal should be *giraffe* gives this particular word any meaning.

These word labels break up our mental image of the world into usable chunks. In this line of reasoning, meaning is the set of parameters we use to place a piece of conceptual reality into categories. Words do not typically refer to individual cases of something, but rather refer to generalized categories, which are likely to contain more variation than a static set of definitional features can contain. Thus words

> provide linguistic labels for concepts, and meanings constitute the associative links between concepts and words. A word's *meaning* refers to the total communicative value expressed by the word (Leech 1974). The instances to which a word's meaning can be applied in linguistic usage define the word's *semantic boundaries.* Within the total linguistic/semantic system of the language a word's semantic boundaries are defined by the word's semantic relations to other words. These, in turn, are defined in terms of different degrees of similarity and contrast.
>
> (Ijaz, 1986, p. 403)

Lehrer (1985) notes that meaning can refer to unambiguous referents in the real world, thus being perception driven. But there are also many words that (although they can refer to real-world entities) draw most of their meaning from their relationship to other worlds in the lexicon. In this case meaning is relative. It can be flexible in describing real world phenomena, but the relationship does not change among words. For example, *cold* can describe a variety of absolute temperatures in the real world, but linguistically, *cold* will always be of a lower temperature than *cool.* In this way, meaning is also network driven.

It is worth noting that concepts consist of more than just representations of real-world phenomena. The concepts-in-theories view, discussed by Keil (1994), emphasizes that concepts are also mediated by an underlying intuition or 'theory' that "some features and relations [are emphasized] over others because of their fit within broader explanatory patterns" (p. 174). Thus we have a sense not only of the details of individual concepts, but also of how concepts in general work.

So it does seem fair to speak of the two kinds of meaning knowledge: core meaning and encyclopedic knowledge. But how does one distinguish between the two, especially with the difficulties of defining core meaning? The answer may lie with prototype theory. Instead of assuming that core meaning consists of a set number of attributes that are essential in defining a concept, we should consider whether the mind uses a prototypical 'best example' of that concept to compare potential meaning extensions against. Rosch (1975) found that people tend to have a good idea of what they consider the best example (exemplar) of a category. For instance, robins were considered the best example of a bird, because they represent the attributes people most commonly associate with birdiness: flying, laying eggs, building nests, and singing. Penguins and ostriches had enough of these features to be considered birds, although not typical ones. Prototype theory can explain how atypical cases can be still considered a part of a category. It should be noted that these experiments used category words like *furniture, fruit,* and *clothing*. However, there seems to be no reason why the idea of prototypes cannot be used for almost any other kind of word as well, since nearly every word has enough variation of meaning for it to behave like a category. For example, there are varying degrees of being *handsome* and different degrees of *walking* and different degrees of doing things *slowly*. Since any non-technical and non-proper noun word relates to an abstract category and not to an individual object, action, condition, or case, it is a category in a sense. Therefore, the core meaning is tied to the prototypical attributes of a word. The core meaning is made up of an interaction of these prototypical attributes, but not all of them are necessary. Wittgenstein (1958) introduced the useful metaphor *family resemblances*. Just as in a family there might be a cluster of distinguishing features (big nose, large ears, red hair), of which all members of the family have some, but no family member has all, a word has a cluster of prototypical core meaning attributes. A meaning sense of a word might have certain of these prototypical attributes, while other related meaning senses have others. These prototypical attributes define the underlying meaning of a word which can be expanded to more figurative senses.

The number of prototypical core meaning attributes are likely to be somewhat limited, but the encyclopedic knowledge for a word is open-ended. The limited nature of the core meaning makes it easier to describe the relationships and organization between words. Discussions of synonymy, autonomy, hyponymy, semantic fields etc. can be found in many places (Carter and McCarthy, 1988; Lyons, 1977; Cruse, 1986). Encyclopedic knowledge, being open-ended, is more problematical. To use this potentially vast amount of knowledge in real time online requires it to be organized and controlled in some way. This is especially true since the mind draws up more meaning information than what is finally used. An experiment by Swinney (1979) showed that when exposed to polysemous words, like *bug*, both the insect meaning and the spy listening device meaning were brought up to the subconscious. (See also Lackner and Garrett, 1972). Context seems to be what sets the parameters of what actually reaches conscious thought. Context works in two directions: it limits what encyclopedic knowledge is finally activated, and it enriches the gaps in our encyclopedic knowledge. Although, as mentioned before, meanings of words in isolation can be recalled, the Anderson and Nagy (1989) 'parachute' sentence shows that, in discourse, the individual information generated in isolation and then strung together does not always result in meaningful discourse. Context is necessary to activate the full resources of word meaning.

It seems that one way context exerts its influence on encyclopedic knowledge is via *schemas* (other terms used for this idea are *schemata*, *frames*, and *scripts*). Schema is knowledge of how things in a specific area of the real world behave and are interconnected, or in other words, a type of organization. Schema can either be activated by a word itself in isolation, or by the context created by groups of words together. Coady (1993) states that schema can be activated by sight vocabulary word-forms which are automatically recognized. This means that schema can be activated from a bottom-up direction, even though it is normally considered a top-down type of knowledge. When a schema is activated, say a skydiving schema, all the encyclopedic knowledge related to this area becomes available, even before the other words related to the schema are encountered. Once the context has activated a certain schema, the schema constrains how each word's core meaning can be utilized in metaphorical meaning and which parts of encyclopedic knowledge are allowed to remain activated. If there is not enough context to induce schema, then the mind must hypothesize a probable context, and the confirmation or refutal of this hypothesis is stored in encyclopedic knowledge for the next time the word(s) is encountered.

Perkins (1983) discusses schema within the framework of *semantic constructivity*. The basic idea is that text does not carry meaning in itself, but rather provides directions to the reader or listener about how they should construct the meaning out of their preexisting knowledge base. His adult ESL subjects exhibited evidence of semantic constructivity in a silent reading task, showing that nonnatives are capable of contributing positively to the reading process, and are not just passive recipients of knowledge. However, this contribution can be constricted by lack of language proficiency, L1 interference, lack of background knowledge, and a weakness in inferencing skills. Perkins believes that semantic content may be more important than syntactic content and advises teachers and materials writers to be aware that ESL students can actively use that semantic content to make inferences.

The term 'schema' can also be applied in a more restricted way to mean the kinds of knowledge about words which can facilitate their learning. Nagy and Scott (1990) distinguish three kinds of *word schemas*: metacognitive and metalinguistic knowledge, knowledge of morphology, and knowledge of the typical patterns of word meaning. They found that by the 7th grade, L1 students could discern plausible meanings for unknown words from implausible ones, and by university could recognize very subtle differences in the semantic content allowable for nouns and verbs. Nagy and Scott conclude that skilled readers use a "myriad of semantic patterns and regularities" (p. 125) in judging possible meanings of new words. Nagy and Gentner (1990) came to a similar conclusion in a slightly earlier study. In addition, Fisher (1994) focused on grammatical patterns, and found there was definite linking between the syntactical behavior of the verbs she studied and their semantic content. She claims awareness of these syntactical constraints helps shape the mapping of meaning onto verb forms.

The Acquisition of Meaning

The next question is how word meaning is acquired. It seems that a great deal of L1 meaning knowledge is transferred to the (roughly) corresponding L2 words. Ijaz (1986) found that the core meanings of words were transferred while nontypical meanings were not. Levenston (unpublished manuscript) postulates from his data that literal meanings are learned first, with metaphorical application meanings being mastered later. There is evidence that "automatic processing of word meaning develops relatively quickly in the course of second-language learning" (Goodman *et al.*, 1985, p. 116). This behavior mirrors what happens when children learn words in their native language: children easily learn the core meanings of words in a kind of 'fast mapping' of meaning, but it may take much longer to learn other more complex aspects (Carey, 1978). Similarly, Miller and Gildea (1987) observe that children acquire word meaning in two stages: a fast initial stage where novel words are fitted into categories, then a slower stage where the words within the categories are differentiated. Aitchison (1987) summarizes the process of meaning acquisition in L1 children. She believes there are three basic stages: 1) labelling - attaching a label (word) to a concept, 2) categorization - grouping a number of objects under a particular label, and 3) network building - building connections between related words.

Once the learner has acquired the prototypical attributes of core meaning, (s)he then learns from additional exposure to the target word in context how far the meaning can be extended and where the semantic boundaries are. This is an ongoing process, as each exposure to a novel usage of a word further defines the boundaries. Even very well-read adult native speakers are learning new meanings and applications for the words they know. But word boundaries are fuzzy, and until enough input is gained to clarify a reasonable boundary, errors can occur.

In defining the boundaries of word meanings, L2 learners resemble children learning their native language, in that L1 children also do not have the advantage of an innate understanding of word boundaries, but have to learn them over time. It seems that L1 children categorize novel words with others of a like kind, but the criteria for 'likeness' seem to change with a child's development. Initially, perceptional similarity (particularly shape similarity) is paramount, but gradually taxonomic relationships become more important (Imai, Gentner, and Uchida, 1994). Children often overextend their first nouns to things outside the category, such as using *dog* to refer to any four-legged animal. They may take as a kind of prototype the first or most frequently heard example of a category, then overextend various aspects of that prototype to different words. They may also underextend objects that are not typical members of a class (deVilliers and deVilliers, 1978). Of course L2 learners seldom over- or underextend with such basic words, because the underlying concepts have already been acquired. The point is that even adult L2 learners may have trouble initially setting the meaning boundary between two or more related words, such as *job*, *career*, and *vocation*.

Dagut (1977) points out that in many cases there is not a one-to-one correspondence between words in different languages. For example, it is often the case where one language will have one word for a semantic space, and another two or more for the same area (this can be exemplified by lexical gridding). So learners may need a fair bit of input to learn to apply the prototypical core attributes in the target language. Learners also have to learn enough additional features to disambiguate similar words which may have almost the same core meaning, but with subtle differences. Sonaiya (1991) contends that this is one of the most difficult and important aspects of learning new words, and that it is continuous (continuous lexical disambiguation). It is especially difficult when the L1 has different conceptual distinctions than the L2. Similarly, a study by Higa (1963) found words that were too close in meaning were often confused and were difficult to learn.

A great deal of the encyclopedic knowledge may also be transferred if both the L1 word and the concept are already known. L1 words presumably already give access to the schema organizing this encyclopedic knowledge; one way to harness this resource is to get L2 words to access the same schema. There is some danger in relying too heavily on already known encyclopedic knowledge, however. Where the core meaning of a word should be fairly universal, encyclopedic knowledge can be largely culturally-based, which may well lead to confusion and error. In many cases it will have to be adjusted to better match the common understanding of the discourse community the learner is trying to enter.

This discussion has suggested that much L1 vocabulary knowledge can be transferred over to L2 vocabulary knowledge. Strick (1980) goes a step further and hypothesizes that semantic development moves from native to L2 semantic structures in L2 learners. During this process, the learner seizes on perceptually salient features to make this shift, since (s)he is not yet able to grasp the higher-order dimensions like register. Later, with increased L2 knowledge, (s)he picks up on these factors. This is similar to children moving from perceptual to abstract representations, with semantic categories being constantly restructured. Again we see the incrementality of word learning.

We often speak of learning a new word as being a single unitary process, but there are several elements to the learning of a word depending on what the learner already knows, and to what depth the word is to be taught. Graves (1987) distinguishes six types of vocabulary learning in an L1 situation, but these should apply equally to an L2 context. They are 1) Learning to read known words, 2) Learning new meanings for known words, 3) Learning new words representing known concepts, 4) Learning new words representing new concepts, 5) Clarifying and enriching the meanings of known words, and 6) Moving words from receptive to productive vocabularies. One can see that Graves' list emphasizes the ongoing learning of a word, as all of these except (4) involve prior knowledge of either the word or associated concept. Graves' list appears in an article on vocabulary instruction, and such a list can remind teachers that learning a word is not a one-off operation which students either complete or fail, but rather an ongoing process where more attention and recylcing eventually lead to a deeper and richer understanding of the word.

But even large amounts of exposure does not guarantee that L2 vocabulary knowledge will reach native-like levels. For example, bilingual children growing up in a L2 immersion environment have massive amounts of exposure, but do not necessarily have the same vocabulary knowledge as L1 students. Verhallen and Schoonen (1993) knew from prior research that bilingual children in The Netherlands generally know fewer words than Dutch L1 children (citing research by Verhoeven and Vermeer, 1989). They studied 9- and 11-year old Dutch and Turkish children and found that the Turkish children knew fewer meaning aspects (definition, description, categorization, constituents, how something is used, etc.) for each Dutch word as well. If vocabulary size was controlled for, there was no difference, but that meant the best 11-year old Turkish students were being compared with the weakest 9-year old Dutch children. Longitudinal studies of this type of children would be interesting to see if they eventually do close the gap.

This survey of meaning has necessarily been a very brief one which could only touch upon a limited number of topics and studies. The reader should be aware that for meaning, perhaps more than any other type of word knowledge, the literature is massive. It would take a thesis in itself to try to cover and synthesize the work that has been previously done in this area. Among the many topics not covered here are the other factors which determine meaning, such as intonation (Mason, 1986), research into more precise ways of understanding key concepts commonly used in our field, such as synonymy (Miller and Charles, 1991), and pedagogical aspects of how to help the learning of meaning, for example, the use of mnemonics and the need to relate new lexical knowledge to known, established knowledge (Stoller and Grabe, 1993). It is hoped that these obvious gaps have not detracted from the limited discussion on meaning which space has permitted.

REGISTER

In a language like English, there is usually more than one word which expresses a concept. Collocation is one way of choosing from among these quasi-synonyms; register is another. Register, as it applies to lexis, is the cover term for all of the stylistic variations which color the basic meaning of a word. It consists of additional (meaning) information beyond the basic denotative meaning, which makes each word more or less appropriate for certain language situations or language purposes. In many ways, the idea of register is similar to that of connotation, the extra implied information which comes with a word in addition to its explicit meaning. Since the extra information can be of several different types, register is a somewhat broad category, as indicated by the variety of labels attached to it (register, stylistic constraints, appropriacy).

There have been several attempts to describe the different types of register variation.[7] Chui (1972) (cited in the better-known paper by Richards, 1976) suggests six areas where there can be register variation. *Temporal variation* covers the continuum of how old-fashioned or contemporary words are. Language is a living thing ever in flux, in which words are constantly falling out of use, while others are being created to take their place. Proficient language users sense this, and words which are archaic or becoming so gain a register marker in peoples' minds to signal the out-of-use status. On the other end of the spectrum, it is possible for words to have a current or cutting-edge feel, as *information technology* (*IT*) has for many people at the time of this writing.

A common language varies according to where it is spoken, in a kind of *geographical variation.* The variation can be divided among countries which speak the same language, in which case the variations are called *language varieties* (Indian English, Australian English). If the divisions are within a country, they are known as *language dialects.* (Wallace (1981) suggests that the next gradation down from country (*variation*) and region (*dialect*) should be situation variation, which he refers to as 'register'.) Geographical variation is probably not consciously manipulated, and is only noticeable when one is exposed to spoken or written discourse from someone outside one's immediate language community.

The third type of variation is *social variation.* It is said that people in privileged classes typically have a somewhat different lexis from people in less privileged classes. The amount of social variation will probably differ from country to country depending on the rigidity of the social class system and on perceptions of its desirability. Richards (1976) gives the example that members of privileged classes refer to a female as a *lady*, where otherwise she is referred to as a *woman* (although this example may be out of date).

Social role variation covers the role of power or social relationship between interlocutors, which directly affects the level of formality each uses. If one is speaking to a social superior, someone it is desirable to impress, or a stranger, polite deference is usually partially indicated by using more formal words (as well as more indirect syntactical structures) than one would use if addressing one's peers or friends. As everyone interacts with numerous people of varying relative power status, this implies that social role variation is routinely and consciously manipulated. Of course we know this is true, but if any evidence were needed that level of register can be mediated consciously, a study by Labov (1970) can be referred to. He found that the number of prestige pronunciation forms increased with the amount of attention. The amount of conscious control a person has is very likely to vary with type of register. Some register types seem to be largely unconscious and therefore likely to be less responsive to deliberate change in any particular situation (geographical variation), while others are obviously more amenable to conscious control (variation stemming from social role).

The topic being discussed can also affect the type of language used. This *field of discourse variation* stems from the fact that many fields have a genre, or expected style of discourse, which determines appropriate language use. This often concerns syntax (using passive voice constructions in academic discourse), but it also involves word choice constraints. In academic discourse, *we* is preferred to *I*. In addition, each field has its own technical vocabulary or jargon, whose use is expected, and whose non-use can be marked. Gregory (1967) suggests that each field has a set of technical words restricted to people familiar with that field, and that there are also non-technical words which are usable in many fields, but with different register ramifications in each one.

Chui's final register area is *mode of discourse variation*; that is, some words are more appropriate to written discourse than oral discourse, as the former is normally more organized and formal than the latter. (See McCarthy and Carter, in press, for an in-depth discussion.)

Gregory (1967) lists a similar set of areas of register variation: temporal, geographical, social, standard vs. non-standard dialects, and dialectal varieties. He makes the point that some of these types of variation are reasonably permanent characteristics of the user, becoming an idiolect, or one's personal idiosyncratic dialect.

Halliday (1978) developed a different description of the components of register variation, one which continues to be influential. He divided register into three basic types. First, there is *idiomatic* register. It deals with the connotative value inherent in the word. Second, there is the *interpersonal* register, which determines how the word is used in context. Third, there is *textual* register, which is the cohesion which binds a text together. It seems that only the first two categories describe psycholinguistic knowledge and processes behind register, while the third describes the physical manifestation (the text) of these processes and knowledge. As such, Halliday's textual register classification is really a redundant reworking of the inherent and idiomatic register knowledge descriptions.

A more widely-used element of Halliday's thinking is his description of the three components of register: *field, tenor,* and *mode*. They try to capture how vocabulary selection is constrained by the complex interactions between "the content of the message, its sender and receiver, its situation and purpose, and how it is communicated..." (McCarthy, 1990, p. 61). Field covers the content and purpose of a message, such as an owners manual explaining how to operate an appliance. Tenor refers to relationship between interlocutors, which is very similar to the social role variation discussed above. Finally, mode describes the channel of communication, that is, whether the message is spoken or written, and how it is physically transferred, eg. via telephone, novel or drum.

Halliday's description of register suggests that we can view register competency as a) knowledge of the various kinds of register marking that a word may have, and b) how to apply that knowledge of a word's register marking to achieve the effect one desires linguistically. To formulate this in a slightly different way, we could say the following. For every familiar word, language users know varying amounts of 1) the above kinds of register information for the word and what the normal applications of the word are, and 2) what the effect(s) are of using the word (with its register marking) in a number of different situations. People choose to use words with a certain type of register marking with the purpose of conforming to or diverging from their interlocutor's expectations.

Most of the time, one would choose to use words with the kind of register marking one's interlocutor expects, because this is the way to maintain communication. Benson and Greaves (1981) partially explain this by stating that in order to have text or communication, we must have a mutually understood field of discourse. Choice of lexis gives an indication of this field (eg. academic discourse or a car repair manual), by utilizing both lexical items which are particular to the field and more general words which have acquired a technical meaning in the field. If this flow of expected specialized vocabulary stops or is changed, then communication breaks down; if the flow is maintained, communication continues. Thus, maintaining register of the *field of discourse variation* type is an important support for continuing communication.

Benson and Greaves also maintain that an even more important way of signalling field is through collocation. Collocation is directly related to register because words with the same register naturally collocate together, forming a mutual interrelationship. So both register marking and collocational ties both help distinguish the topic field so necessary to understand communication.

Members of a language community are likely to have a similar sense of the various kinds of register markings for any particular word, because a consensus must exist for register to be a consistent, useful transactional device. There will, however, be individuals or subgroups which have a different sense of some of these register types for some words. When these words are used in ways perfectly normal to those individuals, the words can seem quite unusual or 'marked' to other members of the language community. These occasional instances of differing senses of register marking can lead to words being unintentionally marked in mild cases, and to misunderstanding and confusion in extreme cases.

Any individual word can carry a number of different kinds of register marking. For example *mosey* is not only old-fashioned, but is also restricted mostly to rural American usage. Different words also carry different levels (strengths) of register marking. Some are very highly marked (*shit, anon*), where others carry little, if any, marking at all. The amount of register marking is connected to the lexical specificity of the word (Cruse, 1977; 1986, p. 153-155). Neutral words are normally the ones with the most basic meaning. In other words, in a group of hyponyms or near-synonyms, the one with a meaning closest to the core meaning of the relevant concept and used at the level of maximum utility will have the minimal amount of register marking. (See Carter, 1987, Chapter 2, for more on coreness and Brown, 1965, for maximal usefulness.) As one moves up or down the hierarchy from the most basic, frequent, usable item, words acquire greater and greater amounts of register marking. For example, *glass* is a neutral, general term. If we become more specific by speaking of its subordinates, like *champagne glass*, or *flute* or *goblet*, we start to gain register marking. Likewise, going in a superordinate direction, toward an item like *drinking vessel*, also increases the marking, as the situations where this term would be naturally used are more restricted. *Offspring↔child↔kid* is another good example.

If we think of register as stylistic constraints, a term it is often referred to as, we can visualize how this works. The core word of the group can be used in the greatest number of situations, but as hyponyms and near-synonyms have increasingly greater levels of register information attached to them, the possible situations where they can be used appropriately decrease accordingly. Thus specific situations require specific vocabulary; and the register information attached to words allow language users to select the best word for each situation.

This view of register as a continuum between the basic core word and the most highly marked specialized word is given some support by research into newspapers. Wallace (1981) examined several factors which affected stylistic differences between sports and news stories in a prestigious and a local paper. He found that no individual feature was always represented, but that it was the cluster of features that determined register. Thus, since register is indicated by features which are typical, but not necessary, there can be varying amounts of register information, making overall register vary along a continuum. This aspect of register parallels word meaning, where there is often no single feature which can separate one conceptual category from another, as in Wittgenstein's (1958, p. 31-34) analysis of games. Likewise, James Hampton (personal communication cited in Miller and Johnson-Laird, 1976) found there was not a single feature which separated 'furniture' from everything else.

Greater register marking may also serve an interactional purpose; Robinson (1988) believes more specific or marked words show liking, 'immediacy', and willingness to continue conversation.

i. Tom let me drive his new *Aston Martin.* / car.

ii. A. I thought the film was good tonight.

B. Yes it was *fantastic.* / nice.

(adapted from Robinson, 1988)

These examples illustrate how the more marked option (in italics) indicates greater involvement and interest in the topic at hand.

The upshot of this is that register is a complex set of information which is affected by a number of different factors, among them: what subject field is being discussed, who one's interlocutors are and what their social relationship is to the speaker or writer, whether the discourse is spoken or written, and what purpose the speaker or writer has in mind. If the speaker or writer is competent, then (s)he will judge the situation and select the word from a group of known hyponyms or near-synonyms which he believes will have the desired effect. When a person is not concerned with register considerations, such as in an informal conversation, then the choice will tend towards the unmarked core words. Register becomes more salient, however, when there is a specific purpose to be achieved by the communication, in an interview, or writing an academic thesis, for example. But in all of these cases, there are lexical choices affected by register constraints, whether they are conscious or not. There are some register aspects which may be almost always totally unconscious, such as geographical variation, but these still carry register information, and may become conscious as the person gains more exposure to a different norms of discourse (living in a foreign country, for example).

A last point is that just as register is determined by a number of factors, it can be indicated in a number of ways. Although this discussion has focused on a lexical perspective of register, it is important to remember that register can be indicated by more than just word choice. Silva and Zwicky (1973) suggest that formality is obtained by at least three factors: 1) syntactic choice, 2) lexical choice, and 3) phonological choice.

GRAMMATICAL KNOWLEDGE

Lexical and grammatical knowledge are inextricably interrelated and perhaps one day we shall discover that they are actually a unitary construct. The sharp dichotomy between the two that has typified much of linguistic discussion must be seen primarily as an expedient for splitting an unmanageably complex notion (knowledge of a language) rather than any proof of their heterogeneity. Bearing this in mind, the expedient does allow us to consider language in more digestible chunks. Anyone conversant with Applied Linguistics realizes that the analysis, description, and pedagogical aspects of grammar have been given a prominent place in the field. In what has become something of a cliché, many commentators have noted that vocabulary is now also receiving a great deal of attention. When lexis and grammar are considered in conjunction, two of the most obvious aspects which emerge are word class and morphology. Because of space considerations, this review will concentrate on only these two grammatical aspects, and not attempt to cover other grammatical notions like countability or valency.

Word Class

Word class describes the category of grammatical behavior of a word. There are a number of potential word classes, but the majority of language research has concentrated on the four major categories of noun, verb, adjective, and adverb. The results from a number of studies suggest that certain word classes are easier to learn than others. In an early study, Morgan and Bonham (1944) looked at nouns, verbs, pronouns, adjectives and adverbs, interjections, and a combined group consisting of prepositions, articles, conjunctions. These were learned by 148 seventh- and eighth-grade junior high school students, who viewed cards containing twenty foreign words (language: Ru-Ro) and their meanings in English which were randomly selected from the pool of words mentioned above (three nouns, pronouns, verbs, adjectives, adverbs and prepositions with two interjections). The results showed that nouns were clearly the quickest to be learned. There were few other statistically significant results, but the ones there were indicated adverbs as being the most difficult part of speech. The subjects in Phillips' (1981) study learned nouns better than verbs or adjectives, but the difference decreased with the increase in the learners' proficiency. For subjects learning Russian-English pairs of words, pairs in which the Russian word was a noun or an adjective were easier to learn than pairs in which the item was a verb or an adverb (Rodgers, 1969).

Thus it would appear that nouns are the easiest word class, adverbs the most difficult, with adjectives and verbs occupying the middle ground. However, Laufer (in press) presents a counterargument, suggesting that the apparent word class effect might be due to other factors. The least-learned verbs and adverbs in Roger's study were more complex morphologically than the better-learned nouns, which may have caused the results instead of word class. Laufer also cites the indirect evidence that Odlin and Natalicio (1982) did not report differences in word class learning in their study of semantic vs. word class knowledge. This lack of reportage does not directly support a 'no word class difference' position, but it is difficult to imagine that the authors would not have at least mentioned a word class difference if one had appeared. Laufer concludes that there is no clear effect of word class on the ease or difficulty of learning a word.

Regardless of whether any particular word class is easier or more difficult than others, there does not seem to be any doubt that word class is involved in the learning and storage of vocabulary. Aitchison (1987, Chapter 9) surveys the psycholinguistic evidence for this statement. She notes that when malapropisms are made, the errors almost always retain the word class of the intended target word.

> I looked in the *calendar* (catalog).
> The tumour was not *malicious* (malignant).
> That model is *extinct* (obsolete).
>
> (Aitchison, 1987: 99)

Similarly, 'tip-of-the-tongue' guesses also tend to retain word class. This suggests that words from the same word class are closely linked, with nouns having the strongest affinity. In contrast, words from different word classes are relatively loosely linked. Certain aphasics retain their use of nouns, but are largely unable to utilize verbs, indicating at the very least that nouns and verbs are stored somewhat differently.

Another kind of evidence appears consistently in association studies in the form of the paradigmatic shift, where children tend towards associations of different word classes (syntagmatic), while adults typically have associations of the same word class (paradigmatic). Brown and Berko (1960: 3) illustrate this in an example table taken from Woodworth (1938).

Stimulus	Response	1000 Children	1000 Men and Women
table	eat	358	63
	chair	24	274
dark	night	421	221
	light	38	427
man	work	168	17
	woman	8	394
deep	hole	257	32
	shallow	6	180

With the paradigmatic shift as a point of departure, they studied the associations of words from six word classes (count nouns, mass nouns, adjectives, transitive verbs, intransitive verbs, adverbs) along with a word class awareness test. Two sentences were read indicating the class of a word and subjects were asked to create a new sentence using the word, which would indicate if it was used in the correct class. Both the paradigmatic tendency of the associations and the scores on the awareness test increased with age and were closely related to another. This suggests that the syntagmatic→ paradigmatic shift may well be a manifestation of a developing appreciation of syntax. Certainly the exposure to grammar and word class categories would increase with age.

So knowledge of word class seems to improve with age in the L1, but what of L2 word class knowledge? The common assumption seems to be that word class easily transfers from the L1 to the L2. Odlin and Natalicio (1982) established that this is not necessarily true. They found that intermediate and advanced ESL students did not always know the word class of words they knew the meaning of. They claim this suggests that "acquisition of the semantic content of target language words does not always entail acquisition of the target language grammatical classification of those words" (page 35). A follow-on study showed that words with an ambiguous classification, ie. *drink* and *correct,* were less likely to be correctly identified for word class than unambiguous words, even if the ambiguous words were cognates. However, when errors were made, they tended to fall into one erroneous word class, rather than be distributed among all the possibilities. They give the example of *I want coffee* <u>*without*</u> *cream.* Most of the incorrect responses identified *without* as an adjective, but few as a noun or verb. The good news to come out of the study is that both the intermediate and advanced students were able to identify the word classes by name about 75% of the time, although it must be noted that the target words were of a relatively high frequency. The upshot is that non-beginner L2 learners are likely to know the word class of at least the more frequent words, but there is also the possibility that they will not know it even if they know the meaning. Also, we cannot assume that the word class of cognates will transfer, especially if the word fits into two or more grammatical classifications in English.

Morphology

Whereas Laufer (in press) suggests that it is not clear how word class affects vocabulary learning, she identifies morphology as an aspect with definite effects. They mainly have to do with the complexity vs. simplicity of affixes and also their transparency. Laufer believes that inflectional features, such as irregularity of plural, gender of inanimate nouns, and noun cases, cause an word to be more difficult to learn than if it did not have such complexity, because these features lead to a greater number of forms which must be learned. She cites a Ph.D dissertation by Stock (1976) which claims that learning inflections is one of the biggest problems for English learners of Hebrew, particularly irregular verb inflections and the inanimate noun/adjective endings for masculine words.

As for derivational affixes, if they are transparent this can facilitate learning. For example, if students know the meaning of *-ful*, it should not be too difficult to recognize the meaning of new words *useful* or *careful* if the base forms are already known. However, a lack of consistency can cause problems even if the affix is transparent. Someone having a special skill is a *specialist*, a person who is pragmatic is a *pragmatist*, but a person who acts on stage is an *actor* not and **actist*. This is more likely to cause problems in production than comprehension, although learners do confuse affixes receptively, eg. thinking *cooker* is a person.

Of course morphology includes more than just affixation, and Laufer points out a similar problem in word compounding. In what she terms 'deceptive transparency', words consisting of apparently meaningful and transparent parts can cause considerable confusion for unwary learners. She gives the example of *outline*, where *out* does not mean *out of*. Yet many students in her experiments (Laufer and Bensoussan, 1982; Bensoussan and Laufer, 1984) interpreted *outline* as *out of line*. Other confusing words were *discourse*, glossed as *without direction* and *falsities* as *falling cities*. The learner assumption in these cases is obviously that the word equalled the sum of meanings of its components. While making this assumption can be a useful strategy in many cases, with deceptively transparent words, it unfortunately leads to incorrect guesses. Unsurprisingly, in a later experiment Laufer (1989) found that subjects made more errors with deceptively transparent words than with non-deceptively transparent words. So words which look simple to analyze, but are in fact not, seem to be more difficult to learn.

Acquisition of Morphology

Nagy and Anderson (1984) estimate that the average upper elementary school student encounters a minimum of 1,000 new word families per year while reading, with better readers being exposed up to a possible 3,000 - 4,000 new vocabulary items. Nagy, Diakidoy, and Anderson (1993) estimate the figure to be closer to 10,000 new words per year. This is a large number, but Nagy and Anderson point out that most are related to words already known through prefixation, suffixation, or compounding. Thus a knowledge of morphology is crucial in handling the flood of novel words being met at this time. But how and when do L1 children gain this morphological knowledge? To some extent it depends on the type of morphology. Inflections and compounding seem to be acquired before derivational suffixes (Berko, 1958). In fact, students are able to recognize the stems within a suffixed word before they understand the function of the derivational suffix itself, with this ability more or less acquired by Year Four (Tyler and Nagy, 1989). On the other hand, knowledge of the suffixes continues to advance into high school (Tyler and Nagy, 1990; Nagy, Diakidoy, and Anderson, 1991). One obvious reason why inflections are learned before derivational suffixes that inflections are more rule-based and consistent. But another reason is that derivational suffixes are more common in the written mode than the oral mode, and are particularly associated with formal and academic discourse (Chafe and Danielewicz, 1987).

Therefore children simply have more exposure to inflections than derivational suffixes. It seems that as their exposure to written language increases, so does their knowledge of derivational suffixes. Nagy, Diakidoy, and Anderson (1993) ran a study which isolated knowledge of derivational suffixes from other confounding factors, such as test-wiseness and reading ability, and found that most of the improvement in suffix knowledge came between Year Four and Year Seven, although it continued to improve into high school. But even by high school, the student's knowledge of the suffixes was not complete, as they still were not able to judge the word class of suffixed items as well as stem items. As with other kinds of language knowledge, morphology seems to be incrementally acquired, and may not be fully mastered until quite late (or perhaps never?).

A somewhat similar situation appears in an L2 context. Hancin-Bhatt and Nagy (unpublished manuscript) studied 196 Latino bilinguals in Years Four, Six, and Eight. They found these students also performed better on inflected than derived words, but that the differences decreased with age. The L2 students rapidly improved their ability to recognize cognates between Year Four and Eight, especially cognate stems in suffixed words. Similar to L1 students, they developed the ability to recognize familiar stems in derivatives before they knew what contribution the suffix makes to the suffixed word. Indeed, the students' knowledge of English suffixes and their relationship to Spanish suffixes was quite low in the youngest group, and even though it did improve over time, it was still low by Year Eight. Hancin-Bhatt and Nagy note that "Students simply do not see all the possible cognate relationships in suffixes" (p. 30).

It has been commonly assumed that once a stem word is known, its inflections and derivations can then be learned with minimum effort. This is exemplified by Bauer and Nation's (1993) vision of word families which states that "once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort". This intuition of the connection between knowledge of a base word and ability to learn other members of its word family does has some empirical backing. In a laboratory setting, Freyd and Baron (1982) found that subjects learned nonword paired associates (*skaffist-thief*) faster when they had previously been exposed to the meaning of the nonword stem (*skaf-steal*). Sandra (1988) found that it did not make any difference whether Dutch learners of English were explicitly told to notice the stems of transparent suffixed words while learning them or not; their level of retention on an immediate and one-week delayed translation test was the same. This suggests that her adolescent L2 learners unconsciously relied on morphological information to learn the novel words. However, when the meaning of the stem to the overall suffixed word meaning was not transparent, cluing subjects in to the connection helped them to learn the suffixed word (Sandra, 1993, cited in Sandra 1994). So it seems that knowing the stem or base word does help facilitate the learning of its derivations.

This does not mean the we can assume that acquiring the derivations is easy or that it will be accomplished almost automatically, however. Indeed, we have seen that even native speakers do not have full mastery over morphology until at least high school (Nagy, Diakidoy, and Anderson, 1993). If it takes that long to develop in natives, with their advantage of maximum exposure, then L2 learners are likely to have difficulties in their morphological acquisition. A study exploring the suffix knowledge of intermediate Japanese learners of English showed that they were able to produce, on average, only 60% of the possible inflected forms of base words (Schmitt and Meara, 1997). Furthermore, they were largely unable to produce the derivations of the base words, with only 15% of the possibilities being given. On a recognition task, they did better, but still generally recognized much less than 50% of the legal derivational suffixes as being allowable. Schmitt and Meara concluded that their subjects did not have very good mastery of suffix knowledge in general, even for words which the Japanese learners rated as known.

Processing of Affixes

Aitchison (1987) briefly surveys the literature and concludes that how the mental lexicon handles affixes depends partly on what kind of affixes they are. Inflections generally seem to be added to base forms on line in the course of speech (see also Prasada and Pinker, 1993). The exceptions are words which are most commonly used in their inflected forms, such as *peas* and *lips*. These words may be 'welded together' and stored as wholes as a result of massive exposure. On the other hand, derivations seem to be stored as single units (*resentful*), which can be analyzed into their components, (*resent* + *ful*) if necessary. As for prefixes, Aitchison concludes that, if they are obligatory (*rejuvenate*), they are stored as part of the word. Non-obligatory prefixes (*unhappy*) probably are as well, or there would be more cases of prefix errors (**dishappy*, **nonhappy*). Also, if cohort versions of word search are correct, the beginnings of words need to be consistent and reliable, suggesting the fixation of prefixes.

Sandra (1994) asks the question of why morphology exists in the first place.

The two main positions are that morphology allows the lexicon to store words most efficiently. The mind would have to store only a limited number of morphological components, which could then be united into a very large number of words (the lexicon would have to store all these individually otherwise). The other position argues that the decomposition of words aids processing efficiency. Taft and Forster (1975) proposed the prefixes are stripped off the stems, which would allow quicker selection of the word. This is because in the lexical search model advanced by Forster (1976), groups of words with similar orthographic shape would be available in 'bins', from which the final selection would be made. Because bins would probably contain words with the same initial letter pattern, prefixed words (*review*) would find themselves with a massive number of other prefixed and nonprefixed words (*remind, recur, remember, restaurant, remedy, repertoire*). Stripping off the prefixes to the stem would generally result in bins containing fewer candidates. Sandra discusses the weakness of both these positions and, noting the research discussed three paragraphs above, presents an interesting hypothesis that the main purpose for morphological relations is to facilitate the process of lexical acquisition. (For more on processing aspects of morphology, see *Language and Cognitive Processes* 9, 1994).

Relative Difficulty of Various Affixes

The idea that morphemes might be learned in a fixed 'developmental order' (thus implying individual degrees of difficulty) took off with the 'morpheme studies' of Dulay and Burt (1973, 1974). They studied about 250 Spanish- and Chinese-speaking children aged six to eight and found a common morpheme order in their speech when it was elicited by a picture-cued sentence-repetition task (Bilingual Syntax Measure). These results were broadly confirmed by Larsen-Freeman's 1975 study of L1 adults. But it was difficult to set the order for individual morphemes, and Krashen (1977), after reviewing more than a dozen studies, hypothesized that the morphemes clustered into the following levels:

ACQUIRED EARLIER

-ing
plural
copula

↓

auxiliary
article

↓

irregular past

↓

regular past
3rd person singular
possessive

ACQUIRED LATER

However, the methodology of the various morpheme studies was soon called in question, casting some doubt over whether the order existed at all (see Long and Sato, 1984, for a review). But the studies continued on, using more advanced techniques, and by 1991, Larsen-Freeman and Long had concluded that there was simply too much evidence of some kind of ordering to be ignored. At the same time, results from studies carried out by Pienemann and his colleagues suggested that the underlying basis for the ordering was cognitive processing constraints (eg. Pienemann, 1984).

Bauer and Nation (1993) use linguistic instead of acquisitional criteria to inform a hierarchy of affixes. They focus on the ease or difficulty of understanding the words containing these suffixes when encountered in written texts. To order the affixes in a principled manner, a set of eight criteria was used.

1. Frequency - the number of words in which an affix occurs. Affixes categorized into lower levels occur in a very large number of words.

2. Productivity - the likelihood that the affix will be used to form new words. Lower level affixes will be highly productive,

3. Predictability - how predictable the meaning of the affix is. Early level affixes will be highly predictable.

4. Regularity of the Written Form of the Base - does the written form of the base word change when the affix is added? Affixes at the beginning levels will leave the base word intact and recognizable.

5. Regularity of the Spoken Form of the Base - the same as (4.) above, but dealing with the base's phonology.

6. Regularity of the Spelling of the Affix - lower level affixes will have one form (*-est*), while upper level ones may have several (*in-* → *in-*, *im-*, *il-*, *and ir-*).

7. Regularity of the Spoken Form of the Affix - the same as (6.) above, but dealing with the affix's phonology.

8. Regularity of Function - the degree to which the affix attaches to a word of known part of speech and produces a word of known part of speech. For example, *-ion* always attaches to a verb and always produces a noun.

Applying these criteria to English affixes results in the following groupings.

Level 1. **Each form is a different word**. Each derivative is counted as a separate type.

Level 2. **Inflectional suffixes**. Base words and their inflections are considered part of the same word family. Affixes include the plural, third person singular present tense, past tense, past particle, *-ing*, comparative, superlative, and possessive.

Level 3. **The most frequent and regular derivational affixes**. The eight criteria above are strictly applied to derivational morphology. The affixes include *-able, -er, -ish, -less, -ly, -ness, -th, -y, non-*, and *un-*.

Level 4. **Frequent, orthographically regular affixes**. The criteria are prioritized to give more weighting to frequency and written form. The affixes are *-al, -ation, -ess, -ful, -ism, -ist, -ity, -ize, -ment, -ous*, and *in-*, all with restricted uses.

Level 5. **Regular but infrequent affixes**. These affixes are not general enough to add greatly to the number of words that can be understood. They include *-age, -al, -ally, -an, -ance, -ant, -ary, -atory, -dom, -eer, -en, -ence, -ent, -ery, -ese, -esque, -ette, -hood, -ian, -ite, -let, -ling, -ly, -most, -ory, -ship, -ward, -ways, -wise, anti-, ante-, arch-, bi-, circum-, counter-, en-, ex-, fore-, hyper-, inter-, mid-, mis-, neo-, post-, pro-, semi-, sub-,* and *un-*.

Level 6. **Frequent but irregular affixes**. These affixes cause major problems in segmentation. Some of these affixes are already listed above; those can be considered the transparent cases, while these are the opaque cases. They include *-able, -ee, -ic, -ify, -ion, -ist, -ition, -ive, -th, -y, pre-* and *re-*.

Level 7. **Classical roots and affixes**. Bauer and Nation do not deal with these roots and affixes, except to suggest that they should be explicitly taught to learners, and to note that many frequent English prefixes belong here, such as *ab-, ad-, com- de-, dis-, ex-,* and *sub-*.

Bauer and Nation have succeeded in creating a useful hierarchy of affixes, but we must be careful not to assume that their ranking of affixes (achieved on solely linguistic criteria) necessarily translates directly into a psycholinguistically- valid statement of learning difficulty. Indeed when Schmitt (1995b) compared his learners' knowledge of affixes with the Bauer and Nation hierarchy, there was little correspondence other than inflections being easier than derivations. What Bauer and Nation have achieved is a useful reference which can help to bring some standardization to work dealing with affixes, thus leading to better comparability between studies. For example, it could help resolve issues of vocabulary size, by helping to standardize the definitions of what constitutes a word.

ORTHOGRAPHICAL KNOWLEDGE

In 1983(b), Meara highlighted the "assumption that many of us make implicitly when we do empirical research on reading in second language: the assumption that we can safely ignore lower level processes and isolate higher level processes easily and simply" (p. 105). Quite recently, there has been an increasing awareness that orthographical knowledge, traditionally considered a 'lower level' type of knowledge, is a key component to both vocabulary knowledge and language processing in general. This awareness comes from both studies which have shown that the eye fixates on most words in a text and from psychological research which has shown the complexity of orthographical decoding. In fact, in the two major collections focusing on vocabulary to be published in 1997, each contains a chapter dedicated solely to orthographic matters (Coady and Huckin, 1997; Schmitt and McCarthy, in press).

Results from reading research have been particularly instrumental in showing the importance of orthographical word form. One illustrative example is research by Huckin and Bloch (1993). The most common cause of unsuccessful guessing from context in their study was that their subjects mistook unknown words for known words which were similar orthographically (eg. *optimal:optional*). Even if the context did not support such erroneous guesses, the subjects often persisted with them all the same, supporting Haynes (1993) assertion that word-shape familiarity can often override contextual information.

What Visual Information is Utilized When Reading?

Perhaps the most basic question that can be asked about orthography is what visual information can be utilized during the reading process. Three possible categories have been identified. The first is the outline of the shape of a word (configuration).

Besner and Johnston (1989) review the literature and conclude there is no strong evidence to support the idea that this kind of visual information is normally used in word identification. Second, there is the entire set of visual features of a word which make up the word-specific visual pattern (WSVP). Although not ruling out the possibility, Besner and Johnston claim that there is no real evidence that WSVP is used either. By the process of elimination, this suggests that the third possibility, use of the individual component letters in a word, is primary in word identification. This is as one might expect, and it has even been shown that the first few letters of the word following the one being focused upon is preprocessed through parafoveal vision (see below).

The use of non-letter identification processes should not be written off completely however. McClelland (1977) found that the use of different scripts had a short-lived effect on word recognition. Eight U.S. undergraduates practiced learning meanings to sixteen nonwords, which were shown either in capitals (BARDREL) or in cursive (*bardrel*) script, but not both. The subjects then had to categorize the words according to meaning category. They reacted more slowly to nonwords presented in the script different from the one they practiced in, but this effect faded quickly after the first exposure, and had disappeared after from 2-6 exposures. McClelland concludes that the global shape configuration of a word is used in addition to pure letter analysis.

Rayner and Hagelberg (1975) studied the word recognition cues for kindergarten and 1st graders. They showed nonsense trigrams and quingrams to the subjects and asked them to choose a trigram or quingram which looked similar to the prompt from either 6 options (trigrams) or 10 options (quingrams). The various options consisted of nonsense clusters systematically varied to correspond with the prompt. One option retained the same first letter as the prompt, but with all other letters different, the second option retained the second letter, etc. In addition, each of the above options were given in one form which replicated the shape of the prompt, and one which differed from the prompt. An example will illustrate this (bold added here to highlight the salient features; note also that the third letter consistently has a 'down' profile and the fifth letter an 'up' profile on the *same shape* options, while the *different shape* options have these features in different positions or not at all):

<u>Prompt</u>	<u>Same Shape</u>	<u>Different Shape</u>
mogad	**m**cqeb	**m**wnir
	n**o**qeb	t**o**nir
	nc**g**eb	tw**g**ir
	ncq**a**b	twn**a**r
	ncqe**d**	twnl**d**

These systematic options allowed the researchers to gauge what features are most important in recognizing a word (and also holding it in short term memory). The kindergarten subjects showed no consistent letter preference when the prompt was a trigram, but used the first letter on the quingrams. The researchers found that the 1st grade subjects choose the options to the trigrams on a basis of overall shape and first letter. The 1st grade subjects were also divided into beginning and more skilled readers. The better 1st grade readers continued to utilize first letter and shape for the more complex quingrams, but the weaker readers used only the first letter. Apparently, children beginning to read in their L1 can utilize word shape if the word is not too complex (three letters in this case), but as it becomes more complex (five letters in this case) the shape becomes less salient and they tend to depend mainly on the first letter. Adults were also tested and consistently used the first and second letters and word shape in recognizing both the tri- and quingrams. The upshot is that most studies point to the importance of the initial letters in word recognition, while aspects of word configuration seem to have a limited and secondary influence.

Results from Eye Movement Studies

The physical way the eye moves and fixates also determines what will be picked up from the page. Rayner and Balota (1989) review the eye-movement literature and arrive at the following conclusions. The average eye fixation during reading is about 200-250 milliseconds, during which the necessary visual information can be obtained within about the first 50 msec. The remaining time (at least 150-175 msecs) is used to program the motor aspects of the next saccade (eye movement). The actual saccade takes from 20-40 msec and covers 7-9 character spaces. Fixations fall on the *preferred viewing location*, normally about halfway between the beginning and the middle of a word. This has the effect of focusing on the more informative beginnings of words. (See Aitchison (1987) for more on the 'bathtub effect' where the beginnings of words are the most salient and information-loaded. Likewise Laufer (1991) states that the initial elements and consonants of words are particularly salient for both L1 and L2 learners.) Most of the words in a text are fixated upon in reading; about 80% of the content words and 40% of the function words. In addition, between 5% and 20% of the content words receive more than one fixation. These figures are averages and reading a difficult text can alter them so that more words are fixated upon for a longer duration.

The width of the viewing span is about 3-4 spaces to the left of the fixation and about 15 spaces to the right. The area 4-8 spaces to the right of the fixation is used to identify the current word of the fixation. Beyond that, the first three or so letters of the next word are preprocessed for the next fixation. This parafoveal preview seems to perceive only the letter identities and does not seem to utilize letter shape configuration, morphemic information, or semantic information. Beyond parafoveal vision, the length of the next word is perceived by peripheral vision, which helps program the length of the next saccade. If the word to the right of the fixated word is short and can be completely identified, then it may be skipped over during the next saccade. Higher-order comprehension influences do affect the programming of the saccade, but appear to come into play later in the programming process, with lexical access having the first influence. Thus, Rayner and Balota argue that lexical access is the trigger which signals the eye to move to the next fixation point.

Word Recognition in the Real World

Most of the above research was done in laboratory contexts, since sophisticated instruments are necessary to make measurements in milliseconds. Meara (1986) wanted to study word recognition in a more natural setting. He made an early attempt to take advantage of home computers in studying vocabulary acquisition by studying how long beginning English-speaking learners of Spanish took to recognize English and Spanish words set in a string of characters. An English example is

```
*****************************
*                           *
*      weolsulusimpletggiha *
*                           *
*****************************
```

where *simple* is embedded. Meara found that this task had latencies which a home computer could measure, rather than the 15-50 msec latencies normal in traditional paradigms. The mean time required to recognize an English (L1) word was 1.47 seconds, but L2 recognition took longer. The Spanish (L2) recognition decreased from 2.15 seconds in the seventh week of a home study course to 1.74 in the twelfth week, to 1.55 in the seventeenth week. So Meara found that his L2 subjects improved from a slow recognition time to one similar to L1 recognition. Unfortunately, there was a large number of very slow recognition responses (>4 seconds) which were not included in the analysis, making it difficult to interpret whether the improvement was gradual or whether there was a stage of rapid improvement. Also, there was an increasingly large number of errors as the course progressed, indicating that the improvement demonstrated above may apply only to the very beginning stages of learning a language. Still, the study suggests that as L2 learners progress, the automaticity of their word recognition improves.

Learning the Orthography of a Word

Ellis and Beaton (1995) review the orthographic factors which effect the ease of learning foreign language vocabulary. The orthography is easier if the L1 and L2 use the same orthographic units and they are read in the same manner (left-to-right, right-to-left, up-to-down in vertical columns). Similarly, use of the same sequential letter probabilities (eg. consonant clusters) facilitate learning. The closer the correspondence between the graphemes and the phonemes they represent, the easier it is to learn. If the L1 and L2 words have similar orthographies based on etymological or loanword reasons (English *hound*; German *hund*), this also facilitates learning. Lastly, shorter words are easier than longer words, partly because shorter words are more frequent (see an explanation of Zipf's law in Crystal, 1987, p. 87). Laufer (in press) makes similar types of claims in her survey of intralexical factors which affect the ease of learning a word, although she finds the effects of word length inconclusive.

Meara (1983b), on the other hand, agrees with Ellis and Beaton that increasing the length of a word makes it harder to handle and decreasing the frequency has the same effect. In his study, the effect of morphological complexity was less clear, but it seems that native speakers decompose words into their morphological components, presumably making them easier to process (in recognition). In addition, he claims that his Spanish-speaking subjects recognized Spanish words more slowly than English speakers recognize English words. Additionally, he did not find the 'bathtub effect' in Spanish, although it is ubiquitous in English. Evidently different languages are processed differently to some extent. As we will see below, these different processes may be transferred to an L2.

Ellis, in a number of articles (1994, 1995, in press), suggests that the input/output lexicons are essentially learned implicitly through exposure. This means receptive and productive mastery of a word's orthographical (and phonological) form is gained from repeated exposures to it. Ellis acknowledges that explicit attention can profitably be brought to bear on a word's orthography, but that for the most part, the learning consists of an implicit tuning into the orthographic regularities which exist in any writing system. Eventually the early conscious control becomes automatized, and this process is driven by the exposure to and manipulation with the very orthographical elements being learned.

Supporting Ellis' view, Thomas and Dieter (1987) found that the act of writing a word enhances memory of its orthographical form. French words were shown to U.S. English-speaking university students three times. For each exposure the students wrote the French word twice (total 6 writing repetitions). They also pronounced the words six times. Then the English equivalents were given and the French words asked for. The oral repetition had no effect on subsequent oral and written production. However, the practice writing improved the spelling of the L2 French words in the written posttest. The experiment was run again but English equivalents were required for the French prompts. This lead to no significant effect. The writing practice affected mainly orthographical word knowledge, since a third experiment showed that it enhanced free recall of French words, but did not have a significant effect on a receptive English-French matching test. Thomas and Dieter suggest that the act of writing down a word when it is being learned may help orthographical recall by drawing attention to the structure of a word, and/or by adding a separate motor trace in addition to the visual memory trace for a word. They note that even the limited writing repetitions used in their experiment were enough to improve recall when accurate spelling was required. In addition, they suggest the reason that oral repetition did not improve performance in their study is because more repetitions were required to have an effect.

L1 Influences on L2 Orthographic Learning

Koda (1997) surveys L2 orthographic knowledge from a cross-linguistic perspective and concludes that a learner's L1 orthographic system plays a strong role in shaping their L2 processing. There are three major types of orthographic systems used in languages around the world- logographic, syllabic, and alphabetic. In logographic systems, the grapheme unit represents a concept, such as in the Chinese writing system. Syllabic systems match the grapheme with syllable, such as in the Japanese *Kana*. The grapheme corresponds to a phoneme in alphabetic systems like English. Each of these systems lead to different processing strategies in the L1, particularly concerning the relative importance of visual versus phonological processing, and Koda argues these strategies are carried over into the L2. She cites research which shows that learners from different orthographic L1 systems process an L2 differently, for example Green and Meara (1987). Students learning an L2 which is similar in orthographic type to their L1 should have less problems with it than if it was different in type. Thus she suggests that L2 instruction be individualized to account for these processing differences, particularly giving explicit instruction in the L2 orthographic system.

Ryan (in press; 1994) and Ryan and Meara (1991) examined this phenomenon with Arabic-speaking learners of English. Apparently these students typically have problems with English orthography. This seems to stem from the fact that Arabic is based on tri-consonantal roots, with vowels being of less importance. When recognition strategies based on these tri-consonants are transferred into English, there can be an 'indifference to vowels' which often results in misrecognized words. Ryan (in press) gives the examples of *moments* being confused with *monuments* (same underlying MNTS structure), and *pulls* for *plus* (PLS). She provides a diagnostic test which can be used to check Arabic-speaking students for such potential problems.

A similar situation has been shown to exist in the case of a single language with two different orthographical systems. Lukatela *et al.* (1978) studied subjects in former Yugoslavia before its dissolution. At that time, the primary language, Serbo-Croatian, could be rendered both by the Cyrillic and Roman alphabets. In eastern Yugoslavia, children first learned to read and write in the Cyrillic alphabet and then later the Roman, in western Yugoslavia, the order was reversed. The researchers' study suggested that university students from eastern Yugoslavia did not process both alphabets in an equal manner, rather they used Cyrillic processing mechanisms even with the Roman alphabet. The researchers "interpret the processing asymmetry and the dependence of its direction on the order of acquisition by saying that whatever the means by which a person has come to read the first-acquired alphabet, those means are adopted to the task of reading the second-acquired alphabet. More precisely, the mechanism for processing the second-acquired alphabet entails the mechanism for processing the first-acquired alphabet, but not vice-versa" (p. 140).

There is some overlap of the Roman and Cyrillic alphabets when rendering Serbo-Croatian. Some words are made up of characters which are included in both alphabets, making them ambiguous if the alphabet system being used is unknown. As would be expected, Lukatela *et al.* (1989) found that preceding an ambiguous word with one which specifies the target alphabet resolves the ambiguity. The preceding prompt did not have to be a complete word; a similar reduction in ambiguity was also achieved from a consonant string. Of course this makes little difference in real-life contexts where ambiguous words are always surrounded by other words, but the authors argue that the results support a connectionist view of word processing.

Just as alternative meanings of a homonym seem to be activated automatically when recognizing a word (the famous *bug*=insect or secret listening device experiment of Swinney, 1979), this also seems to be the case with interlingual homographs. Beauvillain and Grainger (1987) took homographs which had different meanings and pronunciations in English and French (English *coin*=money; French *coin*=corner) and had English-French bilinguals make lexical decisions about them. They found when working in one language, the meaning of the homograph in the other language was initially accessed as well. They also found that the meanings accessed are more affected by the frequency of occurrence in each language than by which language is being used at the time. So for bilinguals who have alternative readings for a homograph in their lexicon, both readings seem to be available to the language processing system, and the access may not be totally dependent on the language of use.

Orthographical versus Phonological Access

Frost, Katz, and Bentin (1987) review the literature on lexical access and the two possible routes: orthographic and phonological. In their review they report that

> There is some agreement that both code types are automatically activated during the process of word recognition and act in parallel (but asynchronously) to mediate lexical access ... The relative use of the orthographic and phonemic codes is determined by factors such as the subject's reading ability, the complexity of the stimuli, and task demands. For example, orthographic codes gain priority when the subjects are fluent readers, when the stimuli are very familiar or phonemically irregular, and when the task emphasizes the graphemic aspects of the printed words. In contrast, phonological codes are used relatively more by inexperienced readers, when the stimuli are more complex, and when the phonemic aspects of the material are emphasized by the task.
>
> (p. 104)

The researchers note that these factors affecting the use of the two codes have derived from studies carried out mainly in English (for another survey, see McCusker, Hillinger, and Bias, 1981). They hypothesized that a person's mother tongue also can affect the use of code. They studied subjects from three languages of increasing 'orthographical depth'. In the 'shallowest' language, Serbo-Croat, there is a close sound symbol relationship, while in the 'deepest' language, Hebrew, there is a much weaker correspondence. English is somewhere in between. Results from lexical decision and word naming tasks indicated that Serbo-Croat subjects (shallower orthography) tended to generate phonology directly from the text, while Hebrew subjects seemed to derive phonology from the internal lexicon. They conclude that both codes probably exist for all languages, but that their weight of usage will depend on the depth of orthography of the particular L1.

Suárez and Meara (1989) studied Spanish learners of English for the path they used. Their learners were of interest because Spanish is a language with highly regular sound-symbol correspondences, and so it is possible that Spanish persons can use a phonological path exclusively when in their L1. The researchers had their subjects read words aloud from a computer screen and measured their pronunciation. Some of the words had regular sound-symbol correspondences in English, and others were exception forms (words with atypical sound-symbol correspondences). The Spanish speakers made more errors reading aloud exception forms than regular forms. This is suggestive that Spanish learners rely on a phonological access route when using English. This implies that they use such a route in their L1, and transfer such processing over to English as an L2.

Spelling

Stubbs (1980, Chapter 3) feels that the English spelling system, although it is not optimal, is reasonably systematic, and that even some of its irregularities have a functional purpose. One example of this is that although the different members of a word family may have different pronunciations, their orthographic shape is likely to highlight their relationship, eg. *finite*, *infinite*; *Christ*, *Christmas*; *crime*, *criminal* (Wallace and Larsen, 1978, p. 364). Upward (1988), on the other hand, argues that the traditional orthography system in place for English is not as organized and systematic as is commonly assumed, and that it is deficient in the sense that its irregularities cause problems in gaining literacy in the language. In fact, Feigenbaum (1958, cited in Wallace and Larsen, 1978) states that there is a minimum of 251 orthographical representations for the 40+ phonemes of English. From among this profligate excess, Upward believes that redundant characters are a major problem, and fall into three classes: 1) silent letters [*b* as in *debt*], 2) post-accentual schwa before liquids and nasals [*e* in *chapel*; *o* in *atom*], and 3) doubled consonants [*committee*]). He suggests a Cut System to remedy the problem, which would result in English being spelled as in the following sentence: An importnt considration behind th Cut Speling systm is that th apearance of words shud not chanje so drasticly that peple uninstructd in th rules of CS find them hard to read (p. 24).

A close look at spelling mistakes reveals that they are not often phonological mistakes. Alper (1942, cited in Tenney, 1980) found over 1000 spelling mistakes in 5,000 college English compositions. Most of these made sense phonetically and followed conventional sound-symbol correspondences. So following sound-symbol rules exclusively does not guarantee accurate spelling of the exceptions in English. Some form of visual appraisal is necessary as well to tell when a word 'looks right'. Barron (1980) studied Year 6 children and found that poorer readers seemed to use primarily a visual-orthographic strategy when reading, but a phonological strategy when spelling, while better readers seemed to use both strategies in reading and spelling. The phonological strategy is useful for phonologically-regular words, while visual strategies are useful for the exceptions. Better readers can use the appropriate strategy as needed.

Spelling strategies tend to change with maturity however. Marsh *et al.* (1980) claims that between Year 2 and Year 5, there is a major change in spelling strategy from reliance on sound-symbol correspondences towards spelling an unknown word by analogy to a known word. This change seems to happen after the child has built up enough known words in his lexicon to use as models.

Imagery may also have a part to play. Ehri (1980) concludes that a mental image of a word's orthography is generated from the visual experience of that word, and that the image can be used to facilitate spelling, especially of familiar words. "Findings indicate that orthographic images can be scanned like real words seen in print, that they include all of the letters in a word's spelling, not just boundary letters or letters mapping onto sounds, that silent letters may have a special status in these images. Findings suggest that the presence of orthographic images in memory increases the likelihood that the spellings produced by readers resemble single conventional forms rather than phonetic variants" (p. 338).

In the end, reading and spelling cannot be simplistically considered two sides of the same coin. Bryant and Bradley (1980) found that L1 children between 7-10 years of age often tried to read and write the same words in different ways. In some cases they could spell out words phonologically which they were not previously able to read (receptive knowledge does not always come before productive knowledge!). Marsh *et al.* (1979, cited in Marsh *et al.* 1980) found that there was no transfer between reading and spelling for words and nonwords presented in the same experiment. This may be because the underlying processes work at cross-purposes to some degree. Where quick and efficient reading takes advantage of minimal cues in word recognition, the downside is that less orthographical information is taken in which could inform spelling (Frith, 1980). Extensive exposure could overcome this deficiency, but at least in the beginning stages, different amounts of attention to orthographical cues may be appropriate depending on whether the immediate purpose is fast reading or accurate spelling.

KNOWLEDGE OF PHONOLOGICAL FORM

Parsing the Speech Stream

Adequate phonological knowledge of a word involves being able to parse out and understand the acoustic representation of a word from a continuous flow of speech, as well as being able to pronounce the word clearly enough in connected speech for other people to do the same when we speak. Being able to manage these input/output processes actually requires a detailed knowledge of not only the pronunciation of the word as a whole, but also its parts. Since phonological word knowledge deals with something tangible, sound, it is not difficult to imagine what the various aspects of the acoustic representation of a word entail--indeed, dictionaries attempt to make just this information explicit. One must know how each consonant and vowel is pronounced when tied together in the consonant clusters and order particular to that word. Unless the word is monosyllabic, it will be divided into a number of syllables. These syllables will not be pronounced with an equal amount of emphasis; rather one or more will be stressed. This stressing can be accomplished by altering the pitch, volume, or length of the syllable, as well as the features of the vowel, especially in unstressed syllables where the vowel is commonly reduced.

Although the individual phonological features of a word might be relatively straightforward, the way they are used in the processes of hearing, speaking, and learning a word are quite complex. Not to minimize the problem of achieving comprehensible pronunciation, but the greater challenge for most language learners lies in the act of listening. Understanding words in continuous speech involves two problems: first isolating the component words from the speechstream, and then using those phonological representations to access lexical knowledge about the corresponding words.

As Cutler and Butterfield (1992) have noted, segmenting the natural flow of continuous speech into the individual component sound groups which represent words is no trivial task. As opposed to written discourse, spoken language does not have clear word boundaries. In fact the words blend together in speech to such an extent that if one does not actually know a language, it is very difficult to pick out any individual words at all. Anne Cutler and her colleagues have been working on the idea that the stress in sentence prosody is an important factor in segmenting the speech stream into its component words. There is evidence from a number of studies which support this view that strong syllables trigger segmentation of the aural speech stream as the first step to lexical access. In a study by Cutler and Norris (1988), words like *mint* were combined with nonsense syllables to produce bisyllabic words with a either a strong/weak stress pattern (*mintesh)* or a strong/strong pattern (*mintayve*). Subjects recognized the real words in the strong/weak bisyllabic words faster than in the strong/strong words. Cutler and Norris argue this is because the second strong syllable in the strong/strong words induce segmentation at that point, making it necessary to assemble speech information across a segmentation point, which slows the recognition process. The crucial point is that the strong stress induced segmentation at that point.

Cutler and Butterfield (1992) examined how erroneous word boundary placement was connected to stress. In natural 'slips of the ear' data, erroneous word boundaries were inserted much more often before strong syllables (containing full vowels) than weak syllables (containing central, or reduced vowels), and boundaries were deleted more often before weak syllables than strong syllables. These results were confirmed by laboratory-induced boundary misplacements. Additionally, when the subjects placed boundaries before strong syllables, a lexical content word followed; when placed before a weak syllable, a function word followed. Cutler and Norris (1988) also cite a doctoral dissertation by Taft (1984a) which used ambiguous strings of syllables. When the first syllable was strong [lɛt'əs], subjects tended to choose one-word readings (*lettuce)* over two-word readings (*let us*); if the second syllable was strong [ɪn vɛsts'], two word readings were normally chosen (*in vests*) over one-word reading (*invests*). This indicates people assume weak syllables do not begin content words.

Cutler and Clifton (1984) found that hearing the correct and expected stress pattern for a word is important for its efficient recognition. When subjects listened to words which in which normally unstressed syllables were stressed, or vice versa, they were difficult to recognize. This was strongly connected to whether the vowels were full or reduced as expected in stressed or unstressed syllables. The subjects were not bothered by weak/strong bisyllabic words being stressed in a strong/weak manner, since this is the kind of natural stress shifting that occurs in natural language. But shifting the opposite way confused subjects, presumably since these kind of shifts violate natural constraints. In fact, there is good reason to believe that people assume strong syllables are word-initial. Cutler and Carter's (1987) analysis of 33,000 entries in a computer-readable English dictionary showed that 73% of them had strong initial syllables. They calculate that there is about a 3 -1 chance of a strong syllable being the onset of a new content word, while weak syllables are likely to be grammatical words.

It does not seem to matter if the words begin with prefixes or not. Tyler, *et al.* (1988) used gating, auditory lexical decision, and auditory naming tasks to explore whether words are accessed as wholes or whether they are decomposed and accessed via word stems with prefixes stripped off. The experiments showed that the presence of a prefix does not delay the recognition times, giving evidence to support the idea that words are accessed from the beginning of the word and that prefixes are not stripped off. The fact of whether the prefix was stressed or not did not seem to have any effect on recognition times.

Grosjean and Gee (1987) put forward a model of oral word recognition which attempts to explain the evidence that people use strong syllables to indicate the beginnings of content words, but not function words. According to this model, in the first stage of segmentation, the mind starts a lexical search at the stressed points in speechstream. At the same time, a parallel system starts analyzing the surrounding unstressed features. This pattern-recognition system knows that grammatical words in English are usually unstressed, helping it to decouple them from the stressed syllables. These two systems work together simultaneously, along with other contextual, semantic, and grammatical information to parse the speechstream.

This model may be rather extreme in the degree that it relies on stress to begin segmentation, but even if it proves too radical, it seems clear stress plays an important role in parsing the speech stream. Once this segmentation is achieved, how are the isolated sound 'chunks' used to access the corresponding lexical word? There have been a variety of theories around for a number of years which try to explain how these processes work, including the logogen model (Morton, 1979), the bin model (Forster, 1976), and the cohort model (Marslen-Wilson and Tyler, 1980). Of these, a modified version of the cohort model seems to be getting the most attention recently. Cohort models say that words are basically recognized in a serial manner, from left-to-right. Let us take the word *candle* as an example. After hearing the initial consonant *c*, the mind activates all words beginning with *c*, such as *carry, chemistry*, and *cup*. After the second letter, *a*, the list is narrowed to words beginning with *ca-*, such as *cam, cassock*, and *cabin*. This culling out continues with *can-* and *candl-* until the candidate list is narrowed to one, the *recognition point.*

As with any model of language processing, there are the inevitable problems which need to be resolved. Garnham (1985) highlights two in particular: the model cannot account for recognition of a word with a mispronounced first syllable (a drunk saying *shigarette* instead of *cigarette*) and second, because "words that do not fit with the context are dropped from consideration, the model predicts that words cannot be recognized in inappropriate contexts" (p. 57). Also, research has thrown up some contra-evidence. Goldstein (1983) found that longer French words (4 syllables) were detected more accurately and more quickly than one-syllable words when heard in fluent connected speech. He suggests this is because learners are able to recognize a longer word as a unit more easily than a monosyllabic word, which is more difficult to decide when it has ended, since it could easily be only the beginning of another word.

This suggestion underscores an interesting situation in phonological processing. In general, the speed of word recognition is very fast. Marslen-Wilson and Tyler (1980) found that native speakers were able to recognize words in about 200 milliseconds, which was usually before the offset of the word. In contrast to this, Grosjean (1985) found that sure identification of a word can sometimes occur only after several subsequent words have already been heard. This is because in connected speech it is difficult to know whether subsequent sounds are part of a longer word or the beginning of new words. This is especially true of unstressed syllables. For example, relying strictly on phonological information, it is impossible to know where to parse the string /ðəlɛgɪslai/. It could be part of the sentence 'The ledge is laden with flowers' or 'The legislature is on recess'. So recognition speed is potentially fast, but is constrained by the parsing process. Of course, context usually comes into play to disambiguate such strings. Thus, it is likely that the cohort model (or any of the others) will have to incorporate a mechanism which describes the mind's ability to use a wide variety of information (including that of contextual ties, word frequency, and grammatical sequencing) in achieving lexical retrieval. The spreading activation model of McClelland and Elman (1986) is a step in this direction. Still, the operation of lexical access may ultimately involve so much parallel processing that a parsimonious explanation is impossible.

Phonological vs. Orthographical Knowledge

As would be expected, knowledge of aural form is closely related to knowledge of its complement, written form. Stahl and Murray (1994) summarize studies which show that phonological awareness is necessary for reading and using the alphabet. Their study provides evidence indicating knowledge of the alphabet (letter names) is necessary for L1 children to separate onsets (initial consonants or clusters) from rimes (vowels and any following consonants), which in turn, seems to facilitate word recognition. This combined knowledge facilitates more complex phonological analysis. The ability to analyze onset/rime structure also fits closely with the ability to spell. So the relationship between phonological awareness and orthographical knowledge is close and interrelated, where knowledge of one facilitates the learning of the other.

Lukatela and Turvey (1994a & b) go so far as to argue that phonological processing is primary in lexical access even when reading. They interpret the results of seven priming experiments as indicating that "the phonological code seems to be ... the earliest constraint on word recognition, creating the circumstances within which the role of orthographical codes is defined ... [and] more pervasive and more empirically demonstratable as the lexical access code than the orthographic code..." (b, p. 350). Segalowitz and Hébert (1990) take the position that phonological recoding of written text is likely to be automatic, although not fast enough to affect lexical access. They suggest it most likely comes into play postlexically in working memory. (For a book-length treatment of lexical accessing, see Marslen-Wilson, 1989.)

Phonological Knowledge and Vocabulary Learning

Phonological awareness is also important for vocabulary learning. Goldstein (1983) suggests lower-level second-language learners must rely more heavily on acoustic clues than native speakers, since they cannot compensate with native-like knowledge of semantic and syntactic constraints to predict and decode words. Thus phonetic similarities between target words and other words and syllables affect L2 listeners far more seriously than native listeners. Additionally, L2 listeners seemed less sensitive to at least one set of phonetic cues, whether words begin with stops or fricatives, something which has been shown to influence how native speakers recognize language.

Lerea and LaPorta (1971) found that young adult American bilingual speakers learned words in a vocabulary list much better aurally than visually when compared to monolingual speakers. They suggest that mode of previous learning may have something to do with this; monolingual speakers have been taught mainly in a visual manner, while bilinguals, especially compound bilinguals, have learned a language primarily in an aural/oral fashion, helping them to learn the words of a new language better aurally. (Note that being bilingual doesn't mean one is *better* at languages; monolinguals learned more words than bilinguals in the visual mode).

The ability to vocalize new words when learning them seems to facilitate that learning. In an study by Taft (1984b), subjects were asked to judge whether two words were homophonic (having the same sound) under four conditions: 1) prompt word was inflected but homophone was not with different morphophonemic structure = FINED - Find, 2) different morphophonemic structure but both were uninflected LAX - Lacks, 3) same structure and both prompt and homophone were inflected HEELED - Healed, 4) same structure but neither were inflected KNEAD - Need. There was no significant difference between whether the subjects vocalized the words aloud or subvocalized them for conditions 3 & 4, but subjects were able to judge homophones in Conditions 1 & 2 much better if they vocalized them aloud. Therefore, vocalizing words aloud appears to aid learning of homophones with different morphological structures. Papagno, Valentine, and Baddeley (1991) found that subjects who were prohibited by articulary suppression from vocally or subvocally repeating a new word to be learned from associated word-pairs were much less able to learn the foreign language words. The learning of L1 words was not affected however. It seems that loss of phonological loop repetition in working memory can be compensated for by using semantic associations for L1 words. A much older paper by Seibert (1927) also suggests that studying aloud is advantageous, but methodological and statistical deficiencies limit the conclusions we can draw from her study.

Not every part of a word is equally easy to learn and remember. Meara and Ingle (1986) found that English-speaking learners of French (age 12 years) could correctly produce the initial consonants of words they had learned much more often than subsequent consonants. In fact, later consonants were relatively unlikely to be produced without error. The authors compare their L2 results to L1 malapropism evidence, and suggest the possibility that when learners meet new words in an exclusively phonological mode, they automatically focus on the initial segments. When words are met in a written mode, on the other hand, other more appropriate strategies may be employed. (See Fay and Cutler (1977) for more on malapropisms.)

Donley (1974) believes that we should utilize a wide range of vocabulary exercise which show how words are related to each other, especially through contrast. He suggests highlighting relationships of pronunciation and spelling in addition to meaning relationships (similar letters, sounds, hyponymy, homographs, homophones, and words of a word family). However, phonological knowledge is likely to be largely learned implicitly through exposure. Ellis (1994, 1995, in press) argues that, although the phonological features of any particular word can be explicitly learned, the speed and overall proficiency of the phonological input (listening) and output (speaking) systems are mainly acquired by the repetitive practice of actually using the target language.

Phonological Knowledge for Languages Other than English

The studies in this section have mainly dealt with English. Since phonological awareness deals with something relatively tangible, a language's sound system, different languages may well elicit different effects. It is unlikely that languages which do not have stress as a prominent feature, such as syllable-timed languages, can be parsed according to stress in the English manner. Native speakers in those languages require some other method of determining where individual words begin. Also, English does not have a particularly good grapheme/phoneme correspondence. Other languages, like French, with a much more regular correspondence enable easy phonological representation for use in postlexical working memory when compared with the cognitive effort required for English, with its grapheme-phoneme irregularities.

Notes

[1]To give some basis of interpreting these correlations, objective frequency and the probability of knowing a word in a language only correlate at .44 (Graves *et al.*, 1980).

[2]This conference paper is widely cited, but impossible to obtain. Its high rate of citation may be a result of a chain- citation, in which only one initial author actually saw or heard the original material. Thus we must be cautious and on our guard against a linguistic urban myth. It seems impossible to track the author down for confirmation; Pierre Arnaud (personal communication) has already tried and failed.

[3]Zero exposure words are ones which appear only on the posttest.

[4]Simpson (1988) believes that native speakers are exposed to a number of spoken and written words that is closer to 1 million per *day*.

[5]Of course this assumes that the corpus is properly balanced to reflect the 'real world', something not necessarily that easy to achieve. See Sinclair (1991) for considerations in compiling a corpus.

[6]I thank Michael McCarthy for introducing me to this useful metaphor.

[7]In addition to the descriptive approach detailed here, there have also been attempts to statistically describe register. Chui (1972) focuses on English and Green (1979) on Spanish. See also Carter (1983) for a discussion on the use of semantic scaling for register measurement.

CHAPTER 3 A STUDY INTO INTUITIONS OF FREQUENCY

Introduction

The literature review has firmly established that full mastery of a word requires much more than knowing a word's meaning, primarily by following Nation's (1990, p. 31) taxonomy of the various kinds of knowledge. Schmitt (1995a) and Schmitt and Meara (1997) have suggested that this 'word knowledge' taxonomy can be used as a framework to inform vocabulary applications and vocabulary research. Nation (interview with Schmitt, 1995c: 5) states that "at the moment we have only the broadest idea of how [vocabulary] acquisition might occur." Since this is partially because it is so complex, one way forward may be to learn more about the component types of word knowledge and their interrelationships, and from that understanding build up a theory of vocabulary acquisition in general.

One hurdle in using a word knowledge framework in this way is our lack of understanding about several of the word knowledge components themselves. This is especially true for intuitions of word frequency. The previous studies discussed in the literature review have demonstrated that native-speakers have reasonably accurate intuitions about the frequency of words in general, but the research designs did not allow researchers to say much more than this. Frequency research reached its peak in the 1960s and 1970s, and unfortunately petered out before any truly useful conclusions were reached. The previous studies stopped short of providing any figures which could be used as benchmarks of native-like performance. This gap limits the use of intuitions of frequency in any practical endeavour. Practical uses hinted at by previous researchers include frequency intuitions as a measure of bilingualism (Cooper and Greenfield, 1969) and as an alternative to corpus word frequency counts (Ringeling, 1984). In addition, it is possible that frequency estimates could be used to partly assess the richness of a learner's knowledge of words on vocabulary tests. Once benchmark figures have been established, then research into applying frequency intuitions to these and other practical matters can begin. This chapter will describe a study which attempts to provide such a quantification of frequency intuitions, which hopefully then can be used in the following longitudinal acquisition and TOEFL studies.

Brief Summary of Findings from Previous Studies

Although the frequency literature has been reviewed in depth in Chapter 1, to help the reader, the main points will be highlighted again here. Previous studies have shown that the assumption that native speakers have intuitions about how frequently words occur in language is correct, with correlations between subjects' Subjective Frequency Estimates (SFE) and Objective Frequency baseline data extracted from corpora (OF) generally being quite high: .74-.78 (Tryk, 1968); .92-.97 (Shapiro, 1969); .92-.97 (Carroll, 1971); .57 (Richards, 1974); .91-.94 (Backman, 1976); .67-.89 (Ringeling, 1984); and .64-.79 (Arnaud, 1989, cited in Arnaud, 1990). Beyond showing that frequency intuitions exist, these high correlations were usually interpreted as indicating that they are relatively accurate as well. However, these correlation figures obviously vary a great deal, and there must be some doubt as to whether all of the cited frequency studies are measuring the same thing. It would be useful to determine more accurately where the 'true' correlation strength lies within the wide range cited above, or to establish whether it is something that truly varies this much.

In addition to indicating the accuracy of SFEs, earlier studies suggested some of their traits. First, two studies suggest that native-speakers as a group tend to give similar SFEs. Carroll (1971) found that the SFEs of 15 lexicographers correlated with each other at .99 and 13 college-educated adults with each other at .97, while Shapiro (1969) found no statistically reliable differences between the SFEs of sixth-graders, ninth-graders, college sophomores, industrial chemists, elementary school teachers, and newspaper reporters. Although it may well be that the younger subjects had had enough exposure to words to form mature frequency intuitions, the finding that all native speakers have very similar frequency intuitions seems somewhat counterintuitive, and needs to be reexamined.

Second, the SFEs seem to be reliable as well. Test-retest correlations after three weeks were .96 and .98 for English native-speakers (Tryk, 1968) and .80 for French native-speakers after five weeks (Arnaud, 1989, cited in Arnaud, 1990). In addition, Arnaud found an interesting pattern where students who provided the most accurate SFEs also had the most stable intuitions, as indicated by higher individual test-retest correlations.

Third, native-speakers are able to judge not only the frequency of words in their own personal situation and context, but also the general frequency of words in society. Instructions to rate personal and public frequency of usage separately resulted in essentially the same SFE/OF correlation figures: .77, .78

[public] and .74, .76 [personal] (Tryk, 1968); .74-.85 [public] and .67-.87 [personal] (Ringeling, 1984).

Fourth, SFEs seem to be more closely related to written discourse than spoken discourse. In a large-scale study, Richards (1974) obtained SFEs from 1000 Canadian university students. They correlated poorly with two oral word counts: *A Study of the Oral Vocabulary of Adults* (Schonell et al., 1956) at .37, and *A Word Count of Spoken English* (Howes, 1966) at .39. The student SFEs correlated much better (.57) with the main written word count used, the *Computational Analysis of Present Day American English* (Kučera & Francis, 1967). Despite this, subjects do not seem to be able to consciously differentiate between spoken and written frequency for most words. Shapiro (1969) found that instructions to rate frequency in terms of spoken language produced results no different from instructions to rate in terms of written language.

Fifth, there is evidence that native-speakers have frequency intuitions for multi-word lexemes as well as single-word lexemes. Swedish university students' SFEs of Swedish 3-word lexemes correlated with OF at .56 (Backman, 1978).

Much less work has been done with nonnative-speakers. Arnaud (1990) suggests that in the process of developing frequency intuitions directly for L2 words, nonnative-speakers may utilize their frequency intuitions from the target words' L1 translation equivalents. To show that this could be possible, he compared the frequency rankings of the translation equivalents in French and English. They correlated at .84-.89, which is in harmony with the figure

(.84) reported by Kirsner, et al. (1984). So nonnative-speakers could rely on their L1 frequency intuitions alone to achieve reasonable estimates of L2 frequency, although this would probably only work when the two languages were closely related. In his 1984 study, Ringeling studied advanced Dutch speakers of English and found that although their SFEs for their private use of English words were much lower than those of native-speakers (.61-.78), their SFEs for public use were quite similar (.75-.90). This indicates the possibility that advanced nonnative-speakers are able to develop native-like frequency intuitions.

Shortcomings of Previous Frequency Research

Previous methodology required subjects to rate unrelated words which varied enormously in frequency, a rather unnatural task. To form a more psycholinguistically-valid task, we must ask ourselves how frequency intuitions are used in real language. In cases where several near-synonyms from a lexical set could be chosen, there must be some determining factor or factors which affect the eventual choice. Frequency is very likely among these factors, since it has close relationships with the register constraints which also affect lexical choice, eg. more formal words are usually less frequent, words typical of written discourse are typically less frequent than words typical of spoken discourse, and words are archaic simply because they have become so infrequent. Thus rating near-synonyms within a lexical set may be a more natural task in frequency experiments, as we often have to choose from among a number of near synonyms in our daily communication. Also, the task may be more natural in that the words in a lexical set are not likely to vary as widely in frequency as some used in other studies, ie. *the* [68,988 occurrences per million words] and *echidna* [about 1 occurrence per 4.5 million words] (Carrol, 1971; Shapiro, 1969).

Other methodological weaknesses in previous studies include the fact that the corpora researchers had available at the time (typically 1-5 million words) were quite small compared to some of today's computer-driven ones (ie. 88 million words for the British National Corpus or 211 million words for the COBUILD Bank of English). As a consequence, the objective frequency figures obtained from the older corpora are not likely to be as accurate as those which can be obtained from the immensely larger modern corpora. Additionally, prior studies have tended to use rather small subject numbers, which severely limits the confidence one can put in their conclusions.

A final problem is the focus on native-speakers. Since the results of most language research can find applications in the area of second language learning and teaching, it would also be useful gather more detailed information on nonnative frequency intuitions as well.

Development of the Elicitation Instrument

The first step in developing the elicitation instrument was isolating a number of lexical set candidates which had at least five nearly synonymous words. (No two words are likely to be completely synonymous and interchangeable.) The *Longman Language Activator* (1993) was used as the source, because it contains lexical sets based around key concepts which have been identified from corpus research. Sixty-sets were originally isolated, but many of these were eliminated because they contained long multi-word lexemes, words which had two or more major meanings within the same word class (*attractive*-beautiful, desirable; *funny*-comical, strange), or words of very different lengths. This left 36 sets. *Webster's New World Thesaurus* (1974) was then consulted to find additional words for these remaining sets.

Once the 36 candidate lexical sets had been isolated, the next step was to fix objective frequency values to the words in the sets. Two large modern corpora were used for this purpose; the British National Corpus (BNC) and the COBUILD Bank of English Corpus, consisting of about 88 million and 211 million words respectively at the time of this study. The two corpora did not always agree, so it was decided to use frequency figures from the newer BNC and confirm them with COBUILD data. No lexical set was used unless both word counts agreed on the rank order of the words in the set. This still left words with quite different frequencies from the two corpora, so sets on which the two corpora had the best agreement were preferred. This reduced the number of lexical sets down to twelve (four noun sets, four verb sets, and four adjective sets).

The final five words for each of these sets were chosen with the following desirable criteria in mind. Given the salience of the beginning of words, words were chosen so that subjects could not produce correct word frequency rankings merely from the frequency of the initial letters. Words of a similar length were preferred. The magnitude of frequency difference between words in a set was kept as uniform as possible, although some sets had relatively small magnitudes and some had relatively large ones. It should be noted that these and the previous criteria were not absolute, as it was necessary to compromise between the various criteria in order to achieve the most unbiased sets. The 12 lexical in their final forms are given in Table 1.

Table 1 Lexical Sets Used In This Study With Number of Occurrences in the 88 Million Word British National Corpus

calamity	134	annihilate	164	awful	2224
cataclysm	30	conquer*	720	bad*	12142
catastrophe*	449	defeat	5875	dreadful	1254
disaster	3077	subdue	648	ghastly	396
tragedy	1873	vanquish	84	terrible	3747
appliance	601	glisten	410	essential*	8235
contrivance	77	shimmer	420	imperative	516
device	4854	sparkle	1230	necessary	17020
gadget*	197	shine*	2962	requisite	297
machine	12343	twinkle	391	vital	4818
chore	386	deflate	199	low-tech	22
duty	7506	diminish	1734	primitive	1781
errand	237	lessen*	546	rudimentary	256
job	26952	minimize	1457	simple	14671
task*	12312	reduce	17982	unsophisticated*	118
emblem*	275	disclose*	1781	flimsy	280
hallmark	328	divulge	150	fragile	418
logo	620	expose	3724	frail	544
sign	12095	impart	364	puny*	99
symbol	2940	reveal	10458	weak	3362

*anchor word

Each lexical set was then fixed to the instrument as in the example below:

calamity _________

cataclysm _________

catastrophe ____1_____

disaster _________

tragedy _________

Each set had one 'anchor word' which served as a benchmark against which other words in the set were to be rated. Thus, for the above set, if a subject judged *disaster* to be ten times more frequent than *catastrophe*, then **10** would be written on the blank. If *calamity* was considered one-third as frequent as *catastrophe*, then **1/3** or **.33** would go on its blank. If a subject (particularly nonnative-speakers) did not know a word, they were instructed to write **X** on the blank. If the anchor word was unknown, then subjects were to write **X** on its blank, choose another word to be the anchor word and write **1** on its blank, and continue rating the set. Two examples were given in the instructions to ensure subjects understood the task. The anchor words were balanced among the sets both in terms of frequency of the anchor word and in terms of its position within the set.

In order to avoid influencing the subjects' SFEs, no parameters of frequency were given. Since previous research has suggested that objective word counts are better criteria for SFEs of *public*, rather than *private,* usage (Ringeling, 1984), the subjects were instructed to rate the words on the basis of their general frequency in society as a whole. Once the instrument was developed, it was piloted on ten native-speakers. The results indicated that it would yield useful information and was an appropriate length. (See Appendix 3.1 for the complete instrument.) It was then administered to the main study subjects to complete at home, with the exception of members from Groups 1 and 6, who filled it out in class. In either case, subjects were under no time constraints and the instrument was completely anonymous.

Subjects

In order to achieve as large and diverse a sample of native-speakers as possible, respondents were included from five groups of various backgrounds from the city of Nottingham. Likewise, nonnative-speakers were drawn from a number of groups representing a variety of L1s. Description of the subjects is presented in Table 2.

Table 2 Description of Subjects

G	N	M	F	L1	EP	Comments
1	41	23	18	English	-	Year 10 & 11 students
2	52	31	21	English	-	Residents of working/middle class neighborhood
3	16	7	9	English	-	Elderly residents of a sheltered apartment complex
4	28	13	15	English	-	Masters students in Education (mostly mature)
5	23	8	15	English	-	TESOL teachers/Lecturers in Linguistics
	160	82	78			Totals (native-speakers)
6	30	11	19	Spanish	A	Mostly postgraduate students at a Spanish university
7	31	6	25	German	I/A	TESOL teachers-in-training
8	60	42	18	Mixed	I/A	ESL students in England
9	11	5	6	Mixed	A	Lecturers/teachers attending British Council Summer Course
10	24	7	17	Mixed	A	Masters students in TESOL at a British university
11	14	7	7	Czech	I/A	Czech Adult Professionals & Teachers
12	27	3	24	Arabic	A	Teachers in Lebanon
13	12	1	11	Arabic	A	Undergraduates/postgraduates at a Lebanese university
	209	82	127			Totals (nonnative-speakers)

G=Group
N=Number
M=Male
F=Female

EP=English proficiency
A=Advanced
I=Intermediate
B=Beginner

The English Proficiency rating is an approximation stemming from teachers' impressions of their students' abilities, since it was impossible to obtain objective language proficiency measures for this diverse a subject group. Overall, the nonnative subjects can be considered at least moderately advanced, with most being involved with either English teaching or having taken classes in an English-medium environment.

Results

Recognizing the Core Word in a Set

A number of progressively more difficult frequency judgements can be made about the lexical sets used in this study. The most fundamental is the ability to pick out the *core* word. Almost every lexical set has a word that is the most central or basic which can be termed core. Carter (1987) discusses some of the tests of coreness, including syntactic substitution (the core word of a lexical set is often used to define the other members), antonymy (it is easier to find antonyms for core words than for those which are less core), collocability (core words generally collocate with more partners than less core words do), extension (core words will tend to have more meaning senses and will be involved in more multi-word-units like idioms or phrasal verbs), superordinateness (core words tend to be general, rather than specific), and the state of being relatively culture-free (basic words that would have little if any connotation if used in other cultures have core tendencies). Although not included by Carter in this particular list of tests, frequency alone can be a very good indicator of coreness. Most of the above tests entail core words occurring more frequently than less core words, making frequency perhaps the best indicator of coreness.

Since core words are so important and widely used, it is sensible to ask how well subjects can discern the core (most frequent) word in a set as the most basic level of frequency intuition. It turns out that the subjects, both native and nonnative, were very good at this task (Table 3).

Table 3 Number of Subjects Able to Discern the Core Word in the Lexical Sets

Lexical Set	Native Speakers N=160		Nonnative Speakers N=209	
	N	%	N	%
calamity*	132	82.5	145	69.4
appliance	122	76.2	155	74.2
chore	151	94.4	175	83.7
emblem	117	73.1	124	59.3
annihilate	153	95.6	170	81.3
glisten	82	51.2	82	39.2
deflate	146	91.2	187	89.5
disclose	123	76.9	130	62.2
awful	55	34.4	48	23.0
essential	127	79.4	186	89.0
low-tech	153	95.6	190	90.9
flimsy	130	81.2	179	85.6
MEAN	124	77.5	148	70.8

*Lexical sets are labeled by first word appearing in the set

An average of 124 out of 160 native subjects (77%) gave the highest frequency rating to the core word in the various sets. This ability to recognize the core word a high percentage of the time is unsurprising, but the moderately advanced nonnative subjects in this study were largely able to do the same.

In fact the nonnative subjects did slightly better than the native subjects on the sets *essential* and *flimsy*.

There are two sets which do not fit with the others, either for native or nonnative speakers -- *glisten* and *awful*. On examination, it was found that these two lexical sets had the core word as the anchor word. There are two possible explanations why this should depress the scores. First, respondents may have been hesitant to rate all of the words in a set as being less frequent than the anchor word. Second, many people are not as comfortable working with fractions as with whole numbers. These explanations would indicate that the low figures are an artifact of the elicitation method. Additionally, the *glisten* set had several words which were very similar in frequency, which surely made it more difficult to judge. This does not account for the low scores for the *awful* set, however, as the degree of frequency difference among its words was no greater than for other sets. Carroll (1971) previously concluded that the objective frequency of a word chosen as an anchor does not affect the accuracy of subsequent frequency judgments, but it seems safest to conclude from this data that core words should not be used as anchor words in future frequency experiments.

If the suspect figures are discarded, we find that both native and nonnative speakers do even better at discerning the core word of a set (Mean: NS 135, 85%; NNS 164, 79%). These figures suggest that a reasonable benchmark for nativeness for this kind of task can be set at around 75%.

Ranking the Words in a Lexical Set According to Frequency

A more complex task is ranking the words in a lexical set according to frequency. We can reduce the subject responses to ordinal information in order to explore the subjects' ability to rank the target words. The results are presented in Tables 4 and 5. In these tables, any set with one or more non-answer (NA) is excluded from the analysis. A transposition occurs when the subject exchanges the ranks of any two words in a set.

Table 4 Ability to Rank Words in Lexical Set According to Frequency (Native Speakers N=160)

Lexical Set	Correct		One Transposition		Correct + 1 Transpos.		NA	
	N	%	N	%	N	%	N	%
calamity	63	39	59	37	122	76	25	16
appliance	18	11	49	31	67	42	25	16
chore	3	2	18	11	21	13	18	11
emblem	12	8	47	29	59	37	20	13
annihilate	15	9	62	39	77	48	19	12
glisten	3	2	9	6	12	8	23	14
deflate	4	2	16	10	20	12	23	14
disclose	13	8	22	14	35	22	26	16
awful	8	5	15	9	23	14	40	25
essential	24	15	49	31	73	46	23	14
low-tech	3	2	19	12	22	14	23	14
flimsy	27	17	54	34	81	51	23	14
MEAN	16	10	35	22	51	32	24	15

Table 5 Ability to Rank Words in Lexical Set According to Frequency (Nonnative Speakers N=209)

Lexical Set	Correct		One Transposition		Correct + 1 Transpos.		NA	
	N	%	N	%	N	%	N	%
calamity	44	21	50	24	94	45	75	36
appliance	24	11	54	26	78	37	90	43
chore	2	1	41	20	43	21	96	46
emblem	11	5	42	20	53	25	67	32
annihilate	9	4	35	17	44	21	88	42
glisten	1	0	4	2	5	2	81	39
deflate	17	8	38	18	55	26	68	33
disclose	16	8	43	21	59	28	92	44
awful	10	5	18	9	28	13	75	36
essential	21	10	51	24	72	34	61	29
low-tech	20	10	22	11	42	20	58	28
flimsy	23	11	37	18	60	29	89	43
MEAN	17	8	36	17	53	25	78	37

We find that the native speakers generally did better than the nonnative speakers. Modelling the above Correct responses (completely correct rankings within a set) using a log-linear model for binomial data (see McCullagh and Nelder, 1989), we can reject the hypothesis that there is no difference between the native and nonnatives for this task (Chi-square (*1df*) = 6.66, $\rho<.01$). An obvious exception is the lexical set *low-tech* where the non-natives noticeably outperform the natives. But even though the native speakers outperformed the nonnative speakers, their figures are still quite low. On average, the native respondents were able to correctly rank the words in the lexical sets only about 10% of the time and with one transposition about 22%. If we accept that a single transposition of ranks within a set is a relatively minor error, then the native speakers were able to rank the words in a set reasonably well about one-third of the time (Correct + 1 Transposition). Clearly, ranking words within a set proved a difficult task, and the level of performance varied greatly among the word sets. For these reasons, it would not be sensible to try to set a native-like benchmark for this task.

Considering how difficult the native speakers found this task, the nonnative speakers did surprisingly well in relative terms. In fact, on four sets (*chore*, *deflate*, *disclose*, and *low-tech*) they did better than the native-speakers, if we take the *Correct + 1 Transposition* results. In sum, although both native and advanced nonnative speakers are able to discern core words relatively well, they can rank the words in a lexical set only with minimal accuracy. Even accepting one transposition as a reasonable performance, native speakers usually achieve less than 50% accuracy. Thus, it does not seem that native intuitions of frequency are robust enough to reliably rank the words in a lexical set.

Some of the sets were ranked better than others and it would be useful to know the reason why. It has already been suggested that using the core word as an anchor word may have caused artificially low scores, but this does not explain other poor performances, for example in sets *chore* and *deflate*. One factor which would presumably affect how well each set was ranked is how close the words in the set were in frequency, with closer words being more difficult to rank correctly. Table 6 gives the percentage difference of the closest pair of words in each set.

Table 6 Magnitude of Frequency Difference for Closest Two Words in each Lexical Set

Set		Occurrences in BNC (88 million words)			Percentage Difference
calamity	tragedy	1873	disaster	3077	39.12
appliance	device	4854	machine	12343	60.67
chore	errand	37	chore	386	38.60
emblem	emblem	275	hallmark	328	16.15
annihilate	subdue	648	conquer	720	10.00
glisten	glisten	410	shimmer	420	2.38
deflate	minimize	1457	diminish	1734	15.97
disclose	disclose	1781	expose	3724	52.17
awful	awful	2224	terrible	3747	40.64
essential	vital	4818	essential	8235	41.49
low-tech	unsophisicated	118	rudimentary	256	53.90
flimsy	frail	544	fragile	843	35.46

The figures in Table 6 were derived in the following way. First the percentage difference was calculated for each pair of words for each set. The formula for percentage difference is:

$$100-\left(\frac{\text{smaller frequency figure}}{\text{larger frequency figure}}\times 100\right).$$

This gives a measure of the relative difference between any two frequency figures, regardless of their absolute magnitude. The lowest percentage figure within a set corresponds to the smallest relative difference in frequency, so it was reported in Table 6. The BNC OF figures are also reported for the reader's interest. Amazingly, if we check the Spearman correlation matrix in Table 7, we see that the magnitude figures in Table 6 do not correlate in a statistically reliable way with the ranking of the sets according to difficulty in

Tables 4 and 5. The nonnative correlations show some tendency toward a relationship, and the correlation between the percentage difference and number of nonnatives correctly ranking the sets approaches significance (ρ=.059). So though we may believe there is *some* relationship, this study does not produce any reliable results which would lend evidence to this belief. So we must conclude that there is no evidence that the magnitude of difference between the closest pairs of words can explain the difficulty in ranking of the sets overall. This seems counterintuitive, but there does not seem to be any ready explanation for this result.

Table 7 Correlations between Percentage Difference of the Closest Pair of Words in a Set and Ranking Performance

Native Speakers

	Correct	One Transposition	Correct + 1 Transpos.	NA
Percentage Difference	.197	.108	.181	.548
Signif.	.539	.737	.572	.065

Nonnative Speakers

	Correct	One Transposition	Correct + 1 Transpos.	NA
Percentage Difference	.559	.496	.370	-.010
Signif.	.059	.101	.236	.974

Spearman correlation, all ρ>.05

Judging The Absolute Frequency of Words

Correlation Results

The above sections discuss intuitions of how frequent words are relative to each other, but we can also ask how good intuitions are regarding the absolute frequency of individual words, ie. how often each actually occurs in language. Almost all previous studies have used correlations between SFEs and corpus word count data (OF) to answer this question. As previously mentioned, the correlations ranged very widely, from .57 (Richards, 1974) to .97 (Carroll, 1971), indicating that native-speakers do have some idea of the absolute frequency of words, but leaving unclear just how accurate these intuitions are. Correlation figures from the present study, which is much larger in scope, may help to clarify this issue.

In our study, the SFEs for absolute frequency were calculated in the following manner. The SFE rating for each word in a set was multiplied by the OF of the anchor word. For example, in the lexical set *calamity*, the anchor word was *catastrophe*, which occurred 449 times in the BNC. If a subject rated *tragedy* as 3 times more frequent than *catastrophe*, then *tragedy* would be given a rating of 1347. These figures were converted to a loglinear scale. The reason for doing this has to do with the wide range of frequency in the English (and presumably any other) language. As mentioned before, previous studies have used words which differed in frequency by many degrees of magnitude, ie. *the* [68,988 occurrences per million words] and *echidna* [about 1 occurrence per 4.5 million words] (Carrol, 1971; Shapiro, 1969). Researchers have often used loglinear scaling to compress the range of frequencies into a more useable range. The best known formula is the *Standard Frequency Index* (Carroll, 1970):

$$SFI = 10((\log_{10}(\text{proportion of word in corpus}))+10).$$

Scaling all frequency judgments to the SFI makes it more convenient to compare figures of such different magnitude. This study did not use words which varied nearly so widely as in previous studies, but logs were still used to maintain a similar methodology to those studies.

The logged SFEs were then correlated to the logged OF data from the BNC. For the total native speaker sample, the correlation was .530, whereas for the total nonnative sample the figure was .577. This suggests surprisingly that the nonnatives are superior, however, the picture becomes clearer if we look at individual and subgroup performance. The correlations for individuals within each subgroup were calculated (again using SFE and OF data on the SFI scale). Table 8 illustrates the mean, median, minimum, and maximum individual correlations.

Table 9 Individual Correlations between BNC Objective Frequency and Subjects' Subjective Frequency Estimates

Native Speakers

Group	1	2	3	4	5
Mean	.580	.542	.499	.645	.618
Median	.597	.550	.526	.648	.590
Minimum	.332	-.128	.142	.371	.340
Maximum	.735	.745	.625	.790	.837

Nonnative speakers

Group	6	7	8	9	10	11	12	13
Mean	.611	.583	.604	.638	.588	.699	.669	.625
Median	.636	.577	.619	.646	.625	.721	.672	.680
Minimum	.409	.245	.257	.424	.158	.506	.550	.331
Maximum	.806	.802	.850	.844	.785	.808	.825	.827

Although Carroll (1971) and Shapiro (1969) found that all of their native-speaking subjects performed similarly, that is clearly not the case here. Using the Kruskal-Wallis test, we find very strong evidence that there are differences in the median correlations among the native-speaking groups ($\rho<.001$). Further, there are differences between the variances within the groups, there being the most variation within Group 3 and the least within Group 5. These results, which have the benefits of a much larger subject population and more accurate objective data than previous studies, indicates that all native speakers should not be considered as a homogenous group when it comes to frequency intuitions. On the other hand, the nonnative speakers showed greater homogeneity with regard to their correlation coefficients. Although the Kruskal-Wallis test shows some evidence of differences between the nonnative-speaking groups ($\rho<.05$), if Group 11 is excluded, the hypothesis for homogeneity is acceptable ($\rho=.18$). The apparently superior performance of the Czech group may simply be an artifact of the relatively small sample size within that group however.

It is interesting to try to explain why the native groups perform differently. The ordering is not due to age or gender. Elderly respondents may well have done more poorly because the frequency of certain words have changed over the years. Words like *chore* and *errand* might have been more current at a time when they were more active in life. However, it seems the most plausible explanation is that education is a factor in better performance on this frequency task. Although the present study was not explicitly designed to test for this factor, the education profile of the groups is highly suggestive. Almost all of the members of Groups 4 and 5 had some form of postgraduate education. Group 1 consisted of students involved in secondary education. Conversely, Groups 2 and 3 had a lower average level of education, with only 23% (Group 2) and 13% (Group 3) reporting any kind of tertiary education and a large number reporting no secondary education qualifications at all. This lower level of education could be expected, since a much lower percentage of students went on to secondary education or university in Britian when these subjects were young. If we compare the educationally-active groups (1,4,5) with the others (2,3), we find a difference which is highly significant (Kruskal-Wallis, $\rho<.001$). The upshot is an apparent correspondence of average group education with performance on this frequency judging task.

Since we have found that native speakers should not be considered as a homogeneous group (see next section for further evidence of this), it makes sense to compare the nonnative speakers (who are all involved in education) with the native speakers who are closest in kind, namely Groups 4 and 5. Making this comparison, then the correlations are similar between natives and nonnatives. This is again suggestive of the importance of education, although it is still surprising that advanced nonnatives would have more accurate SFEs than lesser educated native speakers. One explanation is that this task requires numeracy skills which are integral to much academic study.

As education seems to be a factor in the results, it seems that using academically-inclined native speakers makes the most sense when developing a baseline of native ability for use with educated nonnative students. The results from the correlation analysis indicate that even with this higher baseline standard, the type of advanced nonnative speakers participating in this study have reached a stage where their frequency intuitions are generally native-like.

One advantage of calculating individual correlations is that the results are more directly comparable to earlier studies, such as Carroll (1971), which report much higher correlations. There are a number of possible reasons for the discrepancy. First, most previous frequency studies used small subject populations. Second, previous studies used words which varied very widely in frequency. For example, Carroll (1971) and Shapiro (1969) used *the*, which is the most frequent word in English, occurring about 69,000 times per million words. They also used very rare words like *echidna*, which occur only about once per 4½ million words. With such a range, there are words which obviously fit at the more frequent end of the frequency scale, while others are obviously very rare. Simply by placing these words at the correct end of the frequency scale, reasonable correlations are achieved, even if words in close proximity to each other are not accurately judged. The present study does not contain such an extreme range of frequency, with the most frequent word being *problem* (563 per million words) and the most infrequent being *low-tech* (1 per 4 million words). Importantly, the range in any single lexical set would be much less than this. In a sense, we are asking subjects to make more subtle frequency judgements about sets of words which are relatively close in frequency. Natural language use does not require judgments of unrelated words which vary wildly in frequency, rather of words which are related by meaning and so are likely to be at least somewhat similar in frequency. Thus the lower correlations in this study may well be more representative of intuitions in natural language use than the previously reported higher ones.

Hellinger Distance Modelling

Although correlation analysis has been the procedure of choice in frequency studies, it has several weaknesses and can be perturbed by a number of factors. As mentioned above, the range of frequency of the words used in studies can affect the correlation figures. Likewise, a few subjects behaving quite differently from the rest of the group (outliers) can affect correlations to a greater degree than would be indicated by their numbers. Some procedures of correlation depend on an assumption of normality which is often not strictly valid. These weaknesses are technical, but perhaps the greatest shortcoming is that correlation analysis has been used to give descriptions of group behavior, and not any useful description of what subjects know about individual words. The second kind of description is more useful when establishing benchmarks of performance. These weaknesses, together with the wide and inconclusive range of correlation figures reported for frequency intuitions, suggest that one should look for another analytical approach to compare the objective and subjective frequency data.

One such approach is to convert the estimates of relative frequencies into assessments of relative probabilities, which has the advantage of allowing better descriptions of individual words and sets than correlational procedures. We can consider SFE responses as probability statements, with subjective probability q_w estimating the "true" probability p_w for each word w in a set of five. This is analogous to saying "there is an urn containing balls, each ball having one of five different words written on it. Given the proportion of one kind of ball, estimate the relative proportions of the other balls." For each subject for each set of words, we wish to compare the estimated probability with the actual probability, or in more technical terms, to compare the two discrete probability mass functions $P=\{p_w : w = 1, ..., 5\}$ and $Q=\{q_w : w=1,...,5\}$. Of the various measures of discrepancies between probability distributions P and Q (see Diaconis, 1988, Chapter 3), I have chosen the *Hellinger distance*, defined as:

$$\mathrm{H}(P, Q) := \sum(\sqrt{p_w}-\sqrt{q_w})^2$$

This choice is somewhat arbitrary, though it appears superior in this context to the obvious alternative, the *Variation distance*, defined as $\mathrm{V}(P, Q) := \tfrac{1}{2}\sum|p_w-q_w|$, especially when some of the probabilities are rather small. Furthermore, it seems that sample values of H(P, Q) should (at least approximately) follow a Gamma distribution. This seems justified by Figures 1 and 2. As (natural) logarithms of Gamma variables are approximately Normally distributed, taking logarithms of the Hellinger distances enables the use of tests requiring a Normally distributed sample. The distributional properties of the variation distance, for example, are less convenient.

To illustrate how this procedure works, let us take one nonnative subject's responses to the lexical set *calamity*.

Word	Subject Response	Q	BNC Figure	P
calamity	2	.03770	.29	.02344
cataclysm	.05	.00094	.06	.00485
catastrophe*	1	.01885	1.00	.08084
disaster	40	.75400	6.85	.55375
tragedy	10	.18850	4.17	.33710

*anchor word

The *Q* and *P* numbers indicate the proportion of the individual figures to the totals, ie. 2÷(2+.05+1+40+10)=.03770. These proportions can be inserted into the Hellinger distance formula to obtain a single distance score.

$$H(P,Q) = (\sqrt{.023}-\sqrt{.037})^2 + (\sqrt{.004}-\sqrt{.0009})^2 + \ldots (\sqrt{.337}-\sqrt{.188})^2$$

$$= .00169 + .00152 + .02162 + .01542 + .02144$$

$$= .06169$$

The Hellinger Distances have a range from 0 to 2, with 0 indicating exact correspondence and 2 indicating the maximum possible discrepancy; thus a low value indicates a good performance and a high value a relatively poorer effort. It should be noted that any set of words including at least one missing value is omitted from the Hellinger distance data in what follows. Of course, it would be possible to calculate the score for groups where words are missing, but comparing the distances from four (or fewer) point distributions with the other five point ones would be questionable.

The Hellinger distances were calculated for each subject, for each set of words, and the results plotted for the native and nonnative speakers separately (Figures 1 and 2).

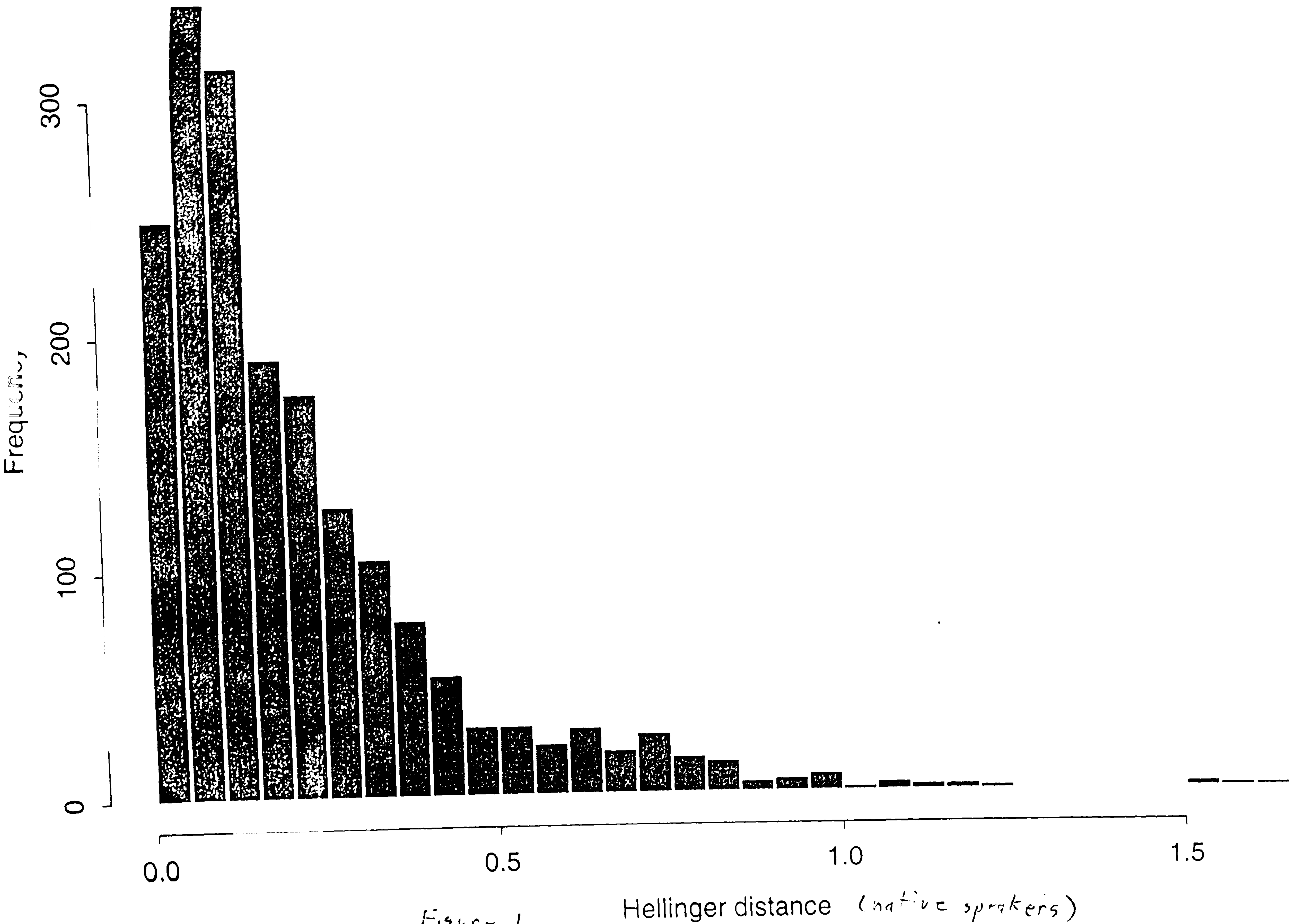
Frequency
0
100
200
300
0.0
0.5
1.0
1.5
Hellinger distance (native speakers)

Figure 1

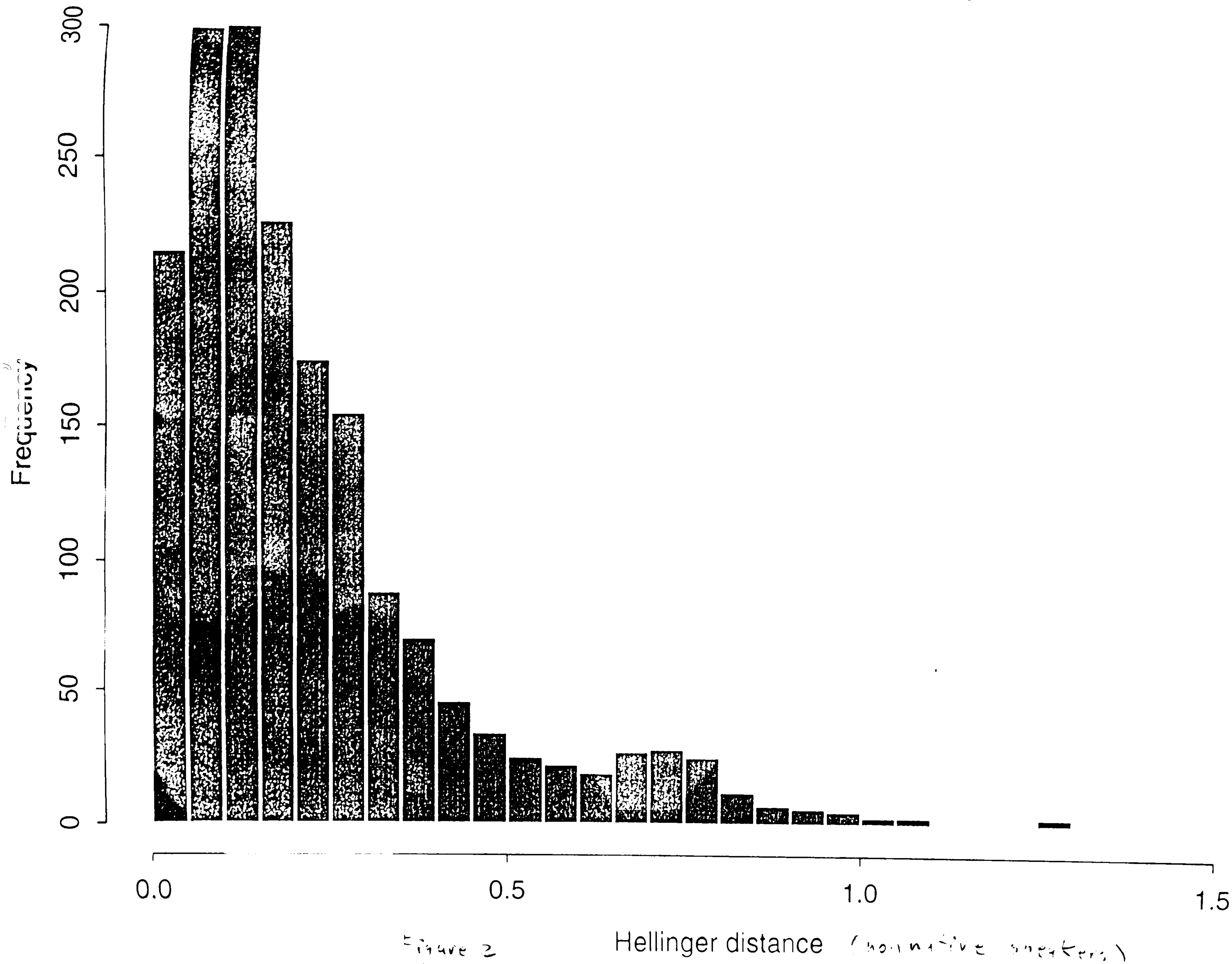
Frequency
0
50
100
150
200
250
300
0.0
0.5
1.0
1.5
Hellinger distance (non-native speakers)

Figure 2

Since scores closest to zero indicate the least disparity between the subjective and objective scores, we can see that in general, both native and nonnative respondents performed comparably. Given that zero constitutes perfect frequency assessment, the skewness towards zero in Figures 1 and 2 suggest a reasonable proportion of accurate performances. The mean Hellinger distance for the native speakers is .229. Unfortunately, there is so much variation in performance when judging the different sets, that it would be unwise to use this mean as any form of benchmark. In addition, this mean is based on all five native-speaking groups combined, which has already been shown to be suspect. Comparison of the Hellinger means for all native speakers (L1-All) versus academic native speakers (L1-Groups 4 & 5) shows clear differences, providing further evidence of the lack of homogeneity among native-speakers. If another group of natives were given this task, then one would expect similar results, but only if these word sets were used.

In order to better facilitate inferential procedures for the samples under study, we take natural logarithms of the raw Hellinger distances for each word set, and for native and nonnative speakers separately. In most cases this leads to a reasonable 'bell-shaped' Normal-like profile, although in some cases there is evidence of mild skewness in both the native and nonnative distributions. (This is due to very different profiles for the responses to the different word sets, not just in mean and variances (the first two moments) but also in shape. In some cases, extreme departures from Normality occur, including bi-modality.) This allows the use of a standard t-test for testing the equality of means between L1 and L2 groups, except in cases of unequal variance, in which Welch's modified two-sample t-test is appropriate. These tests are known to be reasonably robust to modest departures from normality, at least with respect to size. Thus, although the figures in Table 10 are raw Hellinger distances, the t and F (for equality of variances) tests are based on their logarithms.

Table 10 Hellinger Distances for Native and Nonnative Groups

Lexical Set	Mean Hellinger Distances			Variance		
	L1		L2	L1		L2
	All	4&5		All	4&5	
calamity	.085	.066	.116*+	.011	.003	.015
appliance	.277	.201	.191*	.055	.022	.041
chore	.320	.280	.285*	.019	.015	.036*+
emblem	.255	.204	.237	.068	.047	.039
annihilate	.147	.111	.192+	.032	.005	.032*+
glisten	.255	.191	.293+	.057	.049	.067
deflate	.203	.161	.205+	.046	.009	.020*
disclose	.221	.161	.197	.047	.022	.031+
awful	.415	.325	.427	.079	.058	.078
essential	.211	.159	.236+	.039	.012	.021
low-tech	.195	.156	.190	.038	.016	.042
flimsy	.159	.095	.151+	.033	.005	.013*
Overall Average	.229	.176	.229	.050	.104	.043

*L1(All) vs L2 significantly different: $\rho<.05$
+L1(Groups 4 & 5) vs L2 significantly different: $\rho<.05$

The means for the nonnative speakers' performances are not significantly different from those of the combined natives except for lexical sets *calamity*, *appliance*, and *chore*. In two of these sets, the nonnative speakers have lower means, which correspond to better performance. However, when we compare like with like, and compare the academic L1 subjects in Groups 4 and 5 with the academic L2 subjects, then the native speakers perform better. They clearly outperform the nonnatives in half of the sets (*calamity, annihilate, glisten, deflate, essential,* and *flimsy*), and in two of the others (*emblem* and *awful*) there is a suggestion of superior performances, though not quite statistically significant. Though we must be careful in combining multiple t-tests, we note that in only two groups (*appliance* and *chore*) are the means lower for the nonnatives, and then only marginally.

The results from the Hellinger distance method of analysis suggest that educated nonnatives have intuitions of word frequency which are as good or better than natives with less education. However, educated natives have better frequency intuitions than their educated nonnative counterparts.

One caveat is that only lexical sets with all words rated can be used in this analysis. That would make it more appropriate for advanced students which are more likely to be able to do this. In a way, it is an assessment of higher-level knowledge and thus may be appropriate for testing subtle levels of achievement in more advanced learners. For lower level learners, checking to see if they can find the core word in a set would probably be a frequency task much better suited to their level.

An interesting question is why certain word sets are so poorly done compared to others. One would expect that sets which contain words relatively close in frequency would be harder to rank order, but this was not borne out. The fact that the set *glisten* had three words of very similar frequency surely contributed to the subjects' poor performances in judging it, but this explanation does not extend across all sets. The percentage difference of the closest two words in each set did not correlate significantly with ranking performance, either for natives or nonnatives (Spearman, all $\rho > .05$). However, a consistent pattern of overestimation was discovered on examination of the more poorly performed sets, particularly *chore* and *awful*. For example, 94% of the native speakers overestimated the frequency of the word *chore*, 79% of them by more than a factor of 10. The figures for the other words in these two sets are similar: *duty* 69% [overestimated], 29% [overestimated >10 on SFI scale]; *errand* 98%, 85%; *job* 91%, 58%; *awful* 98%, 57%; *dreadful* 92%, 44%; *ghastly* 86%, 50%; and *terrible* 94%, 44%. The nonnative respondents had analogous results. All of these words were overestimated with respect to the anchor words *task* and *bad* respectively. This seems to indicate that the frequency of these anchor words were underestimated, causing all of the other words in the set to be rated poorly. In comparison, the words of sets with more typical performance (such as *calamity* and *appliance*) showed no evidence of systematic overestimation.

Conclusion

This study attempted to arrive at clear statements about native speakers intuitions of frequency which could be used as baselines for future research and vocabulary testing. The hoped-for clarity proved elusive, and the study must be considered a failure in terms of developing those baselines. However, it contributes much to our knowledge of frequency intuitions, for both natives and nonnatives. When considering a number of synonyms in a lexical set (presumably a lifelike task), intuitions of the frequencies of the various words can be taken into account. Native speakers are able to discern the core word in a set about 75% of the time, although they are generally unable to accurately rank the words in terms of frequency. Correlations range from .499 to .645 for native speakers and from .583 to .699 for nonnative speakers, most closely agreeing with Richards (1974) figure of .57, and casting some doubt on the higher figures reported in other studies. Overall, the advanced nonnative subjects used in this study achieved very similar performance to that of the native-speaking subjects. To the extent that these results are generalizable, it seems that learners of English, who have reached the same level of language proficiency as the nonnative-speaking subjects in this study, are able to choose the core word and rate the absolute frequencies of words in a way similar to native-speakers.

In general, it is extremely difficult to tap into intuitions of any kind, and this inevitably caused problems in this study. Until we have a procedure which can adequately handle nonresponses, it will be difficult to come to any firm conclusions about nonnative intuitions of frequency, because nonnative-speakers will always be unlikely to know as many rare words as native-speakers, making comparisons difficult. Still it is hoped that the results from this study will aid future vocabulary research and provide the basis for a renewed discussion into what a truly adequate methodology for capturing frequency intuitions might be.

ACKNOWLEDGEMENTS

Thanks to Della Summers and Keith Mardell for access to the British National Corpus and Gwyneth Fox and Andrea Lewis for access to the COBUILD Bank of English Corpus. Additional thanks to Teresa Brown, Steve Diggle, Burkhard Freund, Michael McCarthy, Adel Sakakini, Diane Schmitt, Pippa Seddon, Lenka Volavková and staff of the Centre for English Language Education (CELE) at the University of Nottingham for facilitating collection of the frequency data. Bruce Dunham kindly provided statistical advice.

CHAPTER 4 DEVELOPING A WEIGHTED METHOD FOR QUANTIFYING ASSOCIATION RESPONSES

In the previous chapter, the gap in knowledge about intuitions of frequency was at least partially filled. This chapter moves on to word associations. Eliciting associations is one of the older methods of probing the mental state and abilities of subjects, beginning with Galton (1879-1880) and gaining momentum with the normative studies in the early and mid-part of this century. More recently, researchers interested in nonnative English competence have adopted the procedure in their investigations, in an effort to determine how well English words are known, by seeing if the nonnative's associations are similar to those of native-speakers. Although using associations in this manner, as a test of L2 vocabulary knowledge, is relatively new, it holds great promise, since much richer information can often be gained from association responses compared to conventional item types. For instance, Schmitt (1995b: 114) gives the example of a Japanese subject's responses to *commit*:

commit - together   meeting   people

Where a traditional vocabulary item would merely show that the meaning of *commit* was not known, the association responses strongly suggest that the source of the problem is a confusion between *commit* and *committee*.

Read (1993, 1994) has taken this idea and developed an experimental test based on word associations, the Word Associates Test. He has chosen a receptive format to eliminate the variation inherent in productive responses. In the original version of the test a stimulus word is given along with eight possible answers. The task is to circle any option which is related in some way to the target word.

<u>edit</u>

arithmetic	film	pole	publishing
revise	risk	surface	text

(Read, 1993)

Test-takers are not told this, but in fact four of the options are semantically related associates of the target word. There are three types of associations: paradigmatic (synonyms), syntagmatic (collocates), and analytical (whole-part relations like *team-member*). The other four are distractors which are unrelated to the stimulus word in any way. (Distractors like this are very difficult to find, since creative minds can find connections between almost any two words!) In the study to validate the format, Read found the test reliable and valid as a measure of vocabulary knowledge, but discovered that guessing by test-takers was a problem. He then revised the format into the following form:

sudden

beautiful quick surprising thirsty	change doctor noise school

(Read, 1994)

The words in the left box are adjectives and in the right box nouns. Again there are four associates, with the adjective associates having a paradigmatic relationship and the noun associates having a syntagmatic relationship. A pilot study of the revised format showed that guessing may still be a problem.

Receptive association tests like these hold much promise. They can be reliable, but until a mechanism is developed to control for guessing, their validity will remain somewhat questionable. In any case, the productive elicitation of word associations is always likely to be more informative than receptive versions, as illustrated in the *commit* example. The problem with productive elicitation is that responses are unpredictable and varied, creating some difficulties in assessing their worth. Word associations cannot be judged as true associations or not, since they are by definition the associations of the individual subjects. Therefore some other criteria must be invoked when judging them, otherwise researchers would only be able to say that any association given by a subject is actually an association, unless the subject has lied. The criteria normally used are whether the associations are either adult-like (for young L1 subjects), or native-like (for L2 subjects). Norms are created by asking large numbers of target respondents and building lists of their responses. The subject responses are then matched to the norming lists.

The standard procedure in previous association research has been to match subject responses one-for-one and give equal credit for any association on the norm list (for an example, see Schmitt and Meara, 1997). This methodology has a serious drawback which does not seem to have been addressed up until now. It can be argued, at least in terms of group norms, that some associations are more central, basic, or native-like than others. For example, for the stimulus *dark*, responses will vary from those given by a large number of respondents to those given by very few or only one (Table 1).

Table 1 Edinburgh Associates Thesaurus Norming List for *DARK*

Number of different answers: 32

Total count of all answers: 99

Response	N		
LIGHT	41	FRESH	1
NIGHT	16	FRIGHTENING	1
FEAR	4	GLOOMY	1
BLACK	3	GOD	1
BRIGHT	3	GROUND	1
ROOM	3	MAN	1
AGES	2	NEGRO	1
ALLEY	2	QUIET	1
BROWN	2	SCARE	1
BENCH	1	SEE	1
BLUE	1	SHADOW	1
BODY	1	SKY	1
LOSE	1	SLEEP	1
CORNER	1	SLOW	1
DARK	1	SUN	1
DARKNESS	1	WINTER	1

(Edinburgh Associates Thesaurus, Internet Resource)

Light is clearly the most frequently given response, certainly much more frequent than *winter* for example, and it would seem uncontroversial to assert that *light* is also a more central, core, or native-like association. The problem lies in the fact that in matching subject associations with associations on the norming list, no consideration for the difference in associations is usually given. What is needed is a method of weighting the various norm list associations in order to give subjects more credit for producing typical or frequent associations than for producing associations given by one or a few respondents. This chapter will describe the development of such a weighting procedure.

Subjects

The first step is to decide on which kind of subjects to use in building the norming lists. Previous studies into L2 word associations have been somewhat indiscriminate in this matter. Often school children have been used, partly because of the ease in gathering a large number of responses, and partly because one of the main utilizations of association norms has been to evaluate L1 children of elementary school age. However, a common finding is that word associations can change depending on age and language proficiency, and possibly a number of other factors (see Chapter 2), making the use of norms based on children's responses questionable for adults. Therefore, it seems best to control for this variability by using norming respondents that are as similar as possible to the target subjects, or at least representing the goal of proficiency that the target subjects aspire to. Since the subjects being studied in the other experiments in this thesis are nonnative-speaking students from various disciplines at the University of Nottingham, the ideal norming group would consist of a variety of native-speaking students at the same university. However, one would expect native-speaking students at other British universities to have very similar, or virtually identical, associations to those at the University of Nottingham. In order to gather sufficient numbers of norming respondents, it was necessary to also use native-speaking students from the other university in Nottingham: Nottingham Trent University. The main advantage of doing this was that respondents from Nottingham Trent University were reading for different majors than the students enlisted at the University of Nottingham, ensuring a slightly broader norming base.

The norming respondents were solicited from three groups. The first consisted of 27 1st year students studying Modern English Language (MEL) at the University of Nottingham. Many of these were joint honors students also studying a parallel major in another department. The second group consisted of 36 **1st and 2nd** year business students at Nottingham Trent University. The third group was made up of 28 1st, 2nd, and 3rd year students from various departments who were taking a French course at Nottingham Trent University. The final group included 9 University of Nottingham students randomly approached on the campus and asked to complete the association task. They were studying various subjects and were at various stages of their study. Thus, there were 100 main respondents in the study. Seven of the business students left a few blanks on their instruments, so the associations for the affected prompt words were taken from extra respondents from the same business group who were not otherwise included in the study. When substitution was necessary, all three associations were taken, so that all three associations for any prompt word always came from the same respondent.

Methodology

The next step in the study was to develop the elicitation instrument. This required selection of the stimulus words. The words were chosen according to requirements of the two main studies in this thesis, the examination of TOEFL test items (Chapter 5) and the longitudinal acquisition study (Chapter

6). In essence, the 11 words which came from the longitudinal study were chosen because they were polysemous and took a number of derivative forms. The 6 words from the TOEFL study were chosen because they were the most polysemous words available from the practice TOEFL tests available (*Practice TOEFL Tests*, 1995). For full details of the word selection criteria, see the relevant chapters. This made a total of 17 prompt words, which piloting showed was at or nearing the maximum number respondents were willing to take the time to answer. They are: *abandon, brood, circulate, convert, dedicate, illuminate, launch, massive, plot, peak, rare, spur, subtle, surging, suspend, trace*, and *trend.*

The 17 stimulus words were placed on an instrument with three blanks attached to each word, in the following manner (see Appendix 4.1 for the complete instrument):

abandon ______________ ______________ ______________

Whereas most association research has asked for only one response per stimulus, this study asks for three, following Schmitt and Meara (1997). This was done mainly in an attempt to capture the richness of the various association connections between the stimulus word and the rest of the lexicon. While a subject may be able to give one native-like response, giving three supplies a more convincing illustration that the stimulus word is incorporated into their lexicon in a way similar to a native speaker. Asking for three responses also gives the subject additional chances to supply more typical associations, in case their first impression produced one which was not particularly typical of native speakers. It was very often the case that respondents would give the primary association as their second or third response. (Note that subjects did not always fill in the three blanks in a left-to-right manner, however.) Since the L2 subjects were asked for three responses, the L1 norming respondents were given an identical task, and so the norming lists were compiled on the basis of three responses per respondent per stimulus word. This point will be touched upon again later.

The respondents were given the instrument with the instructions "Write the first three words you think of when you see each prompt word on the three lines provided." The MEL students and business students were allowed to take it home and were asked to return it within a few days. The French students completed the instrument during class time. The nine respondents approached by the author outside of class completed the instrument while the author waited. A total of 100 respondents each gave three associations for each prompt word, usually resulting in 300 responses per word. Occasionally a response was illegible. In this case the appropriate norming list were consulted to see if the response in question could be discerned. If not, it was counted as an unknown idiosyncratic response. In the case of *spur*, there were 4 illegible words, but in most cases it was 0, 1, or 2 (total: 14). The 296-300 associations were first tallied on lists with the criterion that any response with a different orthographic form counted as a separate association (*religion, religions*). Then the lists were condensed by combining words at Level 1 of the Bauer and Nation (1993) morphological hierarchy. This included any base word and its inflections (eg. *control* + *controls*, *controlled*, and *controlling*). This was done because these words should all have the same underlying meaning and therefore be the same association. In the example above, despite the morphological difference, *religion* and *religions* seem to be the same association for the stimulus word *convert*. However, the situation is not so clear with derivations (*set/setter*, *mood/moody*, *move/movement*). Different members of a word family do not always have the same associations (*massive - huge, attack* ✓; *mass - huge, attack* ?, *massively - huge, attack* ?) so derivations were counted as separate associations on the lists. This seems to be relatively standard procedure, as even early studies lemmatized association responses (an example of this can be seen in the Kent-Rosanoff tally for the prompt *needle,* which was illustrated in the association section of the literature review).

The result at this point was a list of association responses for each prompt word with a tally of how often each response was given. These norming lists are presented in full in Appendix 4.2. An examination of the lists reveals the usual pattern found in association research of a small number of high frequency responses along with a larger number of infrequently given responses, including unique ones. The methodology of previous studies would match nonnative subjects' responses to the associations on this list and give a score from 0 for no matches, all the way to 3 for all matches (eg. Schmitt and Meara, 1997). This type of quantification is somewhat crude, and does not provide an answer to the question "At what point are the associations native-like?" As mentioned before, it also does not capture the difference in the typicality of the norm associations. In order to better quantify association responses, the following procedure was devised.

A Procedure for Weighting Association Responses

The three most frequent responses were identified and their frequency of response added together to make the maximum possible score. For example, for the prompt word *abandon*, the top three responses were *leave* (85), *desert* (28), and *alone* (16). So the most typical performance possible (as in most closely conforming to the norming data) would be to produce the 3 most frequently given associations, which would yield a total of 129 points. This is taken as the maximum score. But of course few norming respondents gave all three top responses. Therefore, each respondent's score was divided by 129 to gain an *association proportion* figure of the number possible. Let us take one subject's responses as an example. Their three responses were *leave*(85), *alone* (16), and *neglect* (7). Summing the associations (85+16+7=108) and dividing the result by the maximum possible score (108÷129=.837) results in an association proportion of about .84. Note that this figure is relatively high and not really representative of all native-speaking respondents. To further illustrate the procedure, let us take the responses from one of the L2 subjects in the TOEFL study (Chapter 6): *surrender* (0), *hopeless* (0), and *forget* (7) for an association proportion of .05 (7÷129). This indicates the subject's associations for *abandon* are somewhat native-like, but only to a minimal degree. The upshot is that using this procedure, it is possible to derive a numerical score which takes into account the typicality of the association responses.

Although the association proportion may capture the typicality of association responses, it is useless without some benchmarks as to what magnitude of association proportion can be considered relatively strong or weak. To provide this guidance, we must look to the behavior of the native-speaking respondents themselves. What type of association proportions do they achieve? The association proportion for each respondent was calculated for each stimulus word. All of these were averaged and the mean association proportion derived, again for each stimulus word. The summary statistics are illustrated in Table 2 (see Appendix 4.3 for a full reporting of the individual association proportions).

Table 2 Association Proportion Scores for Native-speaking Respondents

Stimulus Words	Maximum Raw Score	Mean Prop.	STD Prop.	Number of Different Associations
abandon	129	.68	.25	92
brood	75	.45	.23	93
circulate	71	.43	.19	101
convert	139	.67	.24	95
dedicate	63	.36	.21	144
illuminate	112	.67	.28	89
launch	75	.45	.23	98
massive	187	.71	.24	47
plot	75	.44	.26	102
peak	131	.58	.21	56
rare	76	.47	.25	88
spur	95	.41	.25	109
subtle	56	.36	.23	129
surging	62	.41	.21	109
suspend	104	.59	.31	102
trace	99	.47	.23	93
trend	114	.67	.28	104
AVERAGE	97.82	.52		97.05

The mean proportion for all 17 words for all respondents was .52. This average figure is not particularly informative however, as the proportions vary quite widely depending on the prompt word. For prompt words like *massive* and *convert*, there was a rather high maximum possible score, indicating that there was a high degree of agreement of the respondents' association responses. This also resulted in a rather high proportion score. Other prompt words, like *dedicate* and *subtle*, elicited a wide range of associations, with the most frequent ones being given by a relatively small number of the respondent group. This resulted in lower figures for the maximum possible score and the mean association proportion score. These results make it fairly clear that different stimulus words elicit different group association behavior. It is thus difficult to formulate any blanket association proportion criterion which would work for any stimulus word. It is probably necessary to use norming data collected for each individual stimulus word to evaluate the responses.

The mean association proportion gives us something on which to base our interpretation of responses given by nonnative speakers, since we now have some idea of native respondents' behavior. The continuum from 0 to 1.00 of possible association proportions can be broken most reasonably into 3 obvious levels, at least initially. First, if no responses are given which match those on the norming list (0 score), then that would indicate the subject has demonstrated no native-like associations for that stimulus word. Second, if several very common responses are given, then the word association performance can be considered equivalent to that of an average native-speaker. L2 subjects who achieve an association proportion for a stimulus word equal to or higher than the average native-speaking respondent are clearly at this level. A third level in which the associations are partially, but not typically, native-like exists between the first two.

As is usual with most clines, the extremes of the association proportion continuum are easy to define, but this leaves the really interesting question of whether we can set a threshold criterion for associations which are native-like. The threshold must exist somewhere between 0.00 and the mean association proportion, but before we can place it more precisely, we must define more clearly what we will accept as native-like. This definition must take two things into consideration. First, there is a great deal of agreement among responses given by native speakers. Second, and conversely, native speakers also typically give a number of idiosyncratic responses. This can be seen in the norming lists in Appendix 4.2. The unique responses were tallied for each of the 1700 cases (17 words × 100 respondents) and the results set in Table 3.

Table 3 Breakdown of Idiosyncratic Responses

	Number of Idiosyncratic Associations Within the Three Responses			
	3	2	1	0
Number of Subjects	28 (1.65%)	187 (11.0%)	573 (33.71%)	912 (53.65%)

In over half of the cases, the respondents did not give any unique associations. One unique responses was given in about one-third of the cases. But in less than 2% of the cases were all three responses to a prompt word unique. From these figures it can be seen that giving three idiosyncratic responses is not at all typical of this norming group, and giving two is not all that common either. On the other hand, it is quite usual for native speakers to give one unique response, although giving no unique response is by far the most common category.

These results suggest the following approach concerning the definition of native-likeness. While it is true that native speakers occasionally give three unique responses to a prompt word, this is an unusual situation. Thus it would be unrealistic to use this as a minimal threshold of native-likeness. It is more reasonable to take the native *group* behavior as the criterion instead of any individual native speaker, some of whom may not be typical of the norming group overall. If we accept this premise, then it is necessary to develop numerical definitions of what is *typical* of native speakers overall, but with the caveat that a very limited number of the native speakers will themselves be defined as atypical.

If we decide that individual native respondents who gave three idiosyncratic responses (and perhaps others with very low Association Proportions) are not really typical of native speaker performance, then we need some way of determining a native-like threshold which lies above their scores. One could manually examine all responses to find which respondents gave only three idiosyncratic ones and then set the threshold just above their score. This is a principled method, but it does not address the problem of respondents who achieved only a slightly higher Association Proportion, for example, two idiosyncratic responses and another with a value of 2. A respondent with such a low score would not really be typical of the group behavior either. Determining the threshold by intuitively deciding which sets of association responses are typical and which are not is obviously too subjective, as well as being very time-intensive. The best and most principled method of setting the threshold seems to be the use of descriptive statistics. One could take the mean association proportion and subtract one Standard Deviation to derive a figure which would disregard approximately the bottom one-sixth of the respondent performances. While this method worked satisfactorily for some words, it usually cut too many performances off which were clearly still in the mainstream of group performance. Conversely, subtracting two Standard Deviations set the threshold too low. After some trial and error, it was found that subtracting 1.5 Standard Deviations succeeded in eliminating the responses which seemed atypical (three unique responses and others which added up to a very low association proportion), while not discarding too many responses which seem more in line with group behavior.

To illustrate this, let us take the responses for the stimulus words *abandon, dedicate,* and *rare*, which represent the words with the highest, lowest, and central mean association proportions. They have mean proportions of .684, .356, .467 and standard deviations of .252, .211, .248 respectively. The respondents with the lowest association proportions for each word are listed below.

abandon **(Idiosyncratic responses are in bold)**

1a.	**house**	**junkyard**	**bus**	.02
2a.	hope	**game**	**car**	.04
3a.	child	**refuge**	**homeless**	.05
4a.	**kitten**	child	**family**	.05
5a.	loose	**trip**	**do**	.05
6a.	wild	gay	hope	.07
7a.	neglect	**ignore**	**redundant**	.07
8a.	ship	hope	desolate	.08
9a.	lost	**Moses**	child	.09
10a.	lonely	alone	**frightened**	.17
11a.	lost	left	alone	.19

12a.	lost	island	alone	.19
13a.	**jettison**	reject	desert	.25
14a.	loose	desert	**sacrifice**	.26
15a.	desert	isolate	alone	.36
16a.	leave	**wreck**	**destroy**	.67

dedicate

1b.	**follower**	**poet**	**athlete**	.05
2b.	**resolve**	discipline	**follow**	.06
3b.	**motivation**	**selfless**	sacrifice	.06
4b.	**message**	hardworking	**reward**	.06
5b.	involve	**value**	**constant**	.06
6b.	body	**mind**	**soul**	.06
7b.	**hard worker**	study	**student**	.06
8b.	tribute	**football**	**Bryan Robson**	.06
9b.	**no change**	**stuck with**	heart	.06
10b.	**work hard**	challenge	**success**	.06
11b.	determined	**win**	**succeed**	.06
12b.	assign	**allocate**	**sign**	.08
13b.	baptise	**christen**	**religion**	.08
14b.	passion	**supported**	determined	.08
15b.	sacrifice	**affections**	loyal	.10
16b.	**trust**	monument	poem	.11
17b.	commitment	**D.J.**	**request**	.13

rare

1c.	**none**	**kill**	**melt**	.04
2c.	**unpopular**	distinctive	**inspiring**	.05
3c.	one	never	**whole**	.08
4c.	meat	**lonely**	uncooked	.11
5c.	**banana**	rabbit	different	.11
6c.	**one off**	strange	meat	.11
7c.	species	**occurrence**	**good men**	.13
8c.	old	expensive	**original**	.13
9c.	old	never	expensive	.16
10c.	animal	**ivory**	**diamond**	.16
11c.	exotic	**luxury**	precious	.17
12c.	animal	disease	infrequent	.18
13c.	meat	breed	expensive	.18
14c.	bloody	extinct	limited	.20

Abandon is representative of words which have a high level of agreement among the subjects' responses. It has an exceptionally frequent primary (*leave*) which was given by 85 out of the 100 respondents. Thus the respondents are essentially split into two groups, those who gave *leave* and those who did not. Since such a high percentage of native-speaking respondents produced *leave*, the criteria for native-likeness for the stimulus *abandon* would ideally require this response. Using the formula of *Mean Association Proportion - 1.5 STDs* results in a cut-point of .306. Reaching this score doesn't necessitate the production of *leave*, with the production of the secondary and tertiary responses *desert* and *alone* being sufficient for that (13a). Any subject giving *leave* will clearly be above the threshold however. There are some subjects who gave what appear to be quite reasonable associations (ie. 6a and 8a), but who would not make the native threshold using the '1.5 STD' procedure. We would have to set the cut-point at about .06 to include these and still limit the more idiosyncratic performances. Using less strict criteria, such as subtracting 2.5 STDs to achieve a lower threshold would accomplish this, as would simply shaving off the bottom 5 scores to eliminate them. But as we will see, these methods will not work with stimulus words with lower communality.

At the other end of the spectrum, *dedicate* is a stimulus word with a great deal of diversity of response. There are a large number of idiosyncratic responses, and consequently numerous very low Association Proportions. Looking at examples 1b-17b, it becomes clear that it is quite common to give two idiosyncratic responses to this low communality prompt word. There is no large 'jump' in Association Proportion as there was for *abandon*, just a gradual increase. Thus the only subject whose performance is arguably not native-like is 1b, who gives three unique responses. The '1.5 procedure' provides a cut-point figure of .040, which does not exclude the 1b responses. However, if we consider what a nonnative respondent must achieve in order to reach this threshold, it still seems to work reasonably well. The maximum raw score for *dedicate* is 63, which means that a nonnative matching only one or two idiosyncratic responses on the norming list would not reach the native-like threshold (1÷63=.016; 2÷63=.032). But if the nonnative matches three unique responses (3÷63=.048), their performance is considered native-like. This threshold may seem rather low, but given the degree of diversity and idiosyncrasy of the natives, setting it any higher would not accurately reflect native behavior. Coming back to the point made in the previous paragraph, it is clear that the possible alternative solutions offered for *abandon* above will not work here.

The example of a word in the middle of the Association Proportion range is *rare*. In this case the suggested procedure works quite well, excluding the one subject with only idiosyncratic responses, as well as two other low scores.

The end result is that subtracting 1.5 STDs from the mean Association Proportion provides a threshold of native-likeness that performs very well in the mid-part of the Association Proportion spectrum, and that also gives reasonable cut-points for words at the high- and low-communality extremes. The procedure may classify more native respondents as atypical than might be hoped for, but it does succeed in providing a threshold high enough that one can be confident in the native-likeness of those who achieve it. Analyzing the procedure over all 17 stimulus words shows that the procedure is not ideal in every case, but suggests that it does provide a workable solution to the problem of developing a weighting standard which performs reasonably well for stimulus words with widely different association behavior.

On the assumption that Association Proportion scores on the lower tail of the native group distribution are not typical of native group association behavior, the threshold then represents the minimum level of performance typical of the norming group, in other words, the threshold of native-likeness. We thus have a consistent way of determining what association behavior is native-like for any particular stimulus word.

One problem in setting thresholds and cut-points is that respondents with figures just above the threshold are seldom so dissimilar from respondents with figures just short of the threshold, making it difficult to categorize the two as different. In this case, however, it turns out that most respondents falling above the threshold are comfortably above it. The first few respondents above the threshold may actually be near the boundary, but by the sixth respondent, the clearance is normally at least +.05, and usually much more. Table 4 shows the association proportions of the six respondents above, but nearest, the threshold.

Table 4 Lowest Five Native Association Proportions Which Cleared Native-like Threshold Level

Stimulus Words	Threshold Level	Respondents Nearest, Yet Above, Threshold Level					
		1	2	3	4	5	6
abandon	.306	.36	.67	.67	.67	.67	.67
brood	.094	.11	.13	.13	.15	.15	.16
circulate	.148	.15	.15	.15	.17	.17	.20
convert	.307	.63	.63	.63	.63	.63	.64
dedicate	.040	.05	.06	.06	.06	.06	.06
illuminate	.253	.33	.36	.71	.71	.71	.71
launch	.107	.13	.15	.16	.16	.16	.16
massive	.343	.37	.42	.42	.42	.42	.43
plot	.059	.07	.08	.09	.09	.09	.09
peak	.266	.28	.28	.32	.42	.42	.44
rare	.096	.11	.11	.12	.13	.13	.16
spur	.033	.05	.05	.06	.06	.07	.07
subtle	.021	.05	.07	.07	.07	.07	.09
surging	.094	.10	.10	.13	.13	.13	.15
suspend	.129	.13	.15	.15	.15	.15	.16
trace	.121	.13	.14	.22	.22	.22	.23
trend	.250	.26	.72	.72	.72	.72	.72

A Four-Level Scale of Native-Likeness

Adding the native-like threshold to the previous three levels, we can now derive a principled 4-level scale describing the native-likeness of L2 subjects' responses:

<u>Level 0</u>

*Association proportion = 0

*Produced no native-like associations

Level 1

*Association proportion = >0 and <threshold proportion

*Produced one or more associations which appear on the norming list, but not ones which are typical. Thus the association responses overall are not yet typical of the native norming group

Level 2

*Association proportion = ≥threshold proportion and <mean association proportion

*Native-like productive associations

Level 3

*Association proportion ≥ mean association proportion

*Native-like productive associations similar to those of the top portion of the native norming group

Note that both Levels 2 and 3 are labeled as native-like. Considering that large portion of the native respondents fall under the mean association proportion and into Level 2 (and a few even into Level 1), it is probably unwise to argue that Level 3 performance is any more native-like than Level 2 performance. However, a Level 3 performance necessarily includes more of the most commonly given responses than Level 2, so we can be even more *confident* in labeling a Level 3 performance as native-like.

This 4-level scale has three main advantages over previous methods of determining the nativeness of association responses. First, it quantifies the association responses in a way which results in a clear figure being produced.

Second, it takes into account the differences in typicality of association response. Third, it provides a principled way of determining whether any association performance is native-like or not.

Additional Point

A rationale for requiring three instead of just one response was previously given. At this point it may prove interesting to compare two lists derived from the two different requirements. Table 5 illustrates the norming list for *abandon* which was compiled for this study and a corresponding list from the Edinburgh Associative Thesaurus (Internet) for the same stimulus word. The respondents for both lists were British university students, and the number of subjects is the same (N=100). The main difference is thus the number of responses required.

As can be seen from the comparison, the most frequent associations on the single-response list are also found near the top on the three-response list. The only multiple-response from the Edinburgh list which is not found on the three-response list is LOSE, although LOSS does appear. Out of the 37 different responses on the Edinburgh list, 19 of them are not on the three-response list. However, only the aforementioned LOSE had multiple responses (and then only two); the others were idiosyncratic responses which one would expect to vary anyway. These results suggest that the responses with high commonality on a single-response list will also have high communality on a three-response list, but that the three-response list has the advantage of including more of the idiosyncratic responses which will normally vary considerably. The norming list compiled from a three-response task generated

Table 5 Norming Lists for *ABANDON* from this Study and the Edinburgh Associative Thesaurus

This Study

Response	N
LEAVE	85
DESERT	28
ALONE	16
GO	10
FORGET	7
NEGLECT	7
LOST	6
CHILD	5
LONELY	5
QUIT	5
BABY	4
LOOSE	4
SHIP	4
WILD	4
DESOLATE	3
DISMISS	3
FORSAKE	3
GIVE UP	3
HOPE	3
LEFT	3
REJECT	3
ABORT	2
CHILDREN	2
DEPART	2
DERELICT	2
DISCARD	2
DUMP	2
EMPTY	2
EXIT	2
FREE	2
GAY	2
HOME	2
ISLAND	2
ISOLATE	2
LOSS	2
ORPHAN	2
SCARED	2
STAY	2
VACATE	2
BETRAY	1
BROKEN	1
BUS	1
CAR	1

Edinburgh Associative Thesaurus

Response	N
LEAVE	40
SHIP	7
GIVE UP	6
BABY	5
HOPE	4
LOST	3
FORSAKE	2
LOSE	2
BOAT	1
CARE	1
CAST OFF	1
CHILD	1
COMMIT	1
DEFEAT	1
DESERT	1
DISCARD	1
DOG	1
EJECT	1
FREEDOM	1
ISOLATED	1
JUMP	1
LEFT	1
ME	1
MOTHER	1
NEGLECT	1
NO	1
NOTHING	1
NOW	1
PITY	1
QUIT	1
RELINQUISH	1
SAVE	1
STOP	1
SUBMIT	1
TEAR	1
THROW AWAY	1
WEAVE	1

(Edinburgh Associates Thesaurus, Internet Resource)

CHANGE	1
DEPORT	1
DESTROY	1
DISAPPEAR	1
DISOWN	1
DISREGARD	1
DO	1
DOG	1
DON'T WANT	1
DROP	1
ESCAPE	1
FAMILY	1
FAR	1
FLEE	1
FREEDOM	1
FRIGHTENED	1
GAME	1
GONE	1
HELPLESS	1
HOMELESS	1
HOUSE	1
IGNORE	1
JETTISON	1
JUNKYARD	1
KITTEN	1
LET GO	1
MOSES	1
MOTHER	1
OFF	1
PARK	1
RECKLESSNESS	1
REDUNDANT	1
REFUGE	1
REPLACE	1
RETREAT	1
RID	1
RUNAWAY	1
RUNNING	1
SACRIFICE	1
SENSES	1
STRANDED	1
THROW	1
THROW OUT	1
THROW AWAY	1
TRIP	1
UNCARING	1
VANISH	1
WASTE	1
WRECK	1

a much wider range of responses. This fact is advantageous in association research, since it gives the subjects the greatest chance of matching responses. Since quite a number of L1 respondents give one or two unique responses, it is natural to expect nonnatives to do so as well. A more varied list gives them a fairer chance of receiving credit for responses which, although uncommon, still make associative sense.

Conclusion

The use of word associations holds a great deal of promise in the areas of L2 vocabulary research and measurement. This promise has been rather limited by somewhat unsophisticated methodology. The descriptive procedure proposed in this paper has several advantages over previous methods of determining the nativeness of association responses. First, it quantifies the association responses in a way which results in a tangible figure being produced. Second, it takes into account the differences in typicality of association response. Third, the description of association performance is based on more than a single unit of information. Finally, the procedure provides a principled way of determining whether any association performance is native-like or not, with group typicality as the criterion.

Acknowledgements

Thanks to Robbie Dewa, Sherri Graham, Barry Harrison, Hillary Hillier, Michael McCarthy, Diane Schmitt, and Anoma Siriwardena for facilitating the data collection. The initial idea for this procedure was first generated in a brainstorming session with Paul Meara. Comments from members of the Language Testing Research Group at the University of Lancaster, particularly Charles Alderson and Caroline Clapham, helped to sharpen my thinking on associations prior to embarking on this study.

CHAPTER 5 CREATING A MEASURE OF COLLOCATIONAL COMPETENCE

As mentioned in the literature review, in the last forty years there has been an increased awareness of the importance of collocation as a part of overall language knowledge. This has resulted in some progress being made in describing the different kinds of collocation, but little has been learned about how collocational knowledge is acquired. Thus we are left with only a less-than-complete descriptive account of collocations, with no principled way to define how they are acquired or stored. As a result, even experts working primarily in this area have little or no idea of how to test collocational knowledge. This seems a serious handicap, since a reliable measure of collocational knowledge may well prove useful for more than just the measurement and description of lexis itself. Because the ability to use suitable collocates together is a crucial skill in effective writing, a measurement of this skill would surely contribute to a more accurate evaluation of compositional ability as well.

Unfortunately, almost nothing has been done to develop a collocational measurement procedure. Channell (1981) did research which required subjects to look at sentences and judge whether collocations for words they knew were either correct or incorrect, but this was a task of receptive knowledge only. It is possible that receptive collocational knowledge may aid reading, but it seems intuitive that the most important role for collocational knowledge is in the production of language. I have been unable to uncover any previous attempt to measure productive collocational knowledge, either in the literature or from my talks with colleagues interested in the area of collocation.

Even though no previous work had been done in this area, measuring collocation knowledge was still a necessary part of examining the viability of the word knowledge framework for vocabulary studies. It was therefore necessary to create a collocation measurement procedure. This chapter details its development.

Defining Collocation

Before collocation can be measured, it is necessary first to operationalize what we mean by it. Two main factors contribute to whether we consider words to collocate or not. One factor is the degree to which words *habitually* co-occur together. Almost any two words may occur together by chance; the fact that they do does not necessarily make them collocates (or at least not very strong ones). In the following extract, which is the first sentence taken from a book which happened to be sitting on my desk at the time of writing, some co-occurring words can easily be identified as collocates, while others are unlikely to be found together often.

> At a time when research in language acquisition is more versatile and widespread than ever before, it is hard to bring oneself to the brink of publication. (Entwisle, 1966, p. vii)

For example, evidence from the COBUILD Bank of English Corpus confirms *language* and *acquisition* are frequent collocates. Although not listed among *language*'s most frequent collocates, *research* must surely also be one, at least in Applied Linguistics texts. But *brink* and *publication*, although they fit together and make perfect sense in this sentence, would not be expected to occur together very often at all. Indeed, there is only instance of *brink of publication* in the entire 320 million word COBUILD corpus. Thus not every possible word combination is a collocation; it is the relative frequency of co-occurrence which partially determines whether words are collocates.

The other factor is how *exclusive* the relationship is. *Hair* collocates with many words, such as *nice, long,* and *straight*, but its strength of collocation with *blonde* is said to be strong since *blonde* occurs with virtually no other word except *hair*. Thus the fewer the words a target word habitually occurs together with, the stronger the collocational relationship. So the frequency and exclusiveness of a combination of words interact to determine the strength of collocation.

The collocational relationships between words can derive from two sources: the grammatical dependencies between words and the meaning-based relationships between words (Carter, 1987, p.47-48). Most discussions of collocation focus on lexical collocations, while grammatical collocation (or *colligation*) is normally thought of as part of grammar. This is partly because function words co-occur with most content words, making it difficult to fashion interesting statements about restricted combinationality. For example, it is not very informative to say that *the* collocates with *boy*, since it collocates with almost any other noun as well. By and large, most discussions focus on the content words which make up lexical collocations. The exception is phrasal verbs, where the absolute collocational bonds between content and grammatical words have allowed a separate lexeme with an idiosyncratic meaning to develop.

From this discussion, we can take *collocation* to mean the habitual (relatively frequent) and somewhat exclusive (the words frequently co-occur with a limited number of other words) co-occurrence of words together in discourse. This formulation is slightly more precise than the definition given by McCarthy (1990, p.158):

> **collocation:** the likelihood of co-occurrence between words. It is very likely that 'blonde' will occur with 'hair', but unlikely that it will occur with 'wallpaper'; 'blonde' and 'hair' are said to *collocate.*

Still, it does not take much reflection to conclude that these definitions are of use mainly to students and others interested in a general statement of what collocation is. They are not nearly rigorous enough to provide clear guidance in the development of a measurement procedure. For example, how likely is it necessary for the co-occurrence to be before words are said to collocate? This lack of a definition which can be objectively quantified necessarily imposes limits on any measurement procedure being developed. Without a precise definition, it will be ultimately impossible to design a measurement procedure that is able to unambiguously describe which word combinations are and are not collocations; rather, the best that can be hoped for is a procedure which can indicate typical collocates.

Norming Criteria for Collocations

In any measurement procedure, there must be some criteria of what is 'right' or desirable, even if the accepted definition is somewhat vague, as seen above.

Since collocation represents the relationship between words which habitually co-occur in discourse with a certain level of exclusiveness, we must have some way of determining what those words are. Operationalized, it means when a target word is given, which other words co-occur frequently enough with it in discourse to say that they collocate? At this point we must work on the basis of single target words and their collocates, since we do not yet have any way of visualizing a 'collocational quotient' for the mental lexicon as a whole. Thus, although there may exist some psycholinguistic collocational network which connects most or all of the words in a mental lexicon, we must now be content to make statements about specific pairs or groups of words.

There seems to be two main ways to judge how frequently and exclusively two words occur together. One is to consult large corpora and actually count the co-occurrences, and the other is to obtain native-speaker intuitions of how typical or natural collocations are. While native-speaker judgements may be a perfectly acceptable way of determining collocations, they suffer from a number of drawbacks. Chief among these is the logistical problem of arranging to have a number of native-speakers judge the various collocational candidates that a subject may produce. There is no way of knowing what combinations a subject may come up with, so each one must be individually judged by native-speakers. This may be possible in research paradigms, but it is hardy likely to be conducive to everyday measurement purposes. Also, the judgements must come after the word combinations are given by the subjects; since there is no way of knowing which combinations a subject will give, it is difficult to have a principled *a priori* discussion of possible combinations beforehand. Finally, native-speakers may disagree on the typicality of different combinations. Even if interrater reliability correlations are high, there will always be some discrepancies which are difficult to reconcile.

For all these reasons, criteria derived from corpus data may well be better suited to a norming role in a measurement procedure. Norms derived from corpora have several advantages. First, they have a degree of objectiveness. It is important to note that they are not absolutely objective, since decisions must be made about issues like where to place cut-off points in terms of frequency and exclusiveness, whether to work with individual word forms or lemmas (see below), and how balanced/representative any particular corpus is. Nevertheless, once these decisions have been made, there is a fixed standard to refer to. Second, a list of the most frequent collocates can be derived from corpora, providing a handy reference which can be used before the first subject is tested, both to predict possible answers and to inform the design of a study/test (see the development of sentence prompts below). Finally, computer programs can now quickly provide objective figures which take account of both the frequency and exclusiveness of co-occurrence within a corpus.

While common collocations should cause no problems for either of the above methods, both would likely have trouble with rare combinations, like *brink of publication* above. Groups of native-speaking judges, even if they are trained, are unlikely to agree on these infrequent cases, whereas rare combinations may simply not occur in a corpus. The problem can be minimized with corpora by using a very large ones, since this improves the chances of any particular combination being adequately represented. Corpus data may have the limitation of not being able to capture combinations at the very bottom end of the frequency continuum, but this should not matter, because rare combinations like *brink of publication* are clearly not typical and thus do not need to be catered for in testing learners.

From this discussion, it appears that corpus evidence is the most appropriate way to set norming criteria. We do have to be slightly cautious about what this evidence can tell us however. We have seen that due to an inadequate definition it is almost impossible at present to determine in absolute terms whether a word combination forms a collocation or not. Any measurement procedure is probably on much firmer ground if we interpret corpus evidence in a more restricted way, to determine whether a word combination is typical or not.

Development of the Collocational Measurement Procedure

Since no previous productive collocation measurement procedure existed, it was necessary to start from scratch. The first decision was how to elicit collocational information from the subjects. Following Scholfield (1995) that elicitation tasks should be as close as possible to 'real-life' activities, it seemed desirable to elicit collocations embedded in discourse rather than in isolation. Creating sentences and larger discourse is an everyday language activity, while explicitly thinking of collocations as separate from discourse is not. This line of thought eliminated elicitation formats such as the following:

a) Write words in either or both blanks which naturally occur in sentences with the given word.

______________________massive____________________

b) What words frequently occur together with the word *massive*? Write them on the line below.

____

It seemed a much more natural and realistic task to ask subjects to work with complete sentences. In the end, I decided that having subjects produce sentences with the target word embedded was the most appropriate elicitation method.

Once I had elicited sentences from the subjects, I would need a way to decide if the other words in that sentence collocated with the target word. For the reasons outlined above, I decided to use corpus data for this purpose. A large modern corpus was clearly desirable for my purposes, and I was able to obtain access to the COBUILD Bank of English Corpus. Table 1 gives an indication of its size and composition.

As can be seen, the corpus is heavily biased towards written discourse, with its major sources being newspapers, magazines, and books. For some research purposes this might be undesirable, but for defining the norms of collocation, it is probably an advantage. Since written discourse is generally denser and more complex than spoken discourse (McCarthy and Carter, in press), the fact that the COBUILD corpus has mostly written sources means that it is likely to have a more diverse range of collocation combinations present than if it had a higher percentage of less complex spoken discourse. In addition, subjects in the main part of the study will be asked to give sentences in a somewhat formal interview situation, which means they will likely be working in a register more in tune with written discourse than with casual conversation.

Table 1 The COBUILD Bank of English Corpus

Source	Number of Words (millions)	%	Number of Texts in Corpus
BBC	18.62	5.83	143
Economist	12.31	3.86	178
Guardian	23.90	7.48	271
Independent	19.14	5.99	185
New Scientist	6.34	1.99	138
Australian Newspapers	33.27	10.42	370
Times	20.85	6.53	155
Today	26.76	8.38	794
UK Books	40.87	12.80	572
UK Ephemera	4.79	1.50	2359
UK Magazines	29.65	9.29	760
UK Spoken	17.97	5.63	2669
US Books	32.00	10.02	327
US Ephemera	1.31	.41	939
Wall St Journal	6.06	1.19	75
Palo Alto Weekly	2.73	.85	84
US National Public Radio	22.74	7.12	726
TOTAL	319.31	≈100.00	10745

(COBUILD internal memo regarding the composition of the Bank of England Corpus as of June 25, 1996)

I extracted collocation information for six target words (*massive, peak, rare, surging, subtle,* and *trend*) from the COBUILD corpus. The target words were chosen from a TOEFL practice test as described in Chapter 6. The computer program in use at COBUILD provides two measures of collocational strength. The first is a *Mutual Information* (MI) score. For an MI score, the computer calculates how often a combination of words in the corpus occurs compared to how often the words appear separately. This score is good for highlighting words which may be infrequent in the corpus overall, but which collocate often with a target word when they do occur. The *T-score* factors in a standard deviation measure to gauge the strength of collocation. This results in a better indication of the collocates which most frequently occur in discourse. (For more on the technical details and formulas involved, see Church, Gale, Hanks, & Hindle, 1991; Stubbs, 1995). An example may clarify this. The MI method gives *downward* and *inegalitarian* as the most significant collocates immediately to the left of the target word *trend*. The T-score results give *downward* as the third most significant collocate in the same position, but does not show *inegalitarian* in the top 49 collocates. Thus, although *inegalitarian* collocates with *trend*, this combination does not occur frequently enough to be in common usage. The combination *downward trend*, on the other hand, is both collocationally exclusive <u>and</u> frequently used.

One MI list and one T-score list with a span of ±4 words were generated for each of the target words (Appendix 5.1). According to Sinclair (1991), ±4 ~ ±5 is the usual word span used in collocation research. His current thinking is that the optimal span is -5 to the left of the node word and +4 to the right (personal communication). Each list consisted of the NODE word (target word) in a central column, with four columns on either side with the most frequent collocates in each of the positions -4,-3,-2,-1,+1,+2,+3 and +4. Each column included the 50 most frequent collocates in that position, according to the calculation method used (MI or T-score). Figure 1 illustrates part of the T-score list for the word *trend*.

Figure 1 T-score Collocates for *TREND* from the COBUILD Corpus

-4	-3	-2	-1	Target	+1	+2	+3	+4
there	to	a	the	NODE	in	the	past	continue
part	is	the	a	NODE	towards	continue	towards	is
is	a	reverse	this	NODE	is	been	prices	towards
this	against	an	downward	NODE	for	towards	continue	by
has	this	bucked	growing	NODE	of	has	growth	has
also	part	long	new	NODE	toward	more	rates	to
to	bucked	if	upward	NODE	has	greater	to	in
per	reverse	against	general	NODE	<p>	recent	by	few
last	there	this	term	NODE	which	likely	reversed	increase
seems	long	bucking	recent	NODE	that	from	in	since
reverse	<p>	buck	current	NODE	continues	lower	is	more
continue	has	set	underlying	NODE	was	is	more	states
start	on	<p>	national	NODE	will	clear	last	industry
year	bucking	that	latest	NODE	among	not	companies	of
growth	reversing	started	worrying	NODE	setting	set	upwards	which
continuing	follow	reversing	disturbing	NODE	away	inflation	trend	rates
bucked	reversal	of	that	NODE	growth	continued	has	companies

(Adapted from a computer readout from the Bank of England Corpus, 1/9/97)

The maximum length of the columns that would fit on the computer screen was 50 items, which accounts for the number of collocates in each column. Considering that the 50 items multiplied by eight columns gave 400 possibilities, this was considered sufficient to build a norming baseline. Also, since it is commonly accepted that collocates in closer proximity tend to be stronger than those with wider separations, limiting the span to ±4 would capture the stronger collocates of the target words.

Just as Carter (1987, Chapter 3) had previously observed, it soon became clear when checking the lists of collocates that many of them were related in a kind of semantic field. For example, for the word *massive*, collocates included:

attack, damage, destruction, died, explosion, injuries, launched, military, refugees

amount, billion, budget, companies, debts, deficient, development, dollar, economic, expansion, financial, investment

cause, changes, increase, influx, reduce, rises, turned

These collocates can be used when referring to areas which can be roughly described as *war*, *economics* or *finance*, and *change* respectively.

This tendency for collocates to cluster within certain semantic fields allowed a principled way of constraining the range of collocations elicited, while at the same time allowing subjects a reasonable amount of freedom in creating sentences. If the subjects were allowed to produce *any* sentence, it might be on any subject and would be more likely to include the less common, but still acceptable, kind of combination this procedure is not designed to deal with.

In other words, the range of possible responses would be far too broad. An additional problem with a completely unrestricted task is that many people find it difficult to compose sentences without some parameters to guide them. Since the corpus data suggested semantic fields existed which contained numerous frequent collocates, I decided to require subjects to give sentences using these semantic fields as sentence topics, hoping to elicit collocates from the semantic field clusters already identified.

Because any target word collocates with numerous other words, it was thought desirable to ask for several sentences to better tap into this knowledge. Every target word had collocates which clustered into at least three fields which were somewhat distinct, so three sentences was the number settled upon. Sometimes a semantic field only had a very limited number of collocates (*rare* - *steak, beef*), but in these cases they seemed fairly obvious. Later piloting confirmed that native speakers produced these collocates even though the options were limited in number. Each of the three sentences would include the target word, but was to address a different topic, as determined by the semantic fields evident in the corpus lists. A prompt was given for each sentence to indicate the topic. For example, the fields and respective prompts for the word *massive* were (only the prompts were given to the subjects):

1. [war] If you were talking about war
2. [economics or finance] If you were talking about finance or the economy
3. [change] If you were talking about statistics

The prompts were critical in this task, in that they needed to guide the subject without giving any information away. So far as was possible, the prompts were written so that they did not include any of the collocates on the norming lists. They were also written so as not to give any meaning information about the target word away. The subjects were told they did not have to use the prompts in their sentences, but if they did, they were not penalized.

Once all 18 prompts were devised (6 words × 3 semantic fields) the procedure was piloted. Three international students from a Nottingham university were given the collocation task as part of the pilot for the TOEFL test study in Chapter 6. From the results, the prompts were revised. Then it was piloted on three adult English native speakers, after which the prompts were further refined. At this stage, the prompts seemed to be successful in suggesting the targeted semantic field without being overly restrictive. A complete list of the final prompts is given in Figure 2. Some of the norm collocations are given within parentheses in italics to indicate the kind of collocates the prompt was designed to elicit, but of course these were not given to the subjects.

Figure 2 Prompts Used in Collocation Elicitation Task

---

massive

1. If you were talking about war
   (*attack, destruction, explosion, launched, military*)

2. If you were talking about finance
   (*amount, budget, debts, deficit, economic, investment*)

3. If you were talking about statistics
   (*changes, increase, influx, reduce, rises*)

peak

1. If you were talking about a business
   (*career, demand, levels, season*)

2. If you were talking about a house
   (*roof*)

3. If you were talking about geography
   (*Himalayan, mountain, snow, top*)

rare

1. If you were talking about living things
   (*animals, beasts, breeds, butterflies, species*)

2. If you were talking about cooking
   (*beef, steak*)

3. If you were talking about a special person/entertainer
   (*appearance, gift, talent*)

subtle

1. If you were talking about food
   (*aroma, flavour, sweet, taste*)

2. If you were talking about communication between people
   (*approach, convey, hint, message, nuances, perceive*)

3. If you were talking about a painting, like a Monet
   (*colors, hues, lighting, shades, technique, variation*)

surging

1. If you were talking about the natural world
   (*floodwater, water, tide, waves*)

2. If you were talking about business, finance, or economics
   (*costs, development, exports, growth, inflation*)

3. If you were talking about people at a big sports or entertainment event
   (*crowd, emotions, mob*)

trend

1. If you were talking about economics
   (*business, figures, inflation, market, prices, rates*)

2. If you were talking about the clothing industry
   (*fad, fashion, setter*)

3. Use any topic you like using *trend* as a noun. But you must include an adjective which describes the noun *trend.* (a(n) ____ trend)
   (*current, dangerous, disturbing, growing, increasing*)

The results of the pilot sessions were analyzed according to the corpus norming lists. It soon became obvious that it was best to use both MI and T-score lists as norms for collocates. Since the collocates on the two lists were substantially different, this gave subjects the greatest chance for having words in their sentences match those on the list. Since collocations are not a 'closed set' it seemed better to have this broader criterion as opposed to the narrower criterion which would have resulted from only using one list.

At this point I had a small amount of pilot data which needed to be matched against the norming lists. The next problem was how to match them. Should only the words in the sentence within a proximity of ±4 be used to match with the collocates on the norm list? Did the words in the sentences have to be in the exact same position (ie. -4, -3, +2, etc.) as the collocates as they appeared in their specific columns? Or should a collocate from one of the lists be counted if it occurred anyplace in the sentence? Should the word form match be exact, or would inflected or derivative forms be permissible? The small amount of pilot data suggested that a span of ±4 would be too narrow, and that a collocate should be counted if it existed at some point in the sentence. However, more data was needed to answer this and the other questions with confidence, so the final determination was left until the main study was completed.

The collocation study proper was part of a battery of tests given to subjects to determine what TOEFL vocabulary items indicated about various kinds of word knowledge, including collocation (see Chapter 6 for details). For the purposes of our discussion here, 30 subjects who were studying at summer presessional courses before they entered universities in Britain volunteered to take part. They were interviewed individually on the six target words, and were asked to produce three sentences for each if possible, one for each of the relevant prompts. This would have returned 540 sentences (6 words × 3 sentences each × 30 subjects) if every subject produced 18 sentences, but in many cases the subjects did not know all of the target words. This is largely because the target words were extracted from TOEFL tests, and so were not particularly frequent. In particular, many subjects did not know *subtle* and *surging*. The net result was that 414 sentences were produced, which was 77% of the possible cases (414/540).

Results

These sentences provided a considerable data base from which to develop a formulation of the best criteria for matching a subject's sentence with the norming lists with a view toward most effectively measuring collocational knowledge. Issues which needed to be resolved in order to develop a workable productive collocation measurement procedure are discussed below, along with their solutions.

1. It became clear when checking the sentences that neither norming list had enough diversity of collocates by itself to capture all the collocates which my intuitions said were allowable. It seems using both norming lists (MI and T-score) in conjunction is necessary to provide the broadest and best norming standard. Since the collocates on the two lists are substantially different, this gives subjects the greatest chance for having words in their sentences match those on the list.

2. In many of the sentences, the only matched collocates were function words. This means that the presence of a function word collocate did not imply the presence of a content word collocate. Furthermore, some subjects produced sentences even though they had no idea what the target word meant (subjects were also interviewed for semantic knowledge of the target words). These guessed sentences contained a content word collocate 34% of the time (17/50), but also almost always contained function word collocates which appeared on the T-score norming list. Since function words are so ubiquitous, a number appear on every T-score norming list. When subjects guessed and created a sentence even though they did not know the meaning of the target word (and so presumably knew little collocational knowledge), they almost always included norm-listed function words in the process of building the grammatical structure of the sentence. Therefore, measuring content words as collocates seems to give information which is closer to the construct we are attempting to capture, while measuring function words as collocates seems to give information which is largely grammatical in nature. Of course, collocational and grammatical knowledge is interrelated, but an examination of the

sentences clearly suggested that the focus should be on content word collocates. As a result, it was concluded that function words should be disregarded on the norming lists, leaving only content words as scoreable collocates. It is also probably advisable to disregard other words which do not carry a 'full' content load, such as delexicalized verbs (*take, do*), general nouns like *thing* and *stuff*, modal verbs (*can, will*), and content words which could appear in virtually any sentence, such as temporal words (*now, yesterday*).

3. Typical collocations should correspond to exact lemma forms, unless corpus evidence indicates that several lemma forms are collocates. For example, *losses* appears on the collocate list for *massive*, but *loss* do not. Likewise, *start* and *started* appear for *trend*, but not *starting* or *starts*. On the other hand, the corpus shows that *breed*, *breeding*, and *breeds* all collocate with *rare*. The problem in applying this strictly lies in a weakness of nonnative performance - learners often do not have full control of inflectional suffixes (Schmitt and Meara, 1997) and frequently omit them. Considering the difficulty in determining whether a wrongly-inflected lemma is a competence error or a performance mistake, it is probably necessary to accept any form of a lemma at this point. In the future, it may be possible to develop a grading system giving a higher score for exact lemma form matches and a lower score for inexact matches.

However, occasionally the inflected form of words in English also represent multiple word classes. For example *used* is the verb past form of *use*, but it is also an adjective. As an adjective, it may strongly collocate with words it would not as a verb, ie. "He sold *used cars*" versus "He *used cars* to deliver pizza".[1] But this problem can almost be eliminated by avoiding the use of such words as target words, ie. select *use* or *useful* as a target word, but not *used.*

The various derivatives of a word may take quite different collocates. The most frequent T-score collocates of *massive* in the +1 position are *attack, scale*, *increase* and *amounts.* These seem intuitive, but it is not so easy to think of a sentence which would combine these words, in any position, with the nouns *mass* or *massiveness*. Translating these insights into scoring criteria, if subjects produce a word in any inflected form in their sentences which matches with a word on the norming lists (including irregular past forms like *find→found*), they should be given credit. However, derivative forms must exist separately on the norming lists to be counted. For example, if the norming lists include the collocate *produce*, subjects should gain credit for sentence words *produce, produced, producing, produces*, but not for the words *production* or *producer*, unless they are separately included on the norming lists.

4. Words which are included in the prompt should not be scoreable as collocates, for the simple reason that a test of productive knowledge cannot allow answers which might have been assisted by information previously given on the test.

5. The different sentences must include different collocates to earn a score. If this were not the case, a single collocate could be counted as many as three times. The purpose of using three sentences is to capture in some part the range of a subject's collocational knowledge; allowing multiple scores for a single collocate would defeat this purpose.

6. Since this procedure is not measuring spelling ability, spelling mistakes should be ignored as long as the intended word is clear. In contrast, it is not so obvious how to handle grammatical mistakes. They come into play mainly when collocates are placed in odd positions relative to the target word. For instance, is *massive money* acceptable (because *money* is on the norming lists), even though 1) a more natural and grammatical structuring would be *massive amount of money* and 2) it occurs only nine times in the Bank of English corpus? The expedient of disregarding collocational position has obvious advantages and disadvantages. The major advantage is that it avoids having to make fine judgements about grammaticality or naturalness; all that is required is checking to see if any words on the norming lists exist in the sentence in question. The disadvantage is that it limits the inference one can make about the subject's collocational knowledge. One can only say that the subject knows a common collocate for a target word; one cannot say that the subject can use that collocate in a sentence in a natural and appropriate way. A possible solution is to have separate collocation norming lists for each position in relation to the target word (ie. have a +1 list, a +2 list etc.), but relying on multiple lists would probably make the procedure far too complex and time consuming to be of much practical use. In this study, collocational position was disregarded.

7. If we are to use produced sentences to elicit evidence of collocation knowledge, we cannot control for the number of collocations produced per sentence. Subjects will produce different numbers of norm-listed collocates and there does not seem to be any help in the literature in deciding on how many collocates to require per sentence. It is therefore necessary to refer to the subjects' output in deciding on a required number of collocates per

sentence. The thirty subjects in this study produced the following number of collocates per sentence (Table 2).

Table 2 Number of Collocates Produced Per Sentence N=540

	Collocates Produced						
	1	2	3	4	5	0	Sentence Not Produced
Number of Sentences	153	68	19	12	1	161	126
% of produced sentences (n/414)	37	16	5	3	0	39	

We can see that requiring two or more collocates per sentence would mean that only about 24% of the sentences would be scored positively. This excludes a large number of sentences which seem perfectly appropriate in a collocational sense, but have only one collocate match, such as the following examples from the data:

a) The **peak** <u>season</u> for sale of fruit is in April.

b) The **rare** <u>animals</u> are protected by the government in many countries.

c) The <u>investment</u> in this country is **massive**.

d) The <u>view</u> from the **peak** was beautiful.

***Target words** <u>Collocates</u>

These sentences are the kind that native-speakers produce in everyday communication, albeit with slightly better grammatical construction. So it seems that a requirement of one collocation per sentence is the best criteria we can set from the present evidence.

8. A key question is how broad a span to use in determining whether a word is a collocate or not. Sinclair (1991) suggests a span of ±4 ~ ±5 words, although this seems to come more from experience than any principled rationale. Should words in a sentence matching those on the norming lists be scored positively even though they are quite distant from the target word? If so, how far away can be acceptable? A native-speaker judgement task was used to find workable answers to these questions. First, all words in sentences which appeared on the norming lists, according to the above criteria, were marked. Then I judged whether the marked words seemed appropriate and natural collocates of the target word. Finally, the words were tallied according to their position in the sentences in respect to the target word. This was obviously a very subjective procedure, but it proved useful in supplying at least some indication of an appropriate span length. The results are given in Table 3. Words adjacent to the target word, either before it or after it, are indicated by ±1, and so on.

Although the collocates on the norming lists came from a span of ±4, the table shows that many of these collocates occur outside this span in the subjects' sentences. As Nattinger and DeCarrico (1992, p. 22) believed, collocates can appear a considerable distance from the node word. There seems no reason to score the collocates on the norming list which occur further afield in the sentences as uniformly unacceptable. So although the vast majority of

Table 3 Confirmatory Judgement of Collocates Indicated by the Norming Lists (Number of Words) N=385

	Proximity to Target Word										
	±1	±2	±3	±4	±5	±6	±7	±8	±9	±10	>±10
Judged as Collocates	101	65	55	28	25	15	11	6	5	1	
Not Judged as Collocates	7	9	16	14	9	7	5	3	1		2

matching collocates (80%) occurred within a proximity of ±4 within the subjects' sentences, it may well be the best course is to accept a word matching the norming list, regardless of where it occurs in the sentence, in order to capture the 20% which occur outside the ±4 span.

Unfortunately, there is always the danger that these `distant' collocations are random or accidental. But that can also happen in closer proximity. If fact, the whole notion of using proximity as a criterion may be flawed. The following example illustrates this:

The <u>investment</u> in this <u>country</u> is **massive**.

Whereas the norm list collocate *investment* is clearly connected to *massive*, it lies five words away. However, the closer norm list collocate *country*, as part of a prepositional phrase, does not seem to be directly related, because the investment could be in virtually anything. This suggests that syntactic structure may also have a part to play. It might be worth exploring whether syntactic concepts are useful in more precisely defining the boundaries of collocation.

It is interesting to note that the collocates identified by the norming list procedure outlined in this chapter were largely confirmed by the judgement task. My intuitions indicated that 81% of the collocates identified were indeed natural and appropriate; only 19% were considered misidentified as collocates. Admittedly, these results are only from one rater (myself), but they suggest that this procedure (using norm lists alone as criteria) is able to achieve a reasonable identification of collocates in the subject sentences.

9. Using this procedure, scores can range from 0 to 3. A subject with a 0 score may know the target word's meaning(s), but has not shown the ability to use the word with its common collocates in sentences. On the other hand, a score of 3 indicates the subject probably has a good sense of which words collocate with the target word, even though there may still be problems in using the collocates appropriately. This 0-3 scale is admittedly a rather crude measure, but at least it is a first step in giving some indication of collocational knowledge.

Limitations of the Procedure

Since this is the first attempt to measure productive collocational knowledge, the procedure proposed inevitably suffers from weaknesses. It gives some evidence of a subject's knowledge of the most common collocates for a target word, but is not designed to make statements about less significant collocates.

There is also some element of chance; subjects creating sentences may accidentally include a norm list collocate even if they have no collocational knowledge of the target word. Future measures of collocation will need to develop a principled way to control for this.

Another weakness is indicated by the fact that subjects were able to include a content word collocate in 34% of the cases in which they did not know the meaning of the target word. This shows that the procedure is not foolproof, and that there is an element of guessing involved. Subjects creating sentences have the chance of accidentally including a norm list collocate even if they have no idea of collocation with the target word. For example, one subject produced the following sentence for the word *surging* and the sentence prompt [If you were talking about business, finance, or economics]:

The <u>figure</u> of <u>sale</u> is **surging**.

The subject did not know the meaning of *surging* and evidently came up with the collocates *figure* and *sale* by simply making a general sentence about business and inserting the unknown word at the end. A useful future step would be to develop a way to adjust for guessing like this, ideally being a principled procedure where the criteria are stated beforehand, and which doesn't require a separate test of whether word meanings are known or not.

There is a positive way of viewing the detachment of meaning and collocational knowledge however. It is possible that learners can remember word combinations without knowing the meaning of the constituents, that is, perhaps collocation can be learned without semantics. If so, this procedure can tap that collocational knowledge, even if semantic knowledge is deficient.

Conclusion

The proposed procedure obviously suffers from the lack of a precise definition of what collocation is. Nevertheless, discussion is always easier with something tangible to evaluate and constructively criticise, and this procedure at least provides that. Although it proved informative to some degree in the following research studies (Chapters 6 & 7), its weaknesses were also exposed. Perhaps its main value lies in the fact that it is a pioneering attempt to measure productive collocation knowledge, introducing procedures which can be further developed in the future. If vocabulary continues to be considered increasingly important, then a valid way to measure collocation can only be a valuable addition to the field.

Note

[1]I am grateful to Jeremy Clear for bringing this to my attention.

Acknowledgements

Thanks to Gwyneth Fox and Rosamund Moon for access to the Bank of English corpus. Discussions with Ramesh Krishnamurthy proved helpful in evaluating the strengths and weaknesses of this procedure.

CHAPTER 6 THE WORD KNOWLEDGE APPROACH APPLIED TO VOCABULARY TESTING

People are naturally interested in their progress when they are studying a foreign language. Teachers are likewise interested in their students' improvement. Since one of the key elements in learning a foreign language is mastering the L2's vocabulary, it is probably safe to assume that there has been interest in testing vocabulary from the earliest times in which foreign languages were formally studied.

In modern times, one of the first scholars to concern himself with L2 vocabulary acquisition was Ebbinghaus (1885, cited in Woodworth and Schlosberg, 1955). He ran a self-experiment in which he carefully tried to learn an imitation language and carefully charted his progress. To measure his retention of the nonwords he studied, he tested himself by means of a paired associates list. He looked at a foreign word and if he could give the English equivalent, he considered the word learned. This is one of the earliest accounts of a self-assessment method of testing.

Self-assessment may be fine for a careful researcher like Ebbinghaus, but there are obvious problems, especially the one of people overestimating the vocabulary they know. Institutionalized testing situations require measures which are more verifiable and this involves testees demonstrating their knowledge of words in some manner. Especially in the United States, this need led to an emphasis on objective testing, and the creation of a new field, psychometrics, which attempted to provide accurate measures of human behaviors, such as language learning. Spolsky (1995) believes that the first modern language tests were published by Daniel Starch in 1916. This was the time when psychometrics was beginning to establish itself. Vocabulary was one of the language elements commonly measured in these psychometric tests, and Starch's tests measured vocabulary by having testees match a list of foreign words to their English translations. This is similar to Ebbinghaus' method, except that it was only a test of receptive knowledge. Later vocabulary tests often favored items with a multiple-choice (m-c) format, both for their technical qualities, and for the fact that they were relatively easy to write if the target words were presented in isolation. Standardized objective tests became the norm in the United States from the 1930s, with vocabulary continuing to be one of the components commonly included.

In 1964, this trend culminated in the creation of The Test of English as a Foreign Language [TOEFL](ETS, 1995), which was to become the dominant test of L2 English proficiency. It has achieved worldwide acceptance (except in the UK where the UCLES tests are favored) and is the standard criterion of L2 English proficiency for entrance into American and most other non-British universities. As with other standardized tests of the time, it included a separate vocabulary section until 1976. At that time it was found that results from the TOEFL vocabulary and reading sections correlated so closely that there was arguably no need for both. Thus it was decided to combine the vocabulary and reading sections into one section.

The communicative trend in language pedagogy has influenced linguists' views, and now many (most?) believe that vocabulary should be tested in context. Congruent with this thinking, in the most recent version of the

TOEFL, implemented in 1995, vocabulary items are embedded into a reading passage used to test reading comprehension. Testees are asked the meaning of several words in the passage by means of multiple-choice questions, such as in the example below:

> The word "massive" in line 4 is closest in meaning to
>
> (A) huge
> (B) strange
> (C) cold
> (D) recent
>
> (*TOEFL PRACTICE TESTS*, 1995: 36)

The question now is how good this type of test format is. Of course, the Educational Testing Service (ETS) (the producer of the TOEFL test) is very concerned with ensuring that the TOEFL is both reliable and valid. Reliability is one of the reasons for the multiple-choice answer format; a study by Henning (1991) showed that testing target words embedded in a reading passage with a four-option multiple-choice item was the best format on a technical basis from the eight possibilities looked at. But there is a growing consensus in the testing field that test validity (whether a test actually measures what it claims to measure) is of ultimate importance. And there must certainly be questions about what the TOEFL vocabulary items are actually measuring. The most important lacuna, it seems to me, is that common vocabulary item formats, like that included in the TOEFL, have never been properly validated. There seems to be an assumption that if one particular meaning of a word can be matched to a near-synonym, then the word is adequately known. From the previous review of word knowledge, this assumption must surely be unfounded. To my knowledge (and that of other colleagues I have asked), there has been no study which has carefully explored which kinds of word knowledge this or any other kind of vocabulary item actually addresses. Thus we really have no idea whether the TOEFL vocabulary items give us any information about how well the corresponding words are known, other than about the one meaning sense which is targeted.

The word knowledge approach can provide a framework from which to design such a study. Although it would be difficult to measure all the types of word knowledge in a limited amount of time, measuring four or five would give an initial idea of how well the TOEFL items were measuring the depth of testees' vocabulary knowledge (see Schmitt, 1994, for more on measuring vocabulary breadth vs. depth of knowledge). This study will examine how much the TOEFL vocabulary items tell us about a testees' depth of vocabulary knowledge by first giving subjects a number of TOEFL items, and then measuring what the subjects know about the target words' collocations, grammatical properties, associations, and various meanings. Since there is no information presently available from which to form a hypothesis, this study will be exploratory in nature. There is an underlying intuition to be confirmed that just because a test item is scored correctly, it does not mean that everything is known about that word. Similarly, just because an item is missed does not necessarily mean that absolutely nothing is known about the word. It would be surprising if these intuitions are found to be unsupported, but it remains to be seen to what extent they hold.

Where the target words are embedded in a reading passage in the TOEFL test, it opens another dimension in vocabulary testing. Items which test vocabulary in isolation, or with minimal non-defining context, can be seen to be testing previous knowledge of the target words. But where words are embedded in an extended passage, testees will naturally try to guess the word's meaning from the context if it is not previously known. Inferencing from context is a valuable skill, but it is something different from previous vocabulary knowledge. Should the TOEFL items be seen as measuring this skill or existing knowledge of a word, or both? This study will attempt to at least indirectly inform on this question.

Development of the Study Procedure

Since the main objective of this study was to discover what subjects knew about words given on the TOEFL test, the first step was to decide on a number of TOEFL items. The TOEFL consists of three sections, the third of which is the READING COMPREHENSION section. It is in this section that the vocabulary items reside. The section consists of three short passages, each followed by a number of questions, some of which are vocabulary items like the example above. I received permission from the Educational Testing Service (ETS) to use the TOEFL in this experiment, and they sent me a practice book which included sample TOEFL tests which had previously been used (*TOEFL Practice Tests*, 1995). From the possibilities in this book, I first selected the passages which had at least three vocabulary items attached. Then I narrowed these down to the two which included words which had the most meanings and were likely to have a variety of collocations and word classes. This yielded two passages with three vocabulary items each, for a total of six items.

I decided to look at four kinds of word knowledge: grammatical word class, collocation, word association, and meaning. From previous experience eliciting information for the longitudinal study reported in the next chapter, these four word knowledge types would be all I could hope to elicit for six words in a reasonably-lengthed interview of about one hour. These particular four word knowledge types were chosen for the following reasons. The initial results from the longitudinal study indicated that subjects could usually spell words, even unknown words, so it was felt that interview time could be more usefully utilized measuring other word knowledge types. Since the elicitation procedure was an interview format and the target words would be spoken to the subjects, it was not thought appropriate to test their spoken form (pronunciation). Also, the choice of target words was very limited, and those selected did not seem to have any particularly strong register attributes. We have already seen that even the quite elaborate study on frequency intuitions reported in the previous chapter failed to produce baselines unambiguous enough to use for testing purposes, so frequency could not be assessed either.

The next step was to determine the objective norms for each of these four word knowledge types for the six target words: *massive, peak, rare, subtle, surging,* and *trend.*

I checked the meanings and word classes with three dictionaries: *The Oxford Advanced Learner's Dictionary* (1995), *The Longman Dictionary of English Language and Culture* (1992), and *The COBUILD English Learner's Dictionary* (1989). Multiple meanings for each word were accepted which seemed reasonably distinguishable from one another and which the three dictionaries essentially agreed on. Likewise, the derivative form for each of the four word classes (noun, verb, adjective, adverb) was confirmed for each word, for example, *mass* or *massiveness*, [no verb form], *massive*, and *massively*. (A more detailed description of these procedures are given in the next chapter.)

The association norms were reached as described in Chapter 4. Norms for the six words in this study were developed at the same time as the eleven words to be used in the longitudinal acquisition study in Chapter 7. The sentence data obtained from this TOEFL study was used in the development of the collocation measurement procedure as described in Chapter 5. As a reminder, this means that the collocation norming data was derived from the COBUILD Bank of English Corpus.

Once the norms were set, a provisional elicitation instrument was drawn up to ensure all interviews would be equivalent. It was first piloted in three interviews with three subjects from Nottingham Trent University. The pilot subjects were like the eventual subjects in that they were studying at a presessional course with the goal of entering a British university, and had either already taken the TOEFL or a test used for a similar purpose in England, the IELTS (1996). The results showed that the elicitation guide and interview format were both appropriate and functional, and that the total interview would take about one hour. Responses from the pilot subjects lead to some changes in the meaning and collocational prompts (see below). Since some of the answers seemed to come from knowledge inferred from the passages, a question was added to each TOEFL vocabulary item asking the subjects whether they knew the target word or guessed its meaning from the passage context. The actual question format asked each subject to tick one of two choices for each target word:

_____ I know this word

_____ I don't know this word, but guessed from the text

This additional question would allow exploration of how well the TOEFL items measured previous knowledge versus knowledge inferred from the passage.

Next, I gave the test to three native-speakers. One was an American TESOL teacher, one was a English TESOL teacher, and the last was an English musician. Their results indicated that the elicitation guide and interview were reasonably successful in eliciting their native knowledge of the target words. A key concern during the piloting process was finding the best prompts for the elicitation instrument. Between information gained from the dictionaries, the Bank of English corpus, and the native and nonnative pilot subjects, the prompts were modified to most effectively elicit the desired information without giving any of it away. The final elicitation instrument is shown in Appendix 6.1.

Subjects

This study involved 30 L2 learners of English. Twenty-seven were international students attending a summer presessional course designed to improve their academic English skills, especially composition writing, in preparation for their entrance into the University of Nottingham or another British university. Three were attending a summer course aimed at improving their general English. The subjects were all students who had either taken the TOEFL test before or who would be the type of student who would take the TOEFL test if they had chosen to study in the United States instead of England. They were all volunteers. Their average age was 25.3 (std 4.1, range 18-40), with 16 being male and 14 being female. They came from eight different countries (9 Japanese, 7 Taiwanese, 7 Thai, 3 Turkish, 1 French, 1 Korean, 1 Omani, 1 Spanish). Fifteen subjects had spent less than one month in English-speaking countries before the study, while the other 14 had spent an average of 4.8 months (std 3.6, range 1-12). One student had spent 4.5 years in English-speaking countries. Standardized proficiency scores (TOEFL or IELTS) were available for 25 of the subjects. They are illustrate in Table 1.

Table 1 TOEFL and IELTS Scores of Study Subjects

	TOEFL	IELTS
	503	4.5
	507	4.5
	510	5.0
	513	5.5
	520	5.5
	520	5.5
	527	5.5
	527	5.5
	530	6.0
	530	6.5
	557	
	577	
	600	
	607	
	610	
	----	----
Mean	542.53	5.4
STD	37.77	.61
Median	527	5.5
Number	15	10

Interview Procedure

Interviews were held with individual subjects. The average interview lasted slightly over an hour. The subjects were asked not to discuss the interview with other students. There was no time constraint during the interview, but if an answer was clearly not forthcoming (after perhaps one to two minutes with no response), usually in the collocation section, I went on to the next question.

The subject was first invited to sit down and was briefly made at ease. They were asked to fill in a biodata form which asked their name, mother tongue, total time spent in English-speaking countries, age, sex, and to check one of three possible goals of studying in the summer intensive: 1) entering the University of Nottingham, 2) entering another English-speaking university, and 3) other reason. Then they were told that the interview would focus on six words in particular: *massive, peak, rare, subtle, surging,* and *trend.* They were given a page with these words written in large bold font to refer to and the words were then read to them.

The first task given to them was the association elicitation instrument. The form consisted of the six target words, each followed by three blanks. The subjects were asked to write in the first three words they thought of when the saw the target words. It was stressed that this was not a dictionary definition task, but rather that they should write down any word which occurred to them after seeing the target words, and that the task was of a spontaneous nature. After filling in as many blanks as possible, the sheet was taken back by the researcher. The words were checked while the subject was doing the next section to ensure that all words were legible. If not, the subjects were asked for clarification after the TOEFL task. Spelling errors were ignored, as long as it was obvious what the word in question was.

The second task was taking the TOEFL vocabulary items. This consisted of two texts, both with three vocabulary questions attached. The subjects were verbally instructed to do the test section as they would in an actual TOEFL testing situation. (Many of the more experienced test takers answered the items immediately, and only used the text to confirm their answers.) The TOEFL section was as close a reproduction as possible to the version presented in the *TOEFL Practice Tests* handbook (ETS, 1995), with each text and three questions on a single page. Thus this section consisted of two pages. In addition, after each question, the students were asked to check one of two options probing whether the word was previously known, or guessed from context, as mentioned above. After finishing, the tests were taken away, and the subjects were not allowed to refer back to them in later tasks.

When the TOEFL section was finished and any questions about the associations resolved, I told the subjects that I wanted to find out what they actually knew about each of the words they answered on the TOEFL test. I explained I would ask for three kinds of information: 1) grammatical information concerning each word's word class, 2) using the words in sentences, and 3) giving the meanings of the word.

The subjects were then asked about the target words and their derivatives. First, the subjects were asked if they were comfortable with the concepts ***noun***, ***verb***, ***adjective***, and ***adverb***. Most subjects said they knew these concepts very well from their long years of studying English, but if there was any doubt at all, or if subjects seemed unsure during the subsequent task, they were instructed in the grammatical categories. They were given lists for each word class which contained numerous sample words which could only belong to that word class (Appendix 7.2). Examples of the words on these lists are nouns (*chest, accident, candy*) verbs (*agree, grow, invite*) adjectives (*chilly, clumsy, fierce*) and adverbs (*strongly, truly, always*). An example of a word which would not be on these lists is *stop*, which in the same form could be either a noun or verb. The subjects were then instructed in how the various word classes are used in English if it seemed necessary. The interview continued when the subjects indicated they were confident about the four word class categories.

The subject was then given the example of *stimulate*, which is a verb. They were told that it could be turned into a noun by changing it to *stimulation*, and into an adjective by changing it to *stimulative*. They were told *stimulate* has no common adverb. They were also given the example of *run*, which can be a noun or verb in the same form. They were then told they would be asked about the four word class forms for each target word. Subjects were given credit for cases of a nonexistent word class if they indicated that they did not think it existed. In all other cases they had to give the exact form of the word to be marked correct for that word class. Sometimes more than one form for a word class existed (*rarity* and *rareness* are both noun forms of *rare*); for these, only one form was required for a correct mark.

The subject was first asked what word class the target word was, then asked about the other word classes in this manner "Is there a (noun) form for (massive)? After one or two words, the subjects usually gave the word class information for the other words without further prompting. The only exception was *surging*. Subjects were first asked if they could give the base, root or 'short' form of *surging*. Once *surge* was elicited (almost all subjects were able to do this) then they were asked about the word classes for *surge*. The few subjects who could not do it were asked to do the task as normal with *surging*.

After the word class task was finished the subjects were told they would be asked to verbally compose three sentences, in which the target word would have to be included. It was explained to them that certain words in English naturally occur together, like *blonde* and *hair* or *jump* and *quickly*. Most subjects indicated they realized this was true. I explained to them that I was looking for sentences which contained words that 'fit' naturally with the target words, therefore they should not be creative, but rather should give the most common, normal sentence they could think of, one that they might hear 'out on the street'. I further explained that since giving *any* sentence with the target words might be difficult to do, I would give them three situations or topics to help guide them. The situation prompt words would not need to be included in the sentence, it was only to indicate the general realm the sentence should address. If the subject gave a sentence which did not exactly relate to the situation given, but included the target word, it was accepted without further comment.

As an example, the three sentence prompts for *massive* were 1) *If you were talking about war*, 2) *If you were talking about finance or the economy*, and 3) *If you were talking about statistics*. The prompts represented semantic fields which collocates on the COBUILD lists seemed to cluster around. Thus the first prompt was an attempt to elicit collocates on the list like *attack*, *military*, *explosion*, *retaliation*, and *strikes*. Getting these prompts right was especially tricky, as they had to be informative enough to suggest possible sentences within a semantic field, but without 'giving away' either the collocates themselves, or the meaning of the target words. See Chapter 5 for more on the collocation task.

The last task was to indicate the meaning of the word. All words were polysemous, a fact made clear to the subjects. The subjects were asked to explain the meaning of the words in any manner they could, including giving definitions, using the words in sentences, and drawing pictures or graphs to illustrate knowledge of the meaning. (A notepad was available for use throughout the interview). The subjects' knowledge was probed until the researcher was satisfied that they either knew or did not adequately know the meaning. After the subjects had depleted their productive knowledge of the words, if any meanings were left, they were given prompts to help cue any receptive knowledge they may have had. For example, if they knew that *peak* means the top of a mountain, but could not think of anything else productively, they were given the prompt *peak hours* or *peak season* to see if they knew the meaning in this sense. Again I would probe to see if they really knew this meaning. I was as careful as possible to make sure they did not just repeat the definition word remembered from the TOEFL test, but asked them to expand upon it, or apply the meaning to a different context. After all meanings were exhausted, we went to the next word, beginning with the collocation task.

Scoring

There were 30 subjects who each were tested on six words which make a total of 180 words. The scoring procedures for the association and collocation measures were explained in detail in Chapters 4 and 5 respectively, so they will only be briefly reviewed here. The association categories ranged from 0 to 3. Zero represents no native-like associations given; 1 stands for less than native-like association knowledge demonstrated; 2 for native-like association knowledge demonstrated; and 4 means that the associations given were similar to those given by the top portion of native-speaking norming subjects. The collocation scores also range from 0-3, and indicate the number of sentences produced which contain a collocate as listed on the norming lists.

The word class score is straightforward to interpret. It is possible to have knowledge of a word's derivatives for each of the four major word classes (noun, verb, adjective, adverb), or alternatively, knowledge that a word does not take a certain word class. The word class figure shows how many of the word class derivations the subjects productively knew.

The meaning section was scored by giving 2 points for productive knowledge of a word, 1 point for receptive knowledge of a word, and no points if the subject was not able to demonstrate sufficient knowledge of the word's meaning once the receptive prompt(s) were given. The only exception was the word *massive*. Most subjects were able to give the meaning *huge*, but were unable to add the additional description of *heavy, solid*, or *strong*. Since *huge* is really the main part of the meaning, it did not seem accurate to deny them credit for knowing it but not the additional descriptors. Therefore, subjects were given one point if they knew *huge*, but not the additional descriptors.

A maximum score would represent having productive knowledge of all listed meanings. But as the different words had different numbers of meanings (minimum two meanings, maximum four meanings), it was necessary to report the meaning scores as a proportion of the maximum score for each word. Thus the maximum score for each word was fixed as the number of meanings multiplied by two (2=productive knowledge). For example, *rare* had four meanings, so the maximum score would be 8. A proportion of 1.00 represents productive knowledge of all listed meanings (or a score of 8 in the case of *rare*). A score of .500 would indicate either productive knowledge of half of the listed meanings, or receptive knowledge of all listed meanings, or some combination of the two.

Results and Discussion

The results from the TOEFL test will be analyzed along two main dimensions: 1) whether the appropriate option was chosen and 2) whether the subject indicated they knew the word or whether they did not know the word and had guessed its meaning. Thus there are four possible combinations: a known word combined with an appropriate answer on the test, a known word with an 'incorrect'[1] answer on the test, a word which was guessed with an appropriate answer given, and a guessed word with an 'incorrect' answer given. Table 2 gives the results of the association, word class, collocation, and meaning measures for each of these TOEFL result categories.

Table 2 TOEFL Results vs Association, Word Class, Collocation, and Meaning Results N = 180

TOEFL	Known/ Guessed	N	Assoc (0-3)	Class (0-4)	Colloc (0-3)	Meaning (prop.)
Correct	Known	83	1.386	2.265	1.747	.540
Correct	Guessed	53	.566	1.868	.906	.319
Correct Total		136	1.066	2.110	1.419	.454
Incorrect	Known	14	1.286	2.286	1.929	.455
Incorrect	Guessed	30	.267	1.333	.767	.117
Incorrect Total		44	.591	1.636	1.136	.224
	Known Total	97	1.371	2.268	1.773	.527
	Guessed Total	83	.458	1.675	.855	.246

Meaning and TOEFL Responses

Table 2 supports the intuition that if a TOEFL item is correctly scored, it does not indicate that all of the word's meanings are completely known. The average meaning proportion score for correctly marked TOEFL items shows that subjects know about half of the possible meaning content of a word by this study's criteria. This typically meant that the subject knew the most common meaning sense productively and perhaps one or two other meaning senses receptively. If the word had uncommon meaning senses, few of the subjects knew any of these. Using the word *rare* as an example, subjects usually knew the most frequent meaning sense [unusual or uncommon] productively, and often knew the meaning sense [lightly cooked meat] either productively or receptively, but seldom knew the senses of [thin or light air] or [unusually good, extreme, or remarkable as in a *rare fright*] at all. So, on average, subjects who correctly marked a TOEFL item had at minimum a productive knowledge of the most common meaning sense of that word.

For TOEFL items that were incorrectly marked, the subjects had an average meaning proportion score of .224. This is nowhere close to .000 and indicates that subjects who miss a TOEFL item often have some meaning knowledge of that word. For words with three or four meaning senses, this would mean that the subject knows about one of the meaning senses productively or two senses receptively. For words with only two meaning senses, it would mean that a subject knows one of them receptively. Since if any meaning sense was known, it was usually the most common one, this figure suggests that subjects on average knew the most common meaning sense for the words which they missed the TOEFL items for, at least receptively. An examination of the elicitation instruments shows that this was very often true. For the 44 incorrectly marked TOEFL items, in 19 cases the subjects were able to productively demonstrate knowledge of the most common meaning sense, and in one case receptively. If the results of one particularly weak student were disregarded, the number of incorrect TOEFL items drop to 38, which would mean that about half of the remaining subjects (20/38) knew the most common meaning sense of words they missed on the TOEFL.

Meaning and Self-Judgement of Meaning Knowledge

The subjects indicated that they knew the target words for 97 out of the 180 items. If we take a non-zero meaning proportion to show that the words were indeed at least partially known, then Table 3 indicates that the subjects were fairly accurate in their intuitions of whether they knew a word or not, actually having some meaning knowledge for 91 words. The column reporting all words with a meaning proportion over zero shows that many of the words judged as unknown but guessed were actually known to some extent. This could be because they were known unconsciously or because some learning took place during the execution of the TOEFL test through an inferencing process.

Table 3 Known Words and Corresponding TOEFL Results

TOEFL ANSWERS	Judged as Known by Subject		Judged as Known and Actually Known {meaning prop. > 0}		All Words with meaning prop. > 0	
	N	%	N	%	N	%
Correct	84	86.6	79	86.8	109	83.2
Incorrect	13	13.4	12	13.2	22	16.8
	97	100	91	100	131	100

There is not a great deal of difference in the percentages obtained from any of the three above methods. Taken together, they suggest that when words, in which the meaning(s) are at least partially known, are tested by TOEFL vocabulary items, the appropriate option will be chosen at a percentage somewhere in the mid 80s, while the item will be missed at a percentage in the mid teens.

The same analysis can be done for words which are judged as unknown.

Table 4 Unknown Words and Resulting TOEFL Results

TOEFL ANSWERS	Judged as Unknown by Subject		Judged as Unknown and Actually Unknown {meaning prop. = 0}		All Words with meaning prop. = 0	
	N	%	N	%	N	%
Correct	53	63.9	22	51.2	27	55.1
Incorrect	30	36.1	21	48.8	22	44.9
	83	100	43	100	49	100

While the TOEFL is reasonably good at distinguishing words which are known, it does not seem to work as well with unknown words. For words which were unknown, as determined by a .000 meaning proportion score, over 55% of the corresponding TOEFL items had the appropriate option selected. It has to be a bit disturbing when there are more correct answers for unknown words than incorrect answers.

Of course the subjects probably selected their answers on the TOEFL items from information inferred from the passages, but in these cases that knowledge must not have been retained, or else it would have shown up on the meaning measure. In 27 cases the subjects were able to choose the correct option, but this did not translate into being able to express this meaning in the meaning task. (Merely repeating the option answer from the TOEFL was not accepted as sufficient knowledge of a meaning of a word.) This is not surprising since the subjects indicated they did not know the word and the choice of option was only a guess. Note that 5 words rated as known were in fact totally unknown from a meaning standpoint (22→27).

Association Knowledge and TOEFL Responses

If we look at correctly and incorrectly answered TOEFL items in Table 2 and see what level of association nativeness they represent, we get the following picture. For items scored as correct, the average association score is 1.066. This means the level of association nativeness is about the same as one of the native-speaking norming respondents who gave three idiosyncratic responses. In other words, this is the lowest level of nativeness possible, assuming the respondents filled in all three associations. This score indicates a degree of nativeness, but is not at the level which would indicate associations which are representative of the overall norming population. None of the more common associations given by a number of native-speaker respondents would been given at this level. In short, this is quite a low association figure for words which are supposed to be known.

For incorrectly marked TOEFL items, the association score is .558. As an average, this means that one or two native-like, but idiosyncratic, associations have been given. So overall, if the appropriate option in a TOEFL item is chosen, it cannot be construed to mean that a subject has native-like associations for that word. On the other hand, if the TOEFL item is missed, the subject may well posses one or two associations which only one or a few native-speakers have given in a similar task.

Table 5 Frequency of Occurrence in Association Category N=180

TOEFL ANSWER	0		1		2		3	
	N	%	N	%	N	%	N	%
Correct	56	41	35	26	25	18	20	15
Incorrect	30	68	5	11	6	14	3	7

But as we can see from Table 5, the averages alone give a slightly misleading picture. In terms of raw numbers, 35 out of 136 (26%) subjects who correctly marked the TOEFL items also had an association score in Association Category 1. 45 out of 136 (33%) showed native-like associations by fitting into Categories 2 and 3. But 41% did not produce any native-like associations at all. So although about one-third had native-like associations to match the correct TOEFL items, an even higher number had no native-like association knowledge according to the criteria of this study. This reinforces the implication that we cannot assume native-like association knowledge from the results of the TOEFL vocabulary items.

For the TOEFL items scored as incorrect, 30 of the 44 (68%) items in the *Incorrect Total* category had association scores of 0. The remaining 14 had higher scores which moved the average up. So about two-thirds of the incorrectly answered items corresponded with a complete inability to give a native-like response which appeared on the norming list. It seems we can be somewhat more confident in assuming that a negative TOEFL result indicates lack of association knowledge than we can be in assuming that a positive TOEFL result indicates some degree of associative knowledge.

In sum, if a TOEFL vocabulary item is scored correctly, it cannot be taken to mean that the testee associates that word with others in his or her mental lexicon in a native-like way. Perhaps a third of the correctly-scored words might have native-like associations, but an even greater number are unlikely to have nativelike associations at all, at least as measured by this type of 3-response task. If the TOEFL item is scored incorrectly, we can be more confident that the subject does not have native-like association knowledge for that word, although in about 20% of the cases subjects may indeed possess that knowledge against the test's indication. In short, TOEFL vocabulary items do not seem to be a very good indicator of associative word knowledge.

Word Class Knowledge and TOEFL Responses

If a person knows a word, he or she needs to know its word class if the word is to be used productively. In addition, it is very useful to know how to form the other members of the word's family in order use it in different situations in different parts-of-speech. How well do the TOEFL items measure this grammatical kind of word knowledge? Table 2 shows that, on average, subjects who chose the correct TOEFL option knew how to form the target words in a little over two word classes. In contrast, subjects who missed the TOEFL item knew the words in 1.636 word classes. These figures are informative, but it is perhaps more illuminating to examine the raw frequency data in Table 6.

Table 6 Frequency of Occurrence of Number of Word Classes

	0		1		2		3		4	
	N	%	N	%	N	%	N	%	N	%
Correct	4	3	30	22	59	43	33	24	10	7
Incorrect	7	16	14	32	13	29	8	18	2	5

The frequency distribution table shows that for correct TOEFL items, there is a normal distribution, with the mode being 2 and only a few cases corresponding to 0 or 4 word classes. For incorrect TOEFL items, only 16% of the time were subjects unable to indicate a single word class, and in over half of the cases were able to name two or more word classes.

The subjects were largely relatively successful in naming word class. Only in 6% of the cases (11/180) were they unable to give any word class information. They were usually able to give the word class of the target word (87% 157/180), and in cases with only one word class correct, it was usually the word class of the target word as it was presented (70% 31/44). Usually they were able to give information about one or more of its derivatives as well. In 69% (125/180) of the cases, subjects were able to give the word forms for two or more word classes.

One factor which might have inflated these scores involves the way the word class measurement was structured. Since only *trend* took all four parts-of-speech, the other words allowed subjects to exhibit knowledge of a nonexisting word class by explicitly stating that it did not exist. Although I was as careful as possible to confirm that subjects had a distinct impression that a word did not take a particular word class, often subjects gave one or two derivatives and asserted that there were no other parts-of-speech. In this way the subjects sometimes correctly covered nonexisting word classes in their blanket statement of nonexistence. So it was slightly easier to state that a word did not take a certain word class, than it was to actually give the derivative.

It is also interesting to note that only a very small percentage of subjects were able to give all four word classes for the target words. In the incremental acquisition of words, it seems that learning all members of a word family, as learning native-like associations, is something that is difficult to master by L2 learners, even advanced learners studying to enter British universities.

In sum, if subjects correctly answer a TOEFL vocabulary item, it is very unlikely that they do not know at least the target word's word class, and there is about a 75% chance that they will know its form in two or more word classes. If the TOEFL item is missed, there is still only a 16% chance that the word's part-of-speech is not known, with around a 50% chance of two or more word classes being known. Taken together, this data suggests that knowledge of two classes is a reasonable approximation of the common state of word class knowledge in this level of student.

Collocation Knowledge and TOEFL Responses

The collocation measurement score indicates in how many sentences out of three possible that the subject was able include a content word appearing on the list of most frequent collocates derived from the Bank of English corpus. For words which were correctly answered on the TOEFL test, the subjects were able to give collocates for 1.419 sentences. For words which were missed on the TOEFL test, subjects were able to compose 1.136 sentences which included a collocate. This does not seem to be a great difference, and looking at the frequency of occurrence figures in Table 7 reinforces this feeling.

Table 7 Frequency of Occurrence in Collocational Category N=180

TOEFL ANSWER	Collocational Category 0		1		2		3	
	N	%	N	%	N	%	N	%
Correct	35	26	35	26	40	29	26	19
Incorrect	16	36	9	20	16	36	3	7

Results from TOEFL items do not seem to give any clear indication of collocational ability with the respective target words, especially for TOEFL items which were scored correct. There was no statistically reliable difference between the distribution of results obtained for correct TOEFL items in Table 7 and those that would be obtained by chance. This means that the TOEFL items did not give any information on the degree of collocational knowledge a subject might possess. If a subject misses the TOEFL item for a certain word, then it is quite likely (36%) that s/he will be unable to create even one sentence including a common collocation and very unlikely that s/he will be able to give 3 sentences (7%). However, there is over a 50% chance that s/he would be able to give either one or two sentences. For TOEFL items that were missed, there is not the consistent decay one would expect in number of sentences given; rather, more subjects were able to give two sentences than one.

In sum, the TOEFL test items do not seem to adequately measure collocational knowledge. A correct TOEFL item does not reliably distinguish between the different degrees of collocational knowledge, and while a incorrect item does indicate the likelihood of no collocational knowledge according to the criteria of this study, it does not distinguish between the middle range of collocational knowledge satisfactorily. In the end, the subjects in this study were able to compose between zero and two collocationally appropriate sentences on average with no clear connection to the TOEFL result.

Knowing a Word versus Inferencing from Context

Of the 83 TOEFL items in which the subjects did not know the word and had to guess, 53 were answered correctly. If the subjects were guessing purely at random, we would expect a figure of approximately 21 correct answers from a 4-option multiple-choice format. Clearly, the subjects were exceptionally successful in their guesses (64% correct). The success could be due to a number of possible factors. The subjects may have known the words all along, but were not confident enough in their knowledge to rate them as known. They may also have had some unconscious knowledge of the words. Another likely reason is that they did not know the words beforehand, but inferred their meaning from the attached text. A close look at the passages in question suggests that a person skilled in inferencing could derive the appropriate answer with no prior knowledge of the word. For example, let us look at the TOEFL passage paragraph in which surging occurs.

> Basic to any understanding of Canada in the 20 years after the Second World War is the country's impressive population growth. For every three Canadians in 1945, there were over five in 1966. In September 1966 Canada's population passed the 20 million mark. Most of this surging growth came from natural increase. The depression of the 1930's and the war had held back marriages, and the catching-up process began after 1945. The baby boom continued through the decade of the 1950's, producing a population increase of nearly fifteen percent in the five years from 1951 to 1956. This rate of increase had been exceeded only once before in Canada's history, in the decade before 1911, when the prairies were being settled. Undoubtedly, the good economic conditions of the 1950's supported a growth in the population, but the expansion also derived from a trend toward earlier marriages and an increase in the average size of families. In 1957 the Canadian birth rate stood at 28 per thousand, one of the highest in the world.
>
> (*TOEFL Practice Tests*, p.80, underline mine)

Although the three distractors in the m-c item (*new, extra,* and *surprising*) are plausible, there seem to be enough clues available in the above paragraph to enable testees to come to the appropriate meaning *accelerating* if they are skillful enough in using them. In other words, the contexts seem to be rich enough to enable the inferencing of meaning.

The design of the study allows a principled, if indirect, examination of the question of guessing from context in the TOEFL passages. In order that the association measurement would not be contaminated by the context given in the TOEFL passages, the association task was given first. Thus it gives an indication of the state of the subjects' knowledge of the word before exposure to the TOEFL test. While the association score of .566 for correctly guessed TOEFL items was higher than the score for incorrectly guessed items (.267), it is still low enough to indicate a minimal knowledge of the words at best. In fact, most of the subjects who correctly guessed the appropriate option could not give even a single native-like association (Table 8), indicating little or no knowledge of the word's meaning.

Table 8 Number of Items in each Association Category for Guessed TOEFL Items N = 83

	Association Category			
	0	1	2	3
Guessed/Correct	36	8	6	3
Guessed/Incorrect	26	1	2	1

If we allow the assumption that inability to produce native-like associations indicates lack of knowledge of a word, then subjects did not previously know 68% (36/53) of the words which they were able to chose the appropriate TOEFL option. If the subjects did not have previous knowledge, then the only other source was the TOEFL test itself. Since the multiple-choice items themselves are written in a way not to give away the correct option, then this clearly suggests that the subjects were inferencing the meaning from the texts. In fact, a number of subjects explicitly reported that they did not know the target words, but that they had guessed the meaning from the texts. Although almost everyone agrees that inferencing from context is a positive thing, we may have to reconsider its relationship to the TOEFL vocabulary items. The items may measure existing knowledge about vocabulary in a majority of cases, but this study gives a tentative indication that in a substantial minority (20%, 36/180) of cases the items are measuring inferencing skills. Thus we may have to reconsider what the TOEFL vocabulary items are measuring to include inferencing in addition to prior vocabulary knowledge.

The above argument uses correct TOEFL item results as evidence of having knowledge about a word. In this study we have a much better indication of whether the word's meanings were known, the meaning proportion score. We can thus determine how well previously unknown words (as evidenced by the same zero association score) are known in regards to meaning (meaning proportion score). These results are illustrated in Table 9.

Table 9 Words with Zero Association Scores and their Respective Meaning Proportions
N = 86

Number of Words with the Following Meaning Proportions:

.00	.10-.19	.20-.29	.30-.39	.40-.49	.50.-59	.60-.69	.70-.79
48	3	8	5	0	13	3	6

If we tally all of the non-zero meaning proportions, we find that for 44% (38/86) of the words for which no native-like associations could be given, some form of meaning knowledge could be demonstrated. 40% (35/86) of the words had a meaning of .250 or higher (.250 was the lowest proportion in the .20-.29 range). At this meaning proportion, the subjects clearly had some idea of the meaning(s) of the target word. If we take the meaning proportion of over .50 to show substantial meaning knowledge of a word, then about 25% (22/86) of the words with zero association scores were well known when it came to meaning. Being able to give native-like associations does not equal meaning knowledge, of course. It is extremely unlikely that the .50 or higher meaning proportions for 25% of the words could have been learned exclusively from inferencing. However, it is suggestive that such high percentages of subjects could demonstrate meaning knowledge after they could not give native-like associations. To the extent that the associations given did represent their prior knowledge of the words, it seems likely much of the demonstrated meaning knowledge was gained from an inferencing process of the text, since there was no other source for it. Thus, this gives additional indirect evidence that the subjects relied on inferencing for some words on the TOEFL test.

Limitations of the Study

The results obtained in any study are only as good as the data used to obtain those results. Thus each of the word knowledge elicitation methods needs to be evaluated for its limitations and generalizability. I would suggest that the meaning elicitation method is quite satisfactory. A considerable amount of time was spent determining whether any particular meaning sense was known or not, and the prompts allowed receptive as well as productive knowledge to be measured. Although I feel confident that the interview provided as accurate a result regarding the subjects' knowledge of the meanings of words as possible, it must be said that it was not always a straightforward decision whether a meaning was known or not. Sometimes a subject initially seemed to know a meaning, but on further exploration found their understanding of the meaning was far too broad or narrow.

Similarly, the association task seems sound. The norming group were native-speaking university students who completed exactly the same task as the nonnative subjects. The fact that three associations were asked for each target word both provided a broad pool of native-like associations for the subjects to match, and also gave the subjects three chances to provide the more frequent responses. So the association task gave the subjects every chance to demonstrate their association knowledge. The scoring system is not limited to how many associations a subject gives, but provides a clear indication of how native-like they are.

The measurement of word class is straightforward when the subjects gave word forms for the various word classes; either they matched or they did not. A possible problem exists when a target word does not take a particular word class. With the scoring system used in this study, it seems easier for subjects to believe that a representative of a certain word class does not exist than it is to produce a word form if they believe it does exist. Subjects must be given credit for knowing a target word does not take a certain word class, but it would be useful to avoid subjects giving blanket denials of the existence of word classes and being correct in only one case. This problem could be avoided in the future if only target words were chosen which took all four word classes.

Finally, collocation is the one type of word knowledge which can be considered problematic. This is the first time to my knowledge that anyone has attempted to measure collocational ability. The measurement is obviously crude as it makes no distinction whether the collocate occurs in the usual position regards the target word, or if it exists in a grammatically appropriate position for that matter. The collocational task can be seen to give only the broadest measure of collocational knowledge. Since collocation is becoming recognized as a key type of vocabulary (and language) knowledge, it would be very useful if a truly good measure of it were developed. It is hoped that fellow researchers will see this first attempt with its weaknesses and, believing that they can do better, develop the next generation of collocation test.

Conclusion

This study gives support to the intuition that TOEFL vocabulary items give only a limited amount of information about the wider range of word knowledge necessary to master a word. The items were not particularly strong in indicating a subject's association, word class, and collocation knowledge of target words. They are not as robust as desirable in indicating knowledge of even meaning, if meaning is taken to be all of the main meaning senses of a word. Of course, it is not claimed that these items do give this kind of information. However, if vocabulary is considered an important enough language component to be included in tests like the TOEFL, the question remains of whether items like this are good enough, or whether the field needs to develop items that give a far better measure of the depth of testees' knowledge of words.

NOTE

[1] Although the terms *correct* and *incorrect* are used in this chapter for convenience sake, it is acknowledged that they are problematic in that they imply that there are absolute values to be applied in language tests. Although test designers attempt to write items for which a best answer can be logically argued for with evidence from dictionaries, native intuitions, corpus data etc., it is debatable whether most test items can be infallibly judged as 'right' or 'wrong'.

ACKNOWLEDGEMENTS

The vocabulary passages and items from the TOEFL test were reprinted by permission of Educational Testing Service, the copyright owner. Thanks to Robbie Kantor for facilitating this permission. I am also grateful to Hillary Bool, Rebecca Hughes, Diane Schmitt, and Roger Smith for allowing me to call for volunteers in their language programs. Finally, thanks to the many people who commented on a presentation of this chapter at the 1996 AILA conference in Jyväskylä, Finland.

CHAPTER 7 A LONGITUDINAL STUDY OF VOCABULARY ACQUISITION

The previous chapters have dealt with the measurement of various kinds of word knowledge, and also the evaluation of the TOEFL vocabulary items. These issues are interesting in their own right, but they must take a back seat to what is probably the key question in vocabulary studies, "How are words acquired?" Up until now, most studies have looked at the size and growth of lexicons, and have been concerned with how many new words are gained (or attrited) over time. This line of research is important, but it does not explain how *individual* words are acquired over time. At the moment we have very little knowledge of the stages that words pass through in the acquisition process, or even whether that process is in any way regular or uniform. One obvious reason for this lack of knowledge is that the process of lexical acquisition shows every sign of being complex, and so very difficult to study.

Recognizing the difficulty of the task at hand, this study will nonetheless attempt to begin the description the lexical acquisition process. One way of trying to describe and analyze a complex process is to divide it into more easily handled components, and that is the approach this study will adopt. It will attempt to provide one of the first descriptions of how individual words are acquired by measuring the change in several word knowledge components of a limited number of target words over time. The research design will basically be a multiple case study. This is because the current lack of knowledge about the lexical acquisition process makes it difficult to devise hypothesis testing experiments. It was felt that a good first step would be a qualitative study which would help to better define the issues, and uncover some promising lines of future enquiry. As such, it is exploratory in nature, and although the conclusions reached will no doubt be tentative, it is hoped that they will substantially add to our existing base of acquisitional knowledge.

Designing the Study

It does not seem controversial to state that full mastery of a word is gained incrementally over a period of time. Of course, certain aspects of a word can be learned in one exposure, such as a particular meaning sense and the word's spelling, but it seems impossible to acquire complete knowledge of a word from only one or a few exposures. For one thing, any particular context will only involve a single meaning sense, and most words are polysemous. If indeed words are acquired incrementally, it stands to reason that the best way to track the acquisitional progress of those words is to follow them over time, requiring a longitudinal study. In addition, various word knowledges may be acquired at different rates or at different stages in the acquisition process, so each one will have to be tracked separately. This means that the measurement for each word at each measurement point will be quite involved. Thus only a limited number of words can be studied at a time. The points in this paragraph suggest that the study should be a longitudinal one which examines a limited number of words in great detail. This is in line with Meara's thinking (personal communication) that such an in-depth look at a small number of words over time may provide interesting information.

Target Words

Since the mastery of several different types of word knowledge will be measured for each word, the overall test will be rather time-consuming. It was initially estimated that between ten and twenty words would be the maximum which could be assessed in a session length which would be acceptable to volunteer subjects. This small number of target words meant that word selection was especially important. An ideal target word would have several characteristics. There would have to be a high likelihood that the study subjects would be exposed to it over their course of university study in England. It would have to be amenable to incremental measurement, suggesting polysemy (one way of measuring incremental meaning knowledge is to measure the acquisition of different meaning senses). Frequency was also a factor. Since the subjects would come from different departments, it was decided that words common in a variety of academic texts would be the most suitable, since subjects in any area would be exposed to them. Therefore the University Word List (UWL) was consulted. First, the 800 or so words on the list were scanned for items with multiple meanings. Most of these polysemous words were verbs or adjectives. The resulting word candidate list was further shortened by the criterion of having at least three commonly-known meanings (as judged by the author's intuitions of the various meaning senses given in three dictionaries). This left 15 words. Since the eventual subjects might know all of these words at least partially (their English ability was advanced enough to be admitted to a British university postgraduate program), a few less-well-known words were required. The Brown word list (Francis and Kucera, 1982) was consulted, and 3 words were selected from the 4,000-5,000 word level on the same criterion of having at least three major meanings. The author's intuitions were that these words would be unknown to a majority of international students.

The next step was to ascertain just how familiar the 18 words actually were for a typical international postgraduate student. A form was devised listing the words, each with a 4-point Lickert scale attached (1 indicating not known at all and 4 indicating known very well). This form was completed by 12 international postgraduate students. From the results it was possible to tell which words were likely to be relatively well-known to the eventual postgraduate case-study subjects, which words were unlikely to be known, and which would be somewhere in between. It was deemed desirable to include some words which were relatively well-known, so that improvement of the more advanced types of word knowledge could be measured. For example, the author assumed that a word would have to be relatively well-known before improvements in the subjects' collocational knowledge would occur. If only unknown words were used in the experiment, the subjects might never progress to a point where these two kinds of word knowledge could be studied. At the same time, some words which were virtually unknown were desirable since this opened to the door to insights into the beginning stages of lexical acquisition. Similarly, a certain number of words falling into the partial knowledge range were thought necessary.

Later piloting (below) indicated the test procedure was indeed time-consuming, and that 11 words were all that could be included in a session approximately two hours long. Any longer would be too draining for both the interviewee and researcher, with the chance of words occurring later in the test having spurious results due to fatigue. The final word list consisted of 1) two Brown

Word List words *brood* and *spur* which were relatively unknown [mean Lickert score ≤ 1.75], 2) four UWL words which most of the respondents indicated they knew relatively well - *abandon, dedicate, illuminate,* and *suspend* [≥ 3.5], and 3) five UWL words in between - *circulate, convert, launch, plot,* and *trace.*

Word Knowledge Measurement Procedures

Following Schmitt and Meara (1997), the best measurement procedure would entail both receptive and productive measures of lexical knowledge. But considering that this study was breaking new ground and measuring lexical knowledge in a manner never before attempted, it was felt that the design was complex enough dealing with only a single type of response. It was decided to test the words productively for two main reasons. First, if some type of word knowledge is demonstrated productively, we can be reasonably certain that it is mastered. If it is demonstrated receptively, with our current state of limited knowledge, it is unclear how that demonstrated receptive knowledge relates to and can be scaled against productive mastery. (Although note that the converse is probably also true.) Second, tests measuring productive knowledge can be much easier to develop than those measuring receptive knowledge. For example, to elicit association knowledge productively, one merely has to ask for association responses to a stimulus word. To do the same with receptive association knowledge requires building a list of commonly given responses as well as reasonable distractors, which is quite a bit more involved (see Schmitt and Meara, 1997, for the methodology in detail).

The next step was to decide which types of word knowledge to measure and to fix measurement procedures for each. Ideally, each word would be tested on each of the eight word knowledge types. However, this was impossible or impractical for several reasons. The most obvious constraint was time, which meant that some types of word knowledge had to be excluded. Each word had to be presented in some way, and this would necessarily mean giving away either its phonological or written form. Since reading and writing is so important for academic students, it was decided to present the target words orally, precluding any productive test of phonological form, and test only for the subjects' knowledge of written form (spelling). Most of the target words came from the UWL, and so had an academic register, but little other register information which could be tested for. Also, register is another type of word knowledge for which measurement procedures have not yet been explored. Therefore, it was also decided not to test for register knowledge. The lack of success of the frequency study to establish clear enough baselines to be used in assessment meant that it would be prudent not to involve frequency measurement either. The remaining five types of word knowledge were included in the research design.

Since words do not jump from being totally unknown to totally known (Meara, 1984; Sharwood Smith, 1984), but are rather learned incrementally, the measurement procedures had to be as sensitive as possible to gradually greater control of the different word knowledge aspects, ranging from no knowledge at all to full native-like mastery. In every case, a new measurement procedure had to be devised from scratch to accomplish this. Association and collocation knowledge are somewhat complex to measure and the procedures for these word knowledges were developed as described in their own chapters.

In the case of collocation, evidence from an additional corpus was utilized in this study. The procedures for grammatical knowledge and meaning are essentially the same as used in the TOEFL study, but are described in more detail here. The procedure for written form (spelling) is relatively straightforward and will be described first.

Written Form

The test of written form (spelling) consisted of a 4-point rating system. Zero (0) on the scale indicates that the subject had very little idea of how the word was spelled. One (1) signified that the subject was able to give the initial letters of the target word, but omitted some later letters, added unnecessary letters, or transposed letters. Two (2) indicated that the word was phonologically correct, but perhaps some vowels or consonants were replaced by similar-sounding but erroneous items (*brood - *brud; illuminate - *elluminate*). Three (3) indicated fully correct spelling. It is not claimed that this is an interval scale, with the spacings being equal between all points on the scale, but it was thought to be sufficient to indicate a progression in spelling competence.

Meaning

The measurement of incremental improvement in meaning knowledge can be approached in different ways. One is to try to measure the incremental acquisition of an individual meaning sense. Instruments like the VKS (Paribahkt and Wesche, 1993) attempt to describe lexical competence in this way. Perhaps a more feasible method is to make a dichotomous *knows/doesn't* *know* decision on any particular meaning sense, but attempt to measure knowledge of all the different major meaning senses of a polysemous word. The second method was chosen for this study. Therefore, as described above, each candidate target word was chosen initially because it had at least three different common meaning senses. Once the final target words were determined by the survey of how well postgraduate students knew them, three dictionaries were consulted to determine all of the meaning senses: the *Longman Dictionary of English Language and Culture,* (1992); the *Oxford Advanced Learner's Dictionary*, (1995); and *Webster's Ninth New Collegiate Dictionary*, (1987). Sometimes it was difficult to judge whether two meaning senses were indeed distinct, and the dictionaries themselves disagreed in some cases. Two examples are meaning senses for *launch* and *plot. Launch* can mean 'to begin something', like an attack, or to 'set something in motion', typically a rocket or missile. These two senses are used in somewhat different ways, but the underlying concept is quite similar. In the case of *plot*, one meaning sense is 'to connect a series of points into a curve' as on a graph, but it also means doing something similar on a map in the specialized context of navigation. With some dictionaries giving definitions which overlapped the boundaries of the definitions from other dictionaries, I was forced to make subjective judgements on which meanings senses were indeed distinct and should be included. During the piloting process (below), the responses from both the native and nonnative respondents helped to clarify which meaning senses to include on the final instrument.

During the test interview, each subject was asked to say all the meanings he[1] knew for the word. They were instructed beforehand, that each target word was polysemous. They were told that they were free to use any means to convey their understanding of meaning senses: give definitions, give examples, use the word in sentences, draw sketches or diagrams, use gestures, etc. When the researcher was satisfied that a meaning sense was known, it was checked as 'Unprompted Meaning Knowledge" on the interview instrument (Appendix 7.1). There were often cases of where a subject indicated only a very broad understanding of a particular meaning sense, and where it was not clear that they had a more precise grasp of the concept. An example is the word *circulate*, which has a meaning sense of 'to move about freely' as in 'air circulates'. Subjects often described air circulating in a circular movement, and it was necessary to probe with further questions to discover if their knowledge was restricted to a meaning sense of 'to move around in a circular way, as in a closed system'. Even with repeated questioning, it was sometimes difficult to determine the subjects' knowledge of the differentiation between meaning senses without actually giving away those differences. In the end, it must be admitted that there was a degree of subjectivity in the marking of whether a meaning sense was known or not.

Although the decision had been made to only measure productive knowledge in this study, it was felt that meaning should be an exception. A major part of the incremental acquisition of word meaning is likely to be the move from receptive to productive mastery of different meaning senses. In addition, meaning has traditionally been the type of word knowledge given most weight in vocabulary testing and it was relatively easy to tap into the subjects' receptive knowledge. After the subject could not think of any additional meanings on their own, prompt words were given which were designed to elicit additional meanings which the subject might know, but could not recall independently. The prompts were designed to bring up a situation in the subject's mind which would trigger the related meaning if the subject knew it, but not to give away that meaning if the subject did not know it. For example, for the target word *spur*, the prompt word *horse* was designed to suggest the meaning *metal device worn on the heel of a boot used to guide or encourage a horse the person is riding*. Through the piloting process, the prompt words/phrases were changed and refined until they were effective in helping to retrieve meaning senses known but not previously recalled. Care was also taken to make sure they did not lead to the guessing of unknown meanings. If the prompts led to a satisfactory description of the meaning sense, that sense was marked as 'Prompted Meaning Knowledge'. If the subject was not able to describe the meaning sense after the prompt word was given, that meaning sense was scored as 'Unknown'.

With no real guidance in the literature of how to relate and score productive and receptive meaning knowledge to each other, the researcher had to devise some scoring system. Since the whole study was an exploratory look into a new area, it was felt that a simple, transparent system would be best at this stage. Therefore unprompted explanations of meaning sense were assumed to demonstrate productive knowledge and were awarded 2 points. Prompted explanations were considered to be related to receptive knowledge and were given 1 point. Unknown meaning senses received 0 points. The resulting numbers could be analyzed in a number of ways. Let us take one subject's results from the word *abandon* to illustrate.

Meanings:	U = UNPROMPTED P = PROMPTED X = DON'T KNOW		
V (abandon a baby) leave or desert and not return		P	1
V (abandon a ship) leave because of danger		P	1
V (abandon a project) give up before finishing		U	2
V (abandon a political leader) withdraw help or support from somebody		X	0
V (abandon yourself to despair) allow oneself to be completely controlled by something		X	0
N (gay abandon) state where feelings and actions are uncontrolled or uninhibited		X	0

One way to interpret the results would be to count the percentage of meaning senses which the subject knew productively or knew in some manner (productively + receptively). In this case, the subject knew 17% of the possible meaning senses productively, and had some knowledge of 50%. Using the 0-2 point system, we can also weight the degree of meaning knowledge and obtain a knowledge proportion for all of the meaning senses taken together. For the above results the proportion is .33 (1+1+2=4; 6 meaning senses × a maximum possible score of 2 per meaning sense = 12; 4÷12=.33). One limitation of these scoring methods is that they do not differentiate between common and less common meanings. However, this issue is potentially quite complex, and will not be addressed in this exploratory study. It is felt that the above scoring methods taken together can provide an informative indication a subjects' knowledge of the different meaning senses of the target words.

Grammar

The subjects could be tested on many possible aspects of grammar to ascertain their grammatical knowledge of each of the target words. One of the most obvious is to ask what word class the word belongs to. Following Schmitt and Meara (1997), it was also decided to check for the subjects' knowledge of each word's derivations, since they found that the subjects in their study typically had poor control of this grammatical aspect. The word class and derivational forms were obtained from the dictionaries mentioned above at the same time that the meaning information was extracted. Base or derivative forms were set for each of the four major word class categories: noun, verb, adjective, and adverb. If the dictionaries indicated that no form existed for a certain word class, this was also noted. The norming data taken from the dictionaries was compared with the answers from the three native-speakers taking the pilot test. This was particularly important for adverbial forms, because the dictionaries occasionally listed forms which the native-speaker pilot subjects found very strange. In these cases, the British National Corpus was consulted to check that form's frequency of occurrence. If it was very low, it was still accepted as a possible form for that word class, but an answer that no form existed was also considered acceptable.

Work done at the University of Lancaster has shown that nonnatives often do not have a very good grasp of the metalinguistic terms denoting word class, ie. *noun, verb,* etc. (Caroline Clapham, personal communication). To ensure that a lack of metalinguistic knowledge did not adversely affect this study, subjects were instructed in the concept of word class before the interviews began. They were given four lists to examine, each containing words which only belonged to a single word class (Appendix 7.2). Additionally, the behavior of each word class was described. After looking at the lists and listening to the description, the subjects were asked whether they felt comfortable with each of the four word class categories. If not, additional instruction was given. The subjects kept the lists and could refer to them at any time during the interview.

The researcher had the various base and derivative forms printed on his interview instrument, in the order of noun, verb, adjective, and adverb. The target word *circulate* had the following forms:

Noun:	circulation
Verb:	circulate+
Adjective:	circular / circulatory
Adverb:	circularly / No Form+

+On the actual instrument, the base form was represented by (*) and No Form by (X).

In the course of the interview, the subject was first asked for the word class of the target word (the term *part-of-speech* was used if it seemed the subject was more familiar with it than *word class*). If they responded correctly, that was checked on the instrument. They were then asked if that form could be any other word class (some words like *spur* could be more than one part of speech without derivational alteration). Next, they were asked whether derivations in other word classes existed (in some cases they did not, and that fact could not be given away in the elicitation questions). For example, if a subject indicated that *circulate* is a verb, they were asked whether a noun form of *circulate* exists, then whether an adjective form exists, and finally if an adverb form exists. In cases where two or more forms exists (ie. *circular / circulatory*) then only one of the forms was required. If a subject was able to indicate that a target word does not exist in a certain word class, that was considered positive knowledge and credit was given for that word class. In cases were only a very infrequent derivative form exists, like in the adverb form of *circulate* above, then either answer of *circularly* or *No form exists* was accepted. Thus the possible scores ranged from 0 (knowledge for no word class was demonstrated) up to 4 (forms for all four word classes were demonstrated).

Collocation

The collocation procedure used in this study is the same as described in Chapter 4, but with one difference. In addition to the COBUILD collocate norms, I was also able to obtain corpus evidence from the British National Corpus (BNC). The collocate norms from the BNC for the eleven target words were created with the following parameters. The Mutual Inference statistic was used and only collocates with an MI score of ≥2.00 were taken. The span was ±5. For a subject's sentence to be scored positively for collocation, at least one word from the COBUILD **or** the BNC norm list had to be included. This is because either list lacked words which seemed very natural collocates to me, and so using both seemed to give the subjects the fairest chance of matching collocates in their sentences.

Vocabulary Size Test

The incremental deepening of knowledge of individual words does not happen in isolation from the other words in the lexicon. Meara (in press) suggests the intuitively-reasonable proposition that people with larger vocabularies tend to have deeper knowledge of the individual component words than do people with smaller vocabularies. This is probably due to the longer period of study and exposure which learners with larger vocabularies typically have the benefit of. It was therefore thought worthwhile to measure the subjects' vocabulary size at the same time as their word knowledge. Since the sessions were already long, a relatively brief vocabulary size test was required. The Eurocentres Vocabulary Size Test (EVST) (Meara and Jones, 1990) was selected because it is computerized, quick and easy to take, and because it is becoming established as a measurement instrument in vocabulary research.

Piloting

Once the individual word knowledge tests were completed, the entire battery was piloted on three native-speakers. Two of the pilot subjects were female, educated, English instructors, one from New Zealand and the other from America. The third was a male, educated, English instructor from England. All were about 30 years of age. The piloting was done to make sure that all of the tests were clear and behaved as expected (especially the prompts) and that a native speaker would be able to achieve very high marks. Full marks were not expected, because even native speakers do not know every word in English completely. However, for the target words included in the study, it was anticipated that the native pilot subjects would do very well. As expected, some of the more obscure meaning senses were not known, such as *trace = part of the harness which pulls a cart.* This was not seen as a problem, rather merely as a reflection that native knowledge does not always equal 100% knowledge. The piloting showed that native-speakers could indeed complete the test battery successfully. A few changes were made to the meaning and collocational prompts based on this pilot, and discussions with the subjects after the pilot interview. The complete interview was found to take over 1½ hours (without the EVST test), so it was felt that the eleven words included was the maximum possible, since nonnative subjects would undoubtedly take longer.

Following the initial piloting and revision, three international postgraduate students attending a presessional course (see below) took the test. They included 1 Japanese, 1 Brazilian, and 1 Taiwanese. These nonnative pilot subjects were taken from the same cohort as the eventual study subjects, and gave the appearance of being similar in terms of proficiency. From their results, the tests were again slightly revised and the time required for the eventual nonnative subjects estimated to be a bit over two hours.

Subjects

Four subjects were selected from among the students attending the presessional course at the University of Nottingham in September 1995 (as were the pilot nonnative subjects). The subjects were all of a moderately high English proficiency, as indicated by their acceptance into University of Nottingham postgraduate courses and their TOEFL scores ranging from 530 - 560. The subjects were selected so that they came from different countries and were studying in different departments. In addition, they had previously resided in English-speaking countries for only short amounts of time. It was hoped that the subjects' first encounter with intensive and sustained English exposure would result in tangible language improvement, which would offer the study the greatest chance for success. The subjects voluntarily agreed to participate, and were offered a token fee of £10 per session. The researcher also occasionally helped the subjects with their compositions in appreciation of their cooperation. The details of the four subjects are given in Table 1.

Table 1 Longitudinal Study Subjects

Pseudonym	'Lith'	'Kor'	'Tai'	'Ind'
Native Country	Lithuania	Korea	Taiwan	India
L1	Lithuanian	Korean	Chinese	Tamil
Months*	2	8	3	5
Age	30	24	30	27
Sex	Male	Female	Male	Male
Course of Study	Politics	Neural Science	Insurance	Life Science
TOEFL**	560	530	537	547

*Length of time spent in English-speaking countries at time of first test
**Score as submitted as part of admission requirements

Unfortunately, the best laid plans concerning research studies sometimes go awry, and extended longitudinal studies are particularly prone to subject attrition. The researcher foresaw this and thus used four subjects instead of a single one, which would be more typical for such a detailed case study. In this instance, both Kor and Tai dropped out in the course of the study. In Kor's case, after eight months, it was 'mutually agreed between herself and her department that it would no longer be in her best interests to continue'. This means that only the first test given in November 1996 exists in her data set. The situation was less disruptive with Tai; due to financial reasons, he had to leave the university after ten months.

Test Intervals

Since this was the first longitudinal vocabulary acquisition study of this type to the author's knowledge, there was no way of knowing what a reasonable interval between test sessions would be. The sessions had to be distanced far enough apart so that some change in vocabulary knowledge might occur, but close enough that at least three sessions could be completed within my period of doctoral study. I decided on a six month interval between sessions, mainly on pragmatic terms. This seemed far enough apart that subjects studying in an L2 environment would receive plenty of exposure, and would allow three sessions in a 1 year period. The subjects were interviewed individually. The first session (T1) for all four subjects took place within the first two weeks of November 1995. The subjects became busier in the Spring, and so it was not possible to schedule all of the second sessions (T2) within such a compact period. Ind and Tai took the T2 within the first half of May, but I was not able to interview Lith until June 15th, 1996. As mentioned above, Kor had dropped out just before the second session was scheduled.

For the third session (T3), both Ind and Lith were interviewed during the first week of November, 1996, but Tai had to return home during the summer. He agreed to do the third session in July before he left for Taiwan. Thus his T3 was completed, but only two months after the second session instead of the planned six months. The net result is that the study ended up with three subjects who have completed the series of three test interviews, although the spacing between the sessions is not as uniform as would be hoped. Although the results cannot be included in this thesis, the researcher plans to continue with the study as long as both he and the subjects continue to be at the University of Nottingham.

The Procedure for the Session Interviews

The following description explains the procedure followed in each test session. The pilot sessions were conducted in a similar manner, but without the EVST. The format was a one-on-one interview, with the researcher following a set order of elicitation for each target word. Each session was tape recorded for future reference. The sessions were held in the researcher's office, except for Tai's interviews which were held in his office and the T3 for Lith, which was held at his home.

At the beginning of each session there was a brief exchange of pleasantries which helped to create a relaxed atmosphere for the interview. Then I told or reminded the subject about the format of the session. Next the subject was given the word class lists and checked to see if he was comfortable with the terms. During the T1, the complete test battery was gone through for an example word *secure*. This word was developed for the main study but was not included due to a lack of time. As such, it gave the subjects a very good impression of how the tests worked. The subjects seemed to catch on quite quickly, and so in subsequent sessions, no further example practice was given. At this point, the test proper began. For each word, I worked my way down the interview instrument (Appendix 7.1) in lockstep order.

The first type of word knowledge to be measured was written form. I asked the question "How do you spell ______?" The subject had pencil and paper available throughout the interview, and was allowed to either spell the word out orally or write it out on the paper. In practice, the subjects normally did both simultaneously. I circled the appropriate degree of knowledge on the spelling scale, which was almost always either 'Correct' or 'Phonologically Correct'. If the word was misspelled, I gave the correct spelling to the subject at that point so that he would have it in front of him for the rest of the questions on that word.

Association knowledge was next to be measured. I elicited a maximum of three associations per target word with the following instruction, "Please give the first three words you think of when you hear the word ______." Sometimes the subjects had to be reminded that this was not a definition task, and that the words merely needed to be ones which easily sprung to mind upon the stimulus. I then wrote the responses down on blanks provided on the instrument. If the subject was unable to produce three associations, I recorded whatever number of associations the subject could produce.

The researcher next addressed collocational knowledge. The subjects were asked to form three different sentences, each including the target word. They were told that because creating sentences without any contextual parameters can be rather difficult, the researcher would give them three different situation cues to help them. They were not to use words from the cue in the sentences. They were instructed not to be creative, but rather to try to give the most typical, normal, usual sentence for each context. They were also told that they could use inflected forms of the target word in the sentences. I stressed that what I was looking for was how naturally the other words in the sentence fit with the target word. Because there were several rules for this test, the subjects were given a simplified list of the rules for this and all the other word knowledge tests at the beginning of the interview (Appendix 7.3). It was in front of the subjects at all times. If a subject strayed away from the desired manner of response, I could easily bring them back on task by referring to this list.

I gave the situation cue and waited for the subject to produce a sentence. There was typically a lot of backtracking, changing, and uncertainty, but once the subject had finally decided on a sentence, I repeated it aloud to check it with the subject and to ensure that there was a clear version on the audio tape, since the sentences were transcribed from the tape at a later point to save time during the interview. I had to be especially careful not to partially form the sentences for the subjects, since subjects had the tendency to agree with anything I said at the repetition stage.

The subsequent word knowledge type was grammatical knowledge. I asked the subject "What word class (part-of-speech) is _______?" Depending on the answer, I went on to ask "Is there a (noun, verb, adjective, adverb) form?" for the remaining three word classes. It was important to determine whether a "No" answer meant "No, I believe there isn't a form for that word class" (potentially a correct answer) or "No, I don't know if there is a form for that word class or not" / "No, I think there is a form, but I don't know what it is" (answers receiving no credit). If the subject produced the correct word class form, I circled the appropriate it on the study instrument. If a target word does not exist for a particular word class, and the subject indicated this, credit was given for this answer.

The penultimate word knowledge category was meaning. I prompted productive knowledge of meaning senses by asking the subject something to the effect of "_______ has several meanings. Tell me the ones you can think of." The subjects were reminded to use any means they wanted to in order to convey their understanding of the meaning sense. After it was determined which meaning senses the subject knew productively, he was given the prompts, one-by-one, including the word class of that meaning sense, to elicit any receptive knowledge of meaning. I marked the meaning senses listed on the instrument as 'Unprompted', 'Prompted', or 'Unknown'.

In order to break up the interview session somewhat, the subjects were asked to do the EVST test after the sixth or seventh target word. The computer program showed them their score at the end of the test. Afterwards the rest of the target words were completed.

In the T2 and T3 sessions, after the tests for each word were finished, I went back and explained the different meaning senses to the subject. This was done partially to keep the subjects' interest up during the rather long interview sessions, and to make sure that they received at least some input about the words from which they could learn.

Results

The study's design allows statements to be made about how the five kinds of word knowledge develop over time. It cannot say why the knowledge develops, or what facilitates or inhibits development. Lith and Ind both confirmed that they did not study the words between test sessions, and so the only time the words were explicitly focused upon was in the T2 session. (Also the T3 session, but this has no bearing on the results reported here.) Thus for them, any additional exposure must have been gained in a naturally-occurring context. It is impossible to say what the ratio between explicit and implicit exposure was however. On the other hand, Tai reported explicitly looking up the words in a dictionary and studying them to some small extent. As with Lith and Ind, it is impossible to judge how this amount of explicit focus on the words compares to any implicit exposure he might have received. Thus this report will not spend much time speculating on the effect or amount of implicit vs explicit learning, other than making the assumption that there must have been some exposure to the target words in the course of the six months separating the sessions. It will concentrate on describing how word knowledge changes over time.

Since case studies do not typically produce enough data to allow the use of common statistical procedures, like correlation analysis, the analyses are usually descriptive only. That is the case here as well. In addition, even if there was enough data, it would be unlikely to meet the requirement for normality. Thus the results and discussion will focus on describing the changes in the subjects' word knowledge over time. These changes are summarized in Tables 2 - 5, one for each of the four case study subjects. The spelling, association, and collocation categories show the results on a scale of from 0 - 3. The Grammar results are on a scale from 0 - 4. The Meaning results include three separate kinds of information. First, the three numbers at the top of the box indicate the results in this order: Umprompted meaning senses, Prompted meaning senses, and Unknown meaning senses. Second, the fraction at the lower left hand side of the box is the total number of points (Unprompted=2; Prompted=1; Unknown=0) over the total possible (Number of meaning senses $\times$ 2). Third, the figure in the lower right hand part of the box indicates the proportion equivalent of this fraction (meaning proportion). The EVST scores are given at the bottom of the tables. This test measures words up to the 10,000 word frequency level, so 10,000 words is the maximum possible.

Discussion

Meaning Knowledge

Knowledge of a word's meaning is probably the first thing most people would think of when considering the question "What does it take to know a word?" It is also likely that most specialists would agree that meaning is the primary word knowledge. In addition, this study has generated more detailed data on meaning knowledge than any other. We will therefore begin our analysis with a focus on this area. Table 6 illustrates how the subjects' meaning knowledge of the target words changed over the course of a year. Kor's figures are not reported since she took only the T1 and thus there are no longitudinal results.

Table 2 Longitudinal Study Results for Lith

	SPELLING			**ASSOCIATION**			**COLLOCATION**			**GRAMMAR**			**MEANING**		
SESSION	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Abandon	3	3	3	2	3	2	2	2	2	2	2	3	1 1 4 3/12 .250	1 3 2 5/12 .417	3 1 2 7/12 .583
Brood	2	2	3	0	0	2	1	0	1	2	2	2	0 0 6 0/12 .000	0 0 6 0/12 .000	1 0 5 2/12 .166
Circulate	3	3	3	1	1	2	2	2	2	2	2	3	1 2 1 4/8 .500	1 3 0 5/8 .625	1 1 2 3/8 .375
Convert	3	3	3	1	1	2	3	3	2	2	3	2	1 2 1 4/8 .500	3 0 1 6/8 .750	3 0 1 6/8 .750
Dedicate	3	3	3	2	2	2	3	3	2	3	2	2	1 3 0 5/8 .625	3 1 0 7/8 .875	2 1 1 5/8 .625
Illuminate	2	2	2	3	3	3	3	3	3	3	2	3	2 1 1 5/8 .625	2 1 1 5/8 .625	4 0 0 8/8 1.00
Launch	2	3	3	0	3	3	3	2	3	3	2	4	2 1 5 5/16 .313	2 4 2 8/16 .500	5 3 0 13/16 .813
Plot	3	3	3	0	1	2	0	0	0	3	2	3	0 1 5 1/12 .083	0 2 4 2/12 .166	3 0 3 6/12 .500
Spur	2	2	3	1	1	0	2	0	0	2	2	4	0 0 7 0/14 .000	0 0 7 0/14 .000	0 0 7 0/14 .000
Suspend	3	3	3	2	2	3	2	2	0	2	2	3	2 0 3 4/10 .400	4 0 1 8/10 .800	4 0 1 8/10 .800
Trace	3	3	3	0	2	2	3	2	3	3	2	3	1 3 3 5/14 .357	4 1 2 9/14 .642	4 0 3 8/14 .571
EVST T1		6600			EVST T2			6650			EVST T3		7600		

Table 3 Longitudinal Study Results for Ind

	SPELLING			**ASSOCIATION**			**COLLOCATION**			**GRAMMAR**			**MEANING**		
SESSION	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Abandon	3	3	3	1	0	0	1	1	1	2	2	2	1 1 4 3/12 .250	1 2 3 4/12 .333	1 0 5 2/12 .167
Brood	2	2	2	0	0	0	0	0	1	1	1	0	0 0 6 0/12 .000	0 0 6 0/12 .000	0 0 6 0/12 .000
Circulate	3	3	3	1	3	3	2	2	3	2	2	2	2 1 1 5/8 .625	2 0 2 4/8 .500	1 2 1 4/8 .500
Convert	3	3	3	1	1	3	3	2	2	4	3	3	2 0 2 4/8 .500	2 0 2 4/8 .500	2 0 2 4/8 .500
Dedicate	3	3	3	2	2	2	3	3	3	2	2	2	1 1 2 3/8 .375	2 1 1 5/8 .625	3 0 1 6/8 .750
Illuminate	3	3	3	3	3	3	2	2	1	3	3	1	2 1 1 5/8 .625	2 0 2 4/8 .500	2 0 2 4/8 .500
Launch	2	2	2	3	2	3	3	3	3	2	3	1	3 0 5 6/16 .375	2 2 4 6/16 .375	3 2 3 8/16 .500
Plot	3	3	3	2	2	1	1	2	3	2	3	2	2 2 2 6/12 .500	3 0 3 6/12 .500	2 0 4 4/12 .333
Spur	3	2	2	2	2	3	2	2	2	1	2	1	0 0 7 0/14 .000	0 0 7 0/14 .000	0 0 7 0/14 .000
Suspend	3	3	3	1	1	0	2	2	2	3	2	3	2 0 3 4/10 .400	2 1 2 5/10 .500	2 0 3 4/10 .400
Trace	3	3	3	2	1	1	3	2	3	1	3	1	1 1 5 3/14 .214	2 1 4 5/14 .357	2 1 4 5/14 .357
EVST T1		8750			EVST T2			9800			EVST T3		6750		

Table 4 Longitudinal Study Results for Tai

	SPELLING			ASSOCIATION			COLLOCATION			GRAMMAR			MEANING		
SESSION	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Abandon	3	3	3	1	1	1	2	3	2	2	1	4	1 1 4 3/12 .250	1 2 3 4/12 .333	1 1 4 3/12 .250
Brood	2	2	3	0	0	2	1	1	1	2	2	3	0 0 6 0/12 .000	0 1 5 1/12 .083	1 1 4 3/12 .250
Circulate	3	3	3	2	3	2	3	2	2	2	2	4	1 0 3 2/8 .250	1 0 3 2/8 .250	1 1 2 3/8 .375
Convert	3	3	3	1	2	3	1	2	2	2	3	3	1 1 2 3/8 .375	1 1 2 3/8 .375	2 0 2 4/8 .500
Dedicate	3	3	3	2	0	2	3	3	3	2	2	3	1 0 3 2/8 .250	2 1 1 5/8 .625	1 1 2 3/8 .375
Illuminate	3	3	3	1	3	3	3	2	3	2	3	3	2 0 2 4/8 .500	2 0 2 4/8 .500	3 0 1 6/8 .750
Launch	2	3	3	0	3	3	1	3	2	3	2	4	0 1 7 1/16 .063	2 2 4 6/16 .375	3 1 4 7/16 .438
Plot	3	3	3	1	3	3	1	1	1	2	2	3	0 2 4 2/12 .167	2 2 2 6/12 .500	2 1 3 5/12 .417
Spur	0	3	3	0	2	0	1	0	2	2	4	4	0 0 7 0/14 .000	1 1 5 3/14 .214	4 0 3 8/14 .571
Suspend	3	3	3	2	3	3	2	2	1	1	2	2	2 1 2 5/10 .500	2 2 1 6/10 .600	2 1 2 5/10 .500
Trace	3	3	3	2	2	2	1	3	0	0	3	4	2 1 4 5/14 .357	2 2 3 6/14 .429	2 2 3 6/14 .429
EVST T1		5650			EVST T2		4800				EVST T3		4450		

Table 5 Longitudinal Study Results for Kor

	SPELLING	ASSOCIATION	COLLOCATION	GRAMMAR	MEANING
SESSION	1	1	1	1	1
Abandon	3	1	3	3	0 5 1 5/12 .417
Brood	2	2	1	3	0 1 5 1/12 .083
Circulate	3	2	2	2	1 0 3 2/8 .250
Convert	3	2	2	3	0 3 1 3/8 .375
Dedicate	3	1	1	4	1 0 3 2/8 .250
Illuminate	3	0	2	2	1 0 3 2/8 .250
Launch	3	2	2	3	1 2 5 4/16 .250
Plot	1	0	1	1	0 1 5 1/12 .083
Spur	3	2	2	3	1 2 4 4/14 .286
Suspend	3	0	1	2	0 2 3 2/10 .200
Trace	3	1	1	2	2 2 3 6/14 .429
EVST T1	7700				

Table 6 Average of Meaning Proportions for T1, T2, and T3 Sessions

	T1	T2	T3
Lith	.332	.491	.562
Ind	.351	.381	.364
Tai	.247	.389	.441
Total	.310	.420	.456

We can see that two of the subjects, Lith and Tai, progressed steadily in their meaning knowledge, while Ind remained relatively static overall. Ind seemed to take the sessions seriously and certainly had better language proficiency than Tai, so there is no obvious reason why he did not progress in meaning knowledge.

The meaning proportions in Table 6 give a global picture of meaning knowledge over all the possible meaning senses of a word. Next let us check each of those meaning senses in more detail. Table 7 illustrates the change in meaning knowledge arrived at by tallying the changes in state (unknown, known receptively, known productively).

This table is a rich source of information about the acquisition of a meaning sense. The first noticeable point is that the vast majority of meaning senses stay at the same state of knowledge (72%, 263/366). This suggests knowledge of meaning sense has a certain amount of inertia, and does not change rapidly.

Table 7 Changes in State of Knowledge for the Different Meaning Senses of a Word

	Lith	Ind	Tai	Total
R↗P	14	3	10	27
U↗R	8	9	11	28
U↗P	8	2	9	19
Total ↗	30	14	30	74
P↘R	2	0	5	7
R↘U	4	8	4	16
P↘U	1	3	2	6
Total ↘	7	11	11	29
P→P	28	31	19	78
R→R	11	3	7	21
U→U	46	63	55	164
Total →	85	97	81	263

P=Productive Knowledge of Meaning Sense
R=Receptive Knowledge of Meaning Sense
U=Unknown Mean Sense
↗=Improves to
↘=Deteriorates to
→=Remains in Same State

N=122 per subject (61 meaning senses × 2 intervals, T1-T2 and T2-T3)
N=366 total (122 × 3 subjects)

This is probably to be expected, as having a large number of meaning senses acquired quickly and easily might be too auspicious a result to hope for. (It would be interesting to try this with L1 children to see how rapidly they pick up various meaning senses.) On the other hand, this stability means there is not a large amount of attrition either. In the 103 cases where meaning sense knowledge did change, it improved 2.5 times more than it deteriorated. In fact, there was attrition in only about 8% of the cases, while there was improvement in 20%. Each of the subjects gained more than they lost in terms of number of improvements and attritions.

We can look more closely where this attrition and improvement occurred. It was uncommon for meaning senses in a productive state to slip down to either receptive or unknown states (all as defined by this study). When there was attrition, it was more likely to be from a receptive to unknown state. Tai alone lost more productive knowledge than receptive knowledge, but the overall level of attrition was still quite low. These results suggest that once a meaning sense is known productively, it is not very likely to be forgotten, at least not over a six month period.

When the three subjects are viewed together in the 74 cases of improvement, meaning knowledge moved from receptive↗productive and from unknown↗receptive a similar number of times. As might be expected, there were fewer cases of meaning knowledge making the presumably larger jump from unknown↗productive. If we look at the subjects' results separately, this split becomes less pronounced, and no improvement category shows much dominance or weakness. It is also interesting to note that knowledge of a meaning sense can move from an unknown to a productive state within a period of six months with only natural exposure as input. We know this because there were unknown↗productive improvements from T1 (subjects were not told the meaning senses of the words in this session) to T2.

In addition to looking at the progression of lexical learning over time, we can also explore the degree of meaning knowledge at any one point. In only one case was a meaning proportion of 1.00 attained, signifying that *all* of the meaning senses of that target word were known productively. Thus, in all but that one case, the subjects had only partial meaning knowledge. In addition, we can see from Table 6 that the average meaning proportion was generally under .500, indicating that 'partial' in this usage indicates nowhere close to full meaning knowledge. Tables 2-5 give the most detailed description of the subjects' meaning knowledge, as they show the state of the various meaning senses. In only five cases is there a zero in the 'Unknown' category (and these are all from one subject, Lith), indicating that all the meanings were known either productively or receptively. So in all but five cases in this study, the meanings of the target words were known incompletely. This partial knowledge is somewhat surprising as one might have assumed that advanced subjects like these would know the majority of target words fairly well. The upshot is that even nonnatives at a high enough proficiency level to study in British universities may have mastery over a rather limited number of the possible meaning senses of a word.

Since we have data for several types of word knowledge elicited at the same points in time, it is possible to compare meaning knowledge with association, collocation, and grammar knowledge. To do this I collated the other word knowledge information according to meaning proportion categories in stages of .2. The results are illustrated in Table 8.

Table 8 Meaning Scores vs. Association, Collocation, and Grammar Scores

	Meaning Proportion					
	.000	.000- .200	.201- .400	.401- .600	.601- .800	.801- 1.00
Association Score (Max. 3)	.692	.727	1.514	2.094	2.125	2.667
Mean Collocation Score (Max. 3)	.923	1.000	2.114	1.969	2.313	3.000
Mean Grammar Score (Max. 4)	1.692	2.091	2.486	2.563	2.313	3.000
Number	13	11	35	32	16	3

First we might note that this table again shows the partial nature of the subjects' meaning knowledge. There were only three instances of a meaning proportion higher than .8, with the majority of instances falling in the .2-.6 range. The next point of interest is the fact that subjects possess other kinds of word knowledge even when they can demonstrate no meaning knowledge. The mean association score is relatively low, but still indicates that some native-like associations were produced in the absence of meaning knowledge. The subjects were able to produce nearly one collocate-including sentence on average in the same absence of meaning knowledge. As for grammar knowledge, in only one instance was there a zero grammar score, and subjects usually knew two word classes, which resulted in a mean grammar score of 1.692. Even though there might be a certain amount of luck involved with these scores, it seems that the subjects had some degree of non-meaning knowledge, especially grammar knowledge, about the target words, even though they could not demonstrate any meaning knowledge.

In general, association, collocation, and grammar knowledge seem to increase in line with meaning knowledge. In the relationship between meaning and association knowledge, a meaning proportion level of .4-.6 appears to match the native-like association threshold of 2. As for collocation, the subjects quickly reached the point to where they could produce sentences which included collocates, and even at the .000 level were producing more sentences that included collocates than without. At the .8-1.0 end of the range, they achieved the maximum three sentences per word, although the data is so thin in this case (only three words) that one cannot rely on this figure. Since it was expected that collocation would be one of the more difficult types of word knowledge to master, these results suggest that the criteria in the collocation procedure is a bit too generous and needs refinement in the future. Concerning grammar knowledge, subjects seem to be able to demonstrate the forms for two word classes almost before they know any meaning, but they have trouble moving much beyond knowledge of three word class forms. There will be more detail on each of these three types of word knowledge in the following sections.

Finally, a quick note on the meaning senses used in this study is necessary. In retrospect it was sometimes difficult to tell whether two similar meanings were both actually known, and whether they were known to the point where the subjects could discern the subtle differences in meaning between the two. Meaning senses which sometimes caused this problem included *abandon* (leave or desert and not return/leave because of danger), *dedicate* (devote oneself to a good cause/devote something to a sacred purpose), and *launch* (put something or somebody into action, begin/put something in motion or on its course). In the future, it would be better to use only meaning senses which are clearly distinguishable from each other to ensure more confident marking of meaning knowledge.

Knowledge of Written Form (Spelling)

One of the more noticeable things that comes from the data in Tables 2-5 is that subjects at this level of proficiency do not seem to have much trouble with the spelling of words of this sort of difficulty. It must be said that the target words do not seem to be particularly tricky, there are still cases where they cannot be spelled directly from the phonological rendering, eg. the schwa in the final syllable of *abandon* could be virtually any vowel and the vowel in *launch* could be represented as 'ou'. If the subjects demonstrated any meaning knowledge of the words at all, they were almost always able to spell them as well. But this was not always the case however. It is interesting to examine Lith's results for the word *illuminate*. By T3, he knew all of the meaning senses productively (the only full meaning marks in the study), and also had high scores on the other word knowledge measures. Still, he persisted in spelling *illuminate* with an 'e': T1-elluminate, T2-eluminate, T3-eluminate. So having a good understanding of other types of word knowledge does not *necessarily* mean that one will know how to spell a word correctly.

This skill in spelling seems to extend to when the word is unknown. *Brood* and *spur* were the two words intentionally included because the subjects were unlikely to be familiar with them. In the 13 cases where the subjects had no demonstrated knowledge of meaning for these words, they were able to produce a phonologically-correct spelling 10 times, a completely correct spelling twice, and in only one case did the subject have no idea of how to spell the word. Thus subjects were able to use sound-symbol correspondences to come to a rough approximation of the correct spelling of an unknown word, but these were not reliable for a fully-correct spelling. It seems that comes in conjunction with the improvement of other types of word knowledge. In Tai's case, we can see that as he began to learn the other types of word knowledge for these words, he also consolidated his knowledge of the spelling.

The results show that in most cases the subjects improved their spelling scores over the course of time. Only twice was there any backsliding (Ind/*launch* & *spur*). One might infer from this that once the spelling of a word is mastered it is not usually forgotten. However, it is likely that the subjects had been spelling the vast majority of these words correctly for quite some time. Therefore we probably do not have enough data about words which have been recently learned and consolidated (like *brood* and *spur* for Tai or *plot* for Lith) to make strong claims about words just over the threshold of spelling control.

Association Knowledge

Let us first examine how the subjects' association knowledge of the target words changed over time (Table 9).

Table 9 Means of Association Scores for T1, T2, and T3 Sessions

	T1	T2	T3
Lith	1.636	1.545	1.727
Ind	1.091	2.000	2.182
Tai	1.091	1.727	2.091
Total	1.273	1.758	2.000

The associations given by both Ind and Tai have become more nativelike over time, while Lith seems to be stuck at about the same level. It is difficult to explain why this should be so. From my experience, Lith had a similar level of ESL proficiency as Ind, and was noticeably stronger than Tai. Lith also had the highest TOEFL score. But the result may not have anything to do with general language proficiency, as a weaker subject (Tai) and a comparable subject (Ind) both improved their association averages by a large margin while Lith did not. Perhaps he has a longer latent period before his association score rises. At the moment, the reason must remain a mystery, although results from a future T4 may illuminate the issue.

Of the 33 cases (3 subjects × 11 words), 23 showed stability or improvement, while 10 indicated some backsliding. This seems like a lot of backsliding, but if we consider the structure of the association scale, there is less of a problem. The key break in the scale is between 1 and 2, that is, between nonnative-like and native-like performance. If we count the times that a subject backslid from nativelike knowledge (2 or 3) down to not native-like (0 or 1), then the total number of cases is only four. Three of the remaining six instances of backsliding consists of dropping from 3 to 2, which means these performances were all native-like, but with slightly less typical associates. The other three indicate that subjects had dropped from a level of producing only minimally native-like associates, usually idiosyncratic (1) to not producing any matches at all. In general, subjects' association knowledge seems to progress in an ever-improving direction, and where there are downward fluctuations, they are seldom across the native-like threshold.

Schmitt and Meara (1997) found that when their subjects judged a word as unknown, they were not able to produce native-like associations which were on a norming list. On the face of it, this is not the case in the present study. Of the 12 instances where subjects were unable to demonstrate any knowledge of meaning (.000 meaning proportion), in three they achieved native-like association scores and in two reached a (1) score. However, if we examine the associations themselves, the reason for this soon becomes clear. The nativelike association scores were all scored on the word *spur* by Ind. The meaning section of the interview revealed that he did not know the meaning of the individual word *spur*, but knew the phrase *spur of the moment.* His association performances are as follows:

T1 - moment	suddenly	[no answer]
T2 - moment	events	suddenly
T3 - moment	horse	[no answer]

Since *moment* was on the norming list at a high enough value to put a subject over the nativeness threshold by itself, Ind produced three nativelike scores by knowing only a fixed phrase. Even more interesting is the T3 results. He was able to produce the primary association *horse* even though he had no idea of the meaning of *spur*. This indicates that he was either a lucky guesser, or more likely, that he had some small sense of at least what lexical field the word *spur* connects with. If so, this might be one of the earliest indications that the acquisition of *spur* has begun.

Lith achieved the two (1) scores for *spur*.

T1 - spin	movement	beginning
T2 - water	around	quick

Movement was an idiosyncratic native response, and *quick* was given by two native speakers. The associations taken together suggest that there is some idea of `movement' behind them. This may be the start of the acquisition of *spur*, or it may just be a totally mistaken impression. Taking this into consideration, the data shows that if a subject does not know the meaning of a word, unless it is part of a phrase, they cannot give a number of native-like associations for it.

The next question is how association knowledge relates to positive meaning knowledge. This is illustrated in Table 10. Unsurprisingly, as mean association scores increased, so did average meaning scores. Unfortunately, we are not able to use a correlation analysis to quantify the relationship between meaning and association knowledge which this table illustrates. Still the data is does lend support to Schmitt and Meara's (1997) suggestion that different kinds of word knowledge are interrelated.

Table 10 Association Scores vs. Collocation, Grammar, and Meaning Scores

	Association Category			
	0	1	2	3
Mean Collocation Score (Max. 3)	1.190	1.960	2.026	2.115
Mean Grammar Score (Max. 4)	2.190	2.360	2.421	2.423
Average Meaning Proportion	.168	.372	.409	.523
Number	21	25	38	26

This does not necessarily mean that all types of word knowledge are related, either strongly or weakly, however. A conspicuous point about the collocation and grammar results in this table is the similarity of figures in Association Categories 1, 2, and 3. Although the figures consistently rise from category to category, the differences are so small that we can accept the figures as being virtually the same. Subjects demonstrating no nativelike association knowledge (0) do have somewhat lower scores on the collocation and grammar tasks, however. Of course, each association category has a wide range of collocation and grammar scores. Nonetheless, for all degrees of positive association knowledge, the mean collocation scores cluster around 2.00 and the mean grammar scores around 2.40. This suggests that the development of association knowledge on one hand and collocation and grammar knowledge on the other may not be strongly linked together.

One reason for the lack of parallel progression seems to be that collocation and grammar knowledge are already somewhat advanced when the association knowledge is still at the point where no nativelike associations can be given. Subjects seemed to have a relatively strong sense of the grammatical aspects of words, even when they were unknown according to meaning (see below), which means that the subjects could readily give the word class and at least one derivative form. Likewise, subjects with no nativelike association knowledge (0) usually created at least one sentence which included a collocate from the norm lists. Only five of the 21 were unable to produce a collocate-including sentence. Taken together, it seems that moving from no knowledge (0) to the first minimal knowledge level (1) is more difficult on the association task than on the collocation and grammar tasks. This may be because some of the tasks are inherently easier than others, although this is impossible to evaluate objectively. Certainly the collocation procedure only measures the broadest of collocation knowledge, because it does not require the collocate to be used in the most typical position in a sentence. Evidence from throughout this study is converging on the conclusion that its scoring criteria need to be made more demanding. Still, it must give some indication of collocation knowledge. To the extent that each task is a reasonable measure of its respective type of word knowledge, it does seem that association knowledge lags behind other types at the beginning of the acquisition process. The underlying reason may be that association knowledge is somehow more complex, or perhaps just later in developing, than collocation and grammar knowledge.

Collocation Knowledge

On observation, the collocation results do not appear to have any obvious trend. Of the 33 cases (3 subjects × 11 words), the scores remained the same from T1 to T3 12 times, fluctuated 9 times, went down 8 times, and only increased in what might have been considered the expected manner in 4 instances. Calculating the means of the collocation scores for T1, T2, and T3 sessions resulted in a similar outcome (Table 11).

Table 11 Means of Collocation Scores for T1, T2, and T3 Sessions

	T1	T2	T3
Lith	2.182	1.727	1.636
Ind	2.000	1.909	2.182
Tai	1.727	2.000	1.727
Total	1.970	1.879	1.848

From the table results, there does not seem to be a great deal of change in the overall amount of collocation knowledge over the 11 words, yet we know that the scores remained the same over the three sessions only 36% of the time (12/33). The scores which do change seem to more or less average each other out, making it impossible to chart any consistent route of progression. This may be because collocation knowledge itself is inherently a very variable type of word knowledge, but one must suspect that the results are at least partly an artefact of the experimental collocation procedure. Again it seems that future versions will need to employ much stricter criteria in order to obtain accurate enough results to plot the acquisition of collocation knowledge.

As with the association section, we now compare the results of the collocation task with those of the association and grammar ones (Table 12).

Table 12 Collocation Scores vs. Association, Grammar, and Meaning Scores

	Collocation Category			
	0	1	2	3
Mean Association Score (Max. 3)	1.000	1.125	1.907	1.844
Mean Grammar Score (Max. 4)	2.636	2.000	2.488	2.375
Average Meaning Proportion	.199	.236	.416	.508
Number	11	24	43	32

There seems to be a slight relationship between the number of collocate-including sentences a subject can produce and their association score, although the association score is not very different between Collocation Categories 2 and 3. Also, even for Category 3, the mean association score does not reach the threshold level of 2.00. So, on average, just because subjects can achieve a (3) on the collocation task, this does not mean that they can also demonstrate native-like associations for that word. The grammar score bounces around without a discernible pattern, and it seems that regardless of the collocation score, subjects usually knew the word class of the target words and at least one derivative.

In Table 7, we found that the average meaning proportion paralleled the rising association categories better than the other kinds of word knowledge. Here we find a similar situation, except that the average meaning proportion is tracking the collocation categories. Taken together, this is starting to suggest that meaning knowledge has closer links with the different word knowledges than they have among each other. Still it is somewhat counterintuitive that collocation knowledge does not have a stronger relationship to association and grammar knowledge. At this stage, it is probably prudent to be cautious in our interpretation of the collocation data, but it does seem to suggest that 1) collocation is among the less strongly related of the word knowledge types, and 2) meaning knowledge has relatively stronger ties with the other kinds of word knowledge.

Grammar Knowledge

Next we look at the behavior of grammar knowledge over time. Table 13 illustrates the means of the grammar scores for the three sessions. As for the grammar scores for individual words from T1 to T3, in 13 instances they rose (Tai = 9), in 12 they fluctuated, in 4 they dropped and in 4 they remained the same (Total 33).

Again we find the behavior erratic, with Lith and Ind's mean scores fluctuating from session to session, and only Tai's consistently improving. Tai was the only subject who reported explicitly looking up the target words in a dictionary and studying them, which might have helped improve his performance. This is possible, because grammar information like that required

Table 13 Means of Grammar Scores for T1, T2, and T3 Sessions

	T1	T2	T3
Lith	2.454	2.091	2.909
Ind	2.091	2.364	1.636
Tai	1.818	2.364	3.364
Total	2.121	2.273	2.636

in the grammar task is the kind available in the typical dictionary. His association scores also improved, which could be attributed to better mastery of the words' meaning gained from explicit attention. On the other hand, collocation information is more difficult to derive from a dictionary and his scores on that part of the test did not improve. Although very speculative, one way of interpreting these results is as illustrating the value of explicit study in addition to implicit learning, but only for certain kinds of word knowledge. Tai showed the most consistent improvement on his association and grammar scores, while all three subjects seemed to move laterally in collocation knowledge.

Next let us compare the longitudinal grammar scores to those of the other types of knowledge tested (Table 14). All three types of knowledge increase more-or-less steadily through Grammar Categories 0-3, but then all make an unexpected drop at Category 4. I suspected this drop might have been the result of subjects moving from Category 3 to Category 4 by stating there was no form for a certain word class, since this is somewhat easier than actually

Table 14 Grammar Scores vs. Association, Collocation, and Meaning Scores

	Grammar Category				
	0	1	2	3	4
Mean Association Score (Max. 3)	1.000	1.545	1.538	1.857	1.500
Mean Collocation Score (Max. 3)	1.000	1.818	1.865	2.057	1.500
Average Meaning Proportion	.179	.226	.376	.452	.384
Number	2	11	52	35	10

having to give the form if it exists. If so there might have been a certain amount of guessing occurring. But most subjects actually reached Category 4 by supplying an existing derivative, rather than merely stating there was not one. With this possibility excluded, I can think of no other explanation for this drop. Regardless, we once again find that the word knowledge in focus trends together most closely with meaning knowledge.

Schmitt and Meara (1997) reported that their beginning/intermediate Japanese EFL students did not have very good mastery of the different derivative forms of a word. We find the same situation even with advanced nonnatives who are capable of pursuing postgraduate studies in an English-speaking university. In only 10 cases out of the 110 possibilities were all four word classes known (9%), while in almost 60% of the cases only two or less were known. Even when words were rather well known, the average number of word class forms known was three at best (Table 8). This shows a definite gap in these advanced learners' morphological knowledge which might surprise some people. There is probably an assumption in the field that derivative forms are easy to learn and that if a student can demonstrate the form of one word class, they know all the others in the word family. This research indicates that this is not the case.

If we break the results down into the individual word classes, we can see that some are definitely more readily acquired than others (Table 15).

Table 15 Number of Words Known in Each Word Class

	Noun		Verb		Adjective		Adverb	
	N	%	N	%	N	%	N	%
Lith*	30	91	33	100	11	33	8	24
Ind*	21	64	28	85	6	18	13	39
Tai*	30	91	31	94	10	30	11	33
Kor**	10	91	10	91	7	64	1	9
Total***	91	83	102	93	34	31	33	30

*N=33 **N=11 ***N=110

Nouns and verbs are clearly the best known word classes, and these particular classes are known almost to a degree which would justify the common assumption of mastery. Adjectives and adverbs appear to be learned at a later stage than nouns and verbs, and in these advanced subjects are still not mastered. This suggests that adjective and adverb forms are not so readily learned from general exposure (perhaps due to the lower frequency of occurrence), which means that they might be good candidates for explicit attention. They could be included in classroom material, or, from a learner strategy standpoint, students could be made more aware of the need to focus a bit more on these derivative forms.

Vocabulary Size

The subjects also took a computerized vocabulary size test (EVST) during the interview sessions. The results of those tests are presented along with the corresponding word knowledge scores in Table 16. Since we cannot use a correlation procedure here, we will have to evaluate the table holistically. Suffice it to say that none of the word knowledge types seem to track with vocabulary size. This result does not lend support to Meara's (in press) intuitive suggestion that a larger vocabulary size should correspond to more being known about each individual word. The reason might be because this is in fact not the case, or because some of the tests were not performing adequately. We have already seen that the collocation procedure needs improvement, but it would be hard to find fault with such a detailed interview for meaning and grammar knowledge. This points a finger of suspicion towards the EVST. It is hard to believe that subjects living in an English-speaking country for the first time, and studying full-time at an English-speaking university would not at least maintain their vocabulary size from the T1. However, this is the case in two out of the three longitudinal subjects.

These questionable results are particularly unsettling because the EVST is coming into more frequent use in vocabulary research. There has been continuing research into checklist tests, particularly at the University of Wales, Swansea. Perhaps it is time to take advantage of what has been learned since the EVST was made available, and produce a new improved version.

Table 16 Vocabulary Size vs. Average Meaning Proportion, Association, Collocation, and Grammar Scores

	T1	T2	T3
<u>Lith</u>			
Vocabulary Size	6600	6650	7600
Meaning	.332	.491	.562
Association	1.636	1.545	1.727
Collocation	2.182	1.727	
1.636Grammar	2.454	2.091	2.909
<u>Ind</u>			
Vocabulary Size	8750	9800	6750
Meaning	.351	.381	.364
Association	1.091	2.000	2.182
Collocation	2.000	1.909	2.182
Grammar	2.091	2.364	1.636
<u>Tai</u>			
Vocabulary Size	5650	4800	4450
Meaning	.247	.389	.441
Association	1.091	1.727	2.091
Collocation	1.727	2.000	1.727
Grammar	1.818	2.364	3.364

Conclusion

Although this has been an exploratory study with only three main subjects, a great deal of information has come out of it. Some of the most interesting observations are summarized below. These are obviously somewhat tentative, but they should give vocabulary researchers food for thought, and will probably suggest to the reader a number of new questions and possible lines of inquiry for future research.

1. The subjects, even though advanced English learners studying at a British university, had incomplete knowledge of the different meaning senses of words which it might have been assumed were well known.

2. The state of knowledge of the meaning senses tended to stay the same. When it changed, it improved 2.5 times more than it attrited. The attrition rate overall was very low (8%). It was uncommon for words in a productive state to attrite within the 6-month interval between sessions.

3. There was some evidence that words could be learned to a productive level from only the implicit learning resulting from exposure.

4. Different types of word knowledge do seem to be interrelated, but some more strongly than others. In particular, meaning knowledge appears to have closer links with association, collocation, and grammar knowledge than they do with other types. On the other hand, the links between collocation knowledge and the others appears to be relatively weaker (at least to the extent that the task accurately captured it).

5. The subjects were sometimes able to demonstrate association, collocation, and grammar knowledge in the absence of any demonstrable meaning knowledge.

6. Spelling seems to be one of the easier word knowledge aspects to master.

7. Using this association procedure, knowledge of a phrase (such as *spur of the moment*) can lead to nativelike association scores, even though the subject has no demonstrable knowledge of the individual word's (*spur*) meaning senses. However, if associates related to such phrases are disregarded, subjects were not able to produce native-like associates if they did not know the word's meaning(s).

8. Subjects seemed to know a target word's word class and one other derivative form almost regardless of however else well the word was known. The easiest word class forms seem to be nouns and verbs. But learning all four word class forms was much more difficult. Even advanced subjects like these did not know all of the word classes for target words from the UWL (such as *abandon* or *convert*), words which are not especially infrequent.

Note

[1]Because almost all the test sessions were with male subjects, masculine pronouns are used for convenience.

Acknowledgements

Thanks to Della Summers and Keith Mardell of Longman Publishing for access to the BNC, and Gwyneth Fox and Rosamund Moon of COBUILD for access to the Bank of English corpus.

CHAPTER 8 CHECKING FOR IMPLICATIONAL SCALING IN VOCABULARY ACQUISITION

The previous chapter described the acquisition of five types of word knowledge. This chapter will focus on the *ordering* of that acquisition, ie. whether certain types of word knowledge are typically acquired before others. Intuitively the answer would have to be yes. It seems obvious from experience that learners fully know a word's spelling before they know all of its meaning senses, for example. If, in fact, the different word knowledges are learned in a developmental order, it would have strong implications for both vocabulary teaching and testing. Teachers would then have a much better idea of what information to give students about words and in which sequence. The field of vocabulary testing would benefit because if one type of word knowledge was demonstrated on a test item, then assumptions could then be made about what other kinds of word knowledge were also known about that particular word.

The question is then whether these intuitions of acquisition order can be supported by empirical evidence. Using the word knowledge data from the previous chapter, we can explore this question by borrowing a technique from morpheme and grammar acquisition studies - implicational scaling. This technique ascertains whether certain aspects or skills are acquired in a set order, with 'higher level' aspects implying knowledge of or proficiency in 'lower level' aspects. Applying this technique to word knowledge acquisition allows us to submit our intuitions of ordering to statistical rigour.

Description of Implicational Scaling (Guttman Procedure)

Perhaps the best way to describe implicational scaling is by working our way through an example. Since implicational scaling is most associated with morpheme studies, let us use a fictional morpheme example. Say we have isolated seven morphemes and want to explore whether learners acquire them in any particular order. After eliciting whether they know the various morphemes or not, we place the results on a matrix (called a scalogram) like the following. (1) indicates that a morpheme is known and (0) that it is not known.

	Morpheme						
Subject	M1	M2	M3	M4	M5	M6	M7
S1	0	0	0	0	0	0	0
S2	0	0	0	0	0	0	1
S3	0	0	0	0	0	0	1
S4	0	0	0	0	0	1	1
S5	0	0	0	0	1	1	1
S6	0	0	1	0	1	1	1
S7	0	0	0	1	1	0	1
S8	0	0	1	1	1	1	1
S9	0	1	1	1	1	1	1
S10	1	1	1	1	1	1	1

This scalogram shows a highly idealized result, and in reality the columns and rows would have to be arranged in order of difficulty, because subjects and morphemes will not necessarily fall into the desired pattern simply from the random order that they are entered into the scalogram. Then a line can be drawn through the scalogram at the border between known and unknown forms. In this example, there is a large degree of systematicity in the ordering of the morphemes and one could say that if a 'higher' morpheme (eg. M3) is known, then it is highly likely that the 'lower' morphemes below it (eg. M4, M5, M6, & M7) are also known. But of course, things are not so clear-cut in the real world. There will always be cases in which forms are known were we would not expect it (eg. Subject 6's knowledge of Morpheme 3) and conversely, forms unexpectedly missed (eg. Subject 7's lack of Morpheme 6). These are called *errors* in the scalogram, and are almost inevitable. Since every scalogram will include errors, a statistical approach is needed to determine whether the number present is small enough to still allow an assumption of implicational scalability. The pertinent statistical analysis is called the Guttman procedure. It produces a *Coefficient of Scalability* (C_{scal}) which, if $\geq .60$, indicates that an implicational scale does accurately describe the data. In order to arrive at C_{scal}, one must first arrive at three other figures. The first is the *Coefficient of Reproducibility* (C_{rep}). It tells us how accurately a subject's position in the scalogram predicts their performance. The formula is:

$$C_{rep} = 1 - \frac{\text{number of errors}}{\text{number of subjects} \times \text{number of items}}$$

In the current example, this would translate into

$$C_{rep} = 1 - \frac{2}{10 \times 7}$$

$$= .971$$

The second necessary figure is *Minimum Marginal Reproducibility* (MM_{rep}), which indicates how predictive the implicational scale is without the errors being taken into account. Its formula is:

$$MM_{rep} = \frac{\text{maximum marginals}}{\text{number of subjects} \times \text{number of items}}$$

where maximum marginals means the number of 0s or 1s (whichever is greater) for each morpheme. In the above scalogram the maximum marginal for M2 is 8 (there are eight 0s and two 1s, so eight is the larger) and for M6 it is 6 (four 0s and six 1s, so six is larger). The maximum marginals for all morphemes are totaled to reach the numerator. Thus

$$MM_{rep} = \frac{9+8+6+6+6+6+9}{10 \times 7}$$

$$= .714$$

The last figure necessary to reach C_{scal} is the *Percent Improvement in Reproducibility* (*% improvement*). Since it shows the improvement between the coefficient of reproducibility and the minimum marginal reproducibility, the formula is a simple subtraction:

$$\% \text{ improvement} = C_{rep} - MM_{rep}$$

or $.971 - .714 = .257$

Finally, we can combine these figures to obtain the coefficient of scalability, which tells us if we can be confident that an implicational scale indeed exists. The formula is:

$$C_{scal} = \frac{\%\ \text{improvement in reproducibility}}{1 - MM_{rep}}$$

$$= \frac{.257}{1 - .714} = .898$$

Hatch and Lazaraton[1] (1991) give the commonly accepted parameters for this procedure. They report that mathematicians have determined the coefficient of reproducibility must be over .90 before the scale can be considered valid. If that is the case, then the coefficient of scalability must be above .60 before scalability is claimed. Our example scalogram only contains two errors, and this is reflected in both coefficients being far above these minimum requirements.

Using the Guttman Procedure with Word Knowledge Data

Before the Guttman procedure can statistically explore whether implicational scaling exists, some decisions need to be made. The key one is in setting the cut-points for whether a particular form or skill is known or not. As Anderson (1978) has shown with morpheme studies, the placement of the cut-point can seriously affect the scaling results. The importance of cut-point placement is somewhat problematic in adapting the procedure to lexical research, because in a first-time word knowledge study like this, no clear guidance is available as to where those cut-points should be. The typical criterion of 80% accuracy does little good here, other than suggesting that 100% knowledge/control might be an unreasonably high target. If 100% knowledge/control is not required, how does one set partial knowledge/control criteria for word knowledge? Fortunately, in most cases it is possible to make a principled decision. The association criterion has already been experimentally determined as being Level 2 on the 0-3 scoring system. Likewise judging from available data, the strength of the subjects' performance on the written form task suggests that fully correct spelling is the logical criterion, since the vast majority of spelling performances were at Level 3. As for grammar knowledge, we have seen that the subjects seemed to know a target word's word class plus one derivative form (Level 2) even if they had little or no idea what the word meant. On the other hand, it was quite rare for them to know all four word class forms. This points to knowledge of three word class forms (Level 3) as being the most reasonable criteria.

The remaining two word knowledges, collocation and meaning, are not quite so easy to resolve. Subjects were generally able to produce at least one collocate-including sentence even with low levels of meaning knowledge, and their collocation scores throughout the previous study seemed to cluster around 2. Whether this means that 2 should be taken as the criterion, or whether a more demanding 3 should be required is unclear. I decided upon 2 because, with the exception of spelling, no other criteria was set at the maximum allowable score. In any case, the subjects' collocation performances did not seem to have a strong relationship with their other word knowledges. Since the collocation procedure may be the least robust of the five word knowledge measures, this study will try the Guttman procedure both with and without the collocation data.

Meaning poses a similar problem. No one yet has the answer to what meaning proportion is necessary to 'know a word'. It is clear that the minimum possible is receptive knowledge of a single meaning sense, which would be .125 for words with four meaning senses. However, I feel that reasonable knowledge of such words would be closer to one sense known productively and another receptively, which would equal .375. To attain this proportion for the more polysemous words, more meaning senses would have to be known, but there would be a correspondingly greater chance to be exposed to these meanings. As there seems to be no principled way to set the meaning proportion criterion, I will explore the use of three figures in this study: .250, .375, and .500.

Because this is exploratory research, and because setting appropriate cut-points is not yet something we can be confident in, it was not felt appropriate to apply a single set of criteria and then check for scalability in a single test. Rather, the study looked at a number of different criteria to see if any could produce an implicational scale of word knowledge. The first set of criteria required the maximum score for each of the word knowledge categories. This criteria allows us to have the most confidence that each word knowledge type is actually known. Alternatively, it was thought interesting to check if there is an implicational scale for the first, minimal acquisition of the various word knowledges. In this case, the required score is anything above zero. In between these two extremes, the procedure was run with the rational criteria described above, including a meaning proportion of .375. In addition, the procedure was executed without the collocation component, because previous indications suggested that it may not may relate to the other types of word knowledge on the scale very strongly. These non-collocation runs included meaning proportions of .250, .375, and .500. In total, six attempts were made to discover any underlying scalability. The criteria for these are detailed in

Table 1.

Table 1 Criteria for Implicational Scaling Tests

	Meaning	Grammar	Collocation	Association	Spelling
Maximum Possible Score (MAX)	1.000	4	3	3	3
Minimum Possible Score (MIN)	>.000	1	1	1	2
Rational Criteria (RAT)	.375	3	2	2	3
Rational Criteria w/o Collocation (RAT25-COLL)	.250	3		2	3
Rational Criteria w/o coll. (RAT37-COLL)	.375	3		2	3
Rational Criteria w/o coll. (RAT50-COLL)	.500	3		2	3

Results

The Guttman procedure indicated that the data from the longitudinal study did not fall into an implicational scale. None of the six cut-point criteria allowed the assumption of scalability to be met, although the Maximum Score Criteria came closest. These results are illustrated in Table 2.

Table 2 Results of Guttman Procedure for Word Knowledge Data using Six Cut-point Criteria Sets

	C_{rep}	MM_{rep}	% *improvement*	C_{scal}
(MAX)	.898	.798	.100	.495
(MIN)	.945	.911	.034	.388
(RAT)	.822	*	*	*
(RAT25-COLL)	.868	*	*	*
(RAT37-COLL)	.827	*	*	*
(RAT50-COLL)	.786	*	*	*

*Because the C_{rep} result was not at or approaching .90, these figures were not calculated

In the end, there were simply too many errors for implicational scaling to exist. This can best be realized by examining the various scalograms (Appendix 8.1). In addition to this main feature, the scalograms illustrate other points of interest. Comparing RAT25-COLL, RAT37-COLL, and RAT50-COLL, we see that meaning progresses up the ranks of difficulty. When only a .250 meaning proportion is required, meaning is very close to spelling in the number of times it was achieved (85/110 vs 88/110). This is far above the next category, association, with 64/110. In RAT37-COLL, the meaning and association criteria were achieved an equal number of times. Finally, if when a meaning proportion of .500 was required, meaning became the most difficult word knowledge. Of course, we have seen that these relationships cannot be considered implicational, but nonetheless the raw tallies suggest that the acquisition of multiple meaning senses can be an involved process, and learners will probably know a great deal else about a word before they reach a .500 meaning knowledge level. (As a reminder, a .500 meaning proportion indicates either ½ of possible meanings known productively, all meanings known receptively, or a combination of the two.) In terms of testing, conventional vocabulary tests typically measure only one receptive meaning sense of a target word, which would equate to a .125 meaning proportion for a word with four major meaning senses. A look at the scalograms would suggest that achieving this rather low proportion would not necessarily indicate other word knowledge types had been even partially mastered. This suggestion is congruent with the results from the TOEFL vocabulary item study in Chapter 6, which demonstrated the shortcomings of the receptive multiple-choice format used on the TOEFL test.

If we look at the at the scalogram addressing minimum initial acquisition of lexical knowledge (MIN), the subjects demonstrated some knowledge of all word knowledge categories for the vast majority of words. Being able to spell a word at least phonologically correctly and knowing at least one word class form seem to be the easiest initial knowledges to acquire, and association knowledge the most difficult, but the differences in the raw number results are not great. For the scalogram showing complete word knowledge mastery (MAX), virtually the opposite is true. In no instance did a subject demonstrate complete mastery of every type of word knowledge for a word. In fact, there was only one case in which four word knowledge categories were demonstrated. The order of the columns suggests that it is easiest to learn the spelling of a word, and most difficult to learn all meanings and word class forms. In what might be considered the typical state of lexical knowledge, neither minimal nor complete, the RAT scalogram paints the picture of completely correct spelling being mastered somewhat before partial collocation knowledge, partial meaning knowledge, and nativelike association knowledge. Learning three word class forms seems to be more difficult and lags behind the rest. Combining all of these scalograms, we get the sense that spelling is the easiest type of word knowledge, knowing three word class forms is the hardest, and the rest in between. It must be stressed that there is only weak evidence for this (since the Guttman statistics were not high enough), but these results may prove useful in guiding future lexical acquisition studies.

The last brief point I would like to bring up is the surprising fact that in virtually every case errors on the scalograms came in pairs of two. There seems to be no easy explanation why a subject deviating from the expected pattern for one kind of word knowledge will almost always deviate for a second kind as well on the same word. I do not at present know if this finding has any significance, but it may be something to check for if implicational scaling is again used to study vocabulary learning.

Conclusion

Examining the scalograms has given some weak evidence that our intuitions of ordering in lexical acquisition are well-founded, but the Guttman procedure results show that the ordering is not robust enough to imply any developmental sequence for this set of data. This is the first time that this type of analysis has been done for word knowledge, and so all results must be considered tentative. Just because the Guttman procedure did not indicate implicational scaling in this study, this does not necessarily mean it does not exist in vocabulary learning. Future research may be able to employ more precise measures of the various word knowledge types, with more subjects. If this happens, we might well find statistical analyses illustrating word knowledge sequencing. The present study was more about exploring a new research direction and sharpening questions than ending up with clear-cut answers. I personally believe that implicational scaling has good potential in vocabulary acquisition research, and that this study has gone some of the way toward laying the groundwork for future efforts.

Note

[1]Hatch and Lazaraton (1991) was consulted for the mechanics of the Guttman procedure, and their instructions have been adapted and summarized in this section.

CHAPTER 9 FINAL WORD

After 350 pages and 3 years of work, we reach the end of this thesis. By this time, the reader will have judged the value of the experiments and their results for himself or herself anyway, so I will abandon the convention of scientific objective reporting and finish with my own brief personal opinions of what I have and have not achieved.

Let me start with the experiments to design measurement procedures. I am fairly happy with the association procedure, believing it to be an improvement on existing methods. It gives a principled way to take into account the different strengths of native response, and I like the fact that more than one association is used for input. The cut-point for the native-like threshold will always be a controversial matter, but at least the methodology provides a way at arriving at a principled answer. Future research may indicate that the threshold needs to be shifted somewhat, but I feel that the approach of using native performance to outline *typical*, rather than *absolute*, native association behavior is the most feasible research line to follow. Of course, the robustness of the technique needs to be tested on more and other types of subjects. In this, the study reflects the exploratory nature of the thesis as a whole. In attempting to measure word knowledge in new ways, and in using the word knowledge framework as a basis for designing the larger studies, the thesis has broken new ground. As such, it was always likely to raise new questions and suggest new research directions in addition to providing insights from the individual studies.

I have mixed feelings about the collocation measurement procedure. The idea of consulting major corpora to ferret out likely collocates before the fact and then adjusting sentence prompts to elicit these seems to me to be an approach worth pursuing. This also allows subjects to demonstrate the collocates within the context of a sentence, rather than in isolation. The problem lies in how to score the individual collocates within each sentence. I think it is fairly clear that more stringent criteria need to be developed, and these may well include grammatical notions, or even demand exact positional placement. On balance, I think the procedure was a reasonably successful foray into the completely new domain of quantifying collocational knowledge. Even though the procedure is nowhere near perfect, it lays the groundwork for further thought and research into this potentially important area.

The main purpose of the frequency study was to establish native-speaker baselines which could be used to gauge nonnative performance. On this count, it was only partially successful. The study is far better and more comprehensive than previous frequency intuition studies, in terms of numbers and range of subjects, method of analysis, and size of the English corpus used for the objective frequency counts. I would argue that it gives us the best picture of both native and nonnative intuitions of frequency to date. Nevertheless, the hoped-for clarity of results did not materialize. The native data did not yield clear-cut behavior which could be unequivocally used as a baseline for research. This is partly because intuitions in general are devilishly difficult to explore. It could also be partly because many people are not as numerate as might be supposed. If one were to pursue this frequency research, it might be well to use a non-numerate elicitation procedure, such as the computerized visual elicitation methodology developed at the University of Edinburgh to measure grammaticality judgement intuitions. But in retrospect, some types of word knowledge might resist all attempts to quantify them, and intuitions of frequency may well fit into this category.

On the other hand, I feel that the study into what word knowledge aspects TOEFL vocabulary items measure was a great success. Of course the study did not explain *why* each item exhibited the behavior that it did, but it did send a clear warning about the type of assumptions one can make about what either a *correct* or an *incorrect* response means. No one would have claimed that such multiple-choice vocabulary items provided a comprehensive measurement of all kinds of word knowledge, but at the same time, no one really knew. I think this study introduces an exciting method of validating vocabulary item formats which could dramatically improve our understanding of vocabulary tests and their characteristics. The study also implies the need for new vocabulary test formats which better probe the depth of understanding of a word.

Finally, I think the longitudinal study is important because it represents one of the first times a small number of words were tracked over time for individual subjects to see how vocabulary knowledge develops. (The imposed reality of two subjects dropping out suggests one reason why more longitudinal studies like this have not been carried out.) Some of the results could have been anticipated, eg. that spelling is one of the easier types of word knowledge. Other results were more surprising: eg. the knowledge of all four word classes seems not to be mastered until late in the acquisition process. Once again, I feel that this study demonstrates a research methodology that holds exciting prospects. Understanding the behavior of the word knowledge components cannot do anything but aid our understanding of the more global acquisition processes. Perhaps doing a number of studies like this would provide us with the insights necessary to develop a comprehensive model of vocabulary acquisition.

At the beginning of this thesis, I posed the following questions concerning the usefulness of the word knowledge framework itself.

1. Is the word knowledge framework feasible for use in vocabulary research?
2. Can reasonable tests for the various types of word knowledge be developed?
3. Finally, is the word knowledge framework informative in vocabulary research?

There is no empirical way to answer these, but the judgement of one who has used the framework for five years (including MPhil research) should be worth expressing. As for the word knowledge framework itself, I believe that these studies have shown that it is a viable basis from which to design vocabulary research, and that the results obtained are informative. The main hurdle is developing better measurement procedures for the types of word knowledge which can be quantified, and also realizing which types can't be so easily reduced to numbers, perhaps like frequency. The studies in this thesis were very time- and labor-intensive, and so the framework is probably only useful for research purposes, and not practical ones. At the moment this seems particularly true because there was no statistically-reliable indication of implicational scaling. However, I believe the search for a hierarchy should be explored further, and perhaps with more precise measurement procedures, an implicational scaling will appear.

In sum, it has been a very exciting course of research which has branched out into interesting new areas. It has not answered any questions definitively, but a PhD thesis is supposed to be just the start of a lifetime of research anyway. This exploratory research has certainly provided myself, and any others interested in lexical issues, with many interesting avenues to explore. So much to do and so little time...

REFERENCES

Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon.* Oxford: Blackwell.

Alexander, R. (1982). What's in a four-letter word? Word meaning in English and second language learning. *Die Neueren Sprachen 81*, 2: 219-224.

Allerton, D.J. (1984). Three (or four) levels of word cooccurrence restriction. *Lingua 63*: 17-40.

Alper, T.G. (1942). A diagnostic spelling scale for the college level: its construction and use. *The Journal of Educational Psychology 33*: 273-290.

Anderson, R. (1978). An implicational model for second language research. *Language Learning 28*, 2: 221-282.

Anderson, R.C. and Nagy, W.E. (1989). Word meanings. In *Handbook of Reading Research*, 2nd edition. London: Longman.

Arabski, J. (1979). *Errors as Indications of the Development of Interlanguage.* Katowice: Uniwersytet Slaski.

Arnaud, P.J.L. (1989). Estimations subjectives des fréquences des mots. *Cahiers de Lexicologie 54*, 1: 69-81.

Arnaud, P.J.L. Subjective word frequency estimates in L1 and L2. Paper presented at the 9th World Congress of Applied Linguistics. Thessalonki. April 1990. ERIC Document ED329120.

Backman, J. (1976). Some common word attributes and their relations to objective frequency counts. *Scandinavian Journal of Educational Research 20*: 175-186.

Backman, J. (1978). Subjective structures in linguistic recurrence. ERIC Document ED180195.

Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal 47*, 1: 56-63.

Bahns, J. and Eldaw, M. (1993). Should we teach EFL students collocations? *System 21*, 1: 101-114.

Barron, R.W. (1980). Visual and phonological strategies in reading and spelling. In Frith, U. (Ed.) *Cognitive Processes in Spelling.* London: Academic Press.

Bauer, L. and Nation, I.S.P. (1993). Word families. *International Journal of Lexicography 6*, 3: 1-27.

Beauvillain, C. and Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language 26*: 658-672.

Begg, I. (1974). Estimation of word frequency in continuous and discrete tasks. *Journal of Experimental Psychology 102*, 6: 1046-1052.

Begg, I. and Rowe, E.J. (1972). Continuous judgements of word frequency and familiarity. *Journal of Experimental Psychology 95*, 1: 48-54.

Béjoint, H. (1981). The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills. *Applied Linguistics 2*, 3: 207-222.

Benson, M. (1985). Collocations and Idioms. In R. Ilson (ed.) *Dictionaries, Lexicography and Language Learning*. Oxford: Pergamon.

Benson, M., Benson, E., and Ilson, R. (1986). *The BBI Combinatory Dictionary of English. A Guide to Word Combinations.* Amsterdam/Philadelphia: Benjamins.

Benson, J.D. and Greaves, W.S. (1981). Field of discourse: Theory and application. *Applied Linguistics 2*, 1: 45-55.

Bensoussan, M. and Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading 7*: 15-32.

Berko, J. (1958). The child's learning of English morphology. *Word 14*: 150-177.

Besner, D. and Johnston, J.C. (1989). Reading and the mental lexicon: On the uptake of visual information. In Marslen-Wilson (Ed.) *Lexical Representation and Process*. Cambridge, MA: Bradford.

Biskup, D. (1992). L1 influences on learners' renderings of English collocations: A Polish/German empirical study. In Arnaud, P.J.L. and Béjoint H. (eds.) *Vocabulary and Applied Linguistics*. London: Macmillan.

Blum-Kulka, S. (1981). Learning to use words: Acquiring semantic competence in a second language. In Nahir, M. (Ed.) *Hebrew Teaching and Applied Linguistics*. University Press of America.

Blum, S. and Levenston, E.A. (1978). Universals of lexical simplification. *Language Learning 28*, 2: 399-415.

Brown, D. (1974). Advanced vocabulary teaching: The problem of collocation. *RELC Journal 5*, 2: 1-11.

Brown, R. (1965). How shall a thing be called? In Brown, R. (ed.) *Psycholinguistics.* New York: Free Press, Macmillan Publishing. First published in *Psychological Review 65*, 1: 14-21. 1958.

Brown, R. and Berko, J. (1960). Word association and the acquisition of grammar. *Child Development 31*: 1-14.

Bryant, P.E. and Bradley, L. (1980). Why children sometimes write words which they do not read. In Frith, U. (Ed.) *Cognitive Processes in Spelling.* London: Academic Press.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, and G.A. Miller (eds.) *Linguistic Theory and Psychological Reality.* Cambridge, MA: MIT Press.

Carroll, J.B. (1970). An alternative to Juilland's usage coefficient for lexical frequencies, and a proposal for a Standard Frequency Index (SFI). *Computer Studies in the Humanities and Verbal Behavior* **3**, 61-65.

Carroll, J.B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior 10*: 722-729.

Carroll, J.B., Davies, P., and Richman, B. (1971). *Word Frequency Book.* New York: American Heritage Publishing.

Carroll, J.B., Kjeldergaard, P.M., and Carton, A.S. (1962). Number of opposites versus number of primaries as a response measure in free-association tests. *Journal of Verbal Learning and Verbal Behavior 1*: 22-30.

Carter, R. (1983). "You look nice and weedy these days!" Lexical associations, lexicography and the foreign language learner. *Journal of Applied Language Study 1*, 2: 172-189.

Carter, R. (1987). *Vocabulary: Applied Linguistic Perspectives.* London: Unwin Hyman.

Carter, R. and McCarthy, M. (1988). *Vocabulary and Language Teaching.* London: Longman.

Cattell, J. M. and Bryant, S. (1889). Mental association investigated by experiment. *Mind 14*: 230-250.

Chafe, W. and Danielewicz, J. (1987). Properties of spoken and written language. In Horowitz, R. and Samuels, S.J. (eds.) *Comprehending Oral and Written Language*. New York: Academic Press.

Channell, J. (1981). Applying Semantic Theory to Vocabulary Teaching. *ELT Journal 35*: 115-22.

Chui, R.K. (1972). Measuring register characteristics: A prerequisite for preparing advanced level TESOL programmes. *TESOL Quarterly 6*, 2: 129-141.

Church, K. Gale, W., Hanks, P., and Hindle, D. (1991). Using Statistics in Lexical Analysis. In Zernik, U. (ed.) *Using On-line Resources to Build a Lexicon*. New Jersey: Lawrence Erlbaum.

Clark, H.H. (1970). Word associations and linguistic theory. In Lyons, J. (Ed.) *New Horizons in Linguistics*. Harmondsworth, Middlesex: Penguin.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In Huckin, T., Haynes, M., and Coady, J. (Eds.) *Second Language Reading and Vocabulary Learning*. Norwood, NJ: Ablex.

Cruse, D.A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

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

COBUILD English Learner's Dictionary. (1989). London: Collins.

Cooper, R.I. and Greenfield, L. (1969). Word frequency estimation as a measure of degree of bilingualism. *Modern Language Journal 53*, 3: 163-166.

Courtney, R. (1983). *Longman Dictionary of Phrasal Verbs*. London: Longman.

Cowie, A. P. (1978). Vocabulary exercises within an individualized study programme. *ELT Documents 103*: 37-44.

Cowie, A.P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics 2*, 3: 223-35.

Cowie, T. and Howarth, P. Phraseological competence and written proficiency. Paper read at the British Association of Applied Linguistics (BAAL) Conference, Southampton, England, September 1995.

Cowie, A.P. and Mackin, R. (1975). *Oxford Dictionary of Current Idiomatic English, Volume 1.* (Second edition, entitled *Oxford Dictionary of Phrasal Verbs*, 1993). Oxford: Oxford University Press.

Cowie, A.P., Mackin, R., and McCaig, I.R. (1983). *Oxford Dictionary of Current Idiomatic English, Volume 2.* (retitled *Oxford Dictionary of Phrasal Verbs*, 1993). Oxford: Oxford University Press.

Cronbach, L.J. 1942. An analysis of techniques for diagnostic vocabulary testing. *Journal of Educational Research*, 36, 206-217.

Cruse, D.A. (1977). The pragmatics of lexical specificity. *Journal of Linguistics 13*: 153-164.

Cruse, D.A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruttenden, A. (1981). Item-learning and system-learning. *Journal of Psycholinguistic Research 10*: 79-88.

Crystal, D. (1987). *The Cambridge Encyclopedia of Language.* Cambridge: Cambridge University Press.

Cutler, A. and Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*: 218-236.

Cutler, A. and Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language 2*: 133-142.

Cutler, A. and Clifton, C.E. (1984). The use of prosodic information in word recognition. In Bouma, H. and Bouwhuis, D.G. *Attention and Performance X: Control of Language Processes.* Hillsdale, NJ: Lawrence Erlbaum.

Cutler, A. and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 1: 113-121.

Dagut, M.B. (1977). Incongruencies in lexical "gridding" - An application of contrastive semantic analysis to language teaching. *IRAL 15*, 3: 221-229.

Dale, E. (1965). Vocabulary measurement: techniques and major findings. *Elementary English 42*: 895-901, 948.

Dalrymple-Alford, E.C. and Aamiry, A. (1969). Language and category clustering in bilingual free recall. *Journal of Verbal Learning and Verbal Behavior 8*: 762-768.

Dalrymple-Alford, E.C. and Aamiry, A. (1970). Word associations of bilinguals. *Psychonomic Science 21*, 6: 319-320.

Deese, J. (1962). On the structure of associative meaning. *Psychological Review 69*: 161-175.

Deese, J. (1965). *The structure of associations in language and thought.* Baltimore: John Hopkins Press.

deVilliers and deVilliers. (1978). *Language Acquisition.* Harvard.

Diaconis, P. (1988) *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics lecture notes-monograph series **11**. Haywood, CA: Institute of Mathematical Statistics.

Dolinsky, R. (1972). Clustering and free recall with alternative organizational cues. *Journal of Experimental Psychology 95*, 1: 159-163.

Donley, M. (1974). The role of structural semantics in expanding and activating the vocabulary of the advanced learner: The example of the homophone. *Audio-Visual Language Journal, 12*, 2: 81-89.

Drum, P.A. (1983). Vocabulary knowledge. In Niles, J.A. and Harris, L.A. (Eds.) *Searches for Meaning in Reading/Language Processes and Instruction.* Rochester, NY: National Reading Conference.

Drum, P.A. and Konopak, B.C. (1987). Learning word meanings from written context. In McKeown, M.G. and Curtis, M.E. (Eds.) *The Nature of Vocabulary Acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Dulay, H. and Burt, M. (1973). Should we teach children syntax? *Language Learning 23*: 245-258.

Dulay, H. and Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning 24*: 37-53.

Ebbinghaus, H. (1885). *Über das Gedächtnis.* Leip., Duncker. Translated H. Ruger and C.E. Bussenius. 1913. Cited in Woodworth, R.S. and Schlosberg, H. (1955). *Experimental Psychology.* London: Methuen.

Edinburgh Associate Thesaurus. Internet Address: HTTP://www.cis.rl.ac.uk/proj/psych/eat.html
Background information available in: Kiss, G.R., Armstrong, C., Milroy, R., and Piper, J. (1973). An associative thesaurus of English and its computer analysis. In Aitkin, A.J., Bailey, R.W., and Hamilton-Smith, N. (Eds.) *The Computer and Literary Studies*. Edinburgh: Edinburgh University Press.

Ehri, L.C. (1980). The development of orthographic images. In Frith, U. (Ed.) *Cognitive Processes in Spelling*. London: Academic Press.

Ellis, N. (1994). Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *AILA Review 11*: 37-56.

Ellis, N. (1995). Vocabulary acquisition: Psychological perspectives and pedagogical implications. *The Language Teacher 19*, 2: 12-16.

Ellis, N.C. (in press). Vocabulary acquisition: Word structure, collocation, word-class, and meaning. In Schmitt, N. and McCarthy, M. (eds.) *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

Ellis, N.C. and Beaton, A. (1995). Psycholinguistic determinants of foreign language vocabulary learning. In Harley, B. (Ed.) *Lexical Issues in Language Learning*. Ann Arbor: John Benjamins.

Entwisle, D.R. (1966). *Word Associations of Young Children*. Baltimore, MD: Johns Hopkins Press.

Entwisle, D.R., Forsyth, D.F., and Muuss, R. (1964). The syntactic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior 3*: 19-29.

Ervin, S.M. Grammar and classification. Paper read at the American Psychological Association, New York, September 1957.

Ervin, S.M. (1961). Changes with age in the verbal determinants of word association. *American Journal of Psychology 74*: 361-372.

Farghal, M. and Obiedat, H. (1995). Collocations: A neglected variable in EFL. *IRAL 33*, 4: 315-331.

Fay, D. and Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry 8*, 3: 505-520.

Feigenbaum, L.H. (1958). For a bigger better alphabet. *High Points 40*: 34-36.

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

Fisher, C. (1994). Structure and meaning in the verb lexicon: Input for a syntax-aided verb learning procedure. *Language and Cognitive Processes 9*, 4: 473-517.

Flavell, J.H., Draguns, J., Feinberg, L.D., and Budin, W. (1958). A microgenetic approach to word association. *Journal of Abnormal and Social Psychology 57*: 1-7.

Flexser, A.J. and Bower, G.H. (1975). Further evidence regarding instructional effects on frequency judgements. *Bulletin of the Psychonomic Society 6*, 3: 321-324.

Forrester, W.E. and Tambs, C.A. (1976). Word associations to 19 vowels. *Journal of Experimental Psychology: Human Learning and Memory 2*, 2: 200-207.

Forster, K. (1976). Accessing the mental lexicon. In R.J. Wales and E. Walker, *New Approaches to Languages Mechanisms.* Amsterdam: North-Holland.

Fox, P.W. (1970). Patterns of stability and change in behaviors of free association. *Journal of Verbal Learning and Verbal Behavior 9*: 30-36.

Freyd, P. and Baron, J. (1982). Individual differences in acquisition of derivational morphology. *Journal of Verbal Learning and Verbal Behavior 21*: 282-295.

Frith, U. (1980). Unexpected spelling problems. In Frith, U. (Ed.) *Cognitive Processes in Spelling.* London: Academic Press.

Frost, R., Katz, L. and Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance 13*, 1: 104-115.

Galton, F. (1879-1880). Psychometric experiments. *Brain* 2: 149-162.

Garnham, A. (1985). *Psycholinguistics: Central Topics.* London: Routledge.

Gekoski, W.L. and Riegel, K.F. (1966). A study of the one-week stability of the Michigan restricted association norms. *Psychological Reports 18*: 397-398.

Goldstein, H. (1983). Word recognition in a foreign language: A study of speech perception. *Journal of Psycholinguistic Research, 12*, 4: 417-427.

Goodglass, H. and Baker, E. (1976). Semantic field, naming and auditory comprehension in aphasia. *Brain and Language 3*: 359-374.

Goodman, G.S., Haith, M.M., Guttentag, R.E., and Rao, S. (1985). Automatic processing of word meaning: Intralingual and interlingual interference. *Child Development 56*: 103-118.

Granger, S. (forthcoming). Prefabricated patterns in advanced EFL writing: Collocations and lexical phrases. In Cowie, A.P. (Ed.) *Phraseology: Theory, Analysis and Application.* Oxford: Oxford University Press.

Graves, M.F. (1987). The roles of instruction in fostering vocabulary development. In McKeown, M.G. and Curtis, M.E. (Eds.) *The Nature of Vocabulary Acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Graves, M.F., Boettcher, J.A., Peacock, J.L., and Ryder, R.J. (1980). Word frequency as a predictor of students' reading vocabularies. *Journal of Reading Behavior 12*, 2.

Green, D.W. and Meara, P. (1987). The effects of script on visual search. *Second Language Research 4*: 102-117.

Green, J.N. (1979). Towards a statistical delimitation of register in Spanish. *IRAL 17*, 3: 233-244.

Gregg, V.H., Montgomery, D.C., and Castaño, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior 19*: 240-245.

Gregory, M. (1967). Aspects of varieties differentiation. *Journal of Linguistics 3*, 2: 177-198.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38*: 299-310.

Grosjean, F. and Gee, J.P. (1987). Prosodic structure and spoken word recognition. *Cognition, 25*: 135-155.

Grucza, B. and Jaruzelska, H. (1978). Typowe bledy popelnanie przez kandydatow jezyku niemieckim: Bledy gramatyczne i leksykalne. In Grucza, F. (ed.) *Z Problematyki Bledow Obcojezycznych.* Warsaw: WSP.

Halliday, M.A.K. (1978). *Language as Social Semiotic.* London: Edward Arnold.

Hammerly, H. (1979). Conveying lexical meanings in second-language teaching. *Canadian Modern Languages Review 35*: 567-580.

Hancin-Bhatt, B. and Nagy, W. Lexical transfer and second-language morphological development. Unpublished manuscript.

Hasher, L. and Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior 16*: 173-184.

Hatch, E. and Lazaraton, A. (1991). *The Research Manual.* New York: Newbury House.

Haynes, M. (1993). Patterns and perils of guessing in second language reading. In Huckin, T., Haynes, M. and Coady, J. (Eds.) *Second Language Reading and Vocabulary Learning.* Norwood, NJ: Ablex.

Henning, G. (1991). *A Study of the Effects of Contextualization and Familiarization on Responses to the TOEFL Vocabulary Test Items.* TOEFL Research Reports, No. 35. Princeton, NJ: Educational Testing Service.

Higa, M. (1963). Interference effects of intralist word relationships in verbal learning. *Journal of Verbal Learning and Verbal Behavior 2*: 170-175.

Hintzman, D.L. (1969). Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology 80*, 1 : 139-145.

Hintzman, D.L. (1970). Effects of repetition and exposure duration on memory. *Journal of Experimental Psychology 83*, 3: 435-444.

Hintzman, D.L. and Block, R.A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology 88*, 3: 297-306.

Hintzman, D.L., Block, R.A., and Summers, J.J. (1973). Modality tags and memory for repetitions: Locus of the spacing effect. *Journal of Verbal Learning and Verbal Behavior 12*: 229-238.

Hintzman, D.L. and Summers, J.J. (1973). Long-term visual traces of visually presented words. *Bulletin of the Psychonomic Society 1*: 325-327.

Hirtle, W. (1994). Meaning and referent: For a linguistic approach. *Word 45*, 2: 103-117.

Horton, D.L., Marlowe, D. and Crowne, D.P. (1963). The effect of instrumental set and need for social approval on commonality of word association responses. *Journal of Abnormal and Social Psychology 66*: 67-72.

Howard, D.V., McAndrews, M.P., and Lasaga, M.I. (1981). Semantic priming of lexical decisions in young and old adults. *Journal of Gerontology 36*: 707-714.

Howell, W.C. (1973a). Effects of organization and discrimination of word frequency within and between categories. *Journal of Experimental Psychology 99*, 2: 255-260.

Howell, W.C. (1973b). Storage of events and event frequencies: A comparison of two paradigms in memory. *Journal of Experimental Psychology 98*, 2: 260-263.

Howell, W.C. (1973c). Representation of frequency in memory. *Psychological Bulletin 80*, 1: 44-53.

Howes, D. (1957). On the relation between the probability of a word as an association and in general linguistic usage. *Journal of Abnormal and Social Psychology 54*: 75-85.

Howes, D.H. (1966). A word-count of spoken English. *Journal of Verbal Learning and Verbal Behavior 5*: 572-604.

Huckin, T. and Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In Huckin, T., Haynes, M. and Coady, J. (Eds.) *Second Language Reading and Vocabulary Learning*. Norwood, NJ: Ablex.

Ijaz, I.H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning 36*, 4: 401-451.

Imai, M., Gentner, D., and Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development 9*: 45-75.

Ingersoll, G.M. (1974). Effects of age, form class, and word frequency on homogenous word associations. *Psychological Reports 35*: 59-64.

International English Language Testing System: The IELTS Handbook. (1996). Cambridge: University of Cambridge Local Examinations Syndicate, The British Council, and IDP Education Australia.

Jacoby, L.L. (1972). Context effects on frequency judgements of words and sentences. *Journal of Experimental Psychology 94*, 3: 255-260.

Jenkins, J.J. (1954). Transitional organization: association techniques. In Osgood, C.E. and Sebeok, T.A. (eds.) *Psycholinguistics, Supplement, Journal of Abnormal and Social Psychology 52*: 114-116.

Jenkins, J.J. (1959). Effects on word-association of the set to give popular responses. *Psychological Reports 5*: 94.

Jenkins, J.J. (1970). The 1952 Minnesota word association norms. In L. Postman and G. Keppel *Norms of Word Associations.* New York: Academic Press.

Jenkins, J.J. and Russell, W.A. (1960). Systematic changes in word association norms: 1910-1952. *Journal of Abnormal and Social Psychology 60*, 3: 293-304.

Johnston, M.H. (1974). Word associations of schizophrenic children. *Psychological Reports 35*: 663-674.

Juillard, A., Brodin, D., and Davidovitch, C. (1970). *Frequency Dictionary of French Words.* The Hague-Paris: Mouton.

Kahana, B. and Sterneck, R. (1972). Word-association responses of children as a function of age, sex, and instructions. *The Journal of Genetic Psychology 120*: 39-48.

Katz, J.J. and Fodor, J.A. (1963). The structure of a semantic theory. *Language 39*: 170-210.

Keene, C.M. (1951). Commonality of response on a word-association test: A study of standardization procedures and an attempt to forecast moderate emotional maladjustment. Unpublished doctoral dissertation, Stanford University.

Keil, F.C. (1994). Explanation, association, and the acquisition of word meaning. In Gleitman, L. and Landau, B. (Eds.) *The Acquisition of the Lexicon.* Cambridge, MA: Bradford.

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers in Bilingualism 15*: 60-92.

Kent, G.H. and Rosanoff, A.J. (1910). A study of association in insanity. *American Journal of Insanity 67*: 37-96, 317-390.

Kirsner, K. (1973). An analysis of the visual component in recognition memory for verbal stimuli. *Memory and Cognition 1*, 4: 449-453.

Kirsner, K., Smith, M.C., Lockhart, M.L., King, M.L., and Jain, M. (1984). The bilingual lexicon: Specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior 23*: 519-539.

Kiss, G.R. (1968). Words, associations, and networks. *Journal of Verbal Learning and Verbal Behavior 7*: 707-713.

Koda, K. (1997). Orthographic knowledge in L2 lexical processing. In Coady, J. and Huckin, T. (Eds.) *Second Language Vocabulary Acquisition.* Cambridge: Cambridge University Press.

Kolers, P.A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior* 2: 291-300.

Krashen, S. (1977). Some issues relating to the Monitor Model. In Brown, H. Yorio, C., and Crymes, R. (eds.) *On TESOL '77.* Washington, DC: TESOL.

Kruse, H., Pankhurst, J., and Sharwood Smith. (1987). A multiple word association probe in second language acquisition research. *Studies in Second Language Acquisition* 9: 141-154.

Kučera, H. and Francis, W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Labov, W. (1970). The study of language in its social context. *Stadium Generale 23*: 30-87.

Lackner, J.R. and Garrett, M.F. (1972). Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition 1*: 359-372.

Lado, R. (1964). *Language Teaching.* NY: McGraw-Hill.

Lambert, W.E. and Moore, N. (1966). Word association responses: Comparisons of American and French monolinguals with Canadian monolinguals and bilinguals. *Journal of Personality and Social Psychology 3*, 3: 313-320.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9: 409-430.

Larsen-Freeman, D. and Long, M.H. (1991). *An Introduction to Second Language Acquisition Research.* London: Longman.

Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review 6*: 10-20.

Laufer, B. (1990). Why are some words more difficult than others? - Some intralexical factors that affect the learning of words. *IRAL 28*: 293-307.

Laufer, B. (1991). Some properties of the foreign language learner's lexicon as evidenced by lexical confusions. *IRAL. 29*, 4: 317-330.

Laufer, B. (in press). 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.) *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

Laufer, B. and Bensoussan, M. (1982). Meaning is in the eye of the beholder. *English Teaching Forum 20*, 2: 10-14.

Leech, G. (1974). *Semantics*. Harmondsworth, UK: Penguin.

Lehrer, A. (1985). Is semantics perception-driven or network driven? *Australian Journal of Linguistics 5*: 197-207.

Lerea, L. and LaPorta, R. (1971). Vocabulary and pronunciation acquisition among bilinguals and monolinguals. *Language and Speech, 14*: 293-300.

Levenson, E. The acquisition of polysemic words with both literal and metaphorical meaning. Unpublished manuscript.

Levi, M. (1949). An analysis of the influence of two different cultures on responses to the Rosanoff free association test. Unpublished master's thesis. University of Chicago.

Lewis, M. (1993). *The Lexical Approach*. Hove, England: Language Teaching Publications.

Locke, J. (1700). *An Essay Concerning Humane Understanding*. In 4 books. 4th edition. London: Awnsham and John Churchil.

Long, M. and Sato, C. (1984). Methodological issues in interlanguage studies: An interactionist perspective. In Davies, A. Criper, C. and Howatt, A. (eds.) *Interlanguage*. Edinburgh: Edinburgh University Press.

Long, T.H. and Summers, D. (eds.) *Longman Dictionary of English Idioms*. London: Longman.

Longman Language Activator. (1993). Harlow: Longman.

Longman Dictionary of English Language and Culture. (1992). Harlow: Longman.

Lovell, G.D. (1941). Interrelations of vocabulary skills: Commonest versus multiple meanings. *Journal of Educational Psychology 32*: 67-72.

Lukatela, G., Savić, M.D., Ognjenović, P., and Turvey, M.T. (1978). On the relation between processing the Roman and the Cyrillic alphabets: A preliminary analysis with bi-alphabetic readers. *Language and Speech 21*, 2: 113-141.

Lukatela, G. and Turvey, M.T. (1994a). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones and pseudohomophones. *Journal of Experimental Psychology: General 123*: 107-128.

Lukatela, G. and Turvey, M.T. (1994b). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General 123*: 331-353.

Lukatela, G., Turvey, M.T., Feldman, L.B., Carello, C. and Katz, L. (1989). Alphabet priming in bi-alphabetical word perception. *Journal of Memory and Language 28*: 237-254.

Lyons, J. (1977). *Semantics* (V. 1 & 2). New York: Cambridge University Press.

Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior 6*: 729-736.

Marsh, G., Friedman, M.P., Desberg, P., and Saterdahl, K.A. (1979). Comparison of reading and spelling strategies in normal and reading disabled children. Paper presented at the NATO conference on Intelligence and Learning, July, York.

Marsh, G., Friedman, M., Welch, V., and Desberg, P. (1980). The development of strategies in spelling. In Frith, U. (Ed.) *Cognitive Processes in Spelling*. London: Academic Press.

Marshall, G.R. and Cofer, C.N. (1963). Associative indices as measures of word relatedness: A summary and comparison of ten methods. *Journal of Verbal Learning and Verbal Behavior 1*: 408-421.

Marslen-Wilson, W. (ed.). (1989). *Lexical Representation and Process*. Cambridge, MASS: Bradford Books, The MIT Press.

Marslen-Wilson, W.D. and Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition 8*, 1: 1-71.

Marton, W. (1978). *Dydaktyka Jezyka Obcego w Szkole Sredniej: Podejscie Kognitywne*. Warsaw: PWN.

Mason, K. (1986). The meeting of minds: Linguistics and philosophy of language. In Coulthard, M. (Ed.) *Talking about Text*. University of Birmingham: ELR.

McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.

McCarthy, M. and Carter, R. (in press). Written and spoken vocabulary. In Schmitt, N. and McCarthy, M. (eds.) *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

McClelland, J.L. (1977). Letter and configuration information in word identification. *Journal of Verbal Learning and Verbal Behavior 16*: 137-150.

McClelland, J.L. and Elman, J.E. (1986). The TRACE model of speech perception. *Cognitive Psychology 18*: 1-86.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models* (2nd edition). London: Chapman and Hall.

McCusker, L.X., Hillinger, M.L., and Bias, R.G. (1981). Phonological recoding and reading. *Psychological Bulletin 89*, 2: 217-245.

McNeill, D. (1963). The origin of associations within the same grammatical class. *Journal of Verbal Learning and Verbal Behavior 2*: 250-262.

McNeill, D. (1966). A study of word association. *Journal of Verbal Behavior and Verbal Learning 5*: 548-557.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching & Linguistics: Abstracts 13*, 4: 221-246.

Meara, P. (1983a). Word associations in a foreign language. *Nottingham Linguistic Circular 11*, 2: 29-38.

Meara, P. (1983b). Word recognition in foreign languages. In Pugh, A. and Ulijn, J. (Eds.) *Reading for Professional Purposes*. London: Heinemann.

Meara, P. (1984). The study of lexis in interlanguage. In Davies, A., Criper, C., and Howatt, A.P.R. (eds.) *Interlanguage*. Edinburgh: Edinburgh University Press.

Meara, P. (1986). The DIGAME project. In Cook, V.J. (Ed.) *Experimental Approaches to Second Language Learning*. Oxford: Pergamon Press.

Meara, P. (1992). Network structures and vocabulary acquisition in a foreign language. In Arnaud, P.J.L. and Bejoint, H. (Eds.) *Vocabulary and Applied Linguistics*. London: Macmillan.

Meara, P. (1996). The dimensions of lexical competence. In Brown, G., Malmkjaer, K., and Williams, J. (Eds.) Competence and Performance in Language Learning. Cambridge: Cambridge University Press.

Meara, P. (discussion paper) The Vocabulary Knowledge Framework.

Meara, P. and Ingle, S. (1986). The formal representation of words in an L2 speaker's lexicon. *Second Language Research* 2, 2: 160-171.

Meara, P.M. and Jones, G. (1990). *The Eurocentres Vocabulary Size Test 10K*. Zurich: Eurocentres.

Miller, G.A. and Charles, W.G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes* 6, 1: 1-28.

Miller, G.A. and Gildea, P.M. (1987). How children learn words. *Scientific American 257*: 94-99.

Miller, G.A. and Johnson-Laird, P.N. (1976). *Language and Peception.* Cambridge: Cambridge University Press.

Mintz, S. (1969). Effect of actual stress on word associations. *Journal of Abnormal Psychology 74*, 3: 293-295.

Moran, L.J. (1966). Generality of word-association response sets. *Psychological Monographs: General and Applied 80*, 4: No. 612.

Moran, L.J. (1967a). Objective approximations of idiodynamic associative sets. *Psychological Reports 20*: 827-833.

Moran, L.J. (1967b). Critique of "Objective approximations of idiodynamic associative sets." *Psychological Reports 21*: 240.

Moran, L.J., Mefferd, R.B., and Kimble, J.P. (1964). Idiodynamic sets in word association. *Psychological Monographs 78*, 2 (No. 579).

Moran, L.J. and Murakawa, N. (1968). Japanese and American association structures. *Journal of Verbal Learning and Verbal Behavior* 7: 176-181.

Moran, L.J. and Swartz, J.D. (1970). Longitudinal study of cognitive dictionaries from ages nine to seventeen. *Developmental Psychology 3*, 1: 21-28.

Morgan, C.L. and Bonham, D.N. (1944). Difficulty of vocabulary learning as affected by parts of speech. *Journal of Educational Psychology 35*, 5: 369-377.

Morton, J. (1979). Word recognition. In Morton, J. and Marshall, J.C. (Eds.) *Psycholinguistics Series 2: Structures and Processes*. London: Elek.

Nagy, W. and Anderson, R.C. (1984). How many words are there in printed school English? *Reading Research Quarterly 19*: 304-330.

Nagy, W.E., Diakidoy, I.N., and Anderson, R.C. (1991). *The Development of Knowledge of Derivational Suffixes*. Technical Report No. 536. ERIC Document ED334560.

Nagy, W.E., Diakidoy, I.N., and Anderson, R.C. (1993). The acquisition of morphology: Learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading Behavior 25*, 2: 155-170.

Nagy, W.E. and Gentner, D. (1990). Semantic constraints on lexical categories. *Language and Cognitive Processes 5*, 3: 169-201.

Nagy, W.E. and Scott, J.A. (1990). Word schemas: Expectations about the form and meaning of new words. *Cognition and Instruction 7*, 2: 105-127.

Nation, P. (Ed.). (1994). *New Ways in Teaching Vocabulary*. Alexandria, VA: TESOL.

Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.

Nattinger, J.R. and DeCarrico, J.S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

O'Conner, J. (1928). *Born That Way*. Baltimore: Williams and Wilkins.

Odlin, T. and Natalicio, D. (1982). Some characteristics of word classification in a second language. *The Modern Language Journal 66*: 34-38.

Oxford Advanced Learner's Dictionary. (1995). Oxford: Oxford University Press.

Palermo, D.S. (1963). Word associations and children's verbal behavior. In Lipsitt, L.P. and Spiker, C.C. (eds.) *Advances in Child Development and Behavior.* Vol. 1. New York: Academic Press.

Palermo, D.S. (1971). Characteristics of word association responses obtained from children in grades one through four. *Developmental Psychology* 5, 1: 118-123.

Palermo, D.S. and Jenkins, J.J. (1963). Frequency of superordinate responses to a word association test as a function of age. *Journal of Verbal Learning and Verbal Behavior 1*: 378-383.

Palermo, D.S. and Jenkins, J.J. (1964). *Word Association Norms: Grade School through College.* Minneapolis: University of Minnesota.

Papagno, C., Valentine, T., and Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*: 331-347.

Paribakht, T.S. and Wesche, M.B. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal 11,* 1: 9-29.

Perkins, K. (1983). Semantic constructivity in ESL reading comprehension. *TESOL Quarterly 17*, 1: 19-27.

Peters, A. (1983). *The Units of Language Acquisition.* Cambridge: Cambridge University Press.

Petrey, S. (1977). Word associations and the development of lexical memory. *Cognition 5*: 57-71.

Phillips, T.A. (1981). *Difficulties in Foreign Language Vocabulary Learning and a Study of some of the Factors Thought to be Influential.* M.A. Project. Birkbeck College, University of London.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition* 6, 2: 186-214.

Postman, L. and Keppel, G. (1970). *Norms of word association.* New York: Academic Press.

Prasada, S. and Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes 8,* 1: 1-56.

Preece, P.F.W. (1977). Developmental trends in the continued word associations of physics students. *Journal of Research in Science Teaching 14*, 3: 235-239.

Pressley, M., Levin, J.R., and Miller, G.E. (1982). The keyword method compared to alternative vocabulary learning strategies. *Contemporary Educational Psychology 7*: 50-60.

Rayner, K. and Balota, D.A. (1989). Parafoveal preview and lexical access during eye fixations in reading. In Marslen-Wilson (Ed.) *Lexical Representation and Process*. Cambridge, MA: Bradford.

Rayner, K. and Hagelberg, E.M. (1975). Word recognition cues for beginning and skilled readers. *Journal of Experimental Child Psychology 20*: 444-455.

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

Read, J. Refining the word associates format as a measure of depth of vocabulary knowledge. Paper given at the 19th Annual Congress of the Applied Linguistics Association of Australia, Melbourne, July 1994.

Read, J. (in press). Vocabulary and testing. In Schmitt, N. and McCarthy, M. (Eds.) *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

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

Richards, J.C. (1974). Word lists: Problems and prospects. *RELC Journal 5*, 2: 69-84.

Riegel, K.F., Ramsey, R.M., and Riegel, R.M. (1967). A comparison of the first and second languages of American and Spanish students. *Journal of Verbal Learning and Verbal Behavior 6*: 536-544.

Riegel, K.F. and Zivian, I.W.M. (1972). A study of inter- and intralingual associations in English and German. *Language Learning 22*, 1: 51-63.

Rierdan, J. (1980). Word association of socially isolated adolescents. *Journal of Abnormal Psychology 89*, 1: 98-100

Ringeling, T. (1984). Subjective estimations as a useful alternative to word frequency counts. *Interlanguage Studies Bulletin 8*, 1: 59-69.

Robinson, P.J. (1988). A Hallidayan framework for vocabulary teaching - An approach to organizing the lexical content of an EFL syllabus. *IRAL 26*, 3: 229-238.

Rodgers, T.S. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *International Review of Applied Linguistics 7*: 327-343.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology 104*, 192-233.

Rose, R.J. and Rowe, E.J. (1976). Effects of orienting task and spacing of repetitions on frequency judgements. *Journal of Experimental Psychology: Human Learning and Memory* 2, 2: 142-152.

Rosenzweig, M.R. (1961). Comparisons among word-association responses in English, French, German, and Italian. *American Journal of Psychology 74*: 347-360.

Rosenzweig, M.R. (1964). Word associations of French workmen: Comparisons with associations of French students and American workmen and students. *Journal of Verbal Behavior and Verbal Learning 3*: 57-69.

Rowe, E.J. (1973a). Frequency judgements and recognition of homonyms. *Journal of Verbal Learning and Verbal Behavior 12*: 440-447.

Rowe, E.J. (1973b). Context effects in judgement of frequency. *Bulletin of the Psychonomic Society* 2, 4: 231-232.

Rowe, E.J. (1974). Depth of processing in a frequency judgement task. *Journal of Verbal Learning and Verbal Behavior 13*: 638-643.

Rudzka, B., Channell, J., Ostyn, P., and Putseys, Y. (1981). *The Words You Need.* London: Macmillan.

Rudzka, B., Channell, J., Putseys, Y., and Ostyn, P. (1985). *More Words You Need.* London: Macmillan.

Russell, W.A. and Jenkins, J.J. (1954). *The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test.* Tech. Rep. No. 11, Contract No. N8-onr-66216, Office of Naval Research and University of Minnesota.

Russell, W.A. and Meseck, O.R. (1959). Der einfluss der assoziation auf das Erinnern von Worten in der Duetschen, Französischen, und Englischen sprache. *Z. exp. angew. Psychol. 6*: 191-211.

Ryan A (1994). *'Vowel Blindness' in Arabic learners of English.* Unpublished PhD Thesis. University of Wales, Swansea.

Ryan, A. (in press). Learning the orthographic form of L2 vocabulary - A receptive and productive process. In Schmitt, N. and McCarthy, M. (Eds.) *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press.

Ryan, A. and Meara, P.M. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language* 7, 2: 531-540.

Sandra, D. (1988). Is morphology used to encode derivations when learning a foreign language? *ITL Review of Applied Linguistics 79/80*: 1-23.

Sandra, D. (1993). The use of lexical morphology as a natural aid in learning foreign language vocabulary. In Chapelle, J. and Claes, M.T. (Eds.) *Memory and Memorization in Acquiring and Learning Languages.* C.L.L. Centre de Langues à Louvain-la-neuve et en Woluwe.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes* 9, 3: 227-269.

Schellenberg, P.E. (1930). A group free-association test for college students. Unpublished doctoral dissertation, University of Minnesota. Published in an abridged form in Tinker, M.A. and Baker, K.H. (1938). *Introduction to Methods in Experimental Psychology.* New York: Appleton-Century. pp. 213-214.

Schmitt, N. (1994). Vocabulary testing: Questions for test development with six examples of tests of vocabulary size and depth. *Thai TESOL Bulletin,* 6, 2: 9-16.

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

Schmitt, N. (1995b). *Verbal Suffix and Word Association Knowledge of Japanese Students.* Unpublished MPhil. Dissertation. University of Wales, Swansea.

Schmitt, N. (1995c). The word on words: An interview with Paul Nation. *The Language Teacher* **19**, 5-7.

Schmitt, N. (in press). Vocabulary learning strategies. In Schmitt, N. and McCarthy, M. (eds.) *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press.

Schmitt, N. and McCarthy, M. (Eds.) (in press). *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press.

Schmitt, N. and Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition 19*, 1: 17-36.

Scholfield, P. (in press). Vocabulary reference works in foreign language learning. In Schmitt, N. and McCarthy, M. (eds.) *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press.

Scholfield, P. (1995). *Quantifying Language.* Clevedon, England: Multilingual Matters.

Schonell, F.J., Meddleton, I.G., and Shaw, B.A. (1956). *A Study of the Oral Vocabulary of Adults.* Brisbane: University of Queensland Press.

Segalowitz, N. and Hébert, M. (1990). Phonological recoding in the first and second language reading of skilled bilinguals. *Language Learning, 40*, 4: 503-538.

Seibert, L.C. (1927). An experiment in learning French vocabulary. *Journal of Educational Psychology 18*: 294-309.

Shapiro, B.J. (1969). The subjective estimation of relative word frequency. *Journal of Verbal Learning and Verbal Behavior 8*: 248-251.

Sharp, D. and Cole, M. (1972). Patterns of responding in the word associations of West African children. *Child Development 43*: 55-65.

Sharwood Smith, M. (1984). Discussion of P. Meara's `The study of lexis in interlanguage'. In Davies, A., Criper, C., and Howatt, A.P.R. (eds.) *Interlanguage.* Edinburgh: Edinburgh University Press.

Silva, C.M. and Zwicky, A.M. (1973). Discord. *Working Papers in Linguistics, No. 16.* Paper presented at the 2nd Annual Colloquium on New Ways of Analyzing Variation, Washington, D.C.

Simpson, J. (1988). The vocabulary of English. In Stanley, E.G. and Hoad, T.F. (Eds.) *Words.* Cambridge: D.S. Brewer.

Simpson, P. and Voss, J.F. (1967). Stability of response hierarchies. *Journal of Experimental Psychology 75*, 2: 170-174.

Sinclair, J., *et al.* (1989). *Collins COBUILD Dictionary of Phrasal Verbs.* London: Collins.

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

Söderman, T. (1993). Word associations of foreign language learners and native speakers - A shift in response type and its relevance for a theory of lexical development. In Ringbom, H. (Ed.) *Near-Native Proficiency in English.* Åbo: Åbo Akademi.

Sökmen, A.J. (1993). Word association results: A window to the lexicons of ESL students. *JALT Journal 15*, 2: 135-150.

Sonaiya, R. (1991). Vocabulary acquisition as a process of continuous lexical disambiguation. *IRAL 29,* 4: 273-284.

Spolsky, B. (1995). *Measured Words.* Oxford: Oxford University Press.

Stahl, S.A. and Murray, B.A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 2: 221-234.

Stahl, S.A. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior 15*, 4: 33-50.

Stevens, S.S. (1956). The direct estimation of sensory magnitudes - loudness. *The American Journal of Psychology 69*, 1: 1-25.

Stevens, S.S. (1958). Problems and methods of psychophysics. *Psychological Bulletin 55*, 4: 177-196.

Stock, R.D. (1976). *Some Factors Affecting the Acquisition of Foreign language Lexicon in the Classroom.* Unpublished Ph.D Thesis. Champaign-Urbana, IL: University of Illinois.

Stoller, F.L. and Grabe, W. (1993). Implications for L2 vocabulary acquisition and instruction from L1 vocabulary research. In Huckin, T., Haynes, M., and Coady, J. (Eds.) *Second Language Reading and Vocabulary Learning.* Norwood, NJ: Ablex.

Strick, G.J. (1980). A hypothesis for semantic development in a second language. *Language Learning 30*, 1: 155-176.

Stubbs, M. (1980). *Language and Literacy.* London: Routlege and Kegan Paul.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language 2*, 1: 1-33.

Suárez, A. and 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.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior 18*: 645-659.

Szalay, L.B., Windle, C., and Lysne, D.A. (1970). Attitude measurement by free verbal associations. *Journal of Social Psychology* 82: 43-55.

Taft, M. (1984a). *Prosodic constraints and lexical parsing strategies.* Unpublished doctoral dissertation. University of Massachusetts.

Taft, M. (1984b). Evidence for an abstract lexical representation of word structure. *Memory & Cognition, 12*, 3: 264-269.

Taft, M. and Forster, K.I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Behavior and Verbal Learning 14*: 638-647.

Taylor, I. (1971). How are words from two languages organized in bilinguals' memory? *Canadian Journal of Psychology* 25, 3: 228-240.

Tenney, Y.J. (1980). Visual factors in spelling. In Frith, U. (Ed.) *Cognitive Processes in Spelling.* London: Academic Press.

Thomas, M.H. and Dieter, J.N. (1987). The positive effects of writing practice on integration of foreign words in memory. *Journal of Educational Psychology* 79, 3: 249-253.

Thorndike, E.L. and Lorge, I. (1944). *A Teacher's Word Book of 30,000 Words.* New York: Columbia University Press.

Thrasher, R.H. An experimental measure of lexical knowledge applied to ESL text analysis. Paper presented to the AILA/TESOL Seminar on Language Testing. San Juan, Puerto Rico. May, 11, 1973. Cited in Upsur, J.A. (1975). *Word Frequency Estimation in Foreign Language Learning.* PhD dissertation, University of Michigan. Also cited in Arnaud, P.J.L. Subjective word frequency estimates in L1 and L2. Paper presented at the 9th World Congress of Applied Linguistics. Thessaloniki. April 15-21, 1990.

TOEFL Practice Tests. (1995). Educational Testing Service. Princeton, NJ: ETS.

Tryk, H.E. (1968). Subjective scaling of word frequency. *American Journal of Psychology 81*: 170-177.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology 5*: 207-232.

Tyler, L.K., Marslen-Wilson, W., Rentoul, J., and Hanney, P. (1988). Continuous and discontinuous access in spoken word-recognition: The role of derivational prefixes. *Journal of Memory and Language 27*: 368-381.

Tyler, A. and Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language 28*: 649-667.

Tyler, A. and Nagy, W. (1990). Use of derivational morphology during reading. *Cognition 36*: 17-34.

Underwood, B.J. (1971). Recognition memory. In Kendler, H.H. and Spence, J.T. (Eds.) *Essays in Neobehaviorism*. New York: Appleton-Century-Crofts.

Underwood, B.J. (1972). Word recognition memory and frequency information. *Journal of Experimental Psychology 94*, 3: 276-283.

Upsur, J.A. (1975). *Word Frequency Estimation in Foreign Language Learning*. PhD dissertation, University of Michigan.

Upward, C. (1988). *English Spelling and Educational Progress*. CLIE Working Papers 11.

van Ginkel, C.I. and van der Linden, E.H. (1996). Word associations in foreign language learning and foreign language loss. In Sajavaara, K. and Fairweather, C. (Eds.) *Approaches to Second Language Acquisition*. Jyväskyä.

Van Roey, J. (1990). *French-English Contrastive Lexicography. An Introduction*. Louvain-la-Neuve: Peeters.

Verhallen, M. and Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics 14*, 4: 344-363.

Verhoeven, L. and Vermeer, A. (1989). *Diagnose van Kindertaal* [Diagnosis of Child Language]. Tilburg: Zwijssen.

Wallace, G. and Larsen, S.C. (1978). *Educational Assessment of Learning Problems*. Boston: Allyn and Bacon.

Wallace, W.D. (1981). How registers register: Toward the analysis of language use. *IRAL 19*, 4: 267-286.

Webster's New World Thesaurus. (1974) New York: Warner Books.

Webster's Ninth New Collegiate Dictionary. (1987). Springfield, MA: Merriam-Webster.

Wells, F.L. (1911). *Psychology Review 18*: 229-233.

Wells, F.L. (1927). *Mental Tests in Clinical Practice.*

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

Willis, J. Using concordances in the classroom. Presentation given at MATSDA 97, Dublin, January 1997.

Wittgenstein, L. (1958). *Philosophical Investigations.* Translated by G.E.M. Anscombe. 2nd edition. Oxford: Basil Blackwell.

Woodrow, H. and Lowell, F. (1916). Children's association frequency tables. *Psychology Monographs* 22, 5, No. 97.

Woodworth, R.S. (1938, 3rd edition 1954 with H. Schlosberg). *Experimental psychology.* New York: Holt.

WORD FREQUENCY SURVEY

This survey is part of the research being done at the University of Nottingham English Studies department into how people use language. It contains 12 sets of words. Each set contains 5 words with very similar meanings. Of course, some of these 5 words will be more common and more frequent than the others in the set. This survey is interested in your judgements of how frequently each word in a set occurs in general language compared to the other words in the set. By "general language", we mean you should judge how frequently a word is used *in society as a whole*, not your personal use of that word.

To help you make these frequency judgements, we have randomly given a word in each set a value of one (1). This is the 'anchor' value, or benchmark, by which you should judge the relative frequency of the other words in the set. If you think a word is used twice as frequently as the anchor word, write **2** on the blank beside that word. If you think a word is 500 times more frequent than the anchor word, write **500**. You can also use decimals or fractions. For example, if you think that a word is 1½ times more frequent than the anchor word, write either **1.5** or **1½** on the blank.

If you think a word is *less* frequent than the anchor word, use a decimal figure or fraction to show this. For example, if you think a word is half as frequent as the anchor word, write **.5** or ½ on the blank. If you think it is 8/10ths as frequent as the anchor word, write **.8** or **8/10**; if it is 1/100th as frequent, write **.01** or **1/100**; if it is 1/500th as frequent, write **.002** or **1/500**, etc. The following examples show what sets may look like after you complete them:

amble	1	‹ (anchor word) ›	afraid	1
saunter			frightened	
stroll			petrified	
walk			scared	
wander			terrified	

If you do not know a word, write **X** on the blank. If you do not know the anchor word: write **X** on its blank, choose another word to be the anchor word and write **1** on its blank, and continue rating the set.

Since the anchor words were randomly chosen, they can be anything from extremely frequent to extremely rare. Thus, in some sets, most or all of the words may be more frequent than the anchor word. In other sets, most or all of the words may be less frequent than the anchor word. Also note that some sets may have words which vary widely in their frequency of occurrence, while other sets may contain words which are more closely grouped together in terms of frequency.

Do not discuss this survey with anyone until you are finished.
Please complete every blank.
Your answers will remain anonymous.

calamity	
cataclysm	
catastrophe	1
disaster	
tragedy	

chore	
duty	
errand	
job	
task	1

appliance	
contrivance	
device	
gadget	1
machine	

emblem	1
hallmark	
logo	
sign	
symbol	

Verbs

annihilate	
conquer	1
defeat	
subdue	
vanquish	

deflate	
diminish	
lessen	1
minimize	
reduce	

glisten	
shimmer	
shine	1
sparkle	
twinkle	

disclose	1
divulge	
expose	
impart	
reveal	

awful	____________
bad	_____1_____
dreadful	____________
ghastly	____________
terrible	____________

low-tech	____________
primitive	____________
rudimentary	____________
simple	____________
unsophisticated	_____1_____

essential	_____1_____
imperative	____________
necessary	____________
requisite	____________
vital	____________

flimsy	____________
fragile	____________
frail	____________
puny	_____1_____
weak	____________

[Native-speaker biodata form]

Could we have some information about yourself?

Tick one:

Sex: Male _____ Female _____

Education: O-Levels/GCSE _____ A-Levels _____

University _____ Postgraduate _____

Thank you very much for your cooperation on this survey project.

awful	______
bad	1
dreadful	______
ghastly	______
terrible	______

low-tech	______
primitive	______
rudimentary	______
simple	______
unsophisticated	1

essential	1
imperative	______
necessary	______
requisite	______
vital	______

flimsy	______
fragile	______
frail	______
puny	1
weak	______

[Nonnative speaker biodata form]

Could we have some information about yourself?

Nationality ______________________

How many years have you studied English in school and university?

____________ years

Tick one:

Sex: Male ______ Female ______

APPENDIX 4.1 Association Elicitation Instrument

Results from the following task will help us to better understand the knowledge native-speakers have about the prompt words. This improved understanding will facilitate the creation of better vocabulary tests for international learners of English.

Write the first three words you think of when you see each prompt word on the three lines provided.

1. abandon ____________ ____________ ____________
2. brood ____________ ____________ ____________
3. circulate ____________ ____________ ____________
4. convert ____________ ____________ ____________
5. dedicate ____________ ____________ ____________
6. illuminate ____________ ____________ ____________
7. launch ____________ ____________ ____________
8. massive ____________ ____________ ____________
9. plot ____________ ____________ ____________
10. peak ____________ ____________ ____________
11. rare ____________ ____________ ____________
12. spur ____________ ____________ ____________
13. subtle ____________ ____________ ____________
14. surging ____________ ____________ ____________
15. suspend ____________ ____________ ____________
16. trace ____________ ____________ ____________
17. trend ____________ ____________ ____________

APPENDIX 4.2 Lists of Associations and Frequencies

ABANDON

abort 2
alone 16
baby 4
betray
broken
bus
car
change
child 5
children 2
depart 2
deport
derelict 2
desert 28
desolate 3
destroy
disappear
discard 2
dismiss 3
disown
disregard
do
dog
don't want
drop
dump 2
empty 2
escape
exit 2
family
far
flee
forget 7
forsake 3
free 2
freedom
frightened
game
gay 2
give up 3
go 10
gone
helpless
home 2
homeless
hope 3
house
ignore
island 2
isolate 2
jettison
junkyard
kitten
leave 85
left 3
let go
lonely 5
loose 4
loss 2
lost 6
Moses
mother
neglect 7
off
orphan 2
park
quit 5
recklessness
redundant
refuge
reject 3
replace
retreat
rid
runaway
running
sacrifice
scared 2
senses
ship 4
stay 2
stranded
throw out
throw away
throw
trip
uncaring
vacate 2
vanish
waste
wild 4
wreck

agonize
alone 2
angry
animals 3
annoyed
baby 9
beard
birds 3
birth
book
brew
brood
chicken 20
chicks 3
child
children 12
cogitate
concern
contemplate 6
cry
cute
dank
depression
determined
duck 2
dwell 3
eggs 3
face
family 8
feel
ferment
food
fret 2
frown 2
gin
girl
grieve
group 2
hatch
hen 15
hesitate
ideas
knowledge
lay eggs
like
litter 2
lose
mardy 2
mare 2
maternal 3
meditate 2
melancholy
miserable
moan 2
mood 10
moody 17
mope 2
mother 6
motherly 2
mull over 3
nest 5
nurture
pigs
ponder 7
pregnant
quiet
reflect 3
rough
ruminate
sad 4
sadness
serious
sitting
solipsism
subjective
sulk 22
sulky 2
sullen 2
temper
think 33
thought
thoughtful
together
unforgiving
unhappy 6
upset 2
vipers
want
winge
women
wonder 2
worry 9
young

advertise
air 2
around 18
associate
avoid
blood 30
body
busy
bypass
change
circle 10
circular
circulation 2
communicate 2
community
continues
converse
deliver 4
despatch
disseminate
distribute 22
economics
entertain
flow 9
fluid
flyers
friends
gathering
give 10
go around
go round
go
gossip 3
group of people
gym
hand round
handout 3
heart 2
heat
inform
information 3
issue 2
leaflet 3
letter 4
locomotion
magazines
mail
meet
memo
mingle 8
minutes
mix 6
motion
move 19
move around
move round
movement 2
network 2
news 2
newspaper
notices
oxygen
paper 4
party
pass 12
pass around
pass round 2
pass on
passage
past
people
point
publish 2
pump
return
revolve 3
ring
rotate 6
round 14
rumour 3
run 2
send 3
send around
send round
share
show
socialize 5
socially
spin 2
spread 3
square 2
stir
students
surround 2
take
tell
thoughts
traffic 5
travel
triangle
water

adapt 5
alter 26
amend 3
bi-product
bishop
brainwash
bring
car 2
catholic 4
Catholicism
change 86
charts
Christian 6
church 4
compare
convertible
convince 3
create 3
criminals 3
cult 2
current
develop 4
difference
different 8
difficult
disciple 2
document
effect
electric
electricity
electronics
energy 2
enforce
enlighten
exchange 4
expert
figures
fix
follower
foreign
goal
God
ideas
improve 2
infect
influence 2
invert
involve
Islam
kick 2
loft
make
measures
ministry
money 5
move 4
new
newcomer 2
nun 3
old
opposite
overt
persuade 6
power 2
preach 2
prison
pushy
reaction
realign
rearrange 2
reassign
recruit
reeducate
religion 27
renovate
revamp
revelation
revert
rugby
score
show
Spanish
swap 2
switch 3
thinking 2
thought
transfer 2
transform 4
transition
turn 2
upgrade
win over
zealot

absolve
achieve
achievement
acknowledge 2
address
adore
affections
affirm
allocate
alone
ambitious
apply
assign 3
athlete
attribute 2
author
baby
baptism
baptize 3
belong
bless
body 2
book 12
Bryan Robson
building
by
care 2
certify
challenge 2
cherish
choose
Christian
church
college
commit 23
commitment 6
concentrate 2
consecrate
consign
constant
D.J.
dedicate
dedication
determined 2
devote 12
difficult
discipline 2
donate 2
duty
effort
energy
enthusiasm
faithful 2
follow
follower
football
for 2
fulfill
funeral
generous
gesture
girls
give 25
God-nun
hard worker
hardworking 2
heart 2
honor 11
involve 2
keen
learning
leave 3
life 5
literature
love 9
loyal 3
loyalty 2
martyr
memorial 3
message
mind
mission
monk 2
monument 3
Mother Teresa
motivation
name
no change
novel
nun-God
passion 2
pastime
personalize
pledge 2
poem 3
poet
priest
printer
promise 3
R.I.P.
record 2
religion
remember
request
reserve
resolve
respect
reward
romantic
sacrifice 2
selfless
sign
single-minded
sincere
solo
song 15
soul
special 3
specify
sport
stand by
stubborn
stuck with
student
study 2
succeed
success
supported
swear
testament
time 2
to
to someone
importa
token
tree
tribute 2
true
trust
trustworthy
value
willing
win
work hard
work 6

aglow
attention
Blackpool 10
bright 22
brighten 10
brightness 2
bulb 8
candle 2
Christmas 6
clarify 6
clarity 3
clear 5
clothes
colorful
colors 4
darkness
dazzling
dim
discover
display
educate
electric
electricity 2
enhance
enknowledge
enlighten 8
excited
explain
expose
expressive
fair
fire 2
flood
floodlight
florescent
glass
glow 4
golden
grasp
green
highlight 8
idea 4
inspire
knowledge 2
lamp 5
lazer
learn
light 78
lighten 8
light up 6
lit
make clear
manuscript
meaning
medieval
monks
moon
night
noticeable
open
physics
picture 2
point out
precise
present
question
raise
realization 2
red
reveal 3
rocket
room
see 3
shine 12
show 5
sign 2
stars
sun 3
switch
tacky
thought
torch
tree
ultraviolet
understand 5
unpack
wonder
write
yellow

air balloon
aircraft
airplane
astronaut
attack
away
begin 16
beginning of
blocks
boat 20
book 5
bring out
business
campaign 3
career
cast off
ceremony
champagne 2
commence 2
create
disembark 2
eject
engine
event
exciting
fashion show
fire 4
flight
float
fly 5
fuel
give
go
haul
idea 2
inauguration
initiate 2
introduce
invention
issue
jet
jetty 2
jump
kick
leave 2
let off
lift off 3
lift
lunch
manufacture
media
missile 5
new 5
open
out
pad 13
party 3
place
plane 2
produce
product 6
project 5
promote 2
promotion 2
propel 2
push 6
quay
release 8
rise
rocket 34
run
sail
sea 2
send 5
set off 2
set free
set afloat
ship 18
shoot
shuttle
sky 3
space 8
space shuttle
spacecraft
spaceship
spare
speed
speedboat
stars
start 21
strategy
take off 7
throw 9
thrust 2
up 2
vessel
water
whole-heartedness

attack 13
band
(Massive Attack)
big 62
bills
building
bulk
cast iron
cliff
colossal
country
dominating
enormous 29
extra large
extraordinary
frightening
giant 2
gigantic 6
ginormous 3
grand 3
great 9
haemorrhage
heavy 3
huge 76
Hugh
humungous 2
immense 3
impressive
large 49
little
mall
man
mass
mountain
omnipotent
outgrowth
planet
small
solid 2
star
strong
success
tall
universe
vast 6
weight
weighty
world

action
actors
administrate
adventure
against
allotment 7
area 3
arrange
assassin
assassinate
behind someone's
back
bomb
bonfire
book 4
building 2
calculate
chart
circumference
confer
connive 3
conspiracy 3
conspire 7
crime 4
cunning
data 2
decide
decision
dénouement
design
devious 2
devise 2
dig
discover
drama
establish
flowers
garden 14
graph 10
grass
grave 3
graveyard
ground 3
gunpowder 14
Guy Fawkes 5
hatch
havoc
idea
ink
intrigue 4
kill 2
land 20
lost 3
lot
Machiavelli
map 4
method
murder 3
narrative
novel 4
of TV show
organize 4
overthrow
pat
path
place 2
plan 39
play 4
point 3
position
prepare
research
revenge 3
scam
scheme 16
screen
secret 3
section
seek
set
sinister
site
sneaky
soil 2
sow
space
specific
spot 2
spy 2
square
stay
story 13
storyline 4
table
tale 2
thicken 4
think 2
treason 3
trees
trick
vegetable
weeds
write

ability
activity
air
apex 5
apogee
ash
best 9
cake mixture
cap 5
career
climax 5
crest 3
Derbyshire 2
District 14
excel 2
fares
fitness 2
flow
furthest
goal
graphs
hang gliding
hat
height 13
high 16
highest point
highest 2
hill 11
landscape
look
maximum 5
mountain 56
new height
nosey 2
on form
performance 3
phone
pinnacle 2
point 13
rate
rise
sharp
sky
snow 4
snowcap
spike
stare
success
summit 23
time 5
tip 6
top 52
trough 3
twin 4
ultimate 3
walking 2

abnormal
ancient
animal 10
antique 6
artefact
banana
birds 3
bloody 4
breed 2
coin
common
covert
creature
diamond
different 5
dinosaurs
disease 2
distinctive 2
elephant 2
endangered 9
exotic 2
expensive 7
experience
extinct 7
extraordinary
few 14
go
good men
individual
individuality
infrequent 2
inspiring
ivory
jewels 2
kill
limited 4
lonely
luxury
marbles
meat 5
medium 2
melt
never 3
none
obscure
occasional 3
occurrence
odd 3
old 2
one off
one 2
original
panda
person 2
plants
precious 10
priceless
rabbit 2
raw 2
red
scarce 17
single
small
solitary
solo
special 11
species 8
spider
stamps
steak 20
strange 3
tigers
uncommon 11
uncooked 2
undercooked
underdone
unique 12
unlikely 2
unpopular
unreal
unusual 39
valuable 6
weird
well-done
whole
wild
wildlife
worthy

action
animal
armour 2
army
arouse
Arsenal
athlete
battle
boot 12
break
burst
cheer 2
clap
cock
commence
compel
confidence
contour
cowboy 23
dig
drive on
drive 3
egg on
egg 2
emotion 2
encourage 36
energy
fast 2
fighting 2
flow
foot
football 6
force
game 2
geography
go 2
goad 4
ground
headland
heel 2
height
help 2
hill
hit
horse 36
ignore
immediate
impulsive
incentive 2
incite 2
inspire 4
instinctive
kick 3
knight 5
land
lead on
line
location
map
metal 5
moment 13
motivate 3
mountain
move 4
movement
now
on 2
onwards
pain
painted
persuade
prickly
prod 3
push 4
quick 2
quickly
reaction
reentrant
riding
river
rock 2
rowel
running
saddle 2
second
sharp 5
shiny
shoe 2
soldier
spark
spike 5
spontaneous
star
start
support
surges
throw
time 3
top
topography
Tottenham 8
trace
tree
urge 6
valley
weapon 2
western
Wild West
win

astute
aware
bland 2
blatant
blend
blunt 4
brick
brief
calm 2
canny
careful 9
caring 3
caringly
chess
child
clever 8
color 3
complex 2
conscious
considerate
covert 2
crafty
cunning 3
dangle
deep
delicate 10
denims
devious 3
difference
different
diplomatic
direct 2
discipline
discreet 7
dull
easy
effective
explicit
faint
feather
fine 3
garlic
gentle 24
hammer 2
hard
hidden 2
hint 14
humor
implicit
imply 2
important
indirect
indistinguishable
inference
insignificant
intent
intricate
irony
kind 3
knowing
light 4
lightly
little 2
low
low-key
man
manipulation
nice
not obvious 4
not noticeable
nuance 5
obvious 3
open 2
painting
pale
plain
plan
profound
quick
quiet 18
quietly
quite 2
relaxing
secret
sensitive 3
shades 3
shadowy
shrewd 2
shy 2
simple 2
sledge
sledgehammer
sleight
slight 7
sly
small 7
smart
smooth 3
smoothly
sneaky
so no-one notices
soft 11
sophisticated 3
subdued
sublime
tacit
tact
tactful 4
taste 2
thinking
thought 2
thoughtful 2
unassuming
underhand
underlying 2
understanding 2
understated 4
understatement
undertones
undiscernible
unnoticed 2
unobtrusively
unobvious 5
unspoken
variation
veiled
watchful
water
woman

accelerating
adrenaline
ahead 2
ascending
backwards
beginning
big
blood 3
boats
boost
break
breast
broken
bulge
bulk
bursting
charging 3
clattering
crashing 2
crowd 8
curve
damage
dashing
deluge
determining
diving
driving 2
eager
emotion
energy 3
engulfing
fast 7
flood 3
flow 9
foaming 2
forceful 4
forcing 5
forging
forward 15
galloping
game
glide
glue
going
growing 2
gushing 5
heaving 2
hospital
huge
hydroelectric ideas
illness
invigorating
jerking
lake
lifting
lunge
manoeuvre
mounting
move 3
movement 3
onslaught 2
overwhelming
parking
peaking
people
pouring
power 8
powerful 7
pressure 2
propelling
pulsing
pump 2
pushing 10
quick
quickly
racing 2
raging
rapid 2
riding
rising 4
river 6
rolling
rough
rugby league
running 3
rushing 13
sea 10
speed
spirit
sport
spreading
spurts 2
strength 4
surf 4
sweeping 2
swell 3
thought
thrusting 3
tide 7
torrent 5
uncontrollable
unstoppable
upwards 2
urging
water 17
waterfall 2
wave 30

abandon
abort
above
absent 2
age
air 4
alleviate
animation
back
balloon
banish
banned 3
belief
belt 3
book 2
branch
bridge 7
bungee 2
cancel 2
ceiling 5
circus
criminal
cut short
dangle 21
delay 9
detention
disbelief 2
discipline 2
dismiss
dispense
display
drop 3
elastic 3
end
enforce
exclude 2
expel 5
finish 3
freeze
gallows
get rid of
halt 2
hang 69
height
high 2
hold 8
holiday
hook
hover
in the air
incomplete
lapse
law
lean
leave 3
leg
let go
lift 2
light
linger
loose
lower
mid-air 3
motion
never
nothing
over
pause 3
person
picture
postpone 9
prison
punish 3
relieve
remove
rescheduled
restrain
revoke
rope 3
sack
save
school 9
sentence 5
space
springy
stocking 4
stop 14
suicide
support
suspenders 2
swing 3
temporary 2
tension
time 2
trapeze
unattached
underwear 3
wait 5
withheld
woman
work
years

a bit
amount
Batman
bit
bullet 2
call
chart
chemistry
circuit
clear
clue 2
compass
contents
cook
copy 39
detect 3
detective
draw 20
droplet
element 6
Essex girl
evidence 2
faint
family 2
find 34
find out 2
follow 20
fragment
give
glass
go
hint 3
history
ideas
investigate 5
kids
lead
line 3
little 2
look 3
look for 4
map 4
me
measure
metal
micron
miniature
minimum
minuscule
minute 3
missing
negligible
none
outline 6
overlap
overwrite
paper 26
parents
path
pattern
pencil 5
people
phone 2
picture 5
poison
presence
rare 2
reflect
relative
remains
research
residue
return
sample
scene
scent
score
search 7
see
seek 6
sign 2
sketch
slight 3
small 5
smidgen
speck
study
touch
tracing 2
track 8
track down
transparent
write

APPENDIX 4.3 List of Association Proportions for All Subjects

	ABANDON	BROOD	CIRCULATE	CONVERT	DEDICATE	ILLUMINATE	LAUNCH	MASSIVE	PLOT
1	1.00	0.55	0.76	0.84	0.51	0.90	0.83	0.75	0.80
2	0.75	0.93	0.63	0.84	0.24	0.75	0.27	0.61	0.16
3	0.05	0.31	0.52	0.67	0.38	0.75	0.09	0.63	0.23
4	0.84	0.87	0.35	0.83	0.73	0.82	0.65	1.00	0.77
5	0.72	0.31	0.48	1.00	0.62	0.83	0.65	0.49	0.24
6	0.83	0.09	0.35	0.86	0.40	0.93	0.16	1.00	0.25
7	0.02	0.29	0.45	0.68	0.24	0.77	0.96	1.00	0.25
8	0.71	0.49	0.39	0.64	0.78	0.15	0.61	0.89	0.29
9	0.81	0.07	0.20	0.86	0.19	0.72	0.28	0.79	0.13
10	0.68	0.65	0.49	0.82	0.48	0.98	0.48	0.79	0.12
11	0.05	0.31	0.38	0.68	0.46	0.79	0.79	0.74	0.29
12	0.88	0.39	0.58	0.65	0.6	0.74	0.25	0.89	0.72
13	0.93	0.61	0.75	0.69	0.22	0.09	0.55	0.47	0.27
14	0.09	0.51	0.14	0.04	0.17	0.74	0.27	0.49	0.24
15	0.70	0.21	0.62	0.82	0.41	0.72	0.55	0.74	0.31
16	0.71	0.52	0.14	0.83	0.60	0.79	0.52	0.89	0.59
17	0.04	0.60	0.70	0.67	0.05	0.06	0.87	0.46	0.49
18	0.68	0.11	0.04	0.68	0.57	0.75	0.57	0.89	0.75
19	0.91	1.00	0.59	0.83	0.81	0.79	0.36	0.63	0.88
20	0.93	0.53	0.58	0.67	0.49	0.81	0.73	0.77	0.41
21	0.68	0.24	0.17	0.68	0.40	0.78	0.40	0.43	0.32
22	0.68	0.23	0.35	1.00	0.46	0.77	0.95	0.28	0.04
23	0.91	0.64	0.46	0.82	0.13	0.78	0.77	1.00	0.63
24	0.07	0.29	0.34	0.82	0.30	0.79	0.59	0.74	0.77
25	0.72	0.68	0.39	0.03	0.06	0.21	0.76	0.42	1.00
26	0.74	0.35	0.49	0.82	0.41	0.86	0.37	0.59	0.75
27	0.9	0.87	0.38	0.83	0.62	0.36	0.81	0.63	0.28
28	0.89	0.56	0.30	0.81	0.06	0.79	0.51	0.82	0.71
29	0.72	0.41	0.41	0.82	0.22	0.09	0.31	0.68	0.57
30	0.67	0.15	0.14	0.03	0.21	0.81	0.52	0.52	0.24
31	0.69	0.31	0.15	0.82	0.41	1.00	0.69	1.00	0.17
32	0.80	0.55	0.15	0.65	0.57	0.96	0.97	0.82	0.75
33	0.93	0.69	0.52	0.67	0.78	0.84	0.56	0.77	0.40
34	0.73	0.79	0.76	0.65	0.41	0.15	0.27	0.48	0.55
35	0.70	0.75	0.46	0.85	0.27	0.21	0.39	0.89	0.48
36	0.93	0.33	0.13	0.68	0.81	1.00	0.41	0.82	0.57
37	0.90	0.67	0.17	0.83	0.60	0.86	0.48	0.89	0.36
38	0.74	0.69	0.35	0.66	0.06	0.79	0.49	0.42	0.53
39	0.88	0.27	0.34	0.82	0.59	0.21	0.35	0.45	0.80
40	0.25	0.27	0.62	0.83	0.56	1.00	0.39	0.68	0.60
41	0.71	0.65	0.04	0.82	0.6	0.78	0.37	0.42	0.28
42	0.68	0.05	0.06	0.63	0.41	0.09	0.24	0.27	0.04

	ABANDON	BROOD	CIRCULATE	CONVERT	DEDICATE	ILLUMINATE	LAUNCH	MASSIVE	PLOT
43	0.93	0.44	0.42	0.82	0.4	0.76	0.19	0.61	0.56
44	0.74	0.07	0.63	0.66	0.43	0.73	0.64	0.03	0.49
45	0.71	0.28	0.70	0.82	0.24	0.72	0.16	0.60	0.24
46	0.69	0.55	0.49	0.81	0.56	1.00	0.45	1.00	0.80
47	0.68	0.23	0.85	0.24	0.16	0.78	0.39	0.44	0.16
48	0.74	0.16	0.54	0.03	0.06	0.80	0.28	0.75	0.29
49	0.69	0.71	0.48	0.83	0.22	0.90	0.64	1.00	0.49
50	0.67	0.77	0.46	0.04	0.08	0.11	0.19	0.47	0.77
51	0.72	0.27	0.35	0.67	0.14	0.90	0.13	0.60	0.55
52	0.67	0.27	0.38	0.67	0.14	0.94	0.64	1.00	0.36
53	0.17	0.55	0.51	0.04	0.21	0.8	0.59	0.57	0.19
54	0.71	0.28	0.51	0.22	0.06	0.06	0.52	0.49	0.12
55	0.70	0.45	0.49	0.22	0.22	0.71	0.48	0.89	0.16
56	0.08	0.51	0.46	0.21	0.43	0.76	0.48	0.08	0.13
57	0.88	0.68	0.37	0.66	0.62	0.22	0.31	0.77	0.81
58	0.90	0.55	0.66	0.66	0.78	0.80	0.51	0.75	0.59
59	0.74	0.59	0.76	0.81	0.59	0.71	0.29	0.82	0.32
60	0.81	0.36	0.48	0.67	0.25	0.34	0.59	0.37	0.72
61	0.71	0.75	0.58	0.65	0.17	0.83	0.48	0.61	0.15
62	0.68	0.40	0.48	0.82	0.29	0.98	0.31	0.63	0.75
63	0.68	0.47	0.52	0.63	0.22	0.79	0.55	0.81	0.25
64	0.68	0.63	0.48	0.04	0.33	0.74	0.33	0.02	0.49
65	0.36	0.33	0.34	0.84	0.56	0.72	0.33	0.89	0.28
66	0.79	0.29	0.35	0.69	0.06	0.21	0.04	0.68	0.09
67	0.88	0.05	0.32	0.66	0.10	0.82	0.56	1.00	0.59
68	0.26	0.93	0.44	0.63	0.19	0.13	0.43	0.61	0.60
69	0.91	0.33	0.80	0.84	0.08	0.94	0.23	1.00	0.59
70	0.74	0.76	0.11	0.64	0.30	0.71	0.40	0.75	0.83
71	0.9	0.79	0.46	0.83	0.22	0.21	0.89	1.00	0.84
72	0.68	0.13	0.24	0.63	0.06	0.75	0.04	0.75	0.57
73	0.71	0.59	0.11	0.27	0.44	0.76	0.53	0.77	0.28
74	0.95	0.39	0.31	0.83	0.44	0.84	0.55	0.74	0.21
75	0.95	0.23	0.31	0.86	0.40	0.76	0.60	0.74	0.97
76	0.74	0.63	0.70	0.82	0.44	0.84	0.13	1.00	0.61
77	0.79	0.29	0.21	0.68	0.49	0.77	0.04	0.79	0.07
78	0.80	0.16	0.04	0.23	0.40	0.21	0.09	0.74	0.75
79	0.69	0.57	0.31	0.63	0.22	0.78	0.56	1.00	0.09
80	0.80	0.44	0.45	0.68	0.52	0.79	0.28	0.75	0.37
81	0.86	0.32	0.37	0.65	0.08	0.86	0.71	0.89	0.73
82	0.91	0.13	0.44	0.66	0.11	0.14	0.36	0.11	0.11
83	0.19	0.45	0.56	0.82	0.08	0.81	0.09	1.00	0.55
84	0.75	0.31	0.51	0.68	0.32	0.72	0.32	0.74	0.61
85	0.19	0.07	0.35	0.83	0.06	0.72	0.07	0.75	0.64

	ABANDON	BROOD	CIRCULATE	CONVERT	DEDICATE	ILLUMINATE	LAUNCH	MASSIVE	PLOT
86	0.90	0.83	0.30	0.68	0.44	0.71	0.15	1.00	0.83
87	0.73	0.25	0.69	0.71	0.46	0.11	0.57	1.00	0.19
88	0.67	0.47	0.42	0.83	0.06	0.9	0.16	0.42	0.19
89	0.68	0.16	0.35	0.83	0.65	0.78	0.53	1.00	0.09
90	0.73	0.79	0.58	0.87	0.78	0.82	0.16	0.79	0.08
91	0.91	0.15	0.65	0.83	0.33	0.79	0.35	1.00	0.41
92	0.91	0.36	0.58	0.65	0.27	0.15	0.65	0.68	0.41
93	0.75	0.47	0.37	0.66	0.22	0.78	0.73	0.27	0.29
94	0.89	0.24	0.15	0.82	0.13	0.15	0.07	1.00	0.23
95	0.07	0.53	0.55	0.02	0.40	0.90	0.31	1.00	0.09
96	0.86	0.56	0.51	0.82	0.21	0.94	0.56	1.00	0.73
97	0.79	0.40	0.87	0.67	0.44	0.93	0.6	0.61	0.31
98	0.68	0.59	0.70	0.85	0.44	0.90	0.59	0.63	0.40
99	0.05	0.05	0.48	0.65	0.06	0.81	0.57	0.34	0.24
100	0.67	0.87	0.55	0.81	0.62	0.76	1.00	1.00	0.92
AVG	0.68	0.45	0.43	0.67	0.36	0.67	0.45	0.71	0.44
STD	0.25	0.23	0.19	0.24	0.21	0.28	0.23	0.24	0.26
MIN	0.02	0.05	0.04	0.02	0.05	0.06	0.04	0.02	0.04
MAX	1.00	1.00	0.87	1.00	0.81	1.00	1.00	1.00	1.00

	PEAK	RARE	SPUR	SUBTLE	SURGING	SUSPEND	TRACE	TREND	AVG PROP
1	0.53	0.26	1.00	0.46	0.56	0.94	0.56	0.11	0.67
2	0.50	0.41	0.67	0.14	0.21	0.91	0.52	0.91	0.56
3	0.47	0.55	0.53	0.54	0.58	0.06	0.39	0.76	0.44
4	0.92	0.43	0.80	0.66	0.24	0.70	0.63	0.90	0.71
5	0.71	0.13	0.63	0.09	0.39	0.80	0.14	0.89	0.54
6	0.62	0.55	0.77	0.29	0.60	0.71	0.63	0.72	0.57
7	0.86	0.62	0.43	0.54	0.39	0.41	0.62	0.83	0.55
8	0.61	0.8	0.03	0.54	0.18	0.86	0.09	0.22	0.49
9	0.60	0.28	0.19	0.36	0.68	0.81	0.46	0.75	0.48
10	0.07	0.36	0.34	0.55	0.66	0.73	0.68	0.80	0.57
11	0.54	0.18	0.42	0.52	0.65	0.04	0.37	0.05	0.43
12	0.70	0.46	1.00	0.07	0.66	0.88	0.23	0.78	0.62
13	0.63	0.55	0.77	0.25	0.52	0.72	0.34	0.76	0.54
14	0.53	0.79	0.44	0.11	0.61	0.71	1.00	0.73	0.45
15	0.19	0.29	0.44	0.16	0.35	0.95	0.54	0.75	0.51
16	0.51	0.57	0.44	0.66	0.13	0.15	0.77	0.75	0.56
17	0.47	0.38	0.33	0.32	0.44	0.78	0.68	0.04	0.43
18	0.55	0.74	0.82	0.36	0.08	0.70	0.22	0.89	0.55
19	0.47	0.55	0.43	0.64	0.35	0.94	0.43	0.77	0.67
20	0.54	0.55	0.77	0.64	0.31	1.00	0.81	0.77	0.67
21	0.49	0.75	0.48	0.13	0.15	0.95	0.58	0.82	0.50
22	0.44	0.21	0.42	0.14	0.39	0.8	0.76	0.12	0.47
23	0.56	0.20	1.00	0.07	0.56	0.32	0.44	0.76	0.59
24	0.48	0.33	0.45	0.36	0.34	0.28	0.36	0.75	0.47
25	0.66	0.68	0.53	0.48	0.10	0.05	0.44	0.19	0.44
26	0.93	0.43	0.42	0.64	0.15	0.70	0.47	0.89	0.59
27	0.44	0.8	0.75	0.30	0.13	0.77	0.47	0.72	0.59
28	0.61	0.26	0.45	0.63	0.24	0.91	0.42	0.89	0.58
29	0.19	0.74	0.40	0.23	0.39	0.76	0.03	0.92	0.46
30	0.59	0.38	0.27	0.11	0.08	0.75	0.10	0.11	0.33
31	0.93	0.47	0.44	0.88	0.37	0.77	0.66	0.75	0.62
32	0.59	0.86	0.67	0.16	0.39	0.89	0.80	0.89	0.68
33	0.61	0.46	0.44	0.20	0.24	0.95	0.77	0.94	0.63
34	0.50	0.86	1.00	0.20	0.77	0.75	0.46	0.03	0.55
35	0.53	0.88	0.13	0.09	0.61	0.18	0.76	0.80	0.53
36	0.86	0.66	0.79	0.27	0.26	0.70	0.56	0.93	0.63
37	0.67	0.76	0.08	0.23	0.52	0.84	0.41	0.93	0.60
38	0.51	0.89	0.82	0.11	0.61	0.18	0.42	0.07	0.49
39	0.28	0.79	0.43	0.38	0.26	0.84	0.03	0.81	0.50
40	1.00	0.30	0.41	0.14	0.35	0.88	0.13	0.76	0.54
41	1.00	0.17	0.40	0.50	0.55	0.70	0.09	0.72	0.52
42	0.32	0.75	0.08	0.21	0.35	0.70	0.06	0.21	0.30
43	0.67	0.86	0.44	0.39	0.42	0.89	0.47	0.26	0.56
44	0.12	0.32	0.64	0.13	0.45	0.33	0.33	0.88	0.45
45	0.45	0.22	0.08	0.09	0.29	0.12	0.27	0.84	0.40

	PEAK	RARE	SPUR	SUBTLE	SURGING	SUSPEND	TRACE	TREND	AVG PROP
46	0.06	0.74	0.07	0.95	0.31	0.88	0.61	0.22	0.60
47	0.83	0.11	0.41	0.8	0.69	0.15	0.36	0.77	0.47
48	0.54	0.28	0.53	0.05	0.66	0.68	0.59	0.77	0.46
49	0.42	0.79	0.57	0.88	0.63	0.74	0.3	0.76	0.65
50	0.45	0.21	0.43	0.63	0.37	0.88	0.61	0.76	0.47
51	0.12	0.72	0.03	0.48	0.42	0.11	0.76	0.75	0.45
52	0.69	0.43	0.06	0.18	0.06	0.16	0.75	0.75	0.48
53	0.46	0.53	0.31	0.59	0.23	0.70	0.82	0.74	0.47
54	0.27	0.11	0.22	0.50	0.35	0.70	0.86	0.78	0.40
55	0.47	0.29	0.38	0.07	0.24	0.06	0.67	0.72	0.42
56	0.15	0.41	0.45	0.52	0.21	0.04	0.37	0.07	0.32
57	0.40	0.76	0.45	0.64	0.32	0.77	0.47	0.76	0.58
58	0.55	0.67	0.41	0.32	0.27	0.83	0.65	0.90	0.64
59	0.59	0.70	0.49	0.73	0.55	0.04	0.66	0.82	0.60
60	0.54	0.28	0.67	0.80	0.85	0.68	0.61	0.89	0.58
61	0.65	0.88	0.43	0.23	0.44	0.15	0.46	0.74	0.52
62	0.99	0.18	0.53	0.46	0.34	0.73	0.68	0.77	0.59
63	0.89	0.32	0.34	0.30	0.52	0.09	0.56	0.74	0.51
64	0.50	0.79	0.52	0.16	0.21	0.71	0.54	0.82	0.47
65	0.85	0.41	0.15	0.61	0.63	0.68	0.05	0.22	0.49
66	0.51	0.12	0.03	0.13	0.13	0.03	0.03	0.77	0.29
67	0.62	0.43	0.08	0.46	0.23	0.70	0.06	0.89	0.50
68	0.56	0.75	0.89	0.18	0.39	0.23	0.94	0.81	0.53
69	0.85	0.32	0.82	0.45	0.05	0.84	0.51	0.89	0.61
70	0.42	0.79	0.53	0.25	0.66	0.88	0.59	0.73	0.59
71	0.56	0.86	0.63	0.52	0.58	0.69	0.56	0.72	0.66
72	0.69	0.72	0.47	0.20	0.10	0.11	0.05	0.77	0.41
73	0.62	0.28	0.26	0.21	0.08	0.83	0.61	0.77	0.48
74	0.15	0.25	0.13	0.16	0.53	0.73	0.52	0.73	0.50
75	0.85	0.24	0.52	0.41	0.63	0.68	0.26	0.18	0.56
76	0.56	0.67	0.29	0.16	0.40	0.70	0.67	0.83	0.60
77	0.50	0.08	0.18	0.71	0.06	0.11	0.28	0.15	0.36
78	0.56	0.30	0.20	0.54	0.27	0.18	0.22	0.75	0.38
79	0.91	0.21	0.41	0.2	0.19	0.80	0.46	0.85	0.52
80	0.84	0.28	0.20	0.16	0.37	0.71	0.62	0.73	0.53
81	0.57	0.16	0.17	0.11	0.42	0.17	0.68	1.00	0.51
82	0.85	0.42	0.34	0.84	0.61	0.75	0.27	0.82	0.46
83	0.54	0.61	0.26	0.16	0.32	0.73	0.23	0.93	0.49
84	0.62	0.16	0.22	0.43	0.81	0.67	0.48	0.03	0.49
85	0.53	0.28	0.03	0.29	0.21	0.17	0.09	0.74	0.35
86	0.54	0.63	0.07	0.50	0.27	0.23	0.61	0.18	0.52
87	0.85	0.21	0.08	0.46	0.71	0.13	0.09	0.89	0.48
88	0.95	0.96	0.03	0.39	0.4	0.89	0.22	0.79	0.52
89	0.51	0.22	0.46	0.2	0.74	0.73	0.64	0.77	0.55
90	0.65	0.13	0.05	0.64	0.06	0.88	0.65	0.85	0.56

	PEAK	RARE	SPUR	SUBTLE	SURGING	SUSPEND	TRACE	TREND	AVG PROP
91	0.93	0.30	0.45	0.25	0.53	0.68	0.71	0.87	0.60
92	0.63	0.66	0.22	0.23	0.56	0.72	0.64	0.08	0.49
93	0.52	0.25	0.18	0.23	0.45	0.04	0.52	0.75	0.44
94	0.60	0.21	0.06	0.07	0.79	0.88	0.77	0.73	0.46
95	0.60	0.05	0.03	0.09	0.15	0.15	0.03	0.72	0.34
96	0.53	0.54	0.05	0.38	0.76	0.25	0.45	0.89	0.59
97	1.00	0.33	0.45	0.66	0.76	0.68	0.47	0.93	0.64
98	0.65	0.70	0.67	0.14	0.74	0.73	0.64	0.79	0.64
99	0.50	0.04	0.44	0.14	0.31	0.71	0.28	0.03	0.34
100	0.61	0.50	0.52	0.25	0.55	0.83	0.33	0.93	0.69
AVG	0.58	0.47	0.41	0.36	0.41	0.59	0.47	0.67	0.52
STD	0.21	0.25	0.25	0.23	0.21	0.31	0.23	0.28	0.09
MIN	0.06	0.04	0.03	0.05	0.05	0.03	0.03	0.03	0.29
MAX	1.00	0.96	1.00	0.95	0.85	1.00	1.00	1.00	0.71

Appendix 5.1 Collocates for Target Words taken from the COBUILD Bank of England Corpus. Span: ±4 words

ABANDON   MUTUAL INFORMATION (MI) STATISTIC

700
absurd
accord
acquisition
adopt
advised
advised
agreeing
all
alternative
altogether
altogether
altogether
ambitions
ambitions
ambitious
americans
amid
anc
any
apartheid
approach
aspiration
attempt
attempts
attempts
authorities
babies
barbican
base
basic
become
belief
beloved
blair
Bosnia
boycott
calling
cameras
cannot
career
carefree
careless
cars
catholics
caused
central
centralised
centres
certainly
chancellor
chase
cherished
cherished
chirac
choice
circumstance
clinton
commission
commitment
commitment
commitment
communism
communist
compelled
complete
completely
completely
concentrate
condoms
conservation
conservative
controls
controls
convention
convince
convinced
could
crimes
crippling
crossing
Cuba
decide
decided
decided
decided
decides
deciding
decision
decision
decision
demand
demanded
demanding
demands
democratic
Dionysian
disposal
doctors
does
downing
driven
droves
dump
effectively
efforts
efforts
either
elements
else
embrace
embrace
encourage
encourage
encouraged
entirely
entirely
equal
equidistant
establish
eventually
eventually
everything
failure
fans
farmers
farming
farms
fate
favour
favour
fears
fellow
finally
fishermen
flung
force
force
forced
forced
forced
forcing
forcing
formally
fort
franc
gay
gorbachev
government
government
government
greek
greenpeace
habit
habits
hands
happy
her
his
hobby
homes
hope
hopes
howard
idea
ideology
ideology
immediately
immediately
imperial
imposed
incentive
independent
inflation
insistence
instead
intend
introduce
investors
IRA
israel
its
join
joyous
Korea
Kuwait
Kuwait
labor
labour
led
led
Lenin
Leninism
Leninism
Leninism
liner
Lithuania
loath
loth
loved

Maastricht
mad
mainframe
Marxism
Marxism
Marxism
Marxism
mass
may
mid
might
ministers
modernise
modernise
moments
monetary
motorists
must
mutual
mutuality
nato
never
not
notion
notion
notions
now
nuclear
nuclear
nuclear
nurses
obliged
observation
often
operators
option
or
ordered
organisers
orthodox
our
ourselves
parties
passport
patients
patterns
pensions
persuade
persuade
persuaded
persuaded
persuaded
persuading
persuading
plan
plans
plans
plans
plea
pledge
poised
policies
policy
politics
prejudice
prepared
preparing
pressure
pretence
pretension
principle
principle
principles
principles
principles
privatisation
privatise
probably
profitability
program
programme
programmes
programmes
project
project
prompted
proposing
pursuit
Queensland
quest
readiness
ready
reckless
reckless
reconciliatio
n
reforms
refusal
refused
refuses
refusing
reluctance
reluctant
republican
require
research
retreat
return
rough
route
rsc
rule
rule
ruling
rush
saddam
sanctions
save
scheme
schemes
search
secure
sending
Seoul
serbs
sexual
shell
ship
ship
should
simply
sinking
socialism
socialist
soldiers
solution
soon
status
status
strict
structure
struggle
such
suggesting
suggestion
sweeper
temporarily
temporarily
temptation
tempted
testing
than
their
them
threaten
threatened
threatening
to
total
totally
tradition
tradition
tradition
traditional
traditional
traditional
traditions
unilateral
unilateral
unilateral
unless
unwilling
unwilling
urge
urged
urged
urges
urging
urging
vessel
violence
voted
voters
warned
watching
we
weapons
wild
will
willing
willingness
witchcraft
with
would
ye
your
yourself
zone

ABANDON T-SCORE (T) STATISTIC

<b>
<p>
a
a
about
after
against
agreed
all
also
altogether
altogether
altogether
Americans
an
anc
and
and
and
any
approach
are
armed
as
at
attempt
attempts
attempts
barbican
base
be
because
because
become
been
began
being
bill
Britain
Britain
but
called
calling
campaign
can
cannot
car
career
career
carefree
careless
cars
caused
central
choice
claim
commitment
commitment
communism
communist
complete
completely
completely
concentrate
condoms
controls
convention
could
could
decide
decided
decided
decision
decision
decision
demand
demands
did
did
did
Dionysian
does
dream
effectively
efforts
efforts
either
entirely
entirely
equal
equidistant
even
eventually
everything
failure
fate
favour
favour
fears
fight
finally
fishermen
for
for
for
force
force
forced
forced
forced
forcing
forcing
formally
fort
France
game
game
gay
gleeful
going
Gorbachev
government
government
government
government
had
had
happy
has
has
have
have
he
he
her
him
him
his
his
homes
hope
hope
hopes
hopes
idea
if
if
if
if
if
immediately
immediately
in
independent
instead
intend
introduce
IRA
is
it
its
its
its
its
join
joyous
just
labour
labour
leave
led
led
life
long
mad
made
many
mark
Marxism
Marxism
Marxism
may
may
may
might
might
might
most
move
mr
mr
mr
must
must
mutual
mutuality
my
NATO
need
never
no
no
normal
north
not
not
not
not
notion
notions
now
nuclear
nuclear
nuclear
nuclear
nuclear

of
of
often
old
one
option
or
or
or
order
ordered
orthodox
our
ourselves
own
party
party
party
party
party
patterns
people
people
persuade
persuade
persuaded
persuaded
persuading
persuading
plan
plan
plan
plan
planned
plans
plans
plans
plans
policies
policy
policy
policy
practice
practice
prepared
preparing
present
president
president
pressure
pressure
principle
principles
principles
process
program
programme
programme
programmes
project
project
proposed
punky
rather
ready
reason
reckless
reconciliation
refused
reluctant
research
return
role
role
rsc
rule
rule
run
s
Saddam
said
search
set
sexual
she
ship
ship
should
should
simply
six
so
socialism
solution
some
Soviet
status
strategy
structure
struggle
studies
style
such
such
system
system
t
temporarily
testing
tests
than
that
that
the
the
the
their
their
their
them
them
these
they
they
they
they
they
this
those
threatened
to
to
too
totally
tour
tradition
traditional
traditional
traditional
unilateral
unlikely
unwilling
urged
urged
urging
us
us
violence
voters
want
wanted
war
was
we
we
we
we
weapons
were
what
when
which
which
whole
why
why
wild
will
will
will
willing
witchcraft
with
with
with
without
work
would
would
would
would
ye
your
yourself

BROOD MI

</h>
</h>
</h>
<h>
<h>
<p>
<p>
<p>
<p>
1
10
12
12
14
150
2
20
7
a
a
a
about
about
after
after
after
after
after
again
also
also
an
an
an
an
and
and
and
and
and
and
and
another
any
are
are
are
are
are
are
are
are
as
as
as
as
at
at
at
at
August
be
be
bean
been
been
been
been
been
big
bitch
bitches
but
but
but
but
but
but
but
by
by
by
by
by
by
by
can
can
children
children
children
class
could
country
current
don
down
down
end
European
ever
first
five
followed
for
for
for
for
for
four
four
four
four
French
from
from
from
from
from
from
gathers
German
got
got
growing
grown
had
had
had
had
had
had
happened
has
has
has
hatchery
have
have
have
have
have
have
having
he
he
he
he
he
hen
her
her
her
her
her
his
his
his
his
his
his
i
i
i
i
in
is
is
is
is
it
it
it
it
it
it
it
its
its
its
its
its
its
large
last
last
last
left
like
like
mare
mares
may
members
month
more
more
more
more
more
mother
my
never
new
new
new
new
new
new
next
no
no
no
no
no
not

not
not
not
now
now
now
of
of
of
of
off
off
on
on
on
on
on
on
one
one
one
one
one
only
only
only
or
or
or
or
or
other
our
over
over
over
over
own
past
past
pouch
raise
raising
rather
raven
ravens
rear
rearing
reveals
s
s
s
s
s
s
second
see
seven
she
she
she
she
she
she
since
since
since
sit
so
so
so
some
some
star
still
still
still
stock
t
t
t
t
talented
than
than
than
than
than
that
that
the
the
the
the
the
their
their
their
their
their
them
them
then
then
there
they
they
they
they
they
they
this
this
this
three
time
time
time
tin
to
to
to
too
took
two
two
two
two
under
unemployment
united
until
up
upon
was
was
were
were
were
what
what
when
when
where
where
where
which
which
which
while
while
while
while
while
who
who
whole
will
wings
with
with
with
with
with
with
with
won
world
would
would
would
would
year
year
year
year
year
years
yet
you
you
you
you
young

BROOD T

<p>
14ins
90
a
a
about
about
again
almost
always
among
amongst
an
an
and
and
and
and
and
and
another
archetypes
are
at
auspicious
averaging
balham
bean
began
began
bespoke
best
bitch
bitches
bizarrely
boa
bobbed
boldest
born
bought
braden
bring
brood
burnish
busloads
but
but
buttered
bykov
called
can
can
cannibalism
chalk
chambers
chance
chastise
children
children
children
chubby
class
class
cobh
constrictor
construction
continued
copious
country
cows
crimes
damagingly
daniela
day
derain
desolation
devil
dispensing
doesn
don
don
dowdy
dropouts
ducklings
ducklings
ducklings
ebullient
eggs
endlessly
enervating
enough
entire
even
even
ever
ever
exhorts
expanding
expressing
fangio
fannie
father
feathering
federation
feed
feline
ferencvaro
few
financiers
first
five
flashed
fledged
fledging
fledglings
foals
food
for
four
four
French
friend
fusco
gama
gangly
gathers
generously
ghislaine
gladness
good
gothick
greedily
groupies
growing
had
had
happy
happy
hatchery
having
he
he
he
he
healthy
hellish
help
hen
her
her
her
her
her
his
his
his
his
home
home
hone
horde
how
however
hughes
hunt
immediately
imperturbab
le
including
infestation
inheritor
injustices
insufferable
is
is
it
its
jackson
jealously
Juan
Judy
kids
kids
large
least
left
left
let
let
like
like
little
lombards
long
lowther
lusty
made
mamma
mare
mares
mate
may
may
merciful
mistrustful
mope
more
mother
Mrs
much
mummy
my
nephews
nesbit
nest
never

new
new
next
no
noisy
not
not
nourished
nursemaids
nurture
nuzzling
of
of
of
off
off
off
on
once
one
only
orphaned
our
outre
over
over
overblown
own
own
paddock
paddocks
paradox
parents
parishes
penny
players
poignantly
pouch
princeton
pronounced
protrude
proxy
purposeful
question
rabid
raise
raising
raising
rampart
rather
rather
raven
ravens
rear
rearing
reining
reissued
repentant
replenished
rest
river
roberta
rumbold
ruminate
runt
s
s
sandstorm
scowl
se21
seawater
second
second
seducing
set
she
she
she
she
shielding
shivers
shortfalls
shyer
signed
since
sit
sits
sits
smoulder
snot
snotty
so
so
some
sons
sows
stoats
stock
stuffing
stylists
subtext
successful
t
t
talented
teach
tend
than
the
the
their
their
their
their
them
them
these
these
thespian
they
they
they
things
this
time
time
tinies
tits
to
too
too
took
translator
tribe
tricolour
trysts
uglier
under
undergrowth
unruly
untamed
upon
urchins
usually
ve
ve
vent
very
vipers
walk
want
wayans
weeks
weldon
were
what
what
when
where
while
while
whole
will
wined
wings
with
with
with
woman
woman
women
won
world
would
would
would
wrestles
wrote
years
young
young
young
young
young

CIRCULATE MI

</h>
<h>
<p>
1
11
12
3
5
8
a
a
able
about
about
about
after
after
again
again
air
air
air
air
air
air
all
all
all
all
allow
allow
allowed
allowing
allows
allows
already
also
also
also
also
also
America
America
among
among
an
an
an
and
and
and
and
and
any
any
any
any
are
are
are
are
are
around
around
around
as
as
as
as
as
as
ask
at
at
at
at
at
base
be
been
been
been
before
began
begin
beginning
begun
between
black
blood
blood
blood
body
body
but
but
but
but
by
by
by
by
can
can
can
cannot
cash
cells
cells
coins
community
companies
continue
continued
cook
cooking
cooking
could
currencies
currency
defence
did
did
did
do
do
dollars
don
down
during
during
end
energy
European
evenly
ever
expected
fact
family
fan
far
faster
faster
few
figures
financial
for
for
for
for
four
freely
freely
French
French
fresh
friends
from
from
from
from
from
from
got
had
had
had
had
had
had
has
has
has
has
has
has
have
have
have
have
he
he
heat
help
his
his
his
if
if
in
in
in
in
increased
industry
inside
into
is
is
is
it
it
it
its
its
its
its
its
japan
jobs
let
letter
like
longer
main
market

may
members
members
members
members
mesh
minister
months
months
months
more
more
new
new
new
new
new
new
news
no
no
no
not
not
not
not
now
now
of
on
on
or
or
or
or
or
or
or
other
other
other
other
others
over
over
over
own
own
palm
perforation
petition
petitions
plans
press
price
rarely
received
reference
regularly
report
reports
room
round
rumour
rumours
rumours
rumours
rumours
s
s
same
seven
she
should
so
so
start
started
starting
still
still
still
stores
stories
stories
stories
t
t
terms
than
than
than
that
that
that
that
that
that
that
the
the
the
the
the
the
their
their
them
them
them
them
then
then
there
these
they
they
they
they
they
they
they
this
this
through
throughout
throughout
time
to
to
to
top
two
two
until
until
used
was
was
water
water
we
we
we
we
we
were
were
were
were
were
west
when
where
which
which
which
which
which
which
while
who
widely
will
will
will
will
will
with
with
within
within
would
year
year
years
you
you
you
you
you
your
your
your

10
11
1272
32m
65p
able
about
about
about
adrenaline
again
air
air
air
air
air
air
air
all
all
allegation
allow
allow
allowed
allowing
allows
allows
along
already
also
also
among
among
among
and
and
and
and
and
annexe
antifreeze
appropriate
are
are
around
around
around
around
around
as
as
as
as
ask
base
before
began
begin
beginning
begun
beltway
beneath
bequests
best
between
between
bid
bigwigs
black
black
blacklist
blood
blood
blood
bloodstream
body
body
books
both
boys
breakneck
bullring
can
can
cannot
capital
ceaselessly
cells
cells
chi
city
clockwise
coins
cold
colleagues
community
community
companies
condor
continue
continued
continuous
cooking
cooking
cool
copies
copies
could
coupons
cryptograph
currencies
currency
currents
days
decided
decomposition
defence
demimonde
department
description
did
did
did
digestion
discreetly
dissidents
dissipate
do
do
documents
does
dollar
don
ducted
during
early
efta
electrons
encroach
endlessly
energies
enough
entire
eurocopy
Europe
European
evenly
expected
failed
falcon
fama
family
family
fan
far
faster
favelas
fictions
filtered
fluids
for
foul
four
freely
freely
freely
freely
fresh
fresher
friends
gas
globe
government
had
he
health
heat
help
help
helped
his
house
however
ibj
ideas
ideas
if
in
individual
influenza
information
inside
inside
intended
it
its
jamming
jong
juneau
keep
keep
killing
kintyre
leathery
legends
leicester
let
let
lets
letter
letter
like
little
local
longer
main

Margaret
market
material
mean
meanings
media
meet
meetings
memoranda
memos
mesh
minister
mm
molecules
money
money
monies
monkton
more
much
must
need
new
new
news
night
nightspots
nih
no
not
not
now
nuances
nutrients
off
old
or
or
osservator
other
others
over
own
own
oxygen
palm
pamphlet
perforation
persuaded
petit
petition
petitions
pil
plans
pleasanter
possible

preprints
press
press
pricking
prisoners
prolongs
pump
re
reach
recessed
recognising
reference
references
regularly
religiously
report
reports
resolution
revellers
reynolds
Richards
roast
room
room
round
round
rumour
rumours
rumours
rumours
rumours
s
seconds
see
she
should
should
should
side
silently
singapore
slanderous
so
solution
start
started
starting
still
still
stories
stories
stories
story
story
strains
Sunday

surely
surreptitious
ly
system
t
takeover
taxis
that
that
that
the
the
the
the
the
them
them
them
them
then
then
these
they
those
though
through
throughout
throughout
to
tootsies
top
triglyceride
turkmenistan
turn
unable
unchecked
unhindered
unread
untrue
use
used
waiters
water
water
were
west
wheezing
when
where
where
which
which
which
which
white
widely

will
will
will
will
with
within
within
without
word
work
would
wyld
year
yesterday
your
your
your

CONVERT MI

11
17th
500
ability
able
acid
acid
acid
actually
add
African
aim
alcohol
allow
allowing
allowing
allows
already
am
Andy
another
anybody
anyone
arms
attempt
attempted
attempting
audiences
automaticall
y
bacteria
bank
banks
banks
banks
barn
barn
bath
became
become
billion
body
body
boxes
brother
Buddhism
Buddhism
build
building
building
building
buy
buy
campaign
can
cannot
carbon
carbon
cardinals
carport
cash
cash
cash
catholic
Catholicism
Catholicism
cause
cause
cells
chances
chances
chances
chances
chances
cheaper
cheaply
Christianity
Christianity
Christianity
Christians
Christians
civilian
civilian
clergymen
co2
commercial
common
companies
complete
convention
convertible
corner
corners
cost
could
countries
credit
cross
currency
data
debentures
debt
debt
debt
debt
defence
defence
devices
difficult
digital
digital
digital
dioxide
dollars
dollars
double
drive
eager
early
easily
easily
efforts
electrical
electricity
enable
energy
energy
energy
energy
energy
enthusiast
enthusiast
enzymes
equity
equity
equity
eugene
Europe
everybody
example
existing
existing
factories
failed
failing
failure
faith
famous
fervour
fifth
flats
food
food
format
freely
fund
garage
garrick
gas
glucose
goal
goals
half
harding
harmless
heathen
help
helped
helps
herself
Hindu
his
historic
historic
holders
holders
holders
hoping
house
idea
inability
inches
Indians
industries
instance
instant
instead
intend
intends
intention
into
into
into
into
Islam
it
items
its
itself
Jewish
Jewish
Jews
joys
Judaism
keenan
labour
late
latest
latest
latest
law
league
loans
loft
longer
manage

managed
marinos
marriage
material
may
merge
methane
methane
might
military
missionaries
moment
monoxide
Morgan
mortgage
Muslim
must
mutual
nathan
necessarily
necessary
needs
nitrogen
normal
not
notes
notes
numbers
nurses
obliged
offices
opportunity
opportunity
opportunity
optical
option
ordinary
others
our
oxygen
park
part
penalty
penalty
penalty
penalty
permission
person
plan
plan
plan
planning
plans
plans
playing
plc
points
points
poor
port
possession
possible
potential
pound
pounds
preferred
presidential
pressure
priest
process
process
programs
prospective
protein
quickly
quite
raw
reactions
reality
reality
recent
reduction
refusing
registered
religious
reluctant
reluctant
required
road
Roman
Rome
room
ross
rugby
rugby
rushed
says
seat
seeking
sentiment
set
set
shares
shares
shift
signal
signals
signals
signals
size
socialism
societies
societies
societies
societies
society
society
software
solar
solid
solution
some
sought
south
space
speech
stage
starch
status
status
stepped
stock
struggled
subsequent
sugars
sunlight
symbols
systems
takes
their
them
themselves
then
these
those
those
though
to
to
top
total
tribe
tried
tries
trying
trying
turner
tv
unable
union
urge
usable
useful
variety
varsity
visible
vitamin
vouchers
wallaby
waste
wastes
weapons
were
which
which
whom
will
wished
wishes
woman
would
would
your
zeal
zeal

CONVERT T

11
17th
5
500
a
a
a
a
a
a
ability
able
actually
add
African
agreed
aim
alcohol
all
allowing
allowing
allows
also
American
an
an
an
and
andy
another
another
any
anyone
are
attempt
attempting
automatical
ly
back
bacteria
bank
bank
banks
banks
banks
bath
be
became
became
become
become
been
body
brother
building
building
building
but
can
can
cannot
carbon
carbon
carport
cash
cash
cash
cash
catholic
Catholicism
Catholicism
cause
cause
cells
chances
chances
chances
chances
cheaply
christiani
Christianity
Christianity
Christians
Christians
church
civilian
civilian
clergymen
co2
commercial
common
common
companies
company
convention
convertible
corner
corners
could
could
cross
currency
debt
debt
debt
decided
decision
designed
difficult
digital
digital
dioxide
do
dollars
dollars
don
double
early
easily
easy
economy
effort
efforts
electrical
electricity
enable
energy
energy
energy
energy
enough
enthusiast
equity
equity
existing
existing
factories
failed
failing
failure
faith
famous
fervour
flats
food
food
forced
format
former
free
freely
from
from
from
fund
garrick
gas
gas
glucose
go
goal
has
has
have
he
he
heathen
help
helped
helps
her
him
his
his
holders
holders
house
how
idea
idea
if
inability
inches
industries
instance
instant
interest
into
into
into
into
is
is
Islam
it
items
its
its
its
itself
Jewish
Jewish
Jews
joys
Judaism
keenan
late
latest
latest
law
league
light
likely
loans
loft
made

managed
many
marinos
may
me
merge
methane
methane
might
military
military
million
missionaries
monoxide
more
most
most
Muslim
must
mutual
my
nation
necessarily
necessary
need
needed
new
not
not
notes
notes
numbers
nurses
of
old
one
only
open
option
or
or
ordinary
others
our
Own
paid
part
penalty
penalty
people
per
permission
plan
plan
planning
plans
plans
plc
points
points
potential
pound
pounds
pressure
priest
production
prospective
protein
public
public
quite
reality
reality
recent
reduction
registered
religious
reluctant
required
right
Roman
room
ross
rugby
rugby
s
says
seeking
set
set
shares
shares
shares
shift
signal
signals
signals
signals
size
societies
societies
societies
society
society
solid
solution
some
sought
sound
south
Soviet
space
speech
status
status
stepped
stock
stock
sunlight
symbols
t
takes
than
that
the
their
their
them
themselves
then
then
these
they
they
they
they
third
this
those
those
though
though
to
to
to
to
to
top
total
tried
try
trying
unable
union
union
usable
use
use
used
useful
varsity
visible
vouchers
wallaby
want
wanted
wants
was
waste
were
which
which
which
which
who
who
will
will
world
would
would
would
your
zeal

DEDICATE MI

<p>
<p>
<p>
10
12
1989
1990
30
90
a
a
a
a
after
after
again
agreed
all
all
all
all
all
all
also
am
America
an
and
and
and
and
and
another
any
any
anyone
are
are
are
are
are
as
as
at
band
bank
be
be
be
be
been
being
best
between
book
both
business
but
but
but
but
but
by
by
by
by
called
came
can
can
Canadian
capital
cause
chose
could
could
country
country
current
d
d
day
decided
deficit
did
do
during
each
energies
energy
energy
everything
expected
eyes
fact
family
first
for
for
for
formally
four
fourth
from
from
going
government
grand
gratefully
great
had
had
had
had
half
has
has
have
have
have
he
he
he
help
her
her
her
here
herself
higher
him
himself
his
his
his
his
holocaust
i
i
i
i
if
imports
in
in
industry
inside
instead
is
is
is
it
it
it
it
it
its
its
its
its
itself
July
July
just
just
kids
Korea
largest
last
last
later
less
life
life
life
like
like
lives
lot
love
m
made
making
man
may
memory
modest
money
money
monument
more
more
more
more
more
more
much
much
much
must
my
my
my
my
myself
nearly
need
new
new
next
next
no
no
no
no
noble

not
not
November
November
now
now
now
obviously
of
of
of
on
on
on
one
one
one
one
one
one
only
only
only
or
or
or
or
or
other
other
others
our
our
our
our
ourselves
papers
part
part
period
permission
prepared
producing
production
re
re
regularly
resolved
respectful
rest
s
s
s
s
s
said

said
same
second
she
should
since
since
so
some
some
some
something
song
song
space
speculativ
e
states
still
still
t
than
than
that
that
that
that
that
that
the
the
the
the
their
their
their
their
them
them
themselves
then
then
then
then
they
they
they
they
they
this
this
this
this
this
this
three

time
time
time
time
time
time
to
to
to
to
to
to
training
two
united
us
victory
volume
volume
want
wanted
wants
was
was
was
we
we
we
we
week
week
well
were
what
when
when
where
which
which
which
which
while
who
whole
why
will
will
will
win
with
with
with
work
work
world
world

world
would
would
would
year
years
years
you
you
you
you
you
you
you
your
your
yourself

DEDICATE T

1
11
1863
30
40
640
90
94
a
able
adherents
affectionatel
y
airports
Algeria
all
all
all
America
an
and
and
and
any
anymore
anyone
as
balmain
band
baptism
be
beethoven
best
book
bronson
but
cairn
camellia
can
can
candidates
cathedrals
cause
cemetery
centers
chalet
chanting
chapel
children
chomsky
chose
Christ
cleansed
clinton
clinton
come
commemorate
confess
consecrate
consecrate
could
couldn
crass
cromer
d
d
dam
day
decided
decision
dedicates
deeds
derring
did
didn
digestive
djs
dreaming
dreaming
efforts
elton
encore
energies
energy
enough
ensuring
entire
entirely
everything
exploring
eyes
fact
fans
farris
final
finally
fisher
fittipaldi
flint
football
formally
freddy
fulfilling
full
furnishing
gaining
get
gettysburg
god
going
grammy
grand
gratefully
had
have
he
he
he
headland
help
helpers
her
her
her
here
here
herself
him
him
Himalayas
himself
his
his
historic
holocaust
holocaust
hopes
howes
i
i
I
if
illustrious
imports
infiltration
infinitive
insecure
instead
integrating
it
its
itself
Jehovah
jogger
just
just
kids
Korea
land
last
later
later
life
life
like
liturgical
lives
looking
lot
love
love
lovingly
luckily
m
malik
man
many
margery
may
me
mean
memorial
memory
merits
Methodist
miron
money
monument
more
more
morisot
most
motherland
much
museum
must
my
my
my
my
my
myself
mysticism
nearly
need
new
new
next
nino
noble
nominal
notebook
now
now
numerous
obliging

obviously
oddball
ode
of
officers
one
oneself
operations
other
our
our
our
our
our
ourselves
overturning
palates
papers
park
parliament
part
part
perfecting
permissible
permission
planned
plans
plaque
pleasure
pounds
preaching
prepared
princeton
promised
prosecuting
proudly
Queensland
raoul
re
re
really
rebuilding
record
records
regularly
rehabilitate
remembrance
renew
renounce
renovation
repertory
required
reservation
resolved
resources
resources
respectful
rest
rnli
rsv
sabbatical
said
said
schools
seater
secretariat
several
she
should
sihanouk
simply
sincerity
site
sonata
song
song
space
speculative
st
states
strapping
subcontract
sufficient
superstars
swear
sweat
t
tabernacle
talents
talisman
tearful
telescope
teleworker
than
that
that
the
the
their
their
them
them
themselves
then
then
they
they
thing
this
this
this
thom
thor
those
time
time
time
to
to
today
totally
township
training
translate
tumour
tune
two
undetermined
united
upbringing
uplifting
us
us
victory
victory
victory
violations
volume
voted
walkway
wallenberg
wand
wanda
want
wanted
wants
warhol
was
Washington
we
we
we
weekdays
weekend
were
who
whole
why
will
will
win
wish
without
work
work
would
would
would
wyman
you
you
you
you
your
your
your
your
yourself
yourself
yourselves
zachary
zoning

</h>
<h>
<p>
<p>
<p>
<p>
<p>
<p>
12
1988
2
2
38
6
a
a
a
a
a
able
about
after
after
after
again
again
all
allows
alone
also
also
although
always
America
an
an
an
an
an
an
an
and
and
and
and
and
and
and
another
any
any
are
are
are
are
are
area
area
areas
as
as
as
as
aspects
at
at
at
at
at
be
be
been
been
beginning
being
best
between
big
border
bright
Britain
but
but
but
but
but
but
but
by
by
by
by
by
by
can
can
can
car
character
clearly
closely
company
corners
could
country
current
dark
dark
dark
designed
did
did
different
different
do
does
down
down
down
drama
dropped
each
each
economic
end
end
enough
enough
enough
Europe
even
expected
experience
figures
figures
five
for
for
from
from
from
government
had
had
had
had
has
has
has
has
has
have
have
have
he
help
helped
helps
her
her
him
his
his
his
his
history
how
however
ideally
if
in
in
in
indicators
is
is
is
issue
it
it
it
its
its
its
its
its
just
just
lamps
laser
last
left
life
life
life
life
light
light
light
lighting
lights
little
London
many
Marxism
matters
may
mere
might
mine
month
more
more
more
more
much

much
much
nature
nature
new
next
night
no
no
no
not
not
now
now
now
of
of
of
of
of
on
on
one
one
one
one
one
ones
only
only
only
only
Open
Open
or
or
or
or
or
or
or
or
order
other
other
other
our
over
over
own
parts
past
past
past
period
price
probably
problem
production
rather
rather
real
rendering
s
s
s
same
scene
science
seven
sharp
should
show
show
sky
so
so
some
some
star
still
still
stories
story
subject
such
table
targets
terms
than
than
than
that
that
that
the
the
the
their
their
them
them
then
then
then
there
these
these
they
they
they
they
they
they
this
this
this
this
though
three
three
through
time
to
to
to
to
trade
try
trying
two
two
two
until
used
vision
was
was
way
we
we
week
were
were
were
west
what
when
when
where
where
where
where
which
which
which
which
which
which
while
while
while
who
who
will
with
with
with
with
women
work
world
world
worldwide
would
would
years
years
yet
yet
you
you
your

ILLUMINATE T

a
ability
able
about
acknowledge
adroit
aims
alfresco
all
allows
already
also
although
an
an
and
and
and
and
antonioni
any
anything
apercus
arcane
are
area
area
areas
around
as
as
aspects
aspects
attachment
attempt
attempts
ayer
b
basic
Bavarian
beams
being
being
between
beyond
bits
blackout
bogart
booth
both
breakout
bright
brightly
burglars
but
but
can
can
can
candelabra
candlelight
caprichos
casinos
central
chandelier
character
chime
clarify
cloud
coleman
collection
coloured
competence
configuration
conjure
continue
continues
cornell
corner
corners
could
court
crevices
cultural
curving
dark
dark
dark
deposition
designed
determined
did
different
different
dinah
dioramas
directional
dispassion
displays
do
documents
does
don
done
drama
dropped
duller
each
each
effort
elder
endlessly
engineers
English
enlighten
enough
enough
enough
enough
entertain
even
even
evenly
every
experience
extent
eyre
facades
face
facets
fail
fails
failures
fascinating
features
fiberoptic
fie
field
fireworks
flash
flashlight
floodlight
fully
gawain
generis
gita
go
goal
goya
gratefully
guests
guttering
hallways
harmonies
headlights
help
help
helped
helping
helps
helps
hendricks
his
his
his
historical
history
history
hologram
holzmair
honig
how
how
how
human
human
human
hysterics
ideally
if
illustrate
in
information
inspire
instead
instructive
interiors
ipanema
is
issue
it
it
its
Jungian
just
key
kitchen
knowledge
lamp
lamps
lanterns
laser
lasers
last
life
life
life
life
light
light
light
light
lighting
lights
lights
lights

little
lives
magnify
magnifying
make
mankind
many
Marxism
may
meander
meaning
meant
message
might
momentarily
monkish
moral
motivation
much
narrative
nature
nature
next
night
nihilist
not
nullify
oates
of
of
of
of
one
one
opportunity
or
or
or
order
organ
other
other
our
our
our
ous
over
own
own
particular
past
path
philanthropis
t
pimm
place
position
power
prawer
predicament
problem
processes
pubescent
purifying
qualities
questions
quinton
radar
rather
rather
readily
reality
redefinition
relationship
rendering
scene
science
scriptural
scudding
seems
serves
shafts
shahada
should
simply
skein
sky
sky
skylights
slavish
small
so
so
solos
some
sought
spiritual
spotlights
spotty
squall
stairways
steps
still
story
story
storytelling
stubbornness
style
subject
such
such
sui
sunlight
suppose
swag
tacit
target
targets
tensions
text
text
than
than
that
that
that
the
the
their
them
theories
these
these
they
they
thinkers
thinking
this
those
threadbare
three
throng
through
to
trollope
truth
truths
try
trying
underlining
unfreeze
us
use
used
using
vignettes
virgo
virtuosity
vision
watford
way
ways
ways
ways
weakness
what
which
which
which
which
which
while
why
will
with
within
without
women
women
work
world
would
yet
your

LAUNCH MI

1911
1999
1999
1999
1999
200m
21st
36th
45pm
50m
9th
accidental
announce
announces
ariane
ariane
ariane
armani
assault
assault
assault
assaults
astronauts
astronauts
astronauts
atlantis
atlantis
atlantis
attack
attack
attack
attacks
attacks
attend
attended
attended
attended
attending
awaited
awaited
bi
bi
bid
bid
bid
blistering
blitz
blitz
blockbuster
bombardment
bombers
bombing
branson
brochure
browser
bskyb
bskyb
button
campaign
campaign
campaign
campaign
campaigns
canaveral
canaveral
canaveral
capability
carlton
celebrate
celebrating
champagne
channels
charter
cobe
codes
codes
coincide
coincide
coincided
coincided
coincides
coincides
coinciding
Colombia
Colombia
Colombia
Colombia
commando
commenting
con
concerted
consortium
countdown
countdown
countdown
counter
counterbid
cruise
crusade
crusade
customs
date
dates
delay
delay
delayed
delayed
delaying
delays
delays
digital
disastrous
discovery
disrupt
eagerly
emptive
emptive
endeavour
enquiry
ers
expendable
expenses
exploded
fightback
Flemings
fm
forays
formal
formally
foxtel
foxtel
foxtel
fragrance
fragrance
frontal
gala
gala
Galileo
Galileo
Galileo
gearing
gearing
glitzy
granada
guiana
hailed
hangar
helped
helpline
hopes
hosting
howe
hype
hype
hyped
IBM
imminent
impending
initiative
initiative
inquiry
instants
intending
intends
investigate
jetty
kerry
kirklees
kourou
lavish
lavish
lbn
licences
lifeboat
lottery
mace
mace
manifesto
manifesto
mcmahon
md
md
mg
microsoft
microsoft
microsoft
microsoft
missile
missile
missile
missile
missiles
missiles
missiles
mode
moored
motor
multimillion
multiple
NASA
NASA
NASA
NASA
NASA
NASA
NASA
nintendo
nintendo
observator
offensive
offensive
offensive
offensives
offensives
official
Officially

optus
orbit
orbit
overshadow
packer
pad
pads
pamphlet
paperback
paramedic
patrol
payloads
pep
pep
pep
perfume
persil
persuading
philips
planned
planned
planning
plans
platform
platform
platform
poised
polar
postpone
postponed
postponed
postponement
powerboats
prayers
pre
preemptive
preparation
preparation
preparatory
prepares
preparing
preparing
prestigious
prevention
probe
probe
proceeded
product
prompted
prompted
pronged
prospectus
prospectus
prospectus
publicise
publicised
publicity
raids
raids
ramp
ready
regulator
reschedule
retaliation
retaliator
retaliator
retrospect
reusable
rocket
rocket
rocket
rocket
rocket
rocket
rocket
rocket
rockets
rockets
rockets
rockets
rockets
rockets
russians
saarc
satellite
satellite
satellite
satellite
satellite
satellites
satellites
satellites
satellites
satellites
Saturn
scathing
scheduled
scheduled
scratchcard
scud
scud
scud
scud
scud
scuds
secretive
sees
sega
sensors
shute
shuttle
shuttle
shuttle
shuttle
shuttle
shuttle
since
site
sites
slipper
smoothly
sony
space
spacecraft
spacecraft
spacecraft
spacecraft
spacecraft
speakers
speaking
speaking
splash
springboard
sputnik
sram
standby
strikes
stringent
studded
studded
submarine
subscriber
successful
suitably
tabloid
takeover
teamed
tec
techno
telescope
telescope
tentatively
terrestrial
terrorist
thomson
threatens
tigers
timetable
timing
tirade
Titan
Titan
titanic
titled
tonne
torpedo
tradepoint
transfers
tube
tubes
turin
unofficial
unveiled
unveiled
upcoming
vehicle
vehicles
venture
ventures
vip
vowed
wheelhouse
windows
windows

LAUNCH T

15
95
95
a
a
able
about
action
after
after
after
against
against
agency
ago
aid
air
album
also
an
an
announced
announced
announced
another
any
appeal
appeal
April
are
ariane
army
assault
assault
assault
at
at
at
at
attack
attack
attack
attacks
attacks
attempt
attempt
attend
attended
attended
Australian
base
based
be
before
before
bid
bid
bid
book
book
book
book
Brisbane
Brisbane
Brisbane
business
campaign
campaign
campaign
campaign
can
canaveral
card
card
career
career
career
cd
celebrate
challenge
channel
coincide
Colombia
company
conference
conference
control
costs
could
countdown
counter
currency
date
days
decided
decision
decision
disastrous
discovery
drive
drive
due
during
emptive
expected
February
festival
first
first
first
following
following
for
for
for
for
for
for
formal
Friday
full
fund
gearing
government
government
ground
has
help
helped
helped
helped
her
himself
his
hopes
in
initiative
initiative
inquiry
inquiry
intended
intends
into
into
into
into
investigation
investigation
investment
Iraq
is
issue
its
its
its
its
its
itself
January
last
last
last
latest
latest
league
league
legal
London
London
lottery
lottery
magazine
major
march
market
market
massive
may
media
microsoft
microsoft
might
military
million
million
missile
missile
missile
missile
missiles
missiles
month
month
month
months
motor
mr
multiple
NASA
NASA
national
national
national
new
new
new
new
news
newspaper
next
next
next
next
nuclear
nuclear
of

offensive
offensive
offensive
offer
official
officially
on
on
on
operation
own
pad
pads
party
party
party
party
perfect
plan
planned
planned
planning
plans
plans
plans
platform
point
poised
police
policy
pound
pound
pre
pre
prepared
prepares
preparing
press
press
press
price
product
product
product
programme
project
prompted
Queensland
range
range
range
re
ready
recent
response
rocket
rocket
rocket
rockets
s
s
s
s
said
said
satellite
satellite
satellite
satellite
satellite
satellites
satellites
satellites
Saturday
saw
scale
scheduled
scheduled
scheme
scheme
scud
sea
season
sees
September
series
series
service
service
set
should
shuttle
shuttle
shuttle
shuttle
shuttle
since
since
since
since
site
sites
soon
space
space
space
space
space
speaking
speaking
strike
strike
strikes
successful
system
systems
takeover
television
the
the
the
the
their
themselves
this
this
tigers
time
today
today
today
today
tomorrow
tomorrow
tomorrow
tried
tv
two
two
UK
used
used
using
vehicle
vehicles
venture
version
version
wants
war
was
was
week
week
week
week
week
weeks
whether
which
which
will
will
will
will
windows
windows
windows
with
within
would
year
year
year
years
yesterday
yesterday
yesterday
yesterday

1560
260
2bn
a
aboriginal
absorb
abuses
accompanied
accused
adds
advocated
aerospace
aftermath
aid
airlift
airlift
airpower
amid
amounts
announcing
antibiotic
artillery
astronomer
astrophysics
attract
attracted
attracting
attracts
awarding
bailout
bailout
banking
bases
bcci
belated
bestowed
bhp
bhutto
bhutto
bike
bil
blitz
blitz
blowout
bombardment
bombardment
bombers
bonanza
bongo
borrowings
boulders
bracing
brass
breakwater
bridget
Bristol
buildup
canary
canary
cater
caused
caused
causing
charging
cheating
chechnya
chemotherapy
clot
clues
codenamed
coke
collapsed
collapsed
columns
complexes
conducting
consolidate
constitute
contain
contemplate
cope
creates
creating
crews
crippling
culminating
dams
dangerously
database
dating
deficits
deflection
deliveries
depopulating
despite
despite
destruction
detectives
deterrence
detonated
devastated
died
dislocation
disobedience
disruption
dockland
dogged
dominated
dominates
donors
doses
downsizing
drag
drought
dumped
embarked
embarking
emptive
endured
engage
engaged
engaging
ensured
ensuring
equally
eruption
Europeans
evasion
expenditure
fab
facility
feared
ferris
fireball
firepower
firestone
flintoff
forests
forthe
foundation
fraud
fraud
fraud
fraud
galaxies
galaxy
galaxy
gearing
generate
gravitation
gravity
groaning
haemorrhaging
haemorrhaging
haemorrhaging
halo
halo
Hamburg
handles
handshake
hike
hike
hikes
hindemith
hormone
humanitarian
humanitarian
hunts
hurricane
hydro
imperturbabl
e
incompetence
incur
indoor
inflict
inflicted
inflow
influx
infusion
infusions
ing
initiated
inject
injection
injection
injection
injections
injuries
instantly
insulin
interacting
introducing
involving
irrigated
Italians
jams
Jews
justify
kendall
kurds
lamps
landowners
landslides
launch
launched
launched
launches
launching
laundering
laundering
layoffs
lease
manhunt
massive
massive

mergers
midst
migrations
migrations
misery
momentum
monastery
motors
mounted
mounted
mounting
negotiated
negotiator
nigel
Nottingham
objects
offs
offs
orbit
orchestration
organising
outflow
outpouring
outraged
overcapacity
overdose
overdose
overdoses
overhang
oversee
overstate
oversupply
oxygen
oz
pascal
payouts
payouts
pentagon
perched
perks
perks
perpetrate
pesticide
petty
pillars
placing
plantation
plunge
poorer
popularity
portishead
portishead
portishead
portishead
potentially
Prague
predicting
preparing
preponderance
preventing
printing
privatised
producing
prompt
prompting
proposing
protests
proton
provoke
pullout
pumping
quitter
radiation
raids
railways
rains
raves
rebuilding
rebuilding
received
reclamation
recovering
redevelopment
redundancies
reissue
represents
require
requiring
reserves
responding
restructuring
resulted
resulting
retaliate
retaliation
retaliation
retaliator
retrenchment
reunification
reward
rigging
rigging
rigorous
rises
robbed
rocked
rotating
salyut
sculpture
sharif
signature
soca
sovereign
spark
sparked
sparked
sparked
sparking
spill
spill
spree
spree
staging
staging
starvation
steroids
struggling
stunning
subsidised
subsidy
sudden
suffer
suffered
suffered
suffers
suffers
supernovae
surpluses
sustained
swindle
tailback
tailbacks
technological
tens
tensions
terrorists
threatens
tome
topped
trafford
trains
trident
triggered
triggered
tropical
truly
turbine
uncovered
uncovered
undergo
undergo
undergoing
undergone
undergone
underground
undertaking
undertaking
underway
underway
underwent
unexpected
unfinished
unfinished
unification
unveil
unwittingly
upheaval
verge
violations
volcanoes
weakly
wharf
wharf
wimps
wimps

MASSIVE T

<p>
000
a
absolutely
accompanied
accused
after
against
against
against
aid
aid
aid
aid
air
air
also
American
amid
amount
amounts
and
and
announced
announced
another
anti
are
army
attack
attack
attracted
bank
bases
be
because
been
been
been
began
begin
behind
being
being
big
big
bil
bill
billion
billion
black
boost
boost
budget
build
building
built
built
by
by
by
by
called
campaign
campaign
capital
capital
cash
cause
caused
caused
caused
causing
cent
changes
committee
companies
confidence
confidence
construction
contain
cope
cost
could
could
country
country
country
create
creating
crisis
currently
cut
cuts
damage
debt
debt
debt
debts
deficit
deficit
deficit
demonstration
despite
despite
despite
destruction
development
died
died
dollar
dominated
doses
due
during
economic
economy
economy
education
effort
effort
effort
embarked
energy
equally
exercise
expansion
explosion
face
face
faced
faces
facing
financial
financial
following
food
food
for
for
for
force
force
forces
forces
foreign
fraud
fraud
fraud
fraud
from
from
from
funds
funds
given
going
got
government
government
ground
had
had
haemorrhagin
g
has
has
has
have
have
have
head
heart
his
hit
hit
hole
holes
huge
huge
huge
in
in
in
in
in
increase
increases
increases
influx
information
injection
injuries
interacting
internationa
l
intervention
into
into
investment
investment
investment
involved
involving
Iraq
Iraq
Iraq
is
is
it
it
its
its
its
Japanese
Jews

job
Kuwait
last
launch
launched
launched
launched
lb
lead
lead
led
left
less
likely
lion
loss
losses
losses
massive
massive
military
military
military
military
million
money
more
most
need
need
needed
new
night
number
numbers
of
of
of
off
oil
on
on
on
operation
operation
operation
over
oz
part
past
pay
pay
per
plan
planned
planning
planning
plans
plantation
police
potential
potential
pound
pound
power
presence
presence
prices
problem
problems
produced
producing
profits
program
programme
programme
project
project
projects
provide
public
public
public
queens 1 and
radiation
received
received
reduce
reduce
refugees
require
resources
responsibly
restructuring
result
result
resulted
results
retaliation
rigging
rises
s
scale
security
set
social
Soviet
Soviet
Soviet
Soviet
sparked
spending
spending
spill
stars
strike
strikes
strikes
subsidies
such
such
sudden
suffered
suffered
sun
support
support
support
task
tax
technologies
thanks
the
the
the
the
their
there
there
there
these
they
this
this
those
threatens
through
through
times
to
to
tour
truly
turned
two
under
under
unemployment
united
unless
up
up
using
ve
vote
was
was
weakly
were
where
which
which
which
which
which
will
with
with
without
without
would
would
would
yesterday
yesterday
yesterday

PEAK MI

000ft
000ft
07
1m
10p
106
110
125
184p
1942
1950s
1953
1960s
1962
1970s
1980s
1982
1986
1987
1987
1988
1988
1989
1989
2000
20p
25p
3000
30s
36p
36p
36p
39p
39p
39p
48p
48p
48p
48p
49p
49p
5p
5p
5p
5p
6m
6pm
700
80s
80s
850
87
8p
abilities
absolute
acceleration
achievement
adolescence
adventures
airports
altitude
approaching
approaching
Aquarius
Arizona
Arizona
arnold
arousal
ascent
at
Australia
barrels
bbc1
below
below
below
births
blanc
blanc
blanc
bloom
boom
boom
boom
brightness
broad
buckingham
burnout
burton
canal
canal
capped
career
careers
cavern
central
chardonnay
cheap
cheaper
circulation
climb
climbed
climbed
climbed
climbing
climbing
cloud
clouds
coincided
communities
commuters
commuters
compared
compares
concentratio
n
condition
conservative
considerable
coronet
cortina
Craig
crashed
creative
crowded
curve
cycle
cycle
cycle
cycle
cycles
cyclical
Dakota
darien
declined
deforestatio
n
delays
delays
demand
derbyshire
derbyshire
derbyshire
derbyshire
derbyshire
descended
diameter
diameter
dimensions
district
downs
during
earning
Easter
efficiency
efficiency
eighties
eighty
employer
ensures
entering
epidemic
euphoria
Everest
experience
fallen
fallen
fame
fares
fares
fares
fertility
fever
fisheries
fitness
fitness
flinders
flowering
frenzied
frequency
freshness
Fridays
ft
ft
ft
geyser
glory
graph
guadalupe
guild
guild
halfway
halved
heights
heights
highest
Himalayan
Himalayas
historic
hit
hits
hits
hitting
hitting
hour
hourly
immigration
index
index
inflation
intra
its
itv
itv
itv
k2

k2
k2
k2
kings
kitt
km
knowledgeable
1990s
19th
lassen
lendl
li
load
load
lone
luminosity
lunchtime
mans
matlock
mcdermott
measured
meteor
meter
metre
mid
min
minute
Mondays
monthly
moorland
mount
mount
mountain
mountain
Muslim
navigation
near
nearing
NHS
nightly
nr
observatory
observatory
occur
occurrence
occurs
off
offending
offending
passed
passengers
peak
peak
peak
peak
peak
peak
peaks
pence
pennines
percentage
perfection
perfection
perform
performance
performers
performing
periods
periods
periods
periods
periods
Pg
pierce
pike
pikes
plummeted
poet
pollen
popularity
popularity
popularity
popularity
powers
powers
practice
preston
previous
pricing
pricing
pricing
privatised
produces
profession
projected
pronounced
railway
reach
reach
reached
reached
reaches
reaches
reaching
reaching
reductions
reductions
register
rental
represent
representive
represents
resonance
risen
roads
roofed
routes
rush
sauvignon
season
seasonal
seventies
sheen
shepherd
shingle
shower
ski
slot
slot
slots
slowed
slumped
snow
snowy
soared
soared
spurt
steadily
superannuati
on
supernova
surged
Swiss
tallest
teens
telescope
tends
tent
ticket
tonnes
tons
topping
torque
touched
tourism
towering
traffic
traffic
trains
trains
trains
trains
tram
travelling
trillion
trough
trough
tumbled
twenties
twin
twins
victoria
viewers
viewing
viewing
viewing
viewing
volcanic
volcano
volcano
vulture
waves
wheeler
widow
widows
windy
worksop

PEAK T

</h>
<b>
<h>
<p>
<p>
000
000
000
000
1
10
11
15
1970s
1980s
1980s
1986
1987
1987
1988
1988
1988
1989
1989
1989
1990
1990
1991
1994
1994
2
2
2
20
3
3
3
30
30
36p
36p
36p
40
45
48p
48p
48p
5
5
6
6
7
7
7
8
8
9
9
a
a
about
absolute
after
age
ago
all
almost
and
and
another
approaching
are
Arizona
around
around
at
at
at
at
at
at
at
August
August
August
Australia
average
be
below
below
below
between
billion
body
body
boom
broad
but
calls
calls
career
career
career
careers
cavern
cent
cent
cent
central
chance
cheap
climb
climbed
climbed
close
close
communities
compared
condition
condition
conservation
coronet
cost
creative
current
cycle
cycle
cyclical
demand
derbyshire
derbyshire
derbyshire
derbyshire
district
dollar
dollar
dollar
down
down
during
during
during
during
during
early
early
early
electricity
end
especially
expected
experience
experience
fall
fallen
fallen
fame
fell
fitness
fitness
flow
foot
for
forest
form
form
free
from
from
from
game
group
groups
guild
half
has
have
her
high
highest
Himalayan
his
his
historic
hit
hit
hit
hits
hitting
holiday
hour
hour
hour
hour
hours
in
in
in
in
in
index
inflation
is
is
its
its
its
its
itv
itv
July
June
June
kitt
last
last

late
late
less
level
levels
load
local
lone
low
lower
m
march
market
market
may
mid
mid
mid
mid
million
million
minute
minute
minute
months
morning
mount
mountain
mountain
Muslim
my
nation
national
national
near
new
nightly
observatory
of
off
off
off
on
p
park
park
particular
passed
passengers
past
peak
peak
peak
peak
peak
peak
per
per
per
per
percent
percent
performance
period
period
period
period
periods
periods
Pg
physical
pike
popularity
pound
pound
pound
power
power
powers
powers
practice
pre
preston
previous
prices
probably
production
production
quarter
Queensland
railway
rate
rate
rate
rates
rates
reach
reach
reached
reached
reached
reached
reaches
reaches
reaching
reaching
recent
recession
resonance
rose
s
season
season
season
season
second
sexual
shares
shares
since
since
snow
star
still
still
success
summer
summer
summer
than
than
the
the
their
their
their
third
ticket
time
time
time
time
times
times
times
times
to
to
tonnes
top
tour
traffic
trains
trains
trough
twin
union
until
up
usually
victoria
viewing
vulture
was
week
were
when
when
when
which
which
widow
widows
with
world
year
year
year
year
year
years
yesterday
yet

301
abacha
abortive
abundance
accused
accusing
acre
acre
affecting
afoot
against
allegation
alleged
alleged
alleged
alleged
allegedly
alleging
allotment
anc
anc
anglo
apparent
Arabs
aristide
arrest
arrests
assassinate
assassinate
assassinate
assassination
auteil
ballet
ballet
begins
belongs
betray
bizarre
blacken
blackmail
blow
brad
brassicas
burial
buried
buried
carlos
Castro
cemetery
cemetery
centred
centres
character
characteristic
characteristic
characteristic
characterizati
on
characterizati
on
characters
characters
characters
characters
characters
characters
cheap
churchyard
climax
collapsed
comedy
communists
complicate
complicate
complicity
concerning
concerns
confusing
connection
connection
consciously
conspiracy
conspiracy
conspirator
continuous
contrivance
contrived
contrives
convicted
convoluted
convolution
coup
crops
crystals
cunning
cunning
curious
dastardly
deception
defraud
denigrate
depose
destabilise
destroy
destroy
detail
details
detected
development
device
devices
devious
devised
diabolical
dialogue
dialogue
digging
disclosing
discovery
discovery
discredit
disjointed
distract
doubts
downfall
downfall
dragged
drama
drama
drama
driven
elaborate
elaborate
elderly
elements
entertaining
episode
episode
episodes
episodic
existed
expose
extortion
extremist
farce
fawkes
fetched
fiasco
fiction
fidel
fiendish
foil
foil
foiled
foiled
foiled
foiled
foiled
foiled
forewarned
forgive
framed
gavin
gen
graph
graph
griffiths
gruff
gunpowder
gunpowder
happily
hatch
hatch
hatched
hatched
hatched
hatched
hatched
hatching
hatching
hectare
heroine
hinges
hinges
Hitler
holes
hyde
imperialist
implausible
implicated
implicated
improbable
insurrection
insurrection
integral
integrity
intricate
intriguing
involved
involvement
involvement
involves
involving
jailed
juan
juan
junction
jungle
kidnap
kidnap
kill
kurt

labyrinthine
lacklustre
land
learnt
lines
linked
loose
lose
loses
losing
lost
lover
lover
ma
macabre
Machiavellian
mandela
mastermind
mastermind
mastermind
mathieu
maverick
meaning
melodramatic
melodramatic
mere
michelle
motivating
movie
movie
movie
musketeers
narrative
narrative
narrative
nonsensical
novel
novel
novel
novelist
novels
opera
orchestrate
organised
oust
outline
outline
outrageous
overthrow
owns
paced
papal
part
part
participation
pea
perma
Pg
plot
plot
plot
plot
plot
plot
policeman
pope
pope
potato
predictability
predictable
predictable
predictable
prominent
proportion
punish
rabin
rated
rating
refuge
rehash
remarkably
reminiscent
reserved
restores
revolves
rid
roles
rumours
russell
sabotage
sacrifices
scatter
scenes
script
script
seize
sensational
senseless
sequences
serial
serial
shifts
silly
simplicity
sinister
sinister
smashed
smear
smuggle
snatch
soap
socialism
sooner
sophisticated
spaghetti
spoil
spoiling
stallone
steal
strand
strategies
structure
structures
sub
subvert
Sudan
summaries
summary
sums
suspects
suspects
suspense
suspense
suspenseful
swallow
swallow
swims
swindle
synopsis
synopsis
terrorists
theme
thickened
thickened
thickens
thickens
thriller
thrillers
thwarted
tiffs
topple
trumped
turns
turns
twist
twist
twist
twist
twist
twists
twists
twists
twists
twists
twists
uncovered
uncovered
uncovered
uncovered
uncovered
uncovered
undermine
unfold
unfolded
unfolds
veg
vegetable
veneer
victim
victim
wacky
waterworld
watery
weaken
weakness
weaknesses
worthy
yitzhak
zionist

PLOT T

</h>
</h>

</h>
</h>
<b>
<b>
<h>
<h>
<h>
<p>
<p>
a
a
a
a
about
about
abundance
accused
acre
affecting
against
allegation
alleged
alleged
alleged
alleged
allegedly
along
although
among
an
an
an
an
an
anc
and
anti
any
around
as
assassinate
assassinate
assassination
away
basic
basis
begins
behind
being
blackmail
blow
bomb
bomb
bomb
bought
bring
build
burial
buried
bush
but
but
buy
by
by
by
central
central
centre
centres
character
character
character
character
character
characteristi
cs
characters
characters
characters
characters
characters
characters
cheap
comedy
comment
communist
communists
completely
complex
complicate
complicity
concerns
confusing
connection
connection
conspiracy
control
convicted
convoluted
convolution
country
coup
course
course
death
destabilise
destroy
details
details
details
development
development
device
devices
dialogue
dialogue
director
discovery
discredit
doesn
drama
drama
driven
each
elaborate
elements
entire
episode
essential
evidence
evidence
evil
expose
family
far
film
film
film
film
foiled
foiled
foiled
follow
follow
for
forgive
former
former
garden
gavin
general
get
give
government
griffiths
ground
gruff
gunpowder
hard
harry
has
has
hatched
hatched
hatched
hatched
hatched
have
have
her
him
him
hit
Hitler
implausible
implicated
implicated
in
in
in
in
individual
insurrection
insurrection
integrity
into
intricate
involved
involved
involved
involvement
involvement
involves
involving
IRA
IRA
is
is
its
its
Jewish
john
juan
just
keep
key
kidnap
kidnap
kill
king
kurt

land
land
line
lines
linked
little
lose
loses
losing
lost
ma
main
mandela
many
mastermind
maverick
meaning
michael
michelle
middle
miss
movie
movie
movie
movie
mr
much
murder
murder
murder
murder
next
no
novel
novel
novel
novel
novel
of
of
of
of
off
offers
one
opera
or
organised
original
over
over
overthrow
part
part
partly
pea
play
plot
plot
plot
plot
plot
plot
points
police
pope
pope
potato
power
predictable
president
president
rabin
rated
rather
rating
reminiscent
revolves
rid
right
role
roles
russell
s
s
sabotage
same
scenes
script
seize
senseless
served
setting
silly
simple
simple
sinister
sinister
small
smuggle
so
socialism
some
sooner
stand
steal
story
story
story
strategy
structure
sub
summaries
summary
swims
take
taken
takes
terrorists
that
the
the
the
the
the
theme
there
there
thickened
thickens
this
though
thriller
through
time
tiny
to
too
too
too
topple
turns
turns
twist
twist
twist
twist
twist
twists
twists
two
uncovered
uncovered
uncovered
uncovered
undermine
unfolds
up
up
vegetable
veneer
very
victim
victim
w
was
was
watery
western
where
which
which
while
while
whole
wing
with
with
with
with
woman
young
your

RARE MI

10th
130
1368
a
accidents
accolade
achieves
achieving
actresses
admirable
admittedly
adverse
affaire
affords
ahmed
alleles
amongst
anaemia
antelope
antelope
antique
antiques
antiques
appearance
appearance
appearance
archive
are
artifacts
astonishingly
atomic
automobile
automotive
bacteria
bacterium
baffling
beasts
beckett
beef
belongs
blemish
boyd
breast
breed
breed
breed
breeding
breeding
breeding
breeding
breeds
breeds
breeds
bs
businessman
butterflies
butterflies
butterfly

cancer
cancers
cave
cjd
classified
cockatoo
coffees
collect
collectable
collecting
collection
collection
collection
collectors
collectors
collectors
collects
commodity
commoner
commonplace
comparatively
complicating
complication
condition
conservative
conserve
conserved
containing
contains
contracted
contracting
cooked
cooked
coronary
costly
crested
creutzfeld
creutzfeld
deadly
declining
defect
deficiency
deficiency
deficiency
degeneration
delicacy
delightful
depressing
devotion
diabetes
diagnosed
disease
disease
disorder
disorder
displayed
divine
dolphins
dolphins
drawings
eagles
earliest
earths
eclipse
endanger
endangered
endangered
enjoying
enjoys
epidemics
equally
exceedingly
exception
exceptionally
exceptions
exceptions
exhibits
existent
exotic
exotic
exotic
exotic
experiment
exporting
exquisite
extinct
extinct
extinct
extinction
extraordinari
ly
extremely
fabric
fairly
fatal
fauna
fauna
feat
feat
flora
flutes
footage
foray
fortunately
fortunately
fossils
frog
fungi
funk
funk
gem
gene
genetic
glider
glimpse
glimpse
glimpses
groove
grooves
habitat
habitat
habitat
habitats
habitats
harassment
hen
hen
hens
honesty
honesty
honesty
hostile
hunter
ijaz
illuminate
incidents
incidents
includes
incompetence
increasing
incredibly
infections
insects
insight
instances
iridium
ironically
isotope
isotope
jakob
jakob
jewellery
jiang
kaposi
lapse
lapse
leaved

leukaemia
leukaemia
leukaemia
levity
lining
literature
lively
lowland
lucidity
mammal
mammals
mammals
Manhattan
manuscript
manuscript
manuscript
manuscript
marrow
medium
meningitis
meningitis
mercifully
mesothelioma
meteorites
microbes
migrating
mild
moments
moth
nationally
natterjack
newt
nuclei
obliterate
occasions
occur
occurence
occurrence
occurrence
occurrence
orchid
orchid
orchids
orchids
outing
ovation
owl
owl
panda
paralysed
parrots
particular
pitch
planetary
plants
pneumonia
pneumonia
politician
population
porcelain
portraits
possessed
possesses
praise
precious
precious
precious
pretty
prevents
prey
privilege
prized
provides
queer
quite
ragga
rainforest
rainforest
rare
rare
rarer
rarer
rarest
relatively
relishing
Rembrandt
represents
reprints
requiring
rhino
rhododendrum
rocking
sadly
sadly
sarcoma
savour
savour
sensitivity
sighted
sighting
sighting
sightings
slade
slices
smuggling
snowfall
somewhat
specialist
species
species
species
species
species
specimens
specimens
specimens
specimens
splendid
spontaneous
squandering
stamina
steak
steak
steak
steak
steak
steaks
stealing
struck
stunned
successes
suffered
sufferers
suffering
suffers
suffers
supernovae
surprising
symbiosis
syndrome
syndrome
syndrome
syndrome
takeovers
teeth
temper
terribly
thankfully
thankfully
theft
those
Tibet
togetherness
toxic
tracts
trader
transmission
triple
tss
tss
tumour
ultra
unanimity
unanimity
unforgettable
unheard
unique
unity
unreleased
unseen
unusual
uranium
valuable
vandalism
vegetation
very
vintage
vintages
warmth
waterways
whale
whales
wildlife
wildlife
worthwhile

RARE T

</h>
<h>
<p>
15
a
a
ability
able
accidents
advantage
ago
all
also
also
although
although
among
among
an
an
an
and
and
and
animals
animals
another
apart
appearance
appearance
appearance
are
are
are
are
are
are
as
be
be
beautiful
because
become
become
becoming
beef
been
being
bird
birds
birds
birds
birds
blood
book
book
books
born
breed
breed
breeding
breeds
Britain
Britain
Britain
British
but
but
but
but
but
called
called
called
can
can
cancer
cancer
cancer
cancer
cancer
cars
cases
cases
catch
causes
century
certain
chance
children
collection
collection
collectors
combination
commodity
common
common
commoner
comparatively
condition
condition
condition
condition
condition
country
days
days
died
disease
disease
disease
disease
disease
disease
disease
disorder
disorder
display
does
during
earliest
endangered
enjoyed
enough
equally
even
event
example
example
exceedingly
except
except
except
exception
exceptionally
exceptions
exist
exotic
expensive
experience
extinct
extinct
extremely
fairly
fatal
female
few
few
find
find
find
footage
for
for
for
form
form
fortunately
found
from
from
from
funk
genetic
genuine
gift
glimpse
groove
habitat
habitat
had
happen
has
has
have
having
he
hear
hear
here
his
home
horse
hostile
how
however
however
in
in
in
incidents
includes
including
increasing
incredibly
indeed
indeed
infection
insight
instances
interview
interview
into
is
is
it
it
Japanese
kaposi
known
large
leukaemia
levity
limited
made
makes

making
making
many
may
medium
medium
mild
moment
moment
moments
moments
must
nature
now
now
occasion
occasions
occur
occurrence
of
of
of
offered
offers
old
on
one
one
one
only
only
opportunity
or
otherwise
particular
perfect
pitch
plant
plants
plants
plants
plants
politician
precious
pretty
produced
provided
provides
public
quality
quite
quite
rare
rare
rare
rare
rare
rarer
rather
red
relatively
remain
represents
s
s
sale
scored
see
see
seeds
self
session
several
sexual
she
sheep
show
show
sight
silver
skin
so
some
something
something
specialist
species
species
species
species
species
species
species
statement
still
still
still
struck
successful
such
such
such
such
such
such
such
suffered
suffered
suffering
suffering
suffers
sunshine
syndrome
syndrome
talent
teeth
thankfully
that
that
that
the
themselves
there
these
these
these
these
they
thing
thing
this
this
this
those
though
though
threatened
to
too
took
toxic
treat
treated
trust
ultra
unique
unique
unity
unusual
valuable
very
very
visit
was
was
was
were
were
when
where
where
which
which
which
which
which
which
who
who
whose
wild
wild
wildlife
will
with
with
woman
woman
women
years
yesterday

SPUR MI

</h>

</h>
</h>
</h>
<b>
<b>
<h>
<h>
<h>
<p>
<p>
1
1
25
3
a
act
act
act
act
acted
action
action
action
action
action
acts
added
additional
advance
after
also
also
and
and
another
another
as
attempted
back
be
because
better
between
big
blue
bone
boom
boost
both
both
business
business
business
but
but
by
by
call
can
can
century
change
change
cleanly
coast
come
comment
companies
competition
competition
competition
competitive
congress
constant
consumer
could
create
create
creation
cross
cut
cut
cut
cut
debate
decided
decision
decision
decision
deficit
demand
designed
development
development
diamond
do
do
does
dollar
domestic
done
down
east
economic
economic
economic
economic
economy
economy
economy
economy
efficiency
effort
efforts
enough
European
even
expected
exports
failed
fails
fame
fed
first
first
flying
football
football
for
four
from
further
further
gave
gear
gears
given
golden
golden
golden
government
great
greater
greater
growth
growth
growth
growth
growth
growth
has
he
help
help
help
helped
helped
hill
hill
him
his
home
home
home
hope
hoped
hopes
hopes
horse
house
i
if
immediate
in
incentives
increased
industrial
inflation
innovation
innovation
intended
interest
interest
interest
into
into
into
investment
investment
investment
investment
investment
it
it
its
itself
jobs
jobs
jobs
just
just
land
led
likely
limited
line
little
local
locusts
long
lower
lower
main

make
manufacturing
market
market
may
may
me
memories
merely
might
might
million
moment
money
more
more
most
mount
mountain
move
mr
much
needed
needed
needs
new
new
next
next
northern
not
of
off
oil
on
on
on
only
only
or
order
other
other
other
other
out
package
park
part
people
players
poor
posse
powerful
price
prices
private
production
productivity
program
programme
progress
provide
provided
pruning
purchase
rail
railway
rate
rate
rates
rates
rates
re
recovery
recovery
remains
research
research
ride
road
rock
rocky
rolls
royce
sales
sales
sales
savings
saw
say
scheme
seemed
seems
served
she
short
should
should
side
silver
similar
slipper
slipper
slipper
slipper
something
something
sometimes
south
south
south
southern
spending
spending
success
successful
such
such
sunshine
taken
tax
taxes
tends
than
that
the
the
them
these
thing
things
think
this
those
thought
throated
through
to
to
to
today
top
trainer
trips
truncated
try
trying
turn
up
up
up
us
us
use
use
used
wall
wasn
ways
week
well
went
were
when
which
while
white
will
will
will
will
win
winner
without
without
work
would
year
yesterday
yet
young

SPUR T

</h>

</h>

<b>
<b>
<h>
<h>
<p>
25
a
a
a
abruzzi
achieve
act
act
act
acted
action
action
action
action
action
acts
added
additional
advance
affair
also
also
an
and
another
are
arrangement
as
as
assume
attempted
back
bargains
be
be
been
bits
blue
bone
boom
boost
boost
both
burst
business
business
but
but
by
call
can
can
carl
century
change
change
cleanly
coast
cochran
comment
commodities
companies
compatriot
competition
competition
competition
competitive
concentrate
congress
constant
consumer
copyright
coubertin
could
create
create
creation
cross
cut
cut
cut
cut
cuts
debate
decided
decision
decision
decision
decisions
deficit
definitely
demand
designed
development
development
diagnosed
diamond
dividend
domestic
done
down
driving
earnings
east
economic
economic
economic
economic
economy
economy
economy
economy
efficiency
efficiency
effort
efforts
emotional
enough
European
even
exodus
expansion
expected
exports
failed
fails
fame
fame
fed
flagging
flying
football
football
for
freedman
freedman
fulfil
further
further
gains
gave
gear
gears
golden
golden
golden
golden
government
great
greater
greater
growth
growth
growth
growth
growth
growth
heathrow
heel
help
helped
helped
helps
hill
hill
him
home
home
hope
hoped
hopes
hopes
horse
i
immediate
improve
incentive
incentives
increased
industrial
inflation
innovation
innovation
innovation
inspiration
inspiration
intact
intended
interest
interest
interest
into
into
into
investment
investment
investment
investment
investment
investment
is
is
it
it
its
itself
jobs

jobs
jobs
johanson
just
keeping
land
lane
led
legacy
likely
limited
line
little
local
locusts
lower
lower
main
manufacturing
market
may
may
memories
men
merely
might
might
mighty
migration
moment
more
motorway
mount
mountain
mountain
needed
needed
needs
next
ning
northern
of
oil
on
on
on
on
only
only
order
organisation
other
out
ownership
package
park
part
peking
perpetual
players
poor
posse
powerful
price
prices
principal
priory
private
production
productivity
program
programme
progress
proposed
prospect
provide
provided
providing
pruning
purchase
rail
railroad
railway
rally
rate
rates
rates
rates
recovery
recovery
reject
remains
research
research
restructuring
ride
road
rock
rocky
rolls
royce
saddle
sales
sales
savings
scheme
seemed
seems
served
setbacks
severe
sharp
shocked
short
should
should
side
silver
similar
slipper
slipper
slipper
slipper
sluggish
something
sometimes
south
southern
spending
spending
spur
spurs
stearns
success
successful
sunshine
surge
taken
tax
taxes
tends
that
the
the
them
thing
things
this
throated
through
to
to
to
trainer
trainer
trips
trodden
truncated
try
up
us
violations
wall
was
was
wasn
way
ways
wharf
which
white
will
will
will
will
win
winner
without
without
would
would
yesterday
yet
you

SUBTLE MI

a
abnormality
acting
add
adding
adds
adjustment
ain
allows
allure
allusion
alteration
alteration
altered
altogether
ambiguous
apply
appreciate
appreciate
approach
approaches
argument
aroma
aroma
aroma
attenders
attractive
autumn
aware
aware
balanced
barely
beautifully
bizarre
blackmail
blacks
blend
blend
blend
blues
bold
brilliant
brilliant
Bristol
Camembert
capable
carefully
certain
changes
character
characters
civilised
colour
colour
colour
colour
colour
colour
colouring
colours
colours
colours
colours
colours
combination
complex
complex
complex
complex
complicate
concentrate
connection
consciousness
considerable
contours
controlling
convey
convey
cotton
cream
create
created
creates
creating
cues
culture
cumulative
data
delicate
delicate
delicate
delicate
desert
detail
detail
detailing
detect
detect
detect
deviations
difference
difference
difference
difference
difference
difference
discover
discovered
discriminatio
n
distinction
distinction
dry
earth
effect
effective
effective
effects
effects
effects
effects
elegant
elegant
elegantly
elusive
emotion
emphasis
energies
energy
equally
equally
exactly
excite
exert
exist
experiment
explain
explore
expressed
expression
expression
extremely
fabrics
fairly
fairly
far
fashion
features
female
fine
flavors
flavors
flavour
flavour
flavour
flavour
flavour
flavours
flavours
flavours
flavours
flaw
floral
flowers
fragrance
fragrance
fruit
fruit
fruit
function
gay
gender
genre
gentle
gives
gives
gradation
gradations
gradual
grant
greens
grey
gross
happens
harmonies
highlights
hint
hints
however
hues
hues
humour
humour
imaginative
imparts
increasing
infinitely
infinitely
insidious
instance
intangible
intelligent
interplay
ironies
kinds
language
learn
lemon
less
lies
lies
looks
loud
manipulation
manner
many

master
messages
more
most
nature
none
noted
nothing
notice
notice
noticeable
noticed
novel
nuance
nuances
nuances
oak
objects
observe
obvious
obvious
obviously
occurred
occurring
occurs
offers
often
overt
overt
pace
paint
painted
particular
particular
perceive
perfect
perfectly
performance
persistent
persuasive
physical
pick
pink
politician
potent
powerful
powerful
precision
preschool
pressures
profound
profound
provoking
psychic
psychological
quite
racism
range
rather
reaching
realistic
rebuke
refined
reflected
relationship
remarkably
represents
requires
requires
resist
reveals
rich
rich
rich
search
seemingly
send
senses
sensitive
sensitive
sensitive
sensitive
shades
shades
shading
shadings
shape
shape
sharp
sheer
shift
shift
shift
shifts
shifts
shifts
signals
significant
similarly
simple
sinister
skin
slightly
slow
smart
smell
smooth
smooth
so
soft
soft
soft
some
sometimes
somewhat
sophisticated
sophisticated
sophisticated
sophisticated
sorts
souled
spots
structural
styles
subliminal
subtle
subtle
subtle
subtle
sufficient
surface
surface
surprising
sweet
systems
tactics
tastes
tea
technique
technique
techniques
terribly
texture
textures
these
thousand
tints
tone
tone
tones
tones
tones
too
touches
tradition
turns
unaware
uncompromisin
g
undergone
understand
understand
understate
undulation
unlike
unmistakable
unusually
variation
variation
variations
variations
varied
variety
various
vehicle
very
vibrations
warm
ways
ways
ways
whose
wonderfully
wordplay
yellow
yet

SUBTLE T

15
a
a
about
about
add
add
adds
ain
all
also
also
also
an
and
and
and
and
appreciate
approach
approach
are
are
are
are
are
argument
art
as
as
as
as
aspects
attenders
aware
based
be
be
be
be
become
being
between
between
between
bit
black
blend
body
body
body
both
brilliant
Bristol
but
but
but
but
but
but
camembert
campaign
can
can
can
can
carefully
certain
change
change
change
change
changes
changes
changes
colour
colour
colour
colour
colour
colour
colouring
colours
colours
colours
colours
complex
complex
complex
complex
concentrate
control
create
culture
david
delicate
delicate
delicate
delicate
detail
detect
detect
difference
difference
difference
difference
difference
distinction
distinction
each
earth
effect
effect
effect
effect
effect
effective
effective
effects
effects
effects
effects
elegantly
elusive
emphasis
energies
energy
energy
energy
equally
even
ever
every
example
exist
explain
expression
expression
extremely
face
fairly
far
features
female
few
fine
flavour
flavour
flavour
flavour
flavours
flavours
flavours
flavours
floral
form
form
forms
fragrance
fragrance
from
fruit
fruit
function
genre
give
gives
gross
have
her
hint
hints
his
however
however
humour
humour
imparts
important
in
in
in
insidious
instance
intelligent
is
is
is
it
it
it
it
its
its
itself
kind
language
learn
less
less
less
less
level
lies
light
lighting
likely
little
looks
lot
love
make
make
makes
making
making

manner
many
may
may
may
message
messages
more
more
more
most
much
much
music
nature
need
need
none
not
not
nothing
noticeable
nuances
oak
obvious
obvious
obviously
of
of
often
often
often
or
or
or
or
or
other
other
others
our
own
pace
particular
performance
perhaps
physical
pink
potent
powerful
powerful
powerful
precision
pressure
pressures
profound
psychologica
l
quality
quite
range
range
rather
rather
rather
reaching
remarkable
result
rich
rich
s
s
s
sense
sensitive
series
shade
shades
shape
shift
shift
shift
shifts
shifts
show
significant
simple
simple
skills
slightly
so
so
soft
some
some
some
something
something
sometimes
sometimes
sometimes
sophisticate
d
sophisticate
d
sort
sound
structural
subtle
subtle
subtle
subtle
such
such
such
surface
surprising
sweet
takes
taste
tastes
technique
techniques
than
than
than
than
than
that
that
the
the
their
there
there
there
these
these
these
they
this
this
this
those
through
through
to
to
tone
tones
tones
tones
too
touch
touch
tradition
understand
unmistakable
use
use
using
variations
variety
very
warm
was
way
way
way
ways
ways
which
which
which
which
white
white
whole
with
with
with
wonderful
yet
yet
your

SURGING MI

<h>
<h>
<p>
1
1
1
1
1
2
2
3
3
3
3
4
4
4
4
48
5
5
5
5
6
7
7
7
7
a
a
a
across
adrenaline
after
after
after
after
after
after
after
again
ahead
all
all
all
all
also
also
America
an
an
and
and
and
and
and
and
another
are
are
are
around
as
as
at
at
at
at
away
back
back
be
beach
been
been
been
before
before
but
but
but
by
by
by
by
by
c/h>
came
come
consumer
consumer
consumer
continued
cope
could
countries
crowd
crowd
crowds
crowds
demand
demand
demand
demand
different
dollar
dollar
dollar
down
down
drives
economic
end
European
exports
exports
falls
faster
feel
fell
fell
figures
figures
figures
first
floodwater
for
for
for
for
forward
French
from
from
from
from
from
garrick
goal
got
had
had
had
half
has
has
has
has
have
her
her
high
high
him
him
his
his
his
his
imports
in
in
in
in
in
index
industrial
industrial
industry
inflation
inflation
inflation
inflation
inside
into
into
is
is
is
its
its
its
its
japan
japan
japan
Japanese
Japanese
Japanese
jumped
kinds
lead
leapt
like
like
line
made
many
me
midfield
minutes
mob
month
month
months
months
morgan
my
my
new
new
new
new
new
not
now

o
of
of
of
of
of
of
oil
on
on
on
on
one
only
other
other
out
output
output
output
over
over
over
over
past
points
points
polls
popularity
pound
power
pre
prevent
price
price
price
price
prices
prices
prices
prices
prices
prices
prices
production
production
production
production
profits
rate
retail
rise
river
run
runs
s
s
s
s
s
sales
sales
sales
sales
sales
sent
sent
September
share
shares
shares
show
show
show
since
since
small
so
south
star
still
still
stock
stocks
support
Switzerland
tax
terms
testosterone
than
than
than
than
that
that
that
that
the
the
the
the
the
the
their
their
their
them
then
then
then
there
they
they
this
this
through
through
through
tide
to
to
to
to
to
too
too
toward
towards
two
two
two
two
unemployment
unemployment
unemployment
until
up
up
up
up
was
was
was
water
water
water
wave
waves
were
were
were
were
were
were
when
when
when
which
which
which
which
while
will
will
with
with
with
with
with
with
within
workers
workers
workers
year
year
year
year
year
years
years
years
yen
yesterday

SURGING T

<p>
1
1
10
118
15
1995
2
3
4
48
5
511p
according
across
adrenaline
after
after
after
again
ahead
almost
American
and
and
and
and
another
are
are
around
Asia
at
Australia
Australian
Australian
average
avoid
away
back
back
ball
beach
beating
Beatles
been
before
before
began
best
between
blue
body
Bosnia
brings
brixton
broad
broke
building
burrows
by
by
by
by
came
captured
cars
cars
caused
cellular
cent
cent
centre
China
chip
close
come
communicate
confidence
continued
contribute
cope
copper
costs
costs
could
crowd
crowd
crowds
crowds
cupolas
defenders
demand
demand
dense
derived
despite
despite
development
different
dollar
dollar
down
down
drives
early
economic
emma
emotions
energy
engine
engine
everyone
expected
expecting
exports
exports
face
far
fast
fear
fee]
felt
felt
few
finally
find
first
flank
floodwater
for
forward
forward
Friday
from
from
front
front
full
game
garrick
Germany
goal
green
growth
half
has
high
high
high
highs
him
him
his
hopes
hormones
hundreds
imports
impulses
in
in
in
ince
index
inflation
inflation
inflation
inside
into
into
is
its
its
japan
japan
jersey
johnson
keane
kinds
known
labour
lead
lead
lead
leading
led
left
levels
like
like
line
line
lines
linocut
literacy
low
ltd
made
maelstrom
makes
many
march
market
markets
maroon
massive
massive
me
meet
melody
met
middle
midfield
midfield
miles

minute
minutes
mob
morgan
music
music
narrow
new
non
now
nsw
of
of
oil
oil
on
OPEC
orchestra
other
out
output
outside
over
over
past
paul
per
percent
percent
piano
plummeting
points
points
poll
polls
polls
popularity
pound
power
pre
prevent
prevent
price
price
prices
prices
prices
prices
problems
problems
process
production
production
profit
profits
pushed

quarter
queensland
rates
receding
red
region
related
relentless
reported
riot
river
road
run
run
runs
s
sails
sales
sales
sales
sales
saw
second
send
sent
sent
sent
several
several
sewer
share
shares
shares
shares
shock
show
showed
six
small
sound
south
speculation
still
stock
stock
stocks
stop
stream
street
strong
sums
superb
support
swell
swelling
swept

tax
telephone
testosterone
the
the
the
the
the
their
them
then
third
this
thousands
threw
through
through
through
tide
to
too
too
toward
toward
towards
towards
track
trade
transportatio
n
troops
turned
up
up
veins
victory
violence
was
was
was
water
water
water
water
water
waters
wave
wave
wave
waves
ways
wellcome
were
western
when
which

while
wild
wildly
with
with
with
with
with
within
yen
yen
yesterday
yesterday
yesterday
York

SUSPEND MI

</h>
<h>
11
a
ability
action
activities
activities
Afghanistan
again
against
agree
agreed
agreed
agreeing
agreement
agreements
aid
aid
aid
aid
allies
allow
also
also
American
Americans
anc
anc
anc
and
announced
announced
any
anyway
approved
Arab
Arabia
armed
armed
arms
ask
ask
asked
assistance
association
attacks
audience
authorities
authority
b
bank
banks
banning
because
belief
between
bombing
boris
boycott
boycott
boycott
British
bt
bush
bush
business
campaign
campaign
can
cancel
cannot
certain
charity
combat
commission
commission
commitment
committee
community
competition
congress
constitution
constitution
contacts
contacts
contract
cooperation
corpus
could
county
court
covenants
Croatia
Cuba
cynicism
days
decide
decided
decided
decision
decision
decision
declaration
declaration
declaration
decree
decree
defended
demand
demand
department
dialogue
dialogue
dialogue
dialogue
diplomatic
disbelief
disbelief
disbelief
dividend
dividends
down
ec
economic
effective
eighteen
embargo
enforcement
engineering
enjoy
even
exchange
exercises
expel
expel
exports
exports
expulsions
fa
first
flights
flights
for
force
forced
forced
forced
forcing
forcing
France
France
from
fujimori
gijon
goodwill
Gorbachev
Gorbachev
government
government
government
government
government
governor
habeas
habeas
Hague
has
hostile
hostilities
if
immediately
implementati
on
implementati
on
imports
independent
independent
independent
Israel
issue
its
its
its
judgement
judgment
judgment
just
kaunda
Kenneth
law
laws
laws
laws
leaders
leaders
least
led
led
legislation
legislation
lewis
license
likud
lines
Lithuania
Lithuania
Lithuania
Lithuania
Lithuania
Lithuanian
Lithuanian
m
magic
management

manoeuvres
marriott
may
may
means
membership
membership
membership
might
military
military
modify
Monday
month
months
Moscow
motion
must
negotiation
negotiation
Nigeria
night
no
non
not
notice
notice
now
nuclear
of
of
offensive
offered
offering
only
operation
operations
operations
operations
or
order
order
ordered
ordered
others
oxfam
Palestine
parkhurst
parliament
parliament
parliament
parliament
parliament
parliament
participation
participation
party
passed
payment
payments
payments
payments
peace
peace
pending
pending
persuaded
Peru
PLO
possession
possession
power
powers
prepared
president
pressure
probably
proceeding
process
production
program
programme
proposal
provincial
publication
pupils
rated
readiness
ready
recruitment
refused
relief
reluctant
republic
republic
republican
residents
restrictions
revoke
revoke
rights
rise
rules
Russia
s
sanchez
sanctions
sanctions
Serbia
seriously
shares
shipments
should
sind
six
south
south
Soviet
state
states
states
still
struggle
students
submarine
substitute
supplies
surely
t
talks
talks
temporarily
temporarily
temporarily
temporarily
terminate
testing
testing
than
that
then
they
threatened
threatening
three
throughout
time
to
township
trade
trading
unanimously
union
unions
unita
united
united
until
until
urged
urged
v
voted
voted
warrant
Washington
we
week
weeks
well
which
while
white
will
willing
willingly
willingness
with
withdraw
withdrawal
would
writ
yeltsin
yeltsin
yet
Yugoslavia

SUSPEND T

11
a
ability
action
action
activities
activities
Afghanistan
against
against
agree
agreed
agreed
aid
aid
aid
aid
all
also
also
American
anc
anc
and
and
announced
any
any
any
anyway
approved
Arab
Arab
Arabia
armed
armed
ask
ask
Assam
assistance
authorities
authority
authority
b
ban
ban
bank
banning
be
because
been
belief
between
boris
boycott
boycott
boycott
British
bt
bush
bush
by
campaign
campaign
can
cancel
cannot
censure
committee
communist
community
community
competition
congress
constitution
constitution
contract
contras
could
court
court
Croatia
Cuba
days
debate
decide
decided
decided
decision
decision
decision
declaration
declaration
declaration
decree
decree
demand
demand
dialogue
dialogue
dialogue
dialogue
disbelief
disbelief
disbelief
discontinue
dividend
economic
economic
economic
either
emergency
enjoy
even
exchange
exercises
expel
flights
flights
for
for
force
forced
forced
forced
forcing
forthwith
France
France
from
from
fujimori
further
gijon
Gorbachev
government
government
government
government
government
government
governor
habeas
had
had
had
has
has
have
have
have
him
him
Hindus
his
hostilities
if
in
independence
independent
independent
insubordinate
interest
Israel
it
its
its
its
its
judgement
judgment
judgment
kaunda
kenneth
kitchens
last
law
laws
laws
laws
leaders
least
led
legislation
legislation
license
likud
lines
Lithuania
Lithuania
Lithuania
Lithuania
Lithuanian
Lithuanian
Lithuanian
local
longer
m
magic
may
may
means
membership
membership
membership
might
military
military
military
moment
Monday
month
months
Moscow
motion
move

much
must
necessary
need
negotiation
negotiation
Nigeria
nuclear
of
offensive
offer
offered
offering
old
on
on
operation
operations
operations
operations
or
or
order
order
order
ordered
our
oxfam
Palestine
parliament
parliament
parliament
parliament
parliament
parliament
participation
party
party
passed
payment
payments
payments
payments
peace
pending
pending
Peru
PLO
possession
power
prepared
president
president
pressure
pressure
probably
process
production
program
programme
proposal
provincial
publication
pupils
r
rated
ready
refused
republic
republic
republican
residents
right
right
rights
rights
Russia
s
s
said
said
sanchez
sanctions
schott
seriously
service
shares
shipments
should
simply
sind
six
sociologistic
s
some
south
Soviet
state
state
states
states
states
stay
stay
strike
struggle
students
submarine
substitute
summon
supplies
surely
t
talks
talks
temporarily
temporarily
terminate
testing
that
that
the
the
the
the
the
their
their
them
them
they
they
they
threatened
threatening
throughout
today
township
trade
trading
un
unilateral
unions
unita
united
united
until
until
until
urged
urged
use
v
vary
voted
voted
warrant
was
Washington
we
week
were
whether
which
while
while
white
will
will
willing
willingly
with
words
work
work
would
would
would
writ
yeltsin
you
your
zeroual

TRACE MI

071
100g
38
829
930
945
able
absent
accent
accent
acid
acids
acids
amino
amino
amounts
amounts
amusement
ancestors
ancestors
ancestry
ancestry
animals
animals
anxious
any
any
anyone
anywhere
anywhere
anywhere
apparent
appealed
assets
atmosphere
attempting
attempts
availability
back
back
back
barely
barest
behind
betrays
bitterness
boron
boron
calcium
calcium
calcium
calls
calls
can
cannot
carotene
cars
certain
classic
clip
coated
cocaine
concentrate
concentrate
consciousness
contact
contain
contain
containing
contains
contains
convict
copper
cortina
could
cultural
daily
deficiencies
deficiency
deficiency
descent
desperately
detect
detectable
detectives
detectives
diet
diet
dietary
dietary
difficult
directly
disappear
disappear
disappeared
disappeared
disappeared
disappearing
disappears
drugs
eagle
easier
effects
efforts
efforts
element
element
elements
elements
elements
elliot
emotion
enable
enabled
englishness
erase
erasing
essential
essential
essential
eventually
every
evidence
evolution
excellent
except
existence
extensive
extremely
failed
faint
faintest
fat
fathers
fathers
flowing
ford
ford
fossil
found
fronts
gas
gases
gone
grapes
graph
grease
gruesomely
h
heavier
help
helping
her
him
historical
historical
history
hook
hook
hooks
human
humor
immediately
impossible
inc
incident
intake
investigate
investigate
investigate
iodine
iron
iron
iron
iron
iron
irony
its
Jesus
killer
killers
known
languages
lawyers
lb
lb
lb
leaves
leaves
leaving
leaving
leaving
length
levels
line
line
lineage
lineage
lingering
link
link
little
ltd
magazine
magazine
magnesium
main
maintenance
makeup
managed
mascara
memory
merest
metabolism
metals
mg

milk
mineral
mineral
mineral
mineral
mineral
minerals
minerals
minerals
minerals
minerals
minerals
missing
missing
missing
missing
mitochondria
mono
mouth
movements
movements
movements
mrs
natchez
natural
nine
no
nor
nostalgia
nutrients
nutrients
nutrition
o
o
of
origin
original
origins
origins
origins
outline
outlines
owner
paper
parentage
parents
parkway
particular
passing
path
path
paul
phosphorus
phosphorus
pity
plants
plants
police
police
possibly
potassium
predict
presence
promptly
protein
publication
publication
pushed
quantities
quantities
quantities
records
records
regret
relatives
remains
remorse
remove
required
responsible
riboflavin
rich
risen
road
roots
roots
route
salts
sank
sank
sarcasm
scarcely
scientists
sea
sea
search
selenium
selenium
sentimentalit
y
shall
showing
shyness
sink
sink
sinking
sinks
slight
slightest
smaller
smile
soil
soil
sour
source
source
source
sources
species
specific
stolen
stolen
stores
substance
sunk
sunk
supplement
supplement
supply
their
them
there
thiamine
thousands
to
toxic
toxic
trace
trace
trace
trace
trace
trace
tracks
tree
tree
try
trying
unable
using
vanish
vanish
vanished
vanished
vanished
vanishes
various
vegetation
victims
virtually
virus
visible
visual
vital
vitamin
vitamin
vitamin
vitamins
vitamins
vitamins
vitamins
vitamins
voice
vulgarity
wallet
washed
whereabouts
wire
wire
without
without
without
zinc
zinc
zinc

TRACE T

<p>
<p>
15
a
a
able
able
absent
accent
accent
acid
acids
air
all
all
almost
also
although
amino
amounts
ancestors
ancestry
and
and
animals
animals
anxious
any
any
any
any
are
are
around
as
atmosphere
attempt
attempts
back
back
back
back
barest
be
been
behind
believed
between
bitterness
blood
body
body
but
call
calls
can
can
can
can
cannot
certain
child
clip
cocaine
contain
contain
contains
content
copper
cortina
could
could
could
daily
deficiencies
deficiency
detectable
detectives
detectives
development
development
diet
difficult
directly
disappear
disappear
disappeared
disappeared
disappeared
drugs
eagle
easier
easy
easy
effort
efforts
either
element
elements
elements
elements
elliot
emotion
enable
especially
essential
essential
even
ever
every
evolution
failed
failed
faint
faintest
family
family
father
fathers
find
flowing
ford
found
found
found
from
from
gas
gases
girl
gone
gone
had
had
hard
hardly
has
have
have
he
help
help
help
her
her
her
him
him
his
his
his
history
history
history
hook
hook
hooks
how
human
if
if
impossible
in
inc
including
investigate
iron
iron
iron
iron
iron
irony
is
it
it
its
killer
known
last
lb
lb
lb
leave
leaves
leaving
leaving
left
levels
line
lineage
lineage
lingering
little
magazine
main
make
man
man
managed
many
may
memory
merest
metals
mg
mineral
mineral
minerals
minerals
minerals
minerals
minerals
missing
missing
missing
missing

mitochondrial
money
more
movement
movements
movements
movements
Mrs
much
my
my
natchez
natural
no
not
not
now
nutrients
nutrients
o
o
of
of
of
old
only
or
or
origin
original
origins
origins
other
out
outline
owner
owner
paper
parents
parents
parents
parkway
particular
past
path
path
pattern
paul
person
pity
plants
plants
police
police
police
possible
possibly
presence
present
private
protein
quantities
records
red
relatives
remains
remove
responsible
riboflavin
road
roots
roots
route
s
sank
sank
sarcasm
says
sea
sea
selenium
selenium
self
self
she
she
should
sink
sinking
slight
slightest
smile
so
some
some
source
species
still
stolen
such
sunk
sunk
t
than
the
the
the
their
their
their
them
them
then
then
there
there
they
they
this
those
through
through
to
to
to
to
toxic
trace
trace
trace
trace
tree
tried
try
trying
trying
unable
used
using
vanish
vanished
vanished
very
victims
visible
visual
vital
vitamin
vitamins
vitamins
vitamins
vitamins
voice
want
was
was
wasn
water
we
we
we
were
were
where
whereabouts
which
which
who
who
who
will
wire
wire
with
with
without
without
without
would
would
you
your
your
zinc

TREND MI

13p
1977
1980s
199 Os
2p
4p
5p
abs
accelerate
accelerate
accelerate
accelerate
accelerate
accelerate
accelerate
accelerate
accelerating
accelerating
accelerating
alarming
alarming
alarming
alter
among
amongst
analysts
appears
appointing
argues
artistic
aspect
attempts
away
benefiting
blaming
booking
borrowers
brands
brokers
buck
buck
buck
bucked
bucked
bucked
bucking
bucking
bucks
bucks
bucks
bullish
capitalise
capitalise
carlin
cashing
casual
centralisation
centralisation
chart
clarity
climbing
collectivity
commodity
completion
confirm
confirm
confirmed
confirming
confirming
confirms
confirms
conscious
consistent
consolidation
continuation
continuation
continue
continue
continue
continue
continued
continued
continues
continues
continues
continues
continuing
continuing
continuing
contracting
contrary
contribution
convenience
cooling
counteract
criticism
cso
culmination
currencies
cyclical
dating
dearer
decade
decentralize
declining
declining
declining
deficit
defied
defied
defied
defies
definite
deflationary
defy
demographic
designers
deteriorating
direction
discernible
dismal
distressing
disturbing
disturbing
diversity
downsizing
downward
downward
downward
downward
downward
downwards
downwards
downwards
dramatically
easing
economies
economists
elsewhere
emerged
emerging
emerging
employing
employment
established
estimate
estimate
estimates
evident
evident
evident
evident
evolutionary
exacerbate
exception
exception
exceptions
expects
exploit
extension
extrapolate
fad
failures
firmer
flashy
fluctuation
follow
follow
followers
follows
forefront
fragrances
gdp
gdp
gradual
greater
growing
growth
growth
grunge
grunge
halt
halt
healthier
highlighted
highlights
holidays
hottest
illustrate
implication
improving
indicate
indicated
indicates
indication
indicative
individual
inegalitarian
inequality
inequality
inequality
inevitability
inexorable
inflation
inflation
inflational
influences
innovation
innovation
integration
intervention
involves
irreversible
lending
litigation
machinery

Malaysia
managed
manifestation
manufacturing
medium
mergers
mirrors
momentum
nineties
notable
noted
noticeable
noticed
noticed
noticed
novels
nowadays
obesity
offs
ominous
optimistic
organising
output
outsourcing
overall
overtaken
ownership
packages
part
peaked
persists
planning
poultry
predicted
predicting
predominant
prevailing
prevails
prices
profitability
pronounced
provisional
rate
rates
recessional
reflect
reflect
reflected
reflected
reflecting
reflection
reflects
reflects
reflects
reflects
reflects
regulation
reinforce
reinforce
reinforced
remains
renaissance
rents
replacing
resist
resist
resisted
resisted
resume
resurgence
retail
retailers
retailing
retirement
retro
reveals
reversal
reversal
reverse
reverse
reverse
reverse
reverse
reversed
reversed
reversed
reversed
reversed
reversed
reverses
reverses
reversing
reversing
revival
revivalist
rightward
rising
rising
rising
rspca
scent
sectors
secular
secular
sees
separation
setter
setters
setting
shorter
signs
silage
societal
spiralling
spotter
spotters
spotting
started
statistical
statistician
steadily
subrata
suggests
suggests
superimposed
survey
sweeping
sweeping
sweeping
symptomatic
symptomatic
tabloid
technological
teens
tools
toward
towards
towards
towards
towards
traditional
trend
trend
trend
trend
twins
underlying
underlying
underlying
unemployment
unemployment
unemployment
unhealthy
unmarried
unmistakable
upward
upward
upward
upward
upwards
upwards
upwards
urbanization
variable
vegetarian
viewing
volumes
warming
warned
watchers
whereby
Windsor
witnessing
worldwide
worldwide
worrisome
worrying
worrying
worrying
worrying
worrying
worrying
worsen
worsening

<p>
<p>
<p>
<p>
1980s
1992
2
a
a
about
above
according
added
against
against
alarming
already
also
also
although
although
America
among
among
an
another
another
appears
appears
artistic
Australia
away
be
been
been
began
began
began
beginning
beginning
below
benefit
bigger
buck
buck
buck
bucked
bucked
bucked
bucking
bucking
bucks
business
but
but
but
but
by
by
by
by
cent
changing
city
clear
clear
clear
clearly
companies
companies
concerned
confirm
confirmed
confirmed
confirmed
confirms
continuation
continue
continue
continue
continue
continue
continue
continued
continued
continued
continued
continued
continues
continues
continues
continues
continuing
continuing
continuing
continuing
continuing
cooling
corporate
could
cultural
current
current
current
cutting
dangerous
decade
declining
deficit
defied
definite
despite
developed
direction
disturbing
downward
downwards
economic
economy
elsewhere
emerging
emerging
employment
encouraging
established
estimate
Europe
European
evidence
evidence
example
example
exception
exception
expect
expected
falling
fashion
few
figures
figures
follow
follow
follow
followed
followed
following
following
following
follows
for
from
further
future
general
general
global
greater
greater
growing
growing
growth
growth
growth
growth
half
has
has
has
has
has
has
historical
however
however
if
if
improving
in
in
in
income
increase
increase
increased
increasing
increasing
indicative
industry
industry
industry
inequality
inevitability
inflation
inflation
innovation
integration
international
is
is
is
is
is
is
is
last
last
latest
latest
latest
likely
likely
line
line
long
long

longer
lower
major
making
managed
market
market
market
market
may
michael
modern
months
more
more
more
moving
national
new
not
noticed
now
number
of
of
of
of
on
one
other
outsourcing
over
overall
part
part
past
people
per
planning
positive
present
prices
prices
prices
rate
rate
rates
rates
recent
recent
recent
recent
recent
reflect
reflected
reflects
region
remains
results
reversal
reversal
reverse
reverse
reverse
reverse
reversed
reversed
reversed
reversed
reversed
reversed
reversing
reversing
rise
rising
rising
rising
rising
said
said
said
sales
sales
same
says
sector
secular
see
seems
seems
set
set
set
setter
setters
setting
setting
shares
show
show
showed
shows
signs
similar
since
since
small
smaller
social
some
spending
spot
spotters
start
start
start
started
started
started
started
states
still
survey
technological
term
term
terms
that
that
that
that
the
the
there
there
this
this
this
this
though
throughout
to
to
to
to
toward
toward
toward
towards
towards
towards
towards
trend
trend
trend
trend
UK
underlying
unemployment
unemployment
unemployment
united
up
upward
upwards
upwards
us
violence
warming
warned
was
which
which
will
will
with
women
worldwide
worrying
year
year
year
years

Appendix 6.1 Vocabulary/TOEFL Elicitation Instrument

Name ____________________________________

Mother Tongue __________________

Total time spent in English-speaking countries ____________

Age __________________

Sex: Male Female

Goal of studying at CELE:

_____ Enter the University of Nottingham

_____ Enter another English-speaking university

_____ Other reason

Write the first word you think of when you see each prompt word on the line provided.

1. massive __________ __________ __________

2. peak __________ __________ __________

3. rare __________ __________ __________

4. subtle __________ __________ __________

5. surging __________ __________ __________

6. trend __________ __________ __________

There are two basic types of glaciers, those that flow outward in all directions with little regard for any underlying terrain and those that are confined by terrain to a particular path.

The first category of glaciers includes those massive blankets that cover whole continents, appropriately called ice sheets. There must be over 50,000 square kilometers of land covered with ice for the glacier to qualify as an ice sheet. When portions of an ice sheet spread out over the ocean, they form ice shelves.

About 20,000 years ago the Cordilleran Ice Sheet covered nearly all the mountains in southern Alaska, western Canada, and the western United States. It was about 3 kilometers deep at its thickest point in northern Alberta. Now there are only two sheets left on Earth, those covering Greenland and Antarctica.

Any domelike body of ice that also flows out in all directions but covers less than 50,000 square kilometers is called an ice cap. Although ice caps are rare nowadays, there are a number in northeastern Canada, on Baffin Island, and on the Queen Elizabeth Islands.

The second category of glaciers includes those of a variety of shapes and sizes generally called mountain or alpine glaciers. Mountain glaciers are typically identified by the landform that controls their flow. One form of mountain glacier that resembles an ice cap in that it flows outward in several directions is called an ice field. The difference between an ice field and an ice cap is subtle. Essentially, the flow of an ice field is somewhat controlled by surrounding terrain and thus does not have the domelike shape of a cap. There are several ice fields in the Wrangell, St. Elias, and Chugach mountains of Alaska and northern British Columbia.

Less spectacular than large ice fields are the most common types of mountain glaciers: the cirque and valley glaciers. Cirque glaciers are found in depressions in the surface of the land and have a characteristic circular shape. The ice of valley glaciers, bound by terrain, flows down valleys, curves around their corners, and falls over cliffs.

1. The word "massive" in line 4 is closest in meaning to __________

(A) huge
(B) strange
(C) cold
(D) recent

I know this word ______

I don't know this word, but guessed from the text ______

2. The word "rare" in line 13 is closest in meaning to __________

(A) small
(B) unusual
(C) valuable
(D) widespread

I know this word ______

I don't know this word, but guessed from the text ______

3. The word "subtle" in line 20 is closest in meaning to __________

(A) slight
(B) common
(C) important
(D) measurable

I know this word

I don't know this word, but guessed from the text ______

Basic to any understanding of Canada in the 20 years after the Second World War is the country's impressive population growth. For every three Canadians in 1945, there were over five in 1966. In September 1966 Canada's population passed the 20 million mark. Most of this surging growth came from natural increase. The depression of the 1930's and the war had held back marriages, and the catching-up process began after 1945. The baby boom continued through the decade of the 1950's, producing a population increase of nearly fifteen percent in the five years from 1951 to 1956. This rate of increase had been exceeded only once before in Canada's history, in the decade before 1911, when the prairies were being settled. Undoubtedly, the good economic conditions of the 1950's supported a growth in the population, but the expansion also derived from a trend toward earlier marriages and an increase in the average size of families. In 1957 the Canadian birth rate stood at 28 per thousand, one of the highest in the world.

After the peak year of 1957, the birth rate in Canada began to decline. It continued falling until in 1966 it stood at the lowest level in 25 years. Partly this decline reflected the low level of births during the depression and the war, but it was also caused by changes in Canadian society. Young people were staying at school longer; more women were working; young married couples were buying automobiles or houses before starting families; rising living standards were cutting down the size of families. It appeared that Canada was once more falling in step with the trend toward smaller families that had occurred all through the Western world since the time of the Industrial Revolution.

Although the growth in Canada's population had slowed down by 1966 (the increase in the first half of the 1960's was only nine percent), another large population wave was coming over the horizon. It would be composed of the children of the children who were born during the period of the high birth rate prior to 1957.

4. The word "surging" in line 4 is closest in meaning to ________

(A) new
(B) extra
(C) accelerating
(D) surprising

I know this word ______

I don't know this word, ______
but guessed from the text

5. The word "trend" in line 11 is closest in meaning to ________

(A) tendency
(B) aim
(C) growth
(D) directive

I know this word ___

I don't know this word, ___
but guessed from the text

6. The word "peak" in line 14 is closest in meaning to ________

(A) pointed
(B) dismal
(C) mountain
(D) maximum

I know this word ___

I don't know this word, ___
but guessed from the text

MASSIVE

mass / massiveness X massive massively

(If you were talking about war)

(If you were talking about finance or the economy)

(If you were talking about statistics)

ADJ (building, wall) large + heavy, solid, strong _____

ADJ (crowd, increase) exceptionally large, greater than usual _____

PEAK

peak peak peak / peaked X

(If you were talking about a business)

(If you were talking about a house)

(If you were talking about geography)

NOUN/ADJ/VERB (hours, season, sales, output) Point of highest value, intensity, achievement, activity, etc. ____

NOUN (roof, wave, cap) Any shape, edge, or part that becomes narrow and pointed ____

NOUN (geography) the pointed top of a mountain _____

RARE

rarity/rareness X rare/ rarefied rarely

(If you were talking about living things)[Use rare as adjective:
... a rare _______ ...

(If you were talking about cooking)

(If you were talking about a special person/entertainer)

ADJ (book, species) unusual, uncommon, one of only a few _____
not often happening or seen

ADJ (steak) lightly cooked meat _____

ADJ (air) thin, light (air of the mountains) _____

ADJ (time at a party, fright, gift for comedy) unusually good ____
extreme, or remarkable

SUBTLE

subtlety/subtleness X subtle subtly

(If you are talking about food)

(If you are talking about communication between people)

(If you were talking about a painting)

ADJ (flavor, aroma) difficult to detect or describe, _____
fine, delicate

ADJ (plan, argument) organized in a clever or complex way _____
not openly obvious

ADJ (mind) able to perceive and describe fine differences _____
clever in noticing and understanding

WHAT IS THE BASE FORM OF **SURGING?**

surge surge surging X

(If you were talking about the natural world [land, mountains, forests, oceans, etc])

(If you were talking about business, finance, or economics)

(If you were talking about people at a big sports or entertainment event)

VERB/NOUN (crowd, tide) (a/to) move forward suddenly and powerfully, in a mass or in waves _____

VERB/NOUN (sales, anger, electricity) (a/to) a sudden great/powerful increase in something _____

TREND

trend/trendiness trend trendy trendily

(If you were talking about economics)

(If you were talking about the clothing industry)

([You are right] Trend can be used as a noun. Now give another sentence with trend, using any situation you like. But you must include an adjective which describes the noun trend (a(n) ___ trend)

NOUN (economic,political, financial) a general tendency or direction _____

NOUN (clothing) a fashion or style (ADJ TRENDY = fashionable) _____

1. **massive**

2. **peak**

3. **rare**

4. **subtle**

5. **surging**

6. **trend**

Appendix 7.1

ABANDON

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:
(If you were on the *Titanic*)

S1:

(If someone does something a bit foolish without any thought, caution, or care)

S2:

(A someone changes their mind about something they wanted to do or something they believed in)

S3:

Grammar: abandon/abandonment * abandoned X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (abandon a baby) leave or desert and not return _________

V (abandon a ship) leave because of danger _________

V (abandon a project) give up before finishing _________

V (abandon a political leader) withdraw help or support from somebody _________

V (abandon yourself to despair) allow oneself to be completely controlled by something _________

N (gay abandon) state where feelings and actions are uncontrolled or uninhibited _________

BROOD

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(When things go badly for someone)

S1:

(If you were talking about a farm)

S2:

(If you were talking about a household)

S3:

Grammar: broodiness/brood * broody/brooding broodily/broodingly

Meanings:

V (problem) spend time anxiously thinking about something ________

V (clouds) to hang closely; menacing, threatening _________

N (ducks) a family of young creatures [birds] ________

N (situation in a household) the children of one family _________

ADJ (horse) animal used for breeding _________

V (hen) sit on eggs in order to hatch them _________

CIRCULATE

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(If you were talking about liquids)

S1:

(If you were talking about exchanging information within a company)

S2:

(If you were the host(ess) of a party which included a lot of friends)

S3:

Grammar: circulation * circular/circulatory X/circularly

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (liquid) to move around in a closed system _________

V (information) disseminate or spread widely _________

V (air) to move about freely _________

V (party) move from one person to the next _________

CONVERT

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(If you were talking about finance)

S1:

(If you were talking about a person's religion)

S2:

(If you were talking about using computer programs)

S3:

Grammar: converter/convert/conversion * converted/convertible X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (sofa) change (be able to be changed) from one form or use into another _________

V (religion) change one's beliefs _________

N (religion) person who has changed beliefs _________

V (football) gain extra points after scoring a goal or touchdown _________

DEDICATE

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(Working for some good cause)

S1:

(If you were talking about the author of a publication)

S2:

(If you were talking about religious matters)

S3:

Grammar: dedication * dedicated X / dedicatedly

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (cause) devote oneself to a good cause _________

V (publication) address one's publication to someone _________

V (church) devote something to a sacred purpose _________

V (money) to set aside something for a particular reason _________

ILLUMINATE

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(If someone was driving a car at night)

S1:

(Trying to learn a difficult idea)

S2:

(A piece of writing from medieval times)

S3:

Grammar: illumination * illuminated/illuminating X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (candle) to give light or cast light on something _________

V (festival) to decorate a street or building for a special occasion _________

V (difficult idea) cause to understand, make clear _________

V (manuscript) decorate a book with gold paint and colors ________

LAUNCH

Spelling: Ø most consonants phonological correct

Associations: ___________ ___________ ___________

Collocations:

(If you were talking about war)

S1:

(If you were talking about new merchandise to sell)

S2:

(If you were talking about the Navy)

S3:

Grammar: launch/launcher * X / (NEWLY) launched X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (attack) put something or somebody into action, begin _________

V (merchandise) make a new product publicly available _________

V (navy) put a new ship into the water _________

V (missile) put something in motion or on its course _________

V (oneself into something) begin enthusiastically something important or something that will take a long time _________

V (out into something [career]) to do something new or more exciting or profitable _________

N (yacht) a large motor boat _________

V (baseball) to throw long, high, or very hard _________

PLOT

Spelling: Ø most consonants phonological correct

Associations: ____________ ____________ ____________

Collocations:

(If you were talking about spies)

S1:

(If you talking of a novel)

S2:

(If you were talking about navigation on a ship)

S3:

Grammar: plot * plotted X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (spy) to make a secret plan to do something _________

N (spy) the secret plan _________

V (graph) to connect a series of points into a curve _________

V (navigation) to mark something on a map _________

N (real estate) a small piece of marked land for a special purpose _________

N (novel) outline of events in a play or novel _________

SPUR

Spelling: Ø most consonants phonological correct

Associations: ___________ ___________ ___________

Collocations:

(The Wild West in America)

S1:

(A boss talking to an employee)

S2:

(If you were talking about trains or trucks)

S3:

Grammar: spur * X / spurred X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

N (horse) horse accelerator _________

V (horse) to use these spurs _________

V (sports coach) to urge or encourage forcefully _________

N (complaints) event or influence that encourages action _________

N (mountains) length of high ground coming out of mountains ______

N (trains) track or road that goes away from the main line _______

N (rooster) back part of a bird's foot _________

SUSPEND

Spelling: Ø most consonants phonological correct

Associations: ___________ ___________ ___________

Collocations:

(A light) [You can use this word in the sentence if you wish]

S1:

(If you were talking about the railways)

S2:

(A professional person caught acting unfairly or dishonestly)

S3:

Grammar: suspension/suspender(s) * suspended X

Meanings:

V (gravity) to hang something from above _________

V (rail services) to temporarily stop or prevent from being in effect _________

V (prison sentence) not enforce, delay, or happen at a later time _________

V (policeman) prevent someone from holding usual position [because of misbehavior] _________

PASSIVE V (dust) hold still in liquid or air _________

TRACE

Spelling: Ø most consonants phonological correct

Associations: ___________ ___________ ___________

Collocations:

(If you were talking about solving a crime)

S1:

(Something involved with art)

S2:

(If you were talking about something (or an organization) that started a long time ago)

S3:

Grammar: trace/tracing * traceable /tracing /traced? X

Meanings: **U** = UNPROMPTED **P** = PROMPTED **X** = DON'T KNOW

V (detective) find something by following their course _________

N (clue) mark or sign of the presence or someone or something _________

V (family tree) going back in time to find the origins or proof of something _________

V (book about the monarchy) follow the course, development, or history of something _________

V (picture) to copy something by drawing its lines on transparent paper _________

N (poison) very small amount of something _________

N (horse) part of the harness which pulls a cart _________

A11 1955 lists

Fligelstone S
From: UCREL
Sent: 24 January 1996 09:33
To: S.Fligelstone
Subject: nouns

Nouns

abdomen	absence	accident	accountant	accrual	acetal
acquital	action	actuary	acumen	adjective	adjunct
adultery	adversary	advice	aeronautics	afterthought	agent
agnomen	agriculture	ai	airline	alibi	alkali
allowance	almanac	also-ran	amytal	anchovy	android
anguish	anniversary	anode	anomaly	ant	apartheid
aphid	apiary	apoplexy	apothecary	appearance	application
approval	aqualung	area	armful	arrival	arsenal
artery	artisan	artist	asbestos	ascent	aspect
aspic	assailant	assistance	assize	astatine	asteroid
astronautics	atlas	attic	auger	autopsy	aviary
awning	background	baize	barbital	barbiturate	barque
barrier	basketful	batt	battery	bean	beater
beauty	beer	behaviour	bellyful	bequest	bequethal
beriberi	berliner	best-seller	betrothal	bevy	bible
bicarbonate	biddy	bier	bikini	bile	bird
bisque	bitumen	biz	blackguard	blighty	bloodshed
bluestocking	boatshed	bodybuilding	bookkeeping	booty	bottleful
boundary	bounder	bounty	boutique	bowstring	boxful
bracken	bran	brandy	bravery	brewery	bribery
bric-a-brac	bricklaying	brier	brigantine	brilliantine	brine
britisher	briton	broccoli	broker	bucketful	buddy
build-up	bummer	bund	bunny	burglar	burglary
burial	bursary	bushelful	butterfly	butty	cabby
calibre	cambric	camomile	campaigner	candy	canteen
canvas	capful	capstan	caracal	caravan	carbohydrate
cardigan	cardiomegaly	caries	casque	catamaran	catechumen
caterer	caterpillar	cavy	ceiling	celebrant	cell
cellar	celluloid	celt	centre-half	cereal	certainty
ch	chaos	charcoal	charlatan	cheetal	cheque
chest	chest-of-drawers	chianti	chiccory	chicory	child
child-in-arms	childbed	childhood	chili	chinaman	chitling
christendom	christianity	christmastime	chronicler	churchwarden	cinque
cistern	citizen	claimant	clan	classmate	clay
cleric	client	climate	clique	co-ordination	coagulant
cochineal	coefficient	cognac	collegian	combination	comedy
commentary	commerce	communication	communism	comparison	concentration
concept	concert	confidant	confidence	conquest	conservatory
consideration	constituent	consultant	consumer	contestant	controversy
contumely	convalescent	conversation	coolant	cormorant	cotangent
counterfeiter	courtesan	courtezan	coven	coward	command
cowling	cowpen	cowshed	cracker	cranny	crest
crew	critic	criticism	critique	crocodile	cruelty
crystal	cuisine	cul-de-sac	cupful	curate	curd
currant	curtesy	cuspid	custodian	cutter	cyclamen
cynic	dairy	dal	dane	danger	darky
daughter	daybed	deary	deathbed	debutant	decal
decrial	defence	defendant	defense	deliverer	delivery
dell	den	denial	dependant	deputy	dervish
descendant	descent	desk	destiny	detective	development
device	dhal	dia	diabetes	dialectic	diary

dichromate	dictionary	difference	difficulty	dignitary	dinghy
dinic	dinner	directive	director	directory	disciplinarian
discoteque	discovery	disease	dishful	disinterest	dispensary
ditty	divan	do-gooder	doctorate	doctrine	doer
doggy	domini	dormitory	drapery	dressmaking	drought
druid	dubliner	duckling	dumpling	dung	duty
dynasty	earful	earring	earthling	effigy	effort
elec	electricity	electronics	elegy	elephant	emblic
emboidery	emf	energy	engine	englishman	ennui
enquiry	entirety	entrant	environment	epoxy	equation
ethic	ethos	event	ewer	examination	example
exchequer	existence	expenditure	expense	expiry	explanation
exponent	extent	eyeful	eyelid	fable	fabric
facsimile	fact	factory	faculty	fae	failure
faith	faker	fall-out	fame	family	fantasist
farthing	featherbed	fellowfeeling	fetish	fibre	fibreglass
figurine	filbert	filly	finial	finn	firefly
first-half	fistful	flan	flattery	flesh	floozy
fly-half	foible	folly	food	footrest	forbiddal
foreigner	forethought	forget-me-not	formant	foundation	foundling
fourpence	foursome	frailty	free-for-all	free-style	freebooting
frenchman	frenzy	frequenter	freshman	friary	fromenty
fumitory	furmenty	furniture	fwd	gable	gadfly
gal	gala	galaxy	garagehand	gardener	garlic
gasoline	gelatine	generation	gent	gentian	geriatrics
ghq	gi	girl	glass	glassful	glossary
glutamate	gluten	glycerine	goal	godparent	goer
golf	gosling	gourdful	gp	gradient	graffiti
grain	grammarian	grandparent	granite	granny	gravy
greensward	grid	guardian	guile	gully	guppy
gypsy	haemophiliac	hair	hairdressing	hairtonic	half-time
handful	handmaiden	handshaking	handspring	handwriting	hanky
hara-kiri	harpy	has-been	hat	hatful	hatred
haymaking	headspring	heart	heaven	helpmate	hemophiliac
hen	heraldry	heresy	heretic	hexane	hic
hickory	hide-and-seek	highlander	historian	history	hobby
hoky-poky	holly	homebuilding	homecoming	homily	hominy
hood	hooligan	hotbed	houri	houseful	household
housekeeping	ht	hubby	hurricane	hussy	hymen
hyphen	icelander	id	idea	ifs	ignominy
ill-feeling	imagination	importance	improvement	incentive	incest
independence	industry	infant-in-arms	infighting	infinitive	informant
information	ingredient	inhabitant	initiative	inkling	inning
inquest	inquiry	insomniac	instinct	insurance	interim
interlude	interpretation	intestine	investigation	iq	irishman
irs	ivory	jackal	jaloppy	jarful	jealousy
jelly	jellybean	jeopardy	jew	jewry	jiffy
journal	jugful	kelt	kindergarten	kinetics	kitchen
kitten	kiwi	knickknackator y	knock-out	know-all	know-how
knowledge	kraken	lab	laboratory	ladleful	language
lapful	larceny	latrine	latticing	lavatory	lawyer
leader	left-half	lefty	lens	leprosy	levi
liane	lib	liberty	librarian	lichen	lichi
lid	life	lily	linen	linguistics	literature
location	lockful	logic	lolly	londoner	longhand
lorry	lory	loss	lottery	lp	lumen
luminal	lung	lungful	mac	macaroni	madrigal
magistrate	magnate	mains	mainspring	maintenance	maize
make-believe	make-up	mammal	manager	mandible	mane
manganate	manling	manner	margarine	marketer	marriage

martini	masque	mastery	mathematics	maverick	mayer
mayfly	medal	mede	medlar	memory	merchant
merry-making	merrymaking	mescal	message	messmate	metalworking
method	mezcal	midst	misery	missile	missive
missy	mistery	mistrial	mitten	molasses	monastery
money	monopoly	moorhen	motive	motorboating	mouthful
murderer	music	mutiny	mystery	mystique	nai*?31vety
naivety	nappy	nation	nature	navvy	ne'er-do-well
needleful	nestful	netful	newlywed	news	newsagent
newspaper	nicety	nicknackatory	nicotine	nighty	ninny
nitrate	nitroglycerine	northerner	notary	novelty	numismatics
nursery	nyasalander	observatory	obstetrics	occident	occupant
odal	offal	office	offspring	oilheating	oleate
omen	op	opal	opera	operation	optics
oratory	orcharding	orchid	ordeal	ordinate	organ
orthopedy	outbuilding	outing	outsider	ovary	oven
overflying	owner	paederasty	pailful	pal	palmful
palsy	pantile	paperhanging	parable	parent	parish
participant	partizan	pastern	pate	pathos	peace
peahen	peasant	pedant	pederasty	pedestal	pennant
pentothal	pepperoni	percentile	perfidy	performance	pergatory
permanganate	person	perspective	pessary	pest	petal
ph	pheasant	phenobarbital	phial	philately	phonetics
physics	physique	pic	piccalilli	piccaninny	piggy
piggy-wiggy	pigpen	pigsty	pillar	pillory	pinny
pipeful	pitcherful	pittacal	piw	plagiary	plaid
plaque	plateful	platform	playmate	playpen	plaything
playwriting	pleasure	plebiscite	pocketful	poet	poggy
polaroid	polemic	politicking	pollen	pomegranate	poplar
poppy	population	portal	post-card	postern	posy
potpourri	pottery	poverty	prayer	prelude	premolar
preparation	presence	presenter	price-range	priest	primate
primer	principle	priory	problem	procedure	procuratory
prodigy	product	projectile	promontory	propellant	property
propr	proxy	psoas	puberty	publication	pudding
puppy	purseful	pussy	putty	quad	quadrant
quandary	quartile	quid	quint	quotient	rabies
racemate	radian	radiation	radish	railroad	rascal
rebuttal	recital	recovery	reflectory	refreshant	refutal
regimen	relation	relic	remnant	rental	reporter
repository	reptile	republic	requital	resistance	response
restaurant	retrial	revery	rhetoric	rhinoceros	ribaldry
right-half	rinderpest	riw	roadbed	roadbuilding	robbery
roleplaying	roomful	roommate	rosary	rotodial	rsj
rubber	rubbish	rubric	ruffian	rumen	rutile
sac	sackful	sacristan	sacristy	safari	safekeeping
safety	saki	salami	salary	salmagundi	salmi
sanctuary	sapling	sari	sassafras	saucerful	scaffold
scaffolding	scene	sceptic	schnapps	scholar	schoolmate
science	scotsman	scrum-half	scrutiny	sculduddery	second-half
seconder	sedan	seedbed	seedling	seer	seljuk
semantics	semen	semi-final	seminar	serpent	servant
setter	sextant	sextile	shandy	shanty	shareholding
sheen	shelfful	shilling	shipbuilding	shipmate	shoestring
shortcoming	shorthand	shovelful	showcasing	shrubbery	sial
sibling	significance	simile	sine	situation	sixpence
skeptic	skinful	skivvy	slavery	sleeping-quarter	slogan
smelter	smithy	snobbery	sodomy	softy	sonar
song	sonny	source	southerner	sovreignty	sowbelly
sp	spaghetti	spaniard	spate	specimen	speech

spinning-jenny	spleen	splenomegaly	spoonful	sporran	starboard
starling	starveling	statistic	stepparent	stickful	stile
stillatory	stockholding	stomachful	story	strategy	strid
student	sty	subaltern	subdean	subprincipal	subsidy
subtlety	success	suede	sufferer	sulfate	sultanate
sunrising	sunsetting	supervisor	suppository	suppressant	surety
swami	sward	syllable	symbol	syringeful	system
tableful	tablespoonful	tabloid	tactic	talent	tambourine
tankful	tanner	tariff	tea	teacupful	teammate
teaspoonful	technique	teen	telepic	telly	temperature
tempest	templar	territory	testbed	text	textile
thanksgiving	theic	theologian	theory	therapeutics	thermos
thimbleful	thing	thingummy	thingy	threepence	threesome
throat	thunderpeal	tiding	tier	tigerfly	tiller
timbre	time-scale	timekeeping	tincal	ting	tithing
toddy	toffy	token	toothful	topic	torrent
tragedy	trajectory	treaty	tribunal	tribunate	trimaran
tripos	trooper	trouser-press	trunkful	tubful	tumblerful
tummy	tungsten	tunic	tunny	tureen	turk
turnstile	turntable	turpentine	tv	twopence	twosome
tyranny	tyrant	udderful	uncertainty	underbelly	underclothing
underling	unit	universe	unrest	upbringing	upholstery
uprising	uprootal	upswing	urinal	urine	user
vac	valentine	vane	veal	velveteen	vermicelli
vesselful	vial	victory	villainy	vip	virial
vixen	vol	volume	walloon	warranty	watershed
wc	weakling	wearsider	welfare	well-being	wellbeing
whaling	wharfing	wherewithal	whisky	wholesaler	wick
wicking	wife	wildfowling	window	wineglassful	wing-half
wizardry	wolfling	woman	woodshed	woodworking	workmate
worktable	world	writer	wrongdoing	wt	yachtbuilding
yearling	yogi	yokemate	yorkshireman	zeal	zest
zodiac					

Fligelstone S
From: UCREL
Sent: 24 January 1996 09:32
To: S.Fligelstone
Subject: verbs

Verbs

abhor	accept	accompany	achieve	add	adhere
administer	afford	agree	allow	annoy	apply
apportion	ask	assume	attest	augment	avoid
awaken	backbite	balden	beckon	befriend	beguile
belie	believe	bemoan	beware	blacken	bless
borrow	bring	broaden	bury	cater	cf
choose	cling	coexist	cohere	comest	commence
compete	compile	comply	conceal	congeal	consider
continue	convey	create	darken	deaden	deafen
decant	decide	decry	delete	demean	deny
deplete	deploy	destine	destroy	deter	determine
detest	develop	disable	disagree	disobey	divest
double-cross	eat	embed	embody	enable	enchant
encircle	enfold	enhance	enjoy	enlist	enliven
ensure	enter	envision	err	establish	excite
excrete	exist	expect	expedite	explain	extradite
flee	follow	forbid	forego	foresee	forget
forgo	freshen	galavant	gallivant	get	gird
gladden	glean	go	grow	harden	hear
hearest	hearken	ignite	imagine	impale	implete
imply	imprison	inbring	incite	indicate	indite
infest	inhale	install	interfere	invest	invite
learn	lengthen	lessen	liken	linger	loosen
lose	madden	maintain	mar	marry	molest
obey	obtain	occupy	occur	operate	outdo
outfox	outgeneral	outlie	outnumber	outrival	outspring
outvie	overburden	overdo	overwrite	owe	peeter
persevere	pester	preen	prefer	preoccupy	preside
prevent	prophesy	prosper	provide	quicken	reawaken
rebite	recite	reckon	recommence	recomplete	reconcile
reconnoitre	redden	redeploy	refer	refinance	regale
reignite	reincite	rely	remain	remember	repent
represent	require	requite	reseal	reside	retard
retraverse	retry	reunite	reveal	revile	rewrite
rile	roister	roughen	rove	ruminate	sadden
satisfy	scissor	secrete	see	seek	seem
send	shanghai	side-step	sing	sit	slacken
smite	speak	stiffen	strafe	strengthen	subside
succeed	suggest	sully	summon	sun-bathe	tell
thank	thicken	threaten	toughen	unbalance	uncongeal
undergo	underlie	understand	underwrite	undo	unearth
unfold	unharness	unite	unlock	unmoor	unseal
unstring	untie	vanish	vary	veer	vie
waken	weaken	wean	widen	woo	worsen
wrest	write				

Fligelstone S

From: UCREL
Sent: 24 January 1996 09:33
To: S.Fligelstone
Subject: adjs

Adjectives

+ve	-ve	abeyant	abhorrent	abject	abrupt
abs	absentminded	absurd	abundant	accurate	accursed
acrid	actual	acute	adequate	adjoint	administrative
adroit	affectionate	aforethought	afraid	age-old	agile
airy	aisled	alkaline	ample	angry	anterior
apparent	appetizing	apposite	apt	arbitrary	arcaded
ardent	arid	arrant	arrogant	arrowed	ashen
astute	attractive	attrite	atune	austere	avid
aware	bald	balding	balmy	bandy	barefaced
barefooted	barmy	barren	basic	basilican	batty
beady	beautiful	bebearded	beefy	bellicose	bendy
benign	beribboned	bespectacled	bestselling	bewhiskered	biconcave
bigoted	bigotted	bipartisan	bipartite	bipartizan	bittersweet
bizarre	bland	blatant	bleak	bleary	blithe
blobby	bloodshot	bloodstained	blurry	bogus	bold
bonafide	bony	bossy	bouncy	bovine	brash
brawny	breathtaking	breathy	briery	bright	brilliant
brisk	brittle	broadminded	brocaded	brusque	bugeyed
buoyant	buttony	buxom	cadenced	cagey	cagy
calceate	callow	candid	canine	canny	careful
causal	central	certain	chambered	chaste	cheesy
chilly	chockful	choosy	chubby	circumspect	citybred
civil	clandestine	clayey	clear-cut	clever	clothbound
cloudy	clumsy	coarse	cobwebby	cochlear	cockeyed
cocksure	coercive	cogent	coherent	collegiate	collonaded
collosal	colonnaded	colossal	comely	comfy	complaisant
complementary	complexioned	complimentary	concave	conceited	concise
conducive	considerate	contradictory	contrite	convenient	conversant
convex	coolheaded	coplanar	coppery	corporate	costly
cottony	courtly	cowardly	coy	crabby	crass
creepy	crinkly	crispy	cruel	crumbly	crunchy
crystalline	culinary	cultural	cunning	curly	curt
cushiony	cute	cyclic	daffy	daft	dapper
darkhaired	dastardly	dbl	deaf	debonair	decrepit
defiant	definite	deflationary	deft	delapidated	delicate
democratic	demure	dense	desperate	despondent	destitute
determinate	devoid	devout	different	difficult	dim
dire	disciplinary	discontented	discordant	discreet	discrete
disgruntled	dishonest	disinterested	disobedient	disproportionate	distant
distrait	distraught	dotty	douloureux	dour	downcast
draconian	dreary	drowsy	drunken	dumb	duplex
eager	eared	earsplitting	earthbound	earthen	earthly
earthmoving	earthy	easygoing	eerie	eery	effective
effete	electoral	electronic	elegant	elementary	eminent
equine	erudite	everlasting	evident	exemplary	exorbitant
exp	explicit	exquisite	extant	extravert	exuberant
exultant	facaded	facile	faddy	famed	farfetched
farsighted	faulty	feathery	federal	feeble	fervent
fervid	fibry	fickle	fidgety	fierce	fiery
financial	fingery	finite	firy	fisted	flabby

flaccid	flippant	flirty	florid	floury	flowery
fluffy	folksy	fond	foolhardy	footloose	foreign
foreknown	forlorn	formal	forthcoming	forthright	foxy
fragile	fragmentary	fragrant	frail	fraught	fresh
frigid	frosty	frothy	fully-fledged	funny	fussy
gabby	gangling	gassy	gauche	gaunt	genteel
gentle	genuine	ghastly	ghostly	giddy	gifted
gingery	glad	glady	glassy	glib	glimmery
glossy	gloved	glum	godforsaken	godly	golden
goodly	gory	gossamery	grabby	grandiose	grassy
greasy	greatcoated	grim	grisly	gritty	groovy
grouchy	grubby	gruff	guilty	hackneyed	haggard
haired	hairy	halfhearted	handworking	handy	haphazard
hardboiled	harsh	headachy	headstrong	heady	heartbreaking
hearted	heartfelt	heartrending	heavenly	helmeted	hempen
hereditary	hesitant	hidebound	highspeed	hlf	hoarse
hoary	homebound	homely	honest	honeycoloured	honorary
hopeful	horrid	hot	huge	hulking	humane
humdrum	humid	hungry	hyperfine	icy	iffy
ill-at-ease	illegitimate	illicit	imaginary	immaculate	immanent
immature	immediate	immense	imminent	immobile	immodest
immune	impassionate	impassioned	impending	implicit	impolite
important	imprecise	improper	imprudent	impudent	inaccurate
inadequate	inane	inanimate	inappropriate	inapt	inarticulate
inboard	inbuilt	inc	incertain	incessant	incomplete
incondite	inconsiderate	inconvenient	indebted	indefinite	indelicate
indeterminate	indignant	indiscreet	indiscrete	indiscriminate	industrial
inelegant	inept	inert	inexact	inexperienced	inexpert
infinite	infirm	inflationary	infrared	ingoing	inherent
inhibitory	inhumane	innate	inobservant	inopportune	inordinate
insane	insecure	insincere	intact	intense	intentioned
intercollegiate	intercrystalline	interplanetary	intranasal	intricate	inveterate
irate	irrelevant	irreverent	italianate	jackbooted	jaundiced
jaunty	jocose	jubilant	juicy	lacey	lacy
languid	large	latent	latinate	lax	leaded
leaden	leafy	legal	lemony	lengthy	lenient
levelheaded	lewd	lifelong	literary	lithe	lively
livid	lollopy	lonely	long-term	longstanding	longsuffering
longtime	lordly	lovelorn	lucid	lukewarm	lunar
lurid	luxuriant	macabre	mad	maggoty	makeshift
maladjusted	maladroit	malefic	malformed	malnourished	malty
mangy	manifold	mannered	maritime	matter-of-fact	maudlin
meager	meagre	mealynosed	meaty	mediocre	meek
menopausal	mental	merry	messy	middle-aged	miniscule
minty	miry	misshapen	misty	mod	modest
moist	molten	monetary	monocled	montane	morbid
morose	mottled	mouldy	moustached	muddleheaded	muddy
multicoloured	multipartite	mundane	mustachioed	nai*?31ve	naive
naked	nasal	natured	naval	nearsighted	neat
necessary	nervy	new	nice	noisy	nonchalant
nonchurchgoing	noncommissioned	nondescript	nondrying	nonsegregated	nosy
nubile	null	oaken	obdurate	obedient	obese
obscene	obsolete	obtuse	odd	offbeat	offsaddled
oily	old	old-fashioned	olden	olivefaced	olympic
oncoming	oniony	opportune	ordinary	ornate	ornery
orthodox	otherworldly	outdoor	outgoing	outmoded	outspoken
outspread	outstanding	outstretched	overblown	overdue	overeager
overlarge	overripe	oversoft	overt	painstaking	pallid
paltry	paramount	parliamentary	particoloured	partite	passionate
pearly	peaty	peculiar	pedigreed	pedimented	peppery

peremptory	permanent	perseverant	personal	pert	petite
pettifogging	petty	phantomy	philippine	pigtailed	pimpled
pinkfooted	pithy	placid	planar	pleasant	poignant
pokerfaced	polite	polycrystalline	polysaturated	poor	popular
porridgey	portly	posh	potent	powdery	powerful
prebendal	precise	predominant	pregnant	premature	prickly
priestly	prim	princely	prissy	pristine	profound
profuse	prone	proper	provincial	prudent	pubendal
puckery	pungent	puny	purblind	pure	pursuant
putrid	quaint	queasy	rabbity	rabid	racy
raffish	rampant	ramshackle	rancid	rapt	rare
rawboned	recondite	recumbent	redheaded	redundant	reefy
relevant	reliant	reluctant	remote	renowned	repentant
repugnant	resilient	resplendent	reverberant	reverent	ribald
ricketty	rigid	ripe	ritzy	robust	rockbound
ropy	rosy	rotten	rotund	rougy	royal
rubbery	ruddy	rude	rudimentary	rugged	rural
s/c	sacred	sad	saffrony	salty	sane
sanitaire	satisfactory	saucy	saudi	scabby	scanty
scary	scatterbrained	scholarly	scrawny	scrunchy	secondary
sedate	selfeffacing	selfish	seljukid	semidrying	semisecret
senile	sensitive	serene	severe	sexual	sexy
shabby	shady	shamefaced	shapely	shimmery	shiny
ship-shape	shivery	shoddy	shortsighted	showery	shrewd
shrill	sickening	significant	silken	silly	silvery
similar	sincere	sinister	skinny	sleepy	slender
slimy	slippery	slipshod	slovenly	sly	smarmy
smeary	smug	smutty	snooty	snowbound	soapy
soft	solemn	solidary	somber	sombre	sooty
sophisticated	sordid	sorry	southbound	sparse	spavined
spicy	spiny	splendid	spongy	sporty	spotty
squalid	squelchy	squiffy	stagnant	stale	starchy
starry	stately	stationary	statutory	steadfast	steely
sterile	stolid	stony	storied	stormbound	stormy
strange	stricken	strict	strident	stringent	stripy
stuffy	sturdy	suave	subcuticular	sublime	substandard
substantial	subtle	suburban	successful	sudden	sudsy
sulfury	sullen	sultry	summery	sunburnt	sundry
sunny	superb	supercharged	supine	supple	surgy
surpliced	swarthy	sweaty	sweltry	swoony	syrupy
tacit	taffycolored	talented	tall	tangy	taut
techy	temperate	temporary	tensile	terse	thickset
thirsty	thorough	threadbare	throaty	thunderstruck	tiered
timely	timid	tinny	tiny	tipsy	tolerant
torrid	touchy	tr	tranquil	transfinite	transitory
transparent	trite	triumphant	tubby	turbanned	turbid
turgid	tuxidoed	twangy	ugly	ulterior	ultimate
ultramodern	unaccommodating	unaccounted	unassuming	unavailing	unaware
unbecoming	unbelieving	unbid	unblushing	unbroken	uncandid
uncanny	uncast	unceasing	uncertain	unchanging	uncharged
unchosen	uncivil	unclean	uncommon	uncomprehending	uncompromising
unconvincing	uncouth	uncrossing	undecided	under-age	undiscriminating
undreamt	undriven	undue	undying	unearthly	uneasy
uneaten	unemployed	unending	unerring	uneven	unexciting
unfailing	unfair	unfamiliar	unfeeling	unfelt	unfeminine
unflattering	unflinching	unforgiven	unforgotten	unfriendly	unfrozen
ungainly	ungallant	ungentle	ungodly	unheard	unheard-of
unhearing	unhesitating	unhurt	unimportant	unimposing	uninteresting

unique	unjust	unkempt	unkind	unknowing	unknown
unlikely	unlovely	unmixing	unmoving	unmurmuring	unnecessary
unoccupied	unopen	unorthodox	unpatronizing	unpleasant	unprecedented
unpretending	unpromising	unprovoked	unquestioning	unquiet	unread
unreasoning	unrelenting	unremitting	unrevealing	unrewarding	unripe
unrisen	unsafe	unsatisfactory	unseeing	unselfish	unshrinking
unshriven	unsleeping	unsmiling	unsober	unsold	unsound
unspoilt	unspun	unstaring	unstartling	unsteady	unsubtle
unsung	unsure	unsuspecting	unswept	unswerving	unthinking
untidy	untimely	untiring	untold	untrue	untwinkling
unusual	unvarying	unwarrented	unwavering	unwelcome	unwieldy
unwilling	unwinking	unwitting	unworldly	unwritten	unyielding
up-to-date	upcoming	uppity	urban	urbane	urgent
useful	usual	uterine	vacant	vague	vain
valid	varicolored	varicose	varitinted	vast	vehement
velvety	verbose	vigilant	vile	vinegary	virile
vital	vivid	vulgar	vulpine	warmhearted	wartorn
warty	wary	waterlogged	watery	wavy	weak
weary	webby	weird	well-known	well-off	well-to-do
wellknown	whiskered	whitehaired	wholehearted	wicked	widespread
wifely	wily	windblown	windowed	windy	wintery
wintry	wiry	wispy	witty	woebegone	wolfish
wonderful	wooded	wooden	wordy	worldly	wornout
worth-while	wretched	wrinkly	wry	zesty	

Fligelstone S
From: UCREL
Sent: 24 January 1996 09:33
To: S.Fligelstone
Subject: advs

Adverbs

'most	a'right	afresh	almost	aloud	already
also	always	amok	anew	anyhow	anymore
anyway	apace	approx	aright	askance	aslant
awhile	basically	bitterly	certainly	clearly	deeply
doubtless	ever	everytime	fain	faintly	fairly
finally	forever	forthwith	fro	furthermore	greatly
headfirst	hence	henceforth	hereby	herein	hereinafter
hereof	hereto	hereunto	herewith	hitherto	humanly
incidentally	instead	latterly	maybe	meanwhile	mid-way
moreover	never	nevertheless	nonetheless	obviously	oft
often	often-times	oftentimes	on-stream	onstream	otherwise
overly	perchance	perforce	perhaps	properly	regardless
roughshod	secondly	seldom	seriously	sic	similarly
simply	slightly	slowly	sofar	someday	somehow
sometimes	somewhat	somewheres	soon	specifically	straightaway
strangely	strongly	subsequently	surely	thence	thenceforward
thereabouts	thereafter	thereby	therefor	therefore	therefrom
therein	thereof	thereon	thereto	thereunder	thereupon
therewith	thrice	thus	truly	twice	unawares
underway	utterly	well-nigh	whence	whenever	whereby
wherein	whereof	whereon	wherever	wherewith	whither
wholly	why	willy-nilly	withal		

SPELLING: How do you spell ______?

Writing is OK.

ASSOCIATIONS: First 3 words you think of.

MAKING SENTENCES:

Don't use hint words in sentence.

Typical sentence using **related** words.

Key word plus (**-ed**, **-ing**, **-s**) is OK

If you don't know meaning, tell me.

GRAMMATICAL: What part-of-speech is _____?

(noun, verb, adjective, adverb)

If no grammatical form, tell me.

MEANINGS: Any description is OK.

If you don't know a meaning, tell me.

Appendix 8.1

KEY:

There are 110 cases in each scalogram (3 subjects × 11 words × 3 sessions + 1 subject [Kor] × 11 words × 1 session). Each individual case is coded for the subject, session, and target word.

K 1 BR

Subject Name	Session	Target Word
D = Lith	1 = T1	AB = abandon
K = Ind	2 = T2	BR = brood
M = Kor	3 = T3	CI = circulate
G = Tai		CO = convert
		DE = dedicate
		IL = illuminate
		LA = launch
		PL = plot
		SP = spur
		SU = suspend
		TR = trace

Thus D2CO = Lith's responses for *convert* in the second session (T2).

The bottom two rows indicate the number of subjects demonstrating sufficient knowledge to reach the cut-point criterion (1) for each word knowledge category, and also the number of those who did not (0).

Errors are indicated by being circled. The tally of errors for a scalogram is given at the bottom.

	MEANING	GRAMMAR	COLLOCATION	ASSOCIATION	SPELLING	TOTAL PER WORD
K1BR	0	0	0	0	0	0
K2BR	0	0	0	0	0	0
K3BR	0	0	0	0	0	0
G1BR	0	0	0	0	0	0
G1LA	0	0	0	0	0	0
G1SP	0	0	0	0	0	0
G2BR	0	0	0	0	0	0
D1BR	0	0	0	0	0	0
D1SP	0	0	0	0	0	0
D2BR	0	0	0	0	0	0
D2SP	0	0	0	0	0	0
M1PL	0	0	0	0	0	0
K1AB	0	0	0	0	1	1
K1CI	0	0	0	0	1	1
K1SU	0	0	0	0	1	1
K2AB	0	0	0	0	1	1
K2CO	0	0	0	0	1	1
K2SP	0	0	0	1	0	1
K2SU	0	0	0	0	1	1
K2TR	0	0	0	0	1	1
K3AB	0	0	0	0	1	1
K3SP	0	0	0	1	0	1
K3SU	0	0	0	0	1	1
G1AB	0	0	0	0	1	1
G1CO	0	0	0	0	1	1
G1PL	0	0	0	0	1	1
D1CI	0	0	0	0	1	1
D1LA	0	0	1	0	0	1
D1PL	0	0	0	0	1	1
D2PL	0	0	0	0	1	1
D2CI	0	0	0	0	1	1
M1BR	0	0	0	1	0	1
M1SU	0	0	0	0	1	1
M1TR	0	0	0	0	1	1
M1IL	0	0	0	0	1	1
K1IL	0	0	0	1	1	2
K1LA	0	0	1	1	0	2
K1PL	0	0	0	1	1	2
K1SP	0	0	0	1	1	2
K2CI	0	0	0	1	1	2
K2IL	0	0	0	1	1	2
K2LA	0	0	1	1	0	2
K2PL	0	0	0	1	1	2
K3CO	0	0	0	1	1	2
K3IL	0	0	0	1	1	2
K3LA	0	0	1	1	0	2
K3PL	0	0	1	0	1	2
K3TR	0	0	1	0	1	2
G1IL	0	0	1	0	1	2
G1SU	0	0	0	1	1	2
G1TR	0	0	0	1	1	2
G2AB	0	0	1	0	1	2
G2CI	0	0	0	1	1	2
G2CO	0	0	0	1	1	2
G2DE	0	0	1	0	1	2
G2IL	0	0	0	1	1	2
G2PL	0	0	0	1	1	2
G2SU	0	0	0	1	1	2
G3AB	0	1	0	0	1	2
G3BR	0	0	0	1	1	2
G3CO	0	0	0	1	1	2
G3PL	0	0	0	1	1	2
G3SP	0	1	0	0	1	2
G3SU	0	0	0	1	1	2
D1AB	0	0	0	1	1	2
D1CO	0	0	1	0	1	2
D1IL	0	0	1	1	0	2
D1SU	0	0	0	1	1	2
D1TR	0	0	1	0	1	2
D2AB	0	0	0	1	1	2
D2CO	0	0	1	0	1	2
D2IL	0	0	1	1	0	2
D2LA	0	0	0	1	1	2
D2SU	0	0	0	1	1	2
D2TR	0	0	0	1	1	2
D3AB	0	0	0	1	1	2
D3BR	0	0	0	1	1	2
D3CI	0	0	0	1	1	2
D3CO	0	0	0	1	1	2
D3DE	0	0	0	1	1	2
M1CI	0	0	0	1	1	2
M1CO	0	0	0	1	1	2
M1DE	0	1	0	0	1	2
M1LA	0	0	0	1	1	2
M1SP	0	0	0	1	1	2
D3PL	0	0	0	1	1	2
D3SP	0	1	0	0	1	2
D3SU	0	0	0	1	1	2
M1AB	0	0	1	0	1	2
K1CO	0	1	1	0	1	3
K1DE	0	0	1	1	1	3
K1TR	0	0	1	1	1	3
K2DE	0	0	1	1	1	3
K3CI	0	0	1	1	1	3
K3DE	0	0	1	1	1	3
G1CI	0	0	1	1	1	3
G1DE	0	0	1	1	1	3
G2LA	0	0	1	1	1	3
G2SP	0	1	0	1	1	3
G2TR	0	0	1	1	1	3
G3CI	0	1	0	1	1	3
G3DE	0	0	1	1	1	3
G3IL	0	0	1	1	1	3
G3LA	0	1	0	1	1	3
G3TR	0	1	0	1	1	3
D1DE	0	0	1	1	1	3
D2DE	0	0	1	1	1	3
D3IL	1	0	1	1	0	3
D3TR	0	0	1	1	1	3
D3LA	0	1	1	1	1	4
1s	1	10	32	64	88	
0s	109	100	78	46	22	

MAXIMUM POSSIBLE SCORE (MAX)

ERRORS = 56

	ASSOCIATION	MEANING	COLLOCATION	SPELLING	GRAMMAR	TOTAL PER WORD
K1BR	0	0	0	1	1	2
K2BR	0	0	0	1	1	2
K3BR	0	0	0	1	1	2
G1SP	0	0	0	1	1	2
D3SP	0	0	0	1	1	2
D2BR	0	0	0	1	1	2
G1BR	0	0	1	1	1	3
D1BR	0	0	1	1	1	3
D1PL	0	0	1	1	1	3
D2SP	0	0	1	1	1	3
M1PL	0	0	1	1	1	3
K1SP	0	1	1	1	1	4
K2AB	0	1	1	1	1	4
K2SP	0	1	1	1	1	4
K3AB	0	1	1	1	1	4
K3SP	0	1	1	1	1	4
K3SU	0	1	1	1	1	4
G1LA	0	1	1	1	1	4
G1TR	0	1	1	1	1	4
G2BR	0	1	1	1	1	4
G2DE	0	1	1	1	1	4
G2SP	0	1	1	1	1	4
G3SP	0	1	1	1	1	4
G3TR	0	1	1	1	1	4
D1LA	0	1	1	1	1	4
D1SP	0	1	1	1	1	4
D1TR	0	1	1	1	1	4
D2PL	0	1	1	1	1	4
D3PL	0	1	1	1	1	4
D3SU	0	1	1	1	1	4
M1IL	0	1	1	1	1	4
M1SU	0	1	1	1	1	4
K1AB	1	1	1	1	1	5
K1CI	1	1	1	1	1	5
K1CO	1	1	1	1	1	5
K1DE	1	1	1	1	1	5
K1IL	1	1	1	1	1	5
K1LA	1	1	1	1	1	5
K1PL	1	1	1	1	1	5
K1SU	1	1	1	1	1	5
K1TR	1	1	1	1	1	5
K2CI	1	1	1	1	1	5
K2CO	1	1	1	1	1	5
K2DE	1	1	1	1	1	5
K2IL	1	1	1	1	1	5
K2LA	1	1	1	1	1	5
K2PL	1	1	1	1	1	5
K2SU	1	1	1	1	1	5
K2TR	1	1	1	1	1	5
K3CI	1	1	1	1	1	5
K3CO	1	1	1	1	1	5
K3DE	1	1	1	1	1	5
K3IL	1	1	1	1	1	5
K3LA	1	1	1	1	1	5
K3PL	1	1	1	1	1	5
K3TR	1	1	1	1	1	5
G1AB	1	1	1	1	1	5
G1CI	1	1	1	1	1	5
G1CO	1	1	1	1	1	5
G1DE	1	1	1	1	1	5
G1IL	1	1	1	1	1	5
G1PL	1	1	1	1	1	5
G1SU	1	1	1	1	1	5
G2AB	1	1	1	1	1	5
G2CI	1	1	1	1	1	5
G2CO	1	1	1	1	1	5
G2IL	1	1	1	1	1	5
G2LA	1	1	1	1	1	5
G2PL	1	1	1	1	1	5
G2SU	1	1	1	1	1	5
G2TR	1	1	1	1	1	5
G3AB	1	1	1	1	1	5
G3BR	1	1	1	1	1	5
G3CI	1	1	1	1	1	5
G3CO	1	1	1	1	1	5
G3DE	1	1	1	1	1	5
G3IL	1	1	1	1	1	5
G3LA	1	1	1	1	1	5
G3PL	1	1	1	1	1	5
G3SU	1	1	1	1	1	5
D1AB	1	1	1	1	1	5
D1CI	1	1	1	1	1	5
D1CO	1	1	1	1	1	5
D1DE	1	1	1	1	1	5
D1IL	1	1	1	1	1	5
D1SU	1	1	1	1	1	5
D2AB	1	1	1	1	1	5
D2CI	1	1	1	1	1	5
D2CO	1	1	1	1	1	5
D2DE	1	1	1	1	1	5
D2IL	1	1	1	1	1	5
D2LA	1	1	1	1	1	5
D2SU	1	1	1	1	1	5
D2TR	1	1	1	1	1	5
D3AB	1	1	1	1	1	5
D3BR	1	1	1	1	1	5
D3CI	1	1	1	1	1	5
D3CO	1	1	1	1	1	5
D3DE	1	1	1	1	1	5
D3IL	1	1	1	1	1	5
D3LA	1	1	1	1	1	5
D3TR	1	1	1	1	1	5
M1AB	1	1	1	1	1	5
M1BR	1	1	1	1	1	5
M1CI	1	1	1	1	1	5
M1CO	1	1	1	1	1	5
M1DE	1	1	1	1	1	5
M1LA	1	1	1	1	1	5
M1SP	1	1	1	1	1	5
M1TR	1	1	1	1	1	5
1s	89	97	99	108	108	
0s	21	13	11	2	2	

POSSIBLE SCORE (MIN)   ERRORS = 30

	GRAMMAR	MEANING	ASSOCIATION	COLLOCATION	SPELLING	TOTAL PER WORD
K1BR	0	0	0	0	0	0
K2BR	0	0	0	0	0	0
K3BR	0	0	0	0	0	0
G1BR	0	0	0	0	0	0
G1SP	0	0	0	0	0	0
G2BR	0	0	0	0	0	0
D1BR	0	0	0	0	0	0
D2BR	0	0	0	0	0	0
D2SP	0	0	0	0	0	0
M1PL	0	0	0	0	0	0
K1AB	0	0	0	0	1	1
K2AB	0	0	0	0	1	1
K3AB	0	0	0	0	1	1
G1LA	1	0	0	0	0	1
G1PL	0	0	0	0	1	1
D1SP	0	0	0	1	0	1
D2PL	0	0	0	0	1	1
M1SU	0	0	0	0	1	1
K2SP	0	0	1	1	0	2
K3PL	0	0	0	1	1	2
K3SP	0	0	1	1	0	2
K3TR	0	0	0	1	1	2
G1AB	0	0	0	1	1	2
G1CO	0	1	0	0	1	2
G1TR	0	0	1	0	1	2
G2AB	0	0	0	1	1	2
D1LA	1	0	0	1	0	2
D1PL	1	0	0	0	1	2
D3BR	0	0	1	0	1	2
D3SP	1	0	0	0	1	2
M1BR	1	0	1	0	0	2
M1DE	1	0	0	0	1	2
M1IL	0	0	0	1	1	2
M1TR	0	1	0	0	1	2
K1CI	0	1	0	1	1	3
K1LA	0	1	1	1	0	3
K1PL	0	1	1	0	1	3
K1SP	0	0	1	1	1	3
K1TR	0	0	1	1	1	3
K2SU	0	1	0	1	1	3
K2TR	1	0	0	1	1	3
K3IL	0	1	1	0	1	3
K3LA	0	1	1	1	0	3
G1CI	0	0	1	1	1	3
G1DE	0	0	1	1	1	3
G1IL	0	1	0	1	1	3
G2CI	0	0	1	1	1	3
G2DE	0	1	0	1	1	3
G2PL	0	1	1	0	1	3
G2SP	1	0	1	0	1	3
G3AB	1	0	0	1	1	3
G3BR	1	0	1	0	1	3
G3SU	0	1	1	0	1	3
D1AB	0	0	1	1	1	3
D1CI	0	1	0	1	1	3
D1CO	0	1	0	1	1	3
D2CI	0	1	0	1	1	3
D1TR	1	0	0	1	1	3
D2IL	0	1	1	1	0	3
M1CI	0	0	1	1	1	3
K1CO	1	1	0	1	1	4
K1DE	0	1	1	1	1	4
K1SU	1	1	0	1	1	4
K2CI	0	1	1	1	1	4
K2CO	1	1	0	1	1	4
K2DE	0	1	1	1	1	4
K2LA	1	1	1	1	0	4
K3CI	0	1	1	1	1	4
K3DE	0	1	1	1	1	4
K3SU	1	1	0	1	1	4
G1SU	0	1	1	1	1	4
G2LA	0	1	1	1	1	4
G2SU	0	1	1	1	1	4
G3PL	1	1	1	0	1	4
G3SP	1	1	0	1	1	4
G3TR	1	1	1	0	1	4
D1IL	1	1	1	1	0	4
D1SU	0	1	1	1	1	4
D2AB	0	1	1	1	1	4
D2CO	1	1	0	1	1	4
D2DE	0	1	1	1	1	4
D2LA	0	1	1	1	1	4
D2SU	0	1	1	1	1	4
D2TR	0	1	1	1	1	4
D3CO	0	1	1	1	1	4
D3DE	0	1	1	1	1	4
D3IL	1	1	1	1	0	4
D3PL	1	1	1	0	1	4
D3SU	1	1	1	0	1	4
M1AB	1	1	0	1	1	4
M1LA	1	0	1	1	1	4
M1SP	1	0	1	1	1	4
K1IL	1	1	1	1	1	5
K2IL	1	1	1	1	1	5
K2PL	1	1	1	1	1	5
K3CO	1	1	1	1	1	5
G2CO	1	1	1	1	1	5
G2IL	1	1	1	1	1	5
G2TR	1	1	1	1	1	5
G3CI	1	1	1	1	1	5
G3CO	1	1	1	1	1	5
G3DE	1	1	1	1	1	5
G3IL	1	1	1	1	1	5
G3LA	1	1	1	1	1	5
D1DE	1	1	1	1	1	5
D3AB	1	1	1	1	1	5
D3CI	1	1	1	1	1	5
D3LA	1	1	1	1	1	5
D3TR	1	1	1	1	1	5
M1CO	1	1	1	1	1	5
1s	45	64	64	75	88	
0s	65	46	46	35	22	

CRITERIA (RAT) ERRORS = 98

	GRAMMAR	ASSOCIATION	MEANING	SPELLING	TOTAL PER WORD
K1BR	0	0	0	0	0
K2BR	0	0	0	0	0
K3BR	0	0	0	0	0
G1BR	0	0	0	0	0
G1SP	0	0	0	0	0
G2BR	0	0	0	0	0
D1BR	0	0	0	0	0
D1SP	0	0	0	0	0
D2BR	0	0	0	0	0
D2SP	0	0	0	0	0
M1PL	0	0	0	0	0
M1SU	0	0	0	1	1
K3AB	0	0	0	1	1
K2SP	0	1	0	0	1
K3SP	0	1	0	0	1
G1LA	1	0	0	0	1
G1PL	0	0	0	1	1
D2PL	0	0	0	1	1
M1IL	0	0	1	1	2
M1TR	0	0	1	1	2
M1BR	1	1	0	0	2
D3SP	1	0	0	1	2
K1AB	0	0	1	1	2
K1CI	0	0	1	1	2
K1LA	0	1	1	0	2
K1SP	0	1	0	1	2
K1TR	0	1	0	1	2
K2AB	0	0	1	1	2
K2SU	0	0	1	1	2
K3LA	0	1	1	0	2
K3PL	0	0	1	1	2
K3TR	0	0	1	1	2
G1AB	0	0	1	1	2
G1CO	0	0	1	1	2
G1IL	0	0	1	1	2
G2AB	0	0	1	1	2
G2DE	0	0	1	1	2
D1CI	0	0	1	1	2
D1CO	0	0	1	1	2
D1LA	1	0	1	0	2
D1PL	1	0	0	1	2
D2CI	0	0	1	1	2
D2IL	0	1	1	0	2
D3BR	0	1	0	1	2
D2AB	0	1	1	1	3
K1CO	1	0	1	1	3
K1DE	0	1	1	1	3
K1PL	0	1	1	1	3
K1SU	1	0	1	1	3
K2CI	0	1	1	1	3
K2CO	1	0	1	1	3
K2DE	0	1	1	1	3
K2LA	1	1	1	0	3
K2TR	1	0	1	1	3
K3SU	1	0	1	1	3
K3CI	0	1	1	1	3
K3DE	0	1	1	1	3
K3IL	0	1	1	1	3
G1CI	0	1	1	1	3
G1DE	0	1	1	1	3
G1SU	0	1	1	1	3
G1TR	0	1	1	1	3
G2CI	0	1	1	1	3
G2LA	0	1	1	1	3
G2PL	0	1	1	1	3
G2SP	1	1	0	1	3
G2SU	0	1	1	1	3
G3AB	1	0	1	1	3
G3SP	1	0	1	1	3
G3SU	0	1	1	1	3
D1AB	0	1	1	1	3
D1IL	1	1	1	0	3
D1SU	0	1	1	1	3
D1TR	1	0	1	1	3
D2CO	1	0	1	1	3
D2DE	0	1	1	1	3
D2LA	0	1	1	1	3
D2SU	0	1	1	1	3
D2TR	0	1	1	1	3
D3CO	0	1	1	1	3
D3DE	0	1	1	1	3
D3IL	1	1	1	0	3
M1AB	1	0	1	1	3
M1CI	0	1	1	1	3
M1DE	1	0	1	1	3
K1IL	1	1	1	1	4
K2IL	1	1	1	1	4
K2PL	1	1	1	1	4
K3CO	1	1	1	1	4
G2CO	1	1	1	1	4
G2IL	1	1	1	1	4
G2TR	1	1	1	1	4
G3BR	1	1	1	1	4
G3CI	1	1	1	1	4
G3CO	1	1	1	1	4
G3DE	1	1	1	1	4
G3IL	1	1	1	1	4
G3LA	1	1	1	1	4
G3PL	1	1	1	1	4
G3TR	1	1	1	1	4
D1DE	1	1	1	1	4
D3AB	1	1	1	1	4
D3CI	1	1	1	1	4
D3LA	1	1	1	1	4
D3PL	1	1	1	1	4
D3SU	1	1	1	1	4
D3TR	1	1	1	1	4
M1CO	1	1	1	1	4
M1LA	1	1	1	1	4
M1SP	1	1	1	1	4
1s	45	64	85	88	
0s	65	46	25	22	

CRITERIA WITHOUT COLLOCATION; MEANING >= 260 ERRORS = 58

	GRAMMAR	ASSOCIATION	MEANING	SPELLING	TOTAL PER WORD
K1BR	0	0	0	0	0
K2BR	0	0	0	0	0
K3BR	0	0	0	0	0
G1BR	0	0	0	0	0
G1SP	0	0	0	0	0
G2BR	0	0	0	0	0
D1BR	0	0	0	0	0
D2BR	0	0	0	0	0
D2SP	0	0	0	0	0
M1PL	0	0	0	0	0
D1SP	0	0	0	0	0
K1AB	0	0	0	1	1
K2AB	0	0	0	1	1
K3AB	0	0	0	1	1
G1LA	1	0	0	0	1
G1PL	0	0	0	1	1
D2PL	0	0	0	1	1
M1SU	0	0	0	1	1
K2SP	0	1	0	0	1
K3PL	0	0	0	1	1
K3SP	0	1	0	0	1
K3TR	0	0	0	1	1
G1AB	0	0	0	1	1
G2AB	0	0	0	1	1
D1LA	1	0	0	0	1
M1IL	0	0	0	1	1
G1CO	0	0	1	1	2
G1TR	0	1	0	1	2
D1PL	1	0	0	1	2
D3BR	0	1	0	1	2
D3SP	1	0	0	1	2
M1BR	1	1	0	0	2
M1DE	1	0	0	1	2
M1TR	0	0	1	1	2
K1CI	0	0	1	1	2
K1LA	0	1	1	0	2
K1SP	0	1	0	1	2
K1TR	0	1	0	1	2
K2SU	0	0	1	1	2
K2TR	1	0	0	1	2
K3LA	0	1	1	0	2
G1CI	0	1	0	1	2
G1DE	0	1	0	1	2
G1IL	0	0	1	1	2
G2CI	0	1	0	1	2
G2DE	0	0	1	1	2
G3AB	1	0	0	1	2
D1AB	0	1	0	1	2
D1CI	0	0	1	1	2
D1CO	0	0	1	1	2
D2CI	0	0	1	1	2
D1TR	1	0	0	1	2
D2IL	0	1	1	0	2
M1CI	0	1	0	1	2
K1PL	0	1	1	1	3
K3IL	0	1	1	1	3
G2PL	0	1	1	1	3
G2SP	1	1	0	1	3
G3BR	1	1	0	1	3
G3SU	0	1	1	1	3
K1CO	1	0	1	1	3
K1DE	0	1	1	1	3
K1SU	1	0	1	1	3
K2CI	0	1	1	1	3
K2CO	1	0	1	1	3
K2DE	0	1	1	1	3
K2LA	1	1	1	0	3
K3CI	0	1	1	1	3
K3DE	0	1	1	1	3
K3SU	1	0	1	1	3
G1SU	0	1	1	1	3
G2LA	0	1	1	1	3
G2SU	0	1	1	1	3
G3SP	1	0	1	1	3
D1IL	1	1	1	0	3
D1SU	0	1	1	1	3
D2AB	0	1	1	1	3
D2CO	1	0	1	1	3
D2DE	0	1	1	1	3
D2LA	0	1	1	1	3
D2SU	0	1	1	1	3
D2TR	0	1	1	1	3
D3CO	0	1	1	1	3
D3DE	0	1	1	1	3
D3IL	1	1	1	0	3
M1AB	1	0	1	1	3
M1LA	1	1	0	1	3
M1SP	1	1	0	1	3
G3PL	1	1	1	1	4
G3TR	1	1	1	1	4
D3PL	1	1	1	1	4
D3SU	1	1	1	1	4
K1IL	1	1	1	1	4
K2IL	1	1	1	1	4
K2PL	1	1	1	1	4
K3CO	1	1	1	1	4
G2CO	1	1	1	1	4
G2IL	1	1	1	1	4
G2TR	1	1	1	1	4
G3CI	1	1	1	1	4
G3CO	1	1	1	1	4
G3DE	1	1	1	1	4
G3IL	1	1	1	1	4
G3LA	1	1	1	1	4
D1DE	1	1	1	1	4
D3AB	1	1	1	1	4
D3CI	1	1	1	1	4
D3LA	1	1	1	1	4
D3TR	1	1	1	1	4
M1CO	1	1	1	1	4
1s	45	64	64	88	
0s	65	46	46	22	

CRITERIA WITHOUT COLLOCATION, MEANING >= .376 ERRORS = 76

	MEANING	GRAMMAR	ASSOCIATION	SPELLING	TOTAL PER WORD
K1BR	0	0	0	0	0
K2BR	0	0	0	0	0
K3BR	0	0	0	0	0
G1BR	0	0	0	0	0
G1SP	0	0	0	0	0
G2BR	0	0	0	0	0
D1BR	0	0	0	0	0
D1SP	0	0	0	0	0
D2BR	0	0	0	0	0
D2SP	0	0	0	0	0
M1PL	0	0	0	0	0
K1AB	0	0	0	1	1
K1LA	0	0	1	0	1
K2AB	0	0	0	1	1
K2SP	0	0	1	0	1
K3AB	0	0	0	1	1
K3PL	0	0	0	1	1
K3SP	0	0	1	0	1
K3TR	0	0	0	1	1
G1AB	0	0	0	1	1
G1CO	0	0	0	1	1
G1LA	0	1	0	0	1
G1PL	0	0	0	1	1
G2AB	0	0	0	1	1
D2PL	0	0	0	1	1
D1LA	0	1	0	0	1
M1SU	0	0	0	1	1
M1TR	0	0	0	1	1
M1IL	0	0	0	1	1
K1CI	1	0	0	1	2
K1DE	0	0	1	1	2
K1SP	0	0	1	1	2
K1SU	0	1	0	1	2
K1TR	0	0	1	1	2
D1PL	0	1	0	1	2
D1SU	0	0	1	1	2
D1TR	0	1	0	1	2
D2AB	0	0	1	1	2
D2CI	1	0	0	1	2
K2SU	1	0	0	1	2
K2TR	0	1	0	1	2
K2LA	0	1	1	0	2
K3LA	1	0	1	0	2
K3SU	0	1	0	1	2
G1CI	0	0	1	1	2
G1DE	0	0	1	1	2
G1IL	1	0	0	1	2
G1TR	0	0	1	1	2
G2CI	0	0	1	1	2
G2DE	1	0	0	1	2
G2LA	0	0	1	1	2
G3AB	0	1	0	1	2
D1AB	0	0	1	1	2
D1CI	1	0	0	1	2
D1CO	1	0	0	1	2
D2IL	1	0	1	0	2
D3BR	0	0	1	1	2
D3SP	0	1	0	1	2
M1AB	0	1	0	1	2
M1BR	0	1	1	0	2
M1CI	0	0	1	1	2
M1DE	0	1	0	1	2
K3IL	1	0	1	1	3
K3DE	1	0	1	1	3
K3CI	1	0	1	1	3
K1CO	1	1	0	1	3
K1PL	1	0	1	1	3
K2CI	1	0	1	1	3
K2CO	1	1	0	1	3
K2DE	1	0	1	1	3
G1SU	1	0	1	1	3
G2CO	0	1	1	1	3
G2PL	1	0	1	1	3
G2SP	0	1	1	1	3
G2SU	1	0	1	1	3
G2TR	0	1	1	1	3
G3BR	0	1	1	1	3
G3CI	0	1	1	1	3
G3DE	0	1	1	1	3
G3LA	0	1	1	1	3
G3PL	0	1	1	1	3
G3SP	1	1	0	1	3
G3SU	1	0	1	1	3
G3TR	0	1	1	1	3
D1IL	1	1	1	0	3
D2CO	1	1	0	1	3
D2DE	1	0	1	1	3
D2LA	1	0	1	1	3
D2SU	1	0	1	1	3
D2TR	1	0	1	1	3
D3CI	0	1	1	1	3
D3CO	1	0	1	1	3
D3DE	1	0	1	1	3
D3IL	1	1	1	0	3
M1CO	0	1	1	1	3
M1LA	0	1	1	1	3
M1SP	0	1	1	1	3
K2IL	1	1	1	1	4
K2PL	1	1	1	1	4
K3CO	1	1	1	1	4
G2IL	1	1	1	1	4
G3CO	1	1	1	1	4
G3IL	1	1	1	1	4
D1DE	1	1	1	1	4
K1IL	1	1	1	1	4
D3AB	1	1	1	1	4
D3LA	1	1	1	1	4
D3PL	1	1	1	1	4
D3SU	1	1	1	1	4
D3TR	1	1	1	1	4

1s	44	45	64	88
0s	66	65	46	22

CRITERIA WITHOUT COLLOCATION; MEANING >= .500 ERRORS = 94