ASPECTS OF VOCABULARY TEACHING

Miguel García López
Universidad de Sevilla

Resumen

El objetivo del presente estudio es ofrecer un resumen de las principales conclusiones que la investigación actual sobre aprendizaje y enseñanza de vocabulario de una segunda lengua (en este caso, inglés) ofrece a profesores y estudiantes, especialmente de nivel intermedio-avanzado. Para ello se revisa la investigación sobre el lexicon mental y la adquisición de vocabulario, la investigación originada gracias al uso de programas de ordenadores sobre corpus lingüísticos y la investigación sobre estrategias de aprendizaje de vocabulario.

Abstract

This study summarizes the main conclusions of current research into the learning and teaching of vocabulary in a second language (English in this case) for teachers and students, especially at intermediate and advanced levels. To this end, the research considered covers the mental lexicon, the acquisition of vocabulary, and studies of corpuses using computer programmes and vocabulary learning strategies.

1. INTRODUCTION

«Vocabulary is the Everest of language. There is no larger task than to look for order among the hundreds of thousands of words which comprise the lexicon» (Crystal, 1995).

Learning the vocabulary of a language is a lifelong task. We can assimilate the grammar and phonology of a language in a few months, but the complex net of meanings and relations which constitute the lexicon seems to have no limits.

Since Meara remarked in 1980 the lack of attention to vocabulary learning and teaching, there has been a great amount of research on this field and a change of perspective. The first step ahead was taken by the communicative approach to language teaching, which emphasised the importance of vocabulary for communication. We cannot convey meaning without words. From a communicative perspective, lexis is considered central and not a peripheral aspect to language learning. But this communicative paradigm has given a central role to the implicit acquisition of vocabulary, which is considered to be acquired through interaction and exposition to large quantities of oral and written input.
Nevertheless, some studies have shown the importance of language awareness (for instance, Lennon, 1989) and the cognitive implication of the learner in the learning process, which is not considered a subconscious mechanism, but a process where learning strategies and depth of processing have a great relevance. This cognitive perspective has given a new impetus to the role of vocabulary in foreign language learning, since vocabulary learning is an aspect where learning strategies and learner autonomy may have a significant role to play.

These two waves (communicative and cognitive approaches) have placed vocabulary in the position it deserves, together with the use of computers in language studies and education, a powerful tool that has helped to cope with and give certain order to the vast and endless domain that is the lexicon. The importance of vocabulary can be seen in the number of books published around the topic (for example, McCarthy 1990, Natior 1990, Huckin, Haynes and Coady 1993, Hatch and Brown 1995, Coady and Huckin 1997 Schmitt and McCarthy 1997, Singleton 1999).

We will analyse the most important implications of research on the mental lexicon on vocabulary acquisition, on corpora and concordances and on vocabulary learning strategies.

2. RESEARCH ON THE MENTAL LEXICON AND VOCABULARY ACQUISITION

The research on lexical processing has provided us with interesting insights into vocabulary acquisition (both in L1 and L2). Here are some of the most important lines of research and findings.

2.1. The bilingual lexicon

One of the most important implications of the great amount of research on the bilingual lexicon (Schereuder and Weltens, 1993) is that L1 lexicon and L2 lexicon are highly interconnected. The non-target lexical system is always operational, and this is the reason why there is language interference even in the most monolingual processing situations (Grainger, 1993). The traditional difference between compound, coordinate and subordinate bilinguals may be found in any individual bilingual, depending on particular characteristics of the words (de Groot, 1993).

Hence, there is not a clear division between the first and second language lexicon. Some teaching ideas that recommend having both lexicons in independent compartments go against the findings that show that both lexicons are interconnected in many ways even in complete bilinguals. Singleton (1999) shows a wide spectrum of research that rules out the notion that the L1 and the L2 lexicons are disconnected. They are two systems in communication with each other.
2.2. Cognates, frequent words, vocabulary size and depth

One of the most interesting common points between L1 and L2 lexicons are the cognates. The foreign language learner has to take advantage of this important amount of words between the two lexicons. Learners can develop a large vocabulary on the basis of his L1 vocabulary, just applying some general phonological rules and converting L1 items into L2 items (Meara, 1993). There are learners with a tendency to avoid cognates (see, for example, Kellerman’s (2000) research on ‘homoiophobia’), as they think that similarities between the two languages will lead to confusion. Other learners rely on cognates as much as they can, and for them, cognate recognition is a natural strategy (Holmes and Ramos, 1993).

But the most important problem with the use of cognates is a problem of frequency. Although English shares many cognates with Spanish, these items are often relatively rare in English, and they tend to be low frequency items used in formal situations. This is one of the reasons why secondary and University students have important gaps with very common and frequent English words, but they can control some vocabulary that apparently is more difficult for native speakers. Many advanced students may have more problems with items that appear in the lists of frequent English words than with words of the University Word List (a list of approximately 850 words that, while not frequent in the language, are very frequent in academic texts; see Nation, 1990). As De Groot and Keijzer (2000) have found, word frequency does not affect at all the learning and retrieval of words, but cognate status does.

Meara (1993: 285) suggests that this situation demands a specific teaching strategy:

“students with a huge latent vocabulary among the lower frequency bands, but serious gaps among frequent words, would benefit from a course that was specifically aimed at plugging those gaps. Once the gaps have been filled, then the latent cognate vocabulary comes into its own: until then, it remains something of a liability.”

The reason for these gaps is clear: if language surrounds us as in first language environment, frequent words will appear constantly and repeatedly, but in academic settings, the environment is artificial and the most common situations may never appear. The first time university students visit a foreign language country discover they do not know frequent words because they have never seen themselves in these particular environments. A good teaching approach has to recreate these environments. One of the ways of having access to examples of these common contexts is the use of corpora and concordances (see section 3).

A finding of corpora research is that about 2000 word families (words, inflections and derivational suffixes) account for 80% of the words in English texts (Nation, 1990). The problem is that 80% is not enough to allow reasonably successful guessing; at least 95% is needed (Liu Na and Nation, 1985). This implies that a learner needs about 3000
word families as a basis for comprehension; once these words are learned, he can make use of guessing strategies to learn the low frequency words of the language.

Knowing these 2,000 words is not only a question of breadth, but also a question of depth. We may know the central meaning of these 2,000 words, but each word is a world in itself. There are different definitions, connotations, collocations, metaphors and multiple meanings around these words, which have to be learnt.

This long process is the reason why the terms receptive and productive vocabulary are more a continuum than a dichotomy. Melka (1997) shows us that the distance between these receptive and productive lexicons is more a question of degrees of knowledge or degrees of familiarity than ‘watertight’ compartments. The lexical competence is a continuum; there are words with weak connections that we do not use productively. Vocabulary acquisition is an incremental process, and not a question of dichotomies. This developing knowledge is difficult to measure; there are scales that make an attempt to do it, such as the Vocabulary Knowledge Scale, developed by Paribakht and Wesche (1997).

2.3. Connectionist models

Another line of research that has given interesting insights into vocabulary acquisition and use is the connectionist or parallel distributed processing paradigm (Rumelhart and McClelland, 1986). These researchers see knowledge in terms of connection strength rather than rules or patterns. Units are connected in such a way that the activation of one unit can inhibit or excite others. Learning consists of discovering the right connections strengths from the input. Our psychological models of human language processing are becoming more statistical and less logic-based (Kirsner, 1994; Newell 1990).

These ideas are challenging mentalist theories because they suggest that learning is environmentally driven and that innate knowledge is not so important. Words are collections of connections rather than stable, enduring symbols. Learners need to be constantly creating and reinforcing connections to maintain the lexicon alive.

2.4. Memory, chunks and lexical retrieval

In a typical linguistic model processing system, rules are of great importance, and language is produced by filling these rules with lexical items. But as we mentioned above, connectionist systems are showing that these rules are not really the key for language production. Sinclair (1991) suggests that the endless combinatorial possibilities that grammar offers are in practice ignored, because co-occurrences of words are limited. Pawley and Syder (1983) suggest that the average native speaker knows hundreds of thousands of lexicalised sentence stems, and these chunks are crucial for a fluent performance. According to Skehan’s (1998) analysis of aptitude, the most important factor at advanced levels is memory: a coding system reliant on chunks and based on redundant storage.
There are different types of multi-word items: compounds (carpark, long-haired, navy blue), phrasal verbs (give up, break off), idioms (spill the beans, kick the bucket), fixed phrases (enough is enough, dry as a bone) and prefabs (the things is, that reminds me) (Moon, 1997). Nattinger and DeCarro (1992) show the utility of teaching and learning such units. An important index of nativelike competence is the fluent use of them. There are studies that confirm the idea that these units deserve special pedagogical attention and that learners should have specific strategies to deal with them (Arnaud and Savignon, 1997).

2.5. Collocations

The syntagmatic dimension of lexicon has proved to be a very important factor in acquisition: the combinations that words enter into when discourse is produced and how words typically occur with one another. All foreign language learners know the problem of deviant collocations. Even advanced learners may know the meaning of isolated words, but they do not know the usual patterns of co-occurrence. This knowledge is based on years of experience of masses of data (McCarthy, 1990: 15).

Advanced learners need not only the meaning, but also good examples that show them the most frequent collocations. Gairns and Redman wrote in 1986 that «since there are no rules of collocation this aspect of vocabulary learning is often dealt with on an ad hoc basis» (1986: 37). But computers have helped to give us interesting insights into these patterns; previously it was considered something difficult to grasp. The use of corpora and concordances has made it possible to examine collocations in a scientific way (see section 3).

2.6. Written vs. spoken vocabulary

Corpus based study of vocabulary has shown that frequency lists for spoken language differ significantly from those based on written databases. An analysis of these lists show that almost half of the spoken discourse has no content (McCarthy and Carter, 1997). These authors show that less than 25 percent of words of typical conversations are unequivocally content words. Spoken discourse has a lot of repetition and lexical negotiation. Vocabulary choice in conversations has important interpersonal implications. Vagueness is also ubiquitous in spoken vocabulary as opposed to written, being a constant interpersonal strategy. The spoken lexicon is fluid and a medium through which meanings are negotiated.

The conclusion of all this research is that for teaching and learning spoken vocabulary the emphasis must shift away from content words and must give more importance to discourse markers, vague terms and delexical words. We have to move from declarative and definitive vocabulary activities to procedural and fluctuating word-set activities (Robinson, 1993).
2.7. Intentional vs. incidental vocabulary acquisition

Research on reading and vocabulary acquisition has shown that L2 learners, as well as native speakers, can acquire vocabulary incidentally in the context of reading without any intention to learn (Huckin and Coady, 1999). Nevertheless, incidental vocabulary acquisition has its limitations: it is a very slow process and guessing has a lot of difficulties and problems (Haynes, 1993). In addition, even if guessing is correct, it does not imply that there is long-term retention (Mondria and Wit de-Boer, 1991).

Although nobody denies that incidental vocabulary acquisition is extremely important, it is now generally recognised that the explicit, direct and conscious learning of vocabulary is also necessary. For instance, research by Nick Ellis (1995) shows that the aspects of vocabulary learning that have to do with meaning involve conscious and explicit learning; on the other hand, those aspects related to word forms have to do with implicit learning. The conclusion is that both explicit and implicit learning are necessary for vocabulary development. There are indications that mental effort has a positive effect on the learning of word meanings, and the use of vocabulary learning strategies can be highly beneficial for increasing L2 vocabulary (see section 4).

2.8. Psycholinguistics determinants of vocabulary acquisition

The study of avoidance as a learning strategy is an interesting field of research that can provide useful clues in the field of vocabulary learning and teaching. Learners tend to avoid some lexical items and give priority to others. The study of these regularities is another way of giving some structure to the apparent chaos that seems to be the acquisition of vocabulary.

Levenston suggested in 1979 that there are three main factors that make learners avoid the acquisition of new vocabulary: phonological difficulty, grammatical difficulty and semantic difficulty. He also pointed out that learners avoid words with no direct translation to their mother tongue and polysemes, but prefer words which can be generalized to a large number of contexts.

This line of research has been followed by Laufer (1997), who distinguishes between intralexical factors (features inherent in the word that affect their learning) and interlexical factors (relationship between the word and other words familiar to the learner in L2 and his L1). Her research on intralexical factors shows that some of the most important features that affect vocabulary acquisition are:

1) pronounceability (phonological difficulties related to phonemes, combinations of phonemes and suprasegmental features);
2) orthography (the degree of sound-script correspondence);
3) inflectional and derivational complexity (irregularities are a cause of avoidance, and also deceptive transparency, that is, words that look as if they were combined of meaningful morphemes, for instance outline, where out does not mean out of);
4) synformy (words that sound and/or look alike, which L2 learners confuse; Laufer offers ten different categories of synforms; for instance, words with the same root but different suffixes: imaginary/imaginative/imaginable);

5) semantic features of the words (tendency to use words that are superordinates or general terms; avoidance of idiomatic expressions if there are non-idiomatic equivalents; problems with polisemy and homonymy).

According to Laufer, there are some features which do not have a clear effect on learning, or with contradictory results, which are word length (it is not clear that longer words are more difficult to learn); part of speech (evidence is not conclusive about the belief that nouns are the easiest to learn, adverbs are the most difficult, and verbs and adjectives are in the middle); and abstractness (it is not true that abstract words are more difficult to learn than concrete words in L2, because the concepts have already been acquired in L1).

Ellis and Beaton’s study (1993) on psycholinguistics determinants of foreign language vocabulary learning also concludes that the pronounceability has a strong determinant effect, but they find that nouns are far easier to learn than verbs. Regarding abstractness, de Groot and Keijzer’s research (2000) shows that concrete words are easier to learn and less susceptible to forget. As we can see, the results are not conclusive yet.

There are also interlexical factors: the similarities and differences between the vocabulary of the foreign language and the mother tongue have undeniable effects on vocabulary choice (see section 2.2. for some considerations about the use of cognates).

Another determinant of vocabulary acquisition is the learning experience, the context in which the words are learned. For example, trying to learn synforms together -- affect/effect, cute/acute, prize/price -- is not effective and it can cause more trouble than facilitation (Laufer, 1997). Likewise, learning is more difficult if learners try to learn groups of items of related meaning (for example, opposites), because these words will be associated and the differences between them will interfere with each other (Nation 1990, Tinkham 1997).

3. RESEARCH ON CORPORAs AND CONCORDANCES

Vocabulary has always been elusive to scientific study due to its extension. Now it is possible to search for patterns of word combination, check word frequencies and see examples of all the uses of particular words. The study of vocabulary has received a renewed impetus with the use of computers in the compilation of real language. L2 dictionaries increasingly give more information based on corpus. A corpus is a collection of texts or samples of language. Computers have helped to collect texts for analysis of words, dictionary compilation and language study. New data is acquired continuously and processed into a standard format. Research has shown that very large samples of text are necessary to draw statistically valid conclusions. For instance, the British National Corpus (BNC) consists of 90 million words; the Cambridge International Corpus (CIC) has more
than 300 million words and the Bank of English is a collection of over 400 million words which is held at COBUILD, the University of Birmingham.

However, size in itself is no guarantee of reliability. The corpus has to make a serious attempt to be representative of the whole repertoire of discourse types: literature, magazines, papers, ephemeral material and transcribed oral English texts. Until now a significant quantity of ordinary conversational English was missing. But nowadays corpora also include millions of words of transcribed speech from radio broadcasts, personal audio recordings, casual conversation, interviews and so on, transcribed onto computer by specialist trained staff.

A corpus may be processed by a computer program called a concordancer, whose output is a concordance, an index of all the words in the corpus with their linguistic contexts. The user may call up the words he/she wants to examine and this list of words is displayed in the centre of the screen and shown with parts of the context in which they occur. From these examples the user is able to deduce a great deal of information about how a word is being used.

Corpora and concordance are powerful tools for the analysis and study of lexis by linguists and lexicographers and, at the same time, for learners of the language. Let us consider the implications of this research on both directions.

From the perspective of the nature of the lexicon, corpora and concordances have:

1) provided ample evidence of the extent to which speakers rely on pre-assembled chunks of language (see section 2.5.);
2) demonstrated that 2500 word families account for 80% of the individual words in English texts;
3) revealed an important level of systematicity in surface language: the system is not only in the Language Acquisition Device (LAD) but also in the ambient linguistic evidence;
4) shown that many individual words have their own system of collocations, challenging the distinction between grammar and lexis (Willis 1999) and demonstrating that there is an intricate relationship between them.

From the learners’ perspective, concordances:

1) make students use cognitive and analytical skills to manipulate these databases and solve lexical problems;
2) promote inductive learning, facilitating students’ research into the language without knowing in advance what patterns will be discovered;
3) give students the opportunity to be more active in their vocabulary learning, discovering new meanings and habitual collocations, and relating words to syntax;
4) develop language awareness and encourage learner independence;
5) provide students with examples of authentic usage to demonstrate features of vocabulary. Concordances present words in a variety of contexts and indicate by the
number of occurrences how much learning attention each word deserves. This exposure can be a substitute for the massive amount of input that is necessary for incidental vocabulary acquisition.

All these are values that are considered to be of great importance in learning. The use of concordancing for the learning of vocabulary is a clear application of constructivist ideas about learning.

There are web pages where it is possible to have access to examples of collocations, for example, The Complete Lexical Tutor (http://132.208.224.131/). When the user clicks on a word, a concordance appears on the screen. This page, maintained by Thomas Cobb, is based on his research (Cobb, 1997) on how to use corpus and concordance to learn a foreign language vocabulary.

4. RESEARCH ON VOCABULARY LEARNING STRATEGIES

The research on learning strategies and vocabulary learning are highly connected. On the one hand, it has been found that learners employ more learning strategies with tasks related to vocabulary than with more integrated tasks (for example, listening or speaking) (Chamot, 1987). On the other hand, vocabulary learning strategies lend themselves to experimental investigation, perhaps due to the same reason: the discrete nature of vocabulary, which makes it possible to define strategies more precisely than more integrated activities.

There are different taxonomies for classifying vocabulary learning strategies. Brown and Payne (1994) and Schmitt (1997) attempt to present complete lists, taking into account the different steps followed when learning a new word.

Brown and Payne (1994), in a study reported by Hatch and Brown (1995), develop a model where vocabulary learning strategies used by ESL students fall into five different groups:

1) strategies for encountering new words (for instance, learning new words by reading books);

2) strategies for getting the word form, through a visual and/or auditory image of the form of the vocabulary item (for example, associating a word with a similar sound of an English word the learner knows);

3) strategies for getting the word meaning (for instance, making pictures of word meanings in the mind or using dictionaries);

4) strategies for consolidating word form and meaning in memory (all kinds of memory strategies such as grouping, using keywords or vocabulary cards);

5) strategies for using the word (for instance, through hypothesis testing of the learner’s knowledge of collocations).

Schmitt (1997) establishes a distinction between:
1) Strategies for the initial discovery of a word meaning. They can be determination strategies (for instance, analysing affixes and roots, or checking for L1 cognate) or social strategies (asking someone who knows);

2) Strategies for remembering and consolidating a word once it has been encountered. They can be social strategies (using group work to learn vocabulary), memory strategies (for instance, using semantic maps or using physical action to learn a word), cognitive strategies (using repetition or mechanical means to learn the words) and metacognitive strategies (to control and evaluate vocabulary learning).

These taxonomies help to clarify the field for teachers and researchers. There are two kinds of research on vocabulary learning strategies: descriptive studies that try to identify the strategies that different groups of learners employ, and experimental studies that try to measure the efficacy of some strategies when compared to others.

An example of the first kind of studies is Lawson and Hogben's research (1996). This study found that repetition is the strategy most university students use, giving little attention to elaborative procedures, concluding that there is a need to make these students aware of the advantage of these procedures. Nyikos and Oxford (1993) used factor analyses to identify frequency of use of strategies among university students and found that memory strategies are not frequently used because they are viewed as mere gimmickry and a substitute to pure learning. They conclude that there is a need for integrated strategy training.

There is also research on vocabulary learning strategies' effectiveness. The most researched strategy is the mnemonic strategy called Keywords: the learner has to find an L1 word which sounds like the target L2 word, and then an image that combines the two concepts is created. Most of this research (see Hulstijn 1997 for a review) has found that this strategy is highly effective in enhancing the recall of words.

Some studies show that the combination of different strategies is the key for effectiveness in vocabulary learning. For instance, Brown and Perry (1991) found that the combination of keywords with semantic processing provides the students with more versatility and promotes vocabulary learning. Garcia López (1998) found that the combination of keywords with any other kind of association was the best combination to help secondary students with vocabulary learning problems.

The implication of all this research is that strategy training is necessary and useful to improve vocabulary learning, especially when this training combines different strategies suitable to the learners' ages, styles and proficiency.
5. CONCLUSION

The areas that have been surveyed offer important implications for the learning and teaching of second language vocabulary, which can be summarised as follows:

1) Do not treat L1 and L2 lexicons as if they were in independent compartments. Promote their connections.
2) Take advantage of the cognate vocabulary, but at the same time filling the gap with the 3,000 most frequent words.
3) Recreate the environments where the most frequent words occur.
4) Make use of guessing strategies to learn the low frequency words of the language.
5) Create and reinforce connections to maintain the lexicon alive.
6) Pay attention and try to memorize groups of words (chunks), because they are the key for fast lexical retrieval.
7) Pay attention to the syntagmatic dimension of the lexicon: the combinations of words in discourse and how words typically occur with one another.
8) When working on oral vocabulary, give more importance to discourse markers, vague terms and delexical words than content words. Practice with activities where meaning fluctuates and is negotiated.
9) Work on both implicit and explicit learning.
10) Do not complicate the phonological dimension of vocabulary learning by trying to learn synonyms together.
11) Do not complicate the semantic dimension of vocabulary learning by trying to learn groups of items of related meaning.
12) Take advantage of corpora and concordances to work on vocabulary learning since these tools foster cognitive and analytical skills, inductive learning and awareness of collocations.
13) Be aware of the elaborative strategies that can be employed to learn and memorize vocabulary and combine them to improve lexical storage and retrieval.

BIBLIOGRAPHY

ARNAUD, P.J.L. & SAVIGNON, S.J. (1997). Rare words, complex lexical units and the advanced learner. In Coady & Huckin (eds.)


Concordia University, Montreal, Quebec, Canada. (http://132.208.224.131/ResearchWeb/)


LAUER, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt, & M. McCarthy.(eds.).


