

Vocabulary Acquisition in Language Learning

1. Introduction

There is more in a word than just one meaning: there is sound, graphics, a visual image, metaphor, a connection to other words, and much more. A word is a springboard to the knowledge of the world. “Each word is a world”¹ or, as Italian novelist S. Tamaro said, “*Ogni parola è un seme*”² (‘each word is a seed’).

When considering potential of vocabulary, we should think of the many aspects each lexical item conveys:

- the meaning(s) of the word
- the written form of the word
- the spoken form of the word
- the grammatical behaviour of the word
- the collocations of the word
- the appropriateness of the word
- the associations of the word
- the frequency of the word

This list of types of word knowledge³ shows how complex the nature of vocabulary is, and therefore, how many steps a language

¹ John Morgan and Mario Rinvulcri, *Vocabulary*, Oxford: Oxford University Press, 1986, p. 10.

² Title of Susanna Tamaro’s novel (Milano: Rizzoli, 2005).

³ Paul Nation, *Teaching and Learning Vocabulary*, Boston: Heinle & Heinle Publishers, 1990, p. 31.

learner must go through in order to gain full mastery of a word. It is self-evident that vocabulary acquisition is an incremental process.

On hearing the sound [ˈsoɲno] a learner will associate it with the action everyone performs while sleeping ('dream'), but he will be able to spell it correctly only after noticing the written form of the word, *sogno*. Then he will learn that he must use verb inflections to convey the fact that they *sognano*, or he *ha sognato* ('has dreamt') yesterday. If he is thinking of sunbathing on a sandy beach while sitting in his office, we can say that *sogna ad occhi aperti* ('he is day-dreaming'). Moreover, he could be asked if he *ha fatto un brutto sogno* ('has had a bad dream') – realizing that we do not say **cattivo* ('bad') *sogno*, thus using the appropriate collocate. He might also use a derivative when he claims he is not a *sognatore* ('dreamer').

Of course this is not a linear process and depends on the kind of exposure the student has to a word, as well as on his proficiency in language learning. Very often, we have only a partial knowledge of a word.

Literature on vocabulary has shown that the term **word** is not suitable when defining a single unit of meaning, because any language, Italian in particular, uses many expressions made up of two or more words to convey one single meaning. Therefore, the term **lexical unit** or **lexeme** may be more appropriate.

A proposito ('by the way'), *d'altra parte* ('on the other hand'), *come va?* ('how are you?'), *se fossi in te* ('if I were you'), *fuoco amico* ('friendly fire'), *gratta e vinci* ('scratch and win lottery ticket'): these are just examples of **multiword units (MUWs)** or **lexical chunks**,⁴ called *unità polirematiche* in Italian. Though they are composed of more than one word, they produce one single unit of meaning and are

⁴ "Research using large corpora has shown that these *multi-word lexical units* (MWU) are ubiquitous in language use, at least in English. This is especially true of spoken discourse. Moreover, this same corpus research is now beginning to indicate that there is lexical patterning that exists beyond even the MWU level. Some longer strings of language recur frequently and are often connected with the functional usage of language. These longer strings have been called various names, including *lexical phrases* (Nattinger and DeCarrico, 1992), *lexical chunks* (Lewis, 1993), *lexicalized sentence stems* (Pawley and Syder, 1983), and *ready made (complex) units* (Cowie, 1992)." Norbert Schmitt and Ronald Carter, "Lexical Phrases in Language Learning", *The Language Teacher Online* 24(8)/2000, <http://www.jalt-publications.org/tlt/articles/2000/08/schmitt> (accessed June 19, 2009).

used as set phrases with no variations. These fixed expressions are very common, particularly in spoken language, and our mind processes them as a single item. “Speakers do not construct these items, but simply recall them, direct from memory, as learned wholes.”⁵ Corpus linguistics⁶ has fostered recognition of these lexical chunks, thanks to computerized storage of texts that helps identify strong collocates used as single units of meaning.⁷

Another relevant definition is that of **word family** which includes all the variations of a word due to morphology and grammar, inflections, and its derivatives which are formed by adding affixes. A learner can easily recognize the meaning of *sognano*, *ha sognato* or *sognatore* if he has been trained to learn the patterning of grammar and affix meanings.

When considering how many words L2⁸ students should learn, the so-called **fundamental vocabulary** is necessary in order to understand general language and take part in daily conversations. According to Tullio De Mauro, the average Italian educated person masters 60 to 80 thousand words in the native language, but just 2,000 lexemes make up the fundamental vocabulary, which accounts for 90–95% of communication needs.⁹ The figure is not far from the standard set by the old “Threshold level”, corresponding to the B1 level of the Common European Framework of Reference for Languages. As a learner becomes more proficient and wants to read authentic texts, a vocabulary twice as large is necessary. However, in order to understand a particular subject area, as for special purpose

⁵ Michael Lewis, *Implementing the Lexical Approach*, Hove: Language Teaching Publications, 1997, p. 22.

⁶ Corpus linguistics is an approach to study aspects of language by means of research into language corpora, i.e. huge collections of texts stored in a computerized database, labeled and made accessible for queries.

⁷ “The larger units escape the attention of even experienced and well-trained lexicographers. Before the advent of corpora and of corpus linguistics, we did not even have a methodology to detect them. Neither standard linguistics nor Chomskyan linguistics can identify these units of meaning.” Wolfgang Teubert and Anna Čermáková, *Corpus Linguistics. A Short Introduction*, London: Continuum, 2007, p. 21. Here a full account on *friendly fire* as a single lexical unit is given, with a detailed investigation of its usage.

⁸ L2 indicates both a second and foreign language in this instance.

⁹ Tullio De Mauro, *Linguistica elementare*, Bari: Laterza, 9th edition, 2008, pp. 60–62. For an account on research about vocabulary size related to English, see Norbert Schmitt, *Vocabulary in Language Teaching*, Cambridge: Cambridge University Press, 2000, pp. 142–145.

L2 learners, a foundation of higher-frequency vocabulary is needed, in addition to specialized vocabulary for that area.

Frequency is one of the key criteria for choosing which words to teach or learn. Several lists had been compiled to include the most frequent lexical items; however, nowadays language corpora provide updated information about the most useful words in many natural languages, including frequency and collocation.

Most textbooks for Italian and other foreign languages take into account the frequency feature in selecting the vocabulary items to present and practise in their units. They also consider the relevant topics of interest for different types of learners and the use of classroom language.

2. Importance of vocabulary in language learning

Vocabulary acquisition plays an essential role in language learning, particularly in L2 learning. Communication is an exchange of meanings, and the main carrier of meaning is vocabulary. If a learner does not grasp the sense of the lexical items he is reading or listening to, there is neither real understanding nor personal response, as stated in the frequently quoted sentences below:

Without grammar little can be conveyed; without vocabulary nothing can be conveyed.¹⁰

A lexical mistake often causes misunderstanding, while a grammar mistake rarely does.¹¹

A more radical view was introduced by Michael Lewis in his book *The Lexical Approach*¹² and the follow-up *Implementing the Lexical Approach*, where he pointed out that “language consists of grammaticalised lexis”¹³, emphasizing the vital importance of vocabulary to convey meaning by learning not just words, but multi-word

¹⁰ David A. Wilkins, *Linguistics in Language Teaching*, London: Arnold, 1972, p. 111.

¹¹ John Sinclair, “Speaking Englishes, Speaking Cultures”, 30th IATEFL Conference at University of Keele, April 1996, quoted by Michael Lewis, *Implementing the Lexical Approach*, p. 16.

¹² Michael Lewis, *The Lexical Approach*, Hove: Language Teaching Publications, 1993.

¹³ Michael Lewis, *Implementing the Lexical Approach*, p. 102.

chunks. These lexical chunks reveal patterns that can be identified and used, and favour retrieval from memory. Some pedagogical implications of this approach are quite interesting, such as the focus on collocation and fixed or semi-fixed expressions, which are often used as exponents of communicative functions.

Words and chunks convey meaning and provide reference to concepts. By organizing concepts in our mind, we build up categories to understand the world and to be able to communicate with others.

3. How we learn words

Hunt and Beglar outline three basic approaches to vocabulary learning:

The incidental learning of vocabulary requires that teachers provide opportunities for extensive reading and listening. Explicit instruction involves diagnosing the words learners need to know, presenting words for the first time, elaborating word knowledge, and developing fluency with known words. Finally, independent strategy development involves practicing guessing from context and training learners to use dictionaries.¹⁴

Most of the words we know, both in L1 and L2, have been acquired by **incidental learning**, after being exposed to a large number of texts, where unknown words have been inferred and interpreted, then stored in our memory. Learning words from context is a gradual process; it takes time and requires many exposures to a word for understanding its meaning. This approach is more suitable to proficient learners, because beginners would feel discouraged facing too many new words. A good tool to overcome this problem is offered by **graded readers**,¹⁵ where the vocabulary is accurately selected from the most frequent and common words and used according to the levels of the *Common European Framework of Reference for Languages*.

¹⁴ Alan Hunt and David Beglar, "Current Research and Practice in Teaching Vocabulary", *The Language Teacher Online* 22(1)/1998, <http://www.jalt-publications.org/tlt/articles/1998/01/hunt> (accessed June 10, 2009).

¹⁵ A choice of Italian graded readers is offered by publishers such as CIDEB, Bonacci and Guerra.

Explicit instruction of new words is appropriate to beginners, and includes activities devised to elaborate lexical items, e.g.

- sorting lists of words
- drawing semantic maps
- generating derivatives, inflections, synonyms and antonyms of a word
- making trees for superordinate and coordinate words
- identifying or generating associated words
- combining phrases from several columns
- matching collocations
- completing cloze activities
- playing collocation crossword, puzzles or bingo¹⁶

Word processing involves expanding connections between what is already known and new information. As a result the learner will recognize the word form and its meaning automatically, thus acquiring so-called **sight vocabulary**.

To enhance independent **development**, students need to be trained to make inferences from new words by their context. It is a complex activity to carry out successfully (95% of the words in a text should be already known for a full guessing), but very enriching because it “also includes learning about collocations, associations and related grammatical patterns as well as meaning”.¹⁷

In addition, training in the use of various types of dictionaries (monolingual, bilingual, synonyms, thesaurus, etc.) and corpora will help learners become more independent in the research and consolidation of new lexical items.

Mastery of a word involves several components (written and spoken form, denotative and connotative meanings, inflections, derivatives, collocations, register, etc.) which cannot be learned simultaneously, but develop with practice and exposure in different con-

¹⁶ Alan Hunt and David Beglar, “Current Research”.

¹⁷ Ibid.

texts. Therefore it is an **incremental process**. The following steps have been identified:

1. On the first exposure we pick up some of word sense and form; some pronunciation if exposure is verbal, or the first few letters of its spelling if it is written.
2. After a few more exposures we consolidate these features and notice the word class, and perhaps some other meaning senses.
3. Only after a large number of examples, we develop intuitions about the word frequency, register and collocational behaviour.¹⁸

As many lexical items are polysemous, the core meaning of a word is likely to be learned first. This is the sense with the highest frequency, also called denotation, and allows for other figurative meanings and associations to develop later, after further exposure.¹⁹

Memory plays a key role in vocabulary acquisition. Short-term memory, also known as working memory, retains information for a very short time (a number of seconds) while processing, but it has a small storage capacity. Therefore lexical items are grouped into chunks, or longer units, as they enter the working memory for elaboration. The aim of vocabulary learning is to transfer the lexical information from short-term memory to long-term memory, which is more permanent and has a virtually unlimited storage capacity, but is relatively slow. In other words, the language input proceeds from sensory receptors (sight, hearing, etc.) into short-term memory for processing over a short time. Then, in order to become **intake** (permanently stored in the semantic network of the mind, readily available for retrieval), it has to find some pre-existing information to attach to; a **hook** with which to establish connection and remember. The new lexical item settles more easily in the long-term memory when a link to an old, previously acquired word or phrase is created. The most common type of hook is mean-

¹⁸ Norbert Schmitt, *Vocabulary in Language Teaching*, pp. 117–118.

¹⁹ For example, when noticing the word *cane* ('dog'), its core meaning 'four legged animal, kept as a pet' is learned first, as it is by far the most frequent sense. Later on learners might infer the meaning of the idiom *solo come un cane* ('desperately lonely'). Only after exposure to different types of specific texts, can they meet the less frequent meaning 'cock of a gun'.

ing similarity, but it could also be sound or spelling likeness, as well as a visual image.²⁰

Without this connection, words are more easily forgotten. As this usually occurs soon after the learning session, revision programmes that help retention of the new lexical information should be carried out immediately after the lesson and then repeatedly at gradually increasing intervals. **Expanding rehearsal** technique is based on this principle.²¹

4. Strategies for improving vocabulary acquisition

To facilitate vocabulary learning, many strategies have been devised and their success often depends on the subjective learning attitude of the student. Most learners prefer **shallow strategies**, such as verbal or written repetition and note taking.²² These are quite mechanical strategies which do not require deep mind processing, so they are not long lasting and only concern partial knowledge of the word, but can be suitable for beginners. **Deeper strategies** integrate more linguistic skills (receptive and productive skills) and involve active manipulation of information by use of imagery, inference, or creating associations that enhance retention in memory. As they are more effective but complex activities, they are best addressed to intermediate or advanced learners. However, the deeper the mind elaboration of the lexical item, the better and the longer the retention.

The following table shows the list of strategies that Schmitt compiled, both for discovery of a new word's meaning and for consolidating memorization. They are labelled according to the type of mental activity required.

²⁰ Norbert Schmitt, *Vocabulary in Language Teaching*, pp. 131–132. For further reference, see Mario Cardona, *Il ruolo della memoria nell'apprendimento delle lingue. Una prospettiva glottodidattica*, Torino: UTET Università, 2001.

²¹ This strategy suggests that a given item should be initially tested after a very short delay; if the learner succeeds in recalling it, the delay should be systematically increased.

²² Indicated by a survey on strategies for vocabulary development among students of Italian at Háskóli Íslands (see Appendix); see also Norbert Schmitt, *Vocabulary in Language Teaching*, p. 132.

Strategies for the discovery of a new word's meaning

- DET Analyze part of speech
- DET Analyze affixes and roots
- DET Check for L1 cognate
- DET Analyze any available pictures or gestures
- DET Guess meaning from textual context
- DET Use a dictionary (bilingual or monolingual)
- SOC Ask teacher for a synonym, paraphrase, or L1 translation of new word
- SOC Ask classmates for meaning

Strategies for consolidating a word once it has been encountered

- SOC Study and practice meaning in a group
- SOC Interact with native speakers
- MEM Connect word to a previous personal experience
- MEM Associate the word with its coordinates
- MEM Connect the word to its synonyms and antonyms
- MEM Use semantic maps
- MEM Image word form
- MEM Image word's meaning
- MEM Use Keyword Method
- MEM Group words together to study them
- MEM Study the spelling of a word
- MEM Say new word aloud when studying
- MEM Use physical action when learning a word
- COG Verbal repetition
- COG Written repetition
- COG Word lists
- COG Put English labels on physical objects
- COG Keep a vocabulary notebook
- MET Use English-language media (songs, movies, newscasts, etc.)
- MET Use spaced word practice (expanding rehearsal)
- MET Test oneself with word tests
- MET Skip or pass new word
- MET Continue to study word over time²³

²³ Norbert Schmitt, *Vocabulary in Language Teaching*, p. 134; Schmitt refers to learning English vocabulary, of course.

- DET = determination strategies, involve guessing from one's structural knowledge of language, from context, from an L1 cognate, or using reference materials
- SOC = social strategies, use interaction with other people to improve language learning
- MEM = memory strategies, involve relating the new word to some previously learned knowledge, using some form of imagery or grouping
- COG = cognitive strategies, manipulate or transform the target language
- MET = metacognitive strategies, involve a conscious overview of the learning process.²⁴

Michael Lewis suggests six basic types of lexical exercises, derived from his *The Lexical Approach*, aimed at recognizing word relations and deducing patterns by a process of “elimination, using linguistic clues, the group's shared knowledge, and a small element of plain guesswork”.²⁵ They are:

1. **Identifying chunks:** correctly identifying chunks is the fundamental skill that allows for easy storage in the mental lexicon and prevents learners from making inappropriate L1 word = L2 word assumptions
2. **Matching:** providing lexical focus by asking learners to match parts of collocations, expressions, lines of stereotypic dialogue etc.
3. **Completing:** asking learners to fill in “double gaps” with partner words from strong collocations
4. **Categorising:** sorting words or expressions according to their pattern (e.g. verbs or adjectives which partner one or two given nouns, expressions which are elements of two different dialogues, more formal or informal expressions, expressions with positive or negative connotations, etc.)
5. **Sequencing:** giving five expressions or verbs and asking learners to put them in the most likely order by using pragmatic knowledge (like in “scripts”, where actions are performed according to their cause/effect sequence, for instance)

²⁴ See *ibid.*, pp. 135–136.

²⁵ Michael Lewis, *Implementing the Lexical Approach*, p. 88.

6. **Deleting:** preventing learners from over-generalising by analogy with L1, usually focussing on collocations.²⁶

Successful strategies to improve vocabulary acquisition include the activation of more than one sensory receptor. The simultaneous use of verbal and visual codes allows for the activation of both halves of the brain; the right half which is visual and considers the whole before the details, and the left half which carries out analytic, sequential and verbal tasks. This dual code, which involves both halves of the brain and may also include hearing, smell or feeling, provides better conditions for learning and memorization. Here are some dual code strategies grouped according to learning styles:

- visual: spider diagrams, charts, labeling pictures, drawing
- sound: matching rhyming words, listing words with similar beginning, intonation practice
- movement: miming, pointing at objects, card or domino games
- verbal: scripts, matching, ranking, synonyms and antonyms, crosswords, exploring corpora

As vocabulary is learned incrementally, lexical acquisition requires multiple manipulation of a word. Research has shown that 5 to 16 repetitions of a word are necessary for deep learning to take place.²⁷ As a consequence, recycling is necessary to prevent forgetting (most partially known words are usually forgotten without recycling). Reacting with personal response (such as expressing preference, opinion, prediction, or recalling memory, etc.) to the new information provides an emotional or affective connection which makes learning more memorable.

5. Active, passive and potential vocabulary

“It’s on the tip of my tongue...” Everyone has experienced it at one time: knowing a word, name or number, but not being able to

²⁶ Ibid., pp. 89–91.

²⁷ Paul Nation, *Teaching and Learning Vocabulary*, p. 44.

recall it when required. This occurs when a lexical item is stored in the memory as **passive vocabulary**, but needs a trigger to be retrieved and become available for productive use, although it may easily be recognized when reading or listening. The extent of passive vocabulary is commonly considered to be far larger than **active vocabulary**. We should however take into account the amount of words that we have never come across before, but are able to recognize as derivatives or compounds of known items. This is due to the similarity to **cognates** in other languages, or their predictability from context. This **potential vocabulary** is very difficult to measure and is activated by inference and recycling.

The shift from receptive to productive mastery of a word is not a linear process: Schmitt suggests that “an initial ability to use a word productively, at least in a limited way, precedes full mastery of its receptive aspects. Thus productive knowledge does not occur sequentially after receptive; there seems to be some overlap.”²⁸ The reason lies in the incremental and not linear nature of vocabulary acquisition.

6. Vocabulary and text types

When considering vocabulary for productive skills, it is very important to distinguish the **oral** and **written** codes, as the type of words used varies considerably. According to Dodman,²⁹ spoken language is more dynamic, matter-of-fact, verbal, and focuses on people, actions and processes. Moreover, as the interlocutor is present and takes part in the conversation (or listens to the speech), a short time is given for processing the information; therefore the **lexical density** is usually quite low.³⁰ On the other hand, written

²⁸ Norbert Schmitt, *Vocabulary in Language Teaching*, p. 119.

²⁹ Martin Dodman, “Linguaggio verbale e linguaggio nominale. Rappresentazione dinamica e rappresentazione sinottica”, lecture at Scuola Autunnale DILM, 2006, http://www.roberto-crosio.net/SIS/DODMAN_Linguaggio_nominale_linguaggio%20verbale.pdf (accessed June 20, 2009).

³⁰ “Lexical density” is the proportion between the amount of content words (nouns, adjectives, verbs) compared to functional elements (prepositions, adverbs, conjunctions) in a text. To calculate the index divide the number of content words by the total number of words in a text; the typical figure for everyday speaking index is around 0.3–0.4.

language is more abstract, static, expressing facts, knowledge and products. It is a language created for a distant reader, with little pragmatic reference, and its meaning is decoded only by means of the grammar-lexical and textual resources, so its **index of nominalization** is fairly high, as is the cognitive load.³¹

Also scientific and technical language has a tendency towards nominalization, a higher degree of formality, and emotional detachment both in the oral and written forms. Typical features of formal standard Italian include a large use of passive forms, emphasizing the result rather than the process, with subjunctive and conditional verbs and longer or more complex sentences, connected with relative pronouns, conjunctions and other discourse markers. As for the register, a more formal vocabulary is used.

Awareness of such specific features of texts gives the teacher hints to devise practical activities that involve deep strategies and dual code activation to allow meaningful vocabulary acquisition.

In recent years studies in language corpora have provided valuable information about vocabulary which can be effectively used in language learning. Corpora consist of huge collections of language samples stored on computer database and accessible online for queries.³² As they are constantly being added to, they provide updated current contemporary language: authentic material from different geographical areas, genres, and types of texts, both spoken and written. Corpora can help learners understand more about vocabulary usage in context, showing what is actually said or written in the target language in a variety of text types.

³¹ The index of nominalization is obtained by dividing the number of nouns by the total number of content words in a text; 1 is the highest ratio (which means that all content words are nouns), conveying a huge cognitive load. See Martin Dodman, "Linguaggio verbale e linguaggio nominale. Rappresentazione dinamica e rappresentazione sinottica".

³² Corpora of Italian language include Badip: <http://languageserver.uni-graz.at/badip/badip/home.php> (accessed June 15, 2009), CoLFIS (written) : <http://www.istc.cnr.it/material/database/colfis/> (accessed June 15, 2009), Lablita: <http://lablita.dit.unifi.it/> (accessed June 15, 2009). A full list is provided in [http://www.alphabit.net/ Corsi/IUlinks/CorporaList.htm#italiano](http://www.alphabit.net/Corsi/IUlinks/CorporaList.htm#italiano) (accessed June 15, 2009).

7. Conclusion

A survey among the students of Italian courses, using the questionnaire in the appendix below, revealed that most learners rarely think of their vocabulary acquisition strategies, whereas some of them use consciously only shallow strategies.³³ Awareness about learning strategies will certainly help develop one's own learning style and focus on the most productive activities.

As we know that vocabulary acquisition is an incremental process and many aspects of the lexical unit must be considered, we should plan our teaching including a variety of vocabulary activities. Noticing the new items is not enough, they must be re-encountered in different contexts where different senses of the same word are given, then recycled and manipulated in a creative way. Memory retention is improved by the use of visuals and sound, as well as by creating emotional connection (when asked for personal response or opinion, for example); students with analytical learning style will appreciate verbal games, matching and puzzles. New vocabulary has to be worked through, producing contextualized messages which bring the focus from isolated words onto frequent collocates and multiword phrases; allowing faster memorization of these chunks.

In conclusion, vocabulary learning is a very complex and multifaceted cognitive activity and, by being exposed to varied tasks, students should gain awareness of their own learning styles thus developing more proficient and autonomous language acquisition.

³³ The results show that about 80% take notes in class, try to guess from context and use bilingual dictionaries, 60% ask classmates, 30% connect the word to antonyms and synonyms and keep a vocabulary notebook; but only 20% connect the word to personal experience.

Appendix

QUESTIONNAIRE ABOUT VOCABULARY LEARNING

How do you learn new words? Tick the options you generally use:

- use word lists to study words
- study the sound of target word
- take notes in class about vocabulary
- say the new word aloud when studying
- ask classmates for meaning of a word
- guess word's meaning from textual context
- study the spelling of the target word
- written repetition
- verbal repetition
- bilingual dictionary
- use spider diagrams
- group words in semantic areas
- form a visual image of the word's meaning
- use physical action when learning a word
- connect the word to a previous personal experience
- connect the word to antonyms and synonyms
- keep a vocabulary notebook

ABSTRACT

Vocabulary Acquisition in Language Learning

The article will explore the process of learning vocabulary in a foreign language. The term “word” does not refer just to a unit of meaning, but implies many aspects, such as sound, graphics, metaphor, collocation, that we have to take into account in language teaching. Vocabulary acquisition is an incremental process; therefore new items, including both single words and multi-word units, must be encountered several times, then recycled and employed in productive ways. Therefore, memory activation is crucial and learners select strategies according to their own learning styles. The language teacher should be aware of the process and provide students with activities meant to promote independent development of vocabulary.