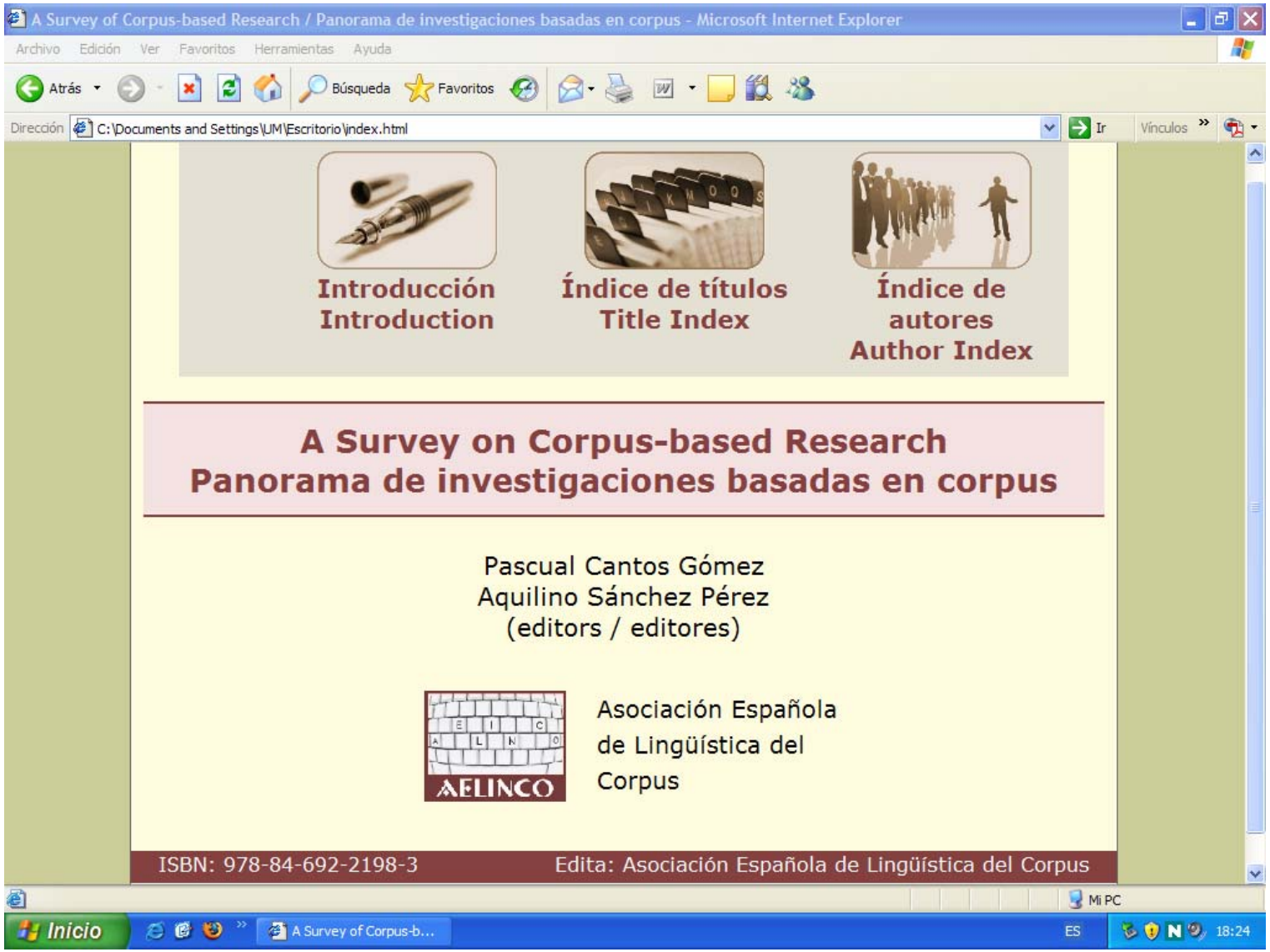
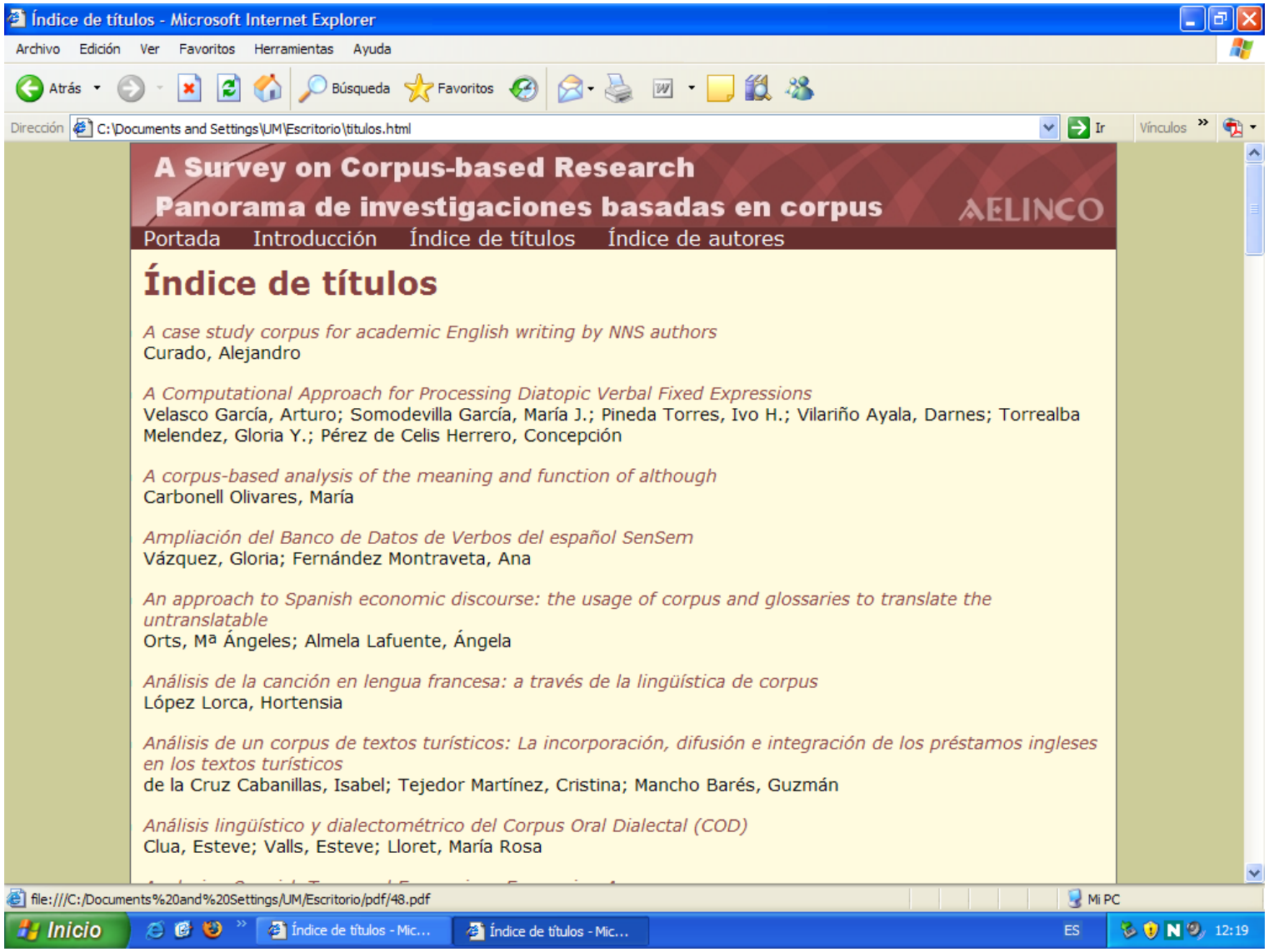

A Survey of Corpus-based Research / Panorama de investigaciones basadas en corpus - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos
Dirección C:\Documents and Settings\UM\Escritorio\index.html
Ir Vínculos
Introducción
Introduction
Índice de títulos
Title Index
Índice de autores
Author Index
A Survey on Corpus-based Research
Panorama de investigaciones basadas en corpus
Pascual Cantos Gómez
Aquilino Sánchez Pérez
(editors / editores)
AELINCO
Asociación Española de Lingüística del Corpus
ISBN: 978-84-692-2198-3
Edita: Asociación Española de Lingüística del Corpus
Mi PC
Inicio
A Survey of Corpus-b...
ES
18:24


Índice de títulos - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos
Dirección C:\Documents and Settings\UM\Escritorio\titulos.html
Ir Vínculos
A Survey on Corpus-based Research
Panorama de investigaciones basadas en corpus
AELINCO
Portada Introducción Índice de títulos Índice de autores
Índice de títulos
A case study corpus for academic English writing by NNS authors
Curado, Alejandro
A Computational Approach for Processing Diatopic Verbal Fixed Expressions
Velasco García, Arturo; Somodevilla García, María J.; Pineda Torres, Ivo H.; Vilariño Ayala, Darnes; Torrealba Melendez, Gloria Y.; Pérez de Celis Herrero, Concepción
A corpus-based analysis of the meaning and function of although
Carbonell Olivares, María
Ampliación del Banco de Datos de Verbos del español SenSem
Vázquez, Gloria; Fernández Montraveta, Ana
An approach to Spanish economic discourse: the usage of corpus and glossaries to translate the untranslatable
Orts, Mª Ángeles; Almela Lafuente, Ángela
Análisis de la canción en lengua francesa: a través de la lingüística de corpus
López Lorca, Hortensia
Análisis de un corpus de textos turísticos: La incorporación, difusión e integración de los préstamos ingleses en los textos turísticos
de la Cruz Cabanillas, Isabel; Tejedor Martínez, Cristina; Mancho Barés, Guzmán
Análisis lingüístico y dialectométrico del Corpus Oral Dialectal (COD)
Clua, Esteve; Valls, Esteve; Lloret, María Rosa
file:///C:/Documents%20and%20Settings/UM/Escritorio/pdf/48.pdf
Mi PC
Inicio
Índice de títulos - Mic...
Índice de títulos - Mic...
ES
12:19


Índice de títulos - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos
Dirección C:\Documents and Settings\JM\Escritorio\titulos.html
Ir Vínculos
Olmos, Carmen
The reanalysis and grammaticalization of nominal type noun constructions with kind/sort of: chronology and paths of change
Davidse, Kristin; Brems, Lieselotte
The use of metaphorical compounds on the web: corpus construction and preliminary analysis
Koteyko, Nelya
The use of specialised corpora: implications for research and pedagogy
Gil Salom, Luz
Usos metafóricos del lenguaje político español. La metáfora estructural en los debates sobre el estado de la nación
Sánchez García, Francisco José
Utilidad y limitaciones de los corpora informáticos en la elaboración de un tesauro fraseológico (alemán-español)
Mellado Blanco, Carmen
"Vengeance" and "Honour" in Shakespeare's Hamlet and Sumarokov's Gamlet: A Collocational Approach
Cantos, Pascual; Keshabyan, Irina
Vocabulary in EFL Textbooks. A Contrastive Analysis against Three Corpus-Based Word Ranges
Criado, Raquel; Sánchez, Aquilino
Vocabulary in EFL textbooks: Frequency levels
Alcaraz Mármol, Gema
Written learner corpora by Spanish students of English: an overview
Díez-Bedmar, María Belén
ISBN: 978-84-692-2198-3
Edita: Asociación Española de Lingüística del Corpus
Design and implementation: Juan Antonio Pastor Sánchez
© AELINCO 2009
Mi PC
Inicio
Índice de títulos - Mic...
Índice de títulos - Mic...
ES
12:19

Vocabulary in EFL Textbooks. A Contrastive Analysis against Three Corpus-Based Word Ranges

Raquel Criado[1]
Aquilino Sánchez
University of Murcia

Abstract
Vocabulary teaching and learning is one the most important components in textbooks. Vocabulary knowledge is also frequently associated to language fluency. It is therefore important to investigate how textbooks present vocabulary, and how they comply with the conditions necessary for vocabulary learning. We will take into consideration three perspectives here: the conditions derived from cognitive processes of knowledge acquisition, the role of frequency in language and vocabulary learning and the distribution of new words throughout textbooks. The analysis of a specific EFL textbook and the comparison of the results against the rationale supporting the three perspectives mentioned above will reveal whether the textbook is suitable for efficiently reaching the goals regarding vocabulary teaching/learning.

Keywords: vocabulary learning, SLA, corpus linguistics, frequency, repetitive practice

I. INTRODUCTION

The history of vocabulary teaching has mainly centred on the teaching of words as isolated or de-contextualized items (Howatt, 2004; Kelly, 1969; Sánchez, 1997, 2009; Schmitt, 2000). It has been clearly so in the traditional *Grammar Translation Method*, but also in other methods not so strongly based on grammar, such as the *Direct Method* and other approaches heavily based on teaching through reading and memorization of dialogues. More specifically, the teaching and learning of vocabulary lists has been one of the pillars in the classroom for centuries. Grammatical rules were followed by classroom practices in which students had to combine the words learned in order to build the kind of sentences required by the rules. It was not explicitly stated, but words were considered or perceived as fully autonomous lexical elements, and they were learned as such. On the other hand, words and grammar constituted the skeleton of the linguistic system, and the tendency was to associate the command of a large vocabulary to language fluency. Some scholars share the same opinion nowadays (Laufer & Goldstein, 2004).

Until recently, the context-dependency of word senses was not a subject of discussion and the learning of words without its corresponding co-textual correlates was the rule. The clearest exponent of this conceptualization of words is found in the *Grammar Translation* method, which pursues the learning of vocabulary thorough the memorization of words presented to the students in large and de-contextualised lists, sometimes grouped in semantic areas. Learning linguistic patterns or language chunks was not explicitly excluded, but this

[1] ***Address for correspondence***: Raquel Criado, Departamento de Filología Inglesa, Campus de *La Merced.* Universidad de Murcia. 30071 Murcia, Spain. Tel: 00 34 868888791; Fax: 00 34 868883185; e-mail: rcriado@um.es Aquilino Sánchez, Departamento de Filología Inglesa, Campus de *La Merced.* Universidad de Murcia. 30071 Murcia, Spain. Tel: 00 34 868883175; e-mail: asanchez@um.es.

was not the kind of learning content favoured by teachers. Teaching activities implied the underlying assumption that words were to be taken as autonomous elements. The absence of any contextual and meaningful elements in the teaching/learning of vocabulary may trigger ambiguity or misunderstanding, and therefore it may lead to failure in communication. However, this question was not even raised at the time.

Vocabulary is of primary importance in language teaching and in linguistic communication, but it has not always been adequately emphasized or adequately highlighted. In particular, the nature of words and its contribution and role in the building of meaning has not been correctly evaluated by most teaching methods. In the last two decades, though, the importance of vocabulary knowledge has been brought to the forefront, especially in the field of vocabulary acquisition research and assessment (Laufer & Hulstijn 2001; Nagy & Scott, 2000; Nation, 2001; Read, 2000, etc.). This movement in favour of vocabulary is rooted in the assumption that the knowledge of words improves the communicative potential, linguistic fluency and accuracy.

Our goal in this paper is in line with the increasing interest in vocabulary studies and will attempt at analysing the lexical component in a specific textbook from three points of view: frequency, distribution and the kind of activities through which vocabulary learning is promoted.

II. VOCABULARY ACQUISITION AND LEARNING

Teaching and learning vocabulary is a most important component in the classroom and in teaching materials. In this section we will briefly highlight three relevant aspects in vocabulary acquisition and learning: (i) the acquisition of vocabulary from a cognitive perspective, i.e., as conditioned by the processes that our brain must necessarily follow; (ii) the role of frequency in vocabulary acquisition or learning and how this question affects textbooks, and (iii) the distribution of the vocabulary to be learnt all along the textbook. Words to be learnt should take into account the learning conditions. These require that new words be introduced progressively and gradually, and at the same time that words previously presented and learnt be re-introduced later again in order to favour repetition, and therefore, proceduralization.

II.1. The learning of vocabulary from a cognitive perspective

Knowledge of the words of a language is the type of knowledge referred to as ‘declarative knowledge’ (*DEC*). *DEC* opposes ‘procedural knowledge’ (*PRO*). The nature of both types of knowledge may imply different strategies for their acquisition. That is the case when we refer, for example, to the role of consciousness (explicit) or implicitness (incidental) in learning. Regarding the consolidation of both types of knowledge, however, the basic strategy is the same: consolidation depends on previous memorization, and memorization is governed by rehearsal. It is true that *DEC* may require only a single stimulus to be acquired at times (Ullman, 2004), while *PRO* will practically always result from repeated action triggered by recurrent stimuli. Nevertheless, the consolidation of both *DEC* and *PRO* share a similar

need for repetition before becoming automatized (Sánchez & Criado, in press). Automatization is the only condition in skill learning that guarantees fluency of performance, which in the case of language will be fluency in communication.

Declarative knowledge is acquired through association. In the case of vocabulary, the acquisition depends on the association of things in the outside world to a concept in our mind. Associations are triggered by stimuli in the neural network (Ullman, 2004). A stimulus may begin at a specific neural node and is transmitted to other neurons by means of neurotransmitters, which result from the release of chemicals that change the electric polarization of the membrane in the neural receptors. The transmission of the electrical signals runs along specific channels, which strengthen under certain conditions. Full consolidation is reached when the same stimulus is able to automatically activate an already shaped channel and produce similar results at the end of the neural circuitry. There is still a long way ahead to fully understand how these initial electrical bits generated by and transmitted through the neural system derive into knowledge. Psycholinguistics firstly and neurolinguistics in the last decades are contributing a better understanding of the cognitive processes that generate what we refer to as 'knowledge' (Anderson, 2005).

One of the most relevant areas of cognitive processes is how data are accessed, transmitted and memorized. Memory is particularly important in cognitive processes, since it is the device responsible for storing data, keeping them at our disposal and accessing them whenever we need them. Our neural system is known to work with two types of memorization devices: *short-term memory* and *long-term memory* (Anderson 2005; Atkinson & Shiffrin 1968). Data captured are first presented to short-term memory, a kind of working memory acting as an interface with the outside world. Input entering the working memory flows very quickly and is immediately lost unless it enters long-term memory. Therefore, it can be stated that our working memory is the main entrance for input data; it is equipped with a filter for evaluating and selecting only the data considered relevant or necessary.

From the point of view of efficiency in vocabulary learning, what matters is the amount of lexical information entering and consolidating in long-term memory. Neurologists and psycholinguists tell us that long-term memory is activated and strengthened mainly (i) through rehearsal or repetitive practice and activation, (ii) when attention is drawn to specific data, and (iii) when new data are associated in some way to already consolidated information. The three options are accessible to learners and teachers.

Repetitive practice has been present all throughout the history of school teaching and there is no doubt on its efficacy as a teaching and learning technique (Sánchez & Criado, in press). The efficacy of repetition is due to the structural changes that take place in the neural synapses (or connections among neurons). Repeated connections strengthen the channel, and so the task is rendered easier. When the task becomes so easy that you can perform it with less effort or attention, it is because a certain degree of proceduralization of the process has been reached. At this point in the process, structural changes in the synapses affected apparently cease and become stable. In addition, more practice implies more efficient execution. Facts regarding the two types of memory and the consolidation of data may be synthesized in the following way: most of the information which flows through the short-term

memory is usually lost, pressed by the permanent flow of incoming data, unless repetitive iteration and/or attention favours its selection to enter long-term memory. Iteration or *repetition*, together with attention, is therefore the habitual mechanism, which guarantees permanence and avoids oblivion in information storing.

Cognitive processes in knowledge acquisition imply some conditions, which teachers and teaching materials must meet. One of them asks not only for the presentation of new words, but also for opportunities to encourage and facilitate repetition. To reach such a goal, textbooks must be adequately planned regarding vocabulary distribution, firstly allowing for the presentation of new lexical items and promoting, in a second stage, abundant instances for repetition of the items previously presented. In addition to that, course books should also provide opportunities for explicit and implicit vocabulary acquisition. The role of consciousness is emphasized in explicit activities, while language usage (both receptive and productive) is at the base of implicit activities. Both types should be at work in vocabulary learning; the method claimed by the textbook will define the prevalence of one or the other type of activities.

II.2. Frequency and vocabulary learning

Frequency matters in vocabulary learning for two reasons: (i) most often used lexical items should be learned first, since they contribute towards a communicative efficiency more significantly, and (ii) frequency of occurrence offers opportunities for repetitive practice. This is one of the necessary conditions for vocabulary consolidation, as is required by cognitive processes in knowledge acquisition.

Studies in vocabulary frequency abound nowadays, hand in hand with the increasing interest in corpus linguistics and the computational facilities available (D'Anna, Zechmeister and Hall, 1991; Nation, 1993a, 1993b; Sánchez, 2000; Schmitt, 2000, among others). We have easy access today to frequency lists of language use, and consequently reliable information on which words are preferred by the speakers in different domains and communicative situations. Information on lexical frequency was already used in the *Audio-lingual* method to take decisions on which words to select for the different teaching levels (Sánchez, 2009). It was then assumed that the most frequent 800-1,000 words were to be learnt by elementary level students, while intermediate and advanced students would be presented the next 800-1,000 or 2,000 words in the frequency list. Research in this field has refined the tools of analysis and offers reliable and very useful results for teachers and textbooks. The work by Nation (2001, 2006) is particularly relevant in this respect. Nation takes a classical classification of words: as *tokens* -every word form in the text, be it repeated or not-, *types* -different words in the text, such as *friend* and *friends,* which are two types- and *word families* -the headword, its inflected forms and its closely related derived forms. He then establishes three consecutive vocabulary ranges, one thousand words each, based on corpus frequency data.

The analysis of the relationship among the three previous classes of words in a coursebook contributes towards significant information regarding the words really presented in it, their frequency and distribution along the book and the opportunities for repetition

directly depending on frequency. Furthermore, the grouping of words in word families adds useful and complementary information on the formal and semantic relationships among the lexical items presented as learning targets. The three basic ranges of words are roughly representative of the three basic levels in language teaching materials concerning the amount of lexical items usually introduced in each level: beginners (first 1,000 most frequent words), intermediate (second 1,000 most frequent words) and advanced (third 1,000 most frequent words). Therefore, those word ranges may be taken as a reference against which teaching materials (and specifically textbooks) can be compared. The results of such a comparison will reveal if a specific coursebook complies with the expectations regarding vocabulary usage and the conditions governing knowledge acquisition in general and vocabulary acquisition in particular. The parallelism between real word usage and frequency and the vocabulary offered by textbooks can be taken as a positive value, since students may benefit from it. The lack of such a parallelism would be considered a rather negative factor, given that teaching should look for more efficiency in communication. This correlates with the amount of words learned and their relevance for communication.

II.3. The distribution of vocabulary in textbooks

The three corpus-based word ranges defined by Nation (2001, 2006) are indicative of three consecutive stages in language command, as reflected in vocabulary use: beginners, intermidiate and advanced. This is precisely the organizational scheme of most textbooks. It is therefore to be expected that coursebooks for each one of those levels keep close to the three vocabulary ranges mentioned above. Beginner courses should roughly include range 1 words, intermediate courses should cover word range 2 and advanced courses should include word range 3. Textbooks can be analysed against this expected pattern and the vocabulary they contain evaluated according to the model described in each word range.

The analysis will reveal how textbooks adjust to the word range they have been designed for. Moreover, a closer look at the vocabulary presented will allow us to discover and define with accuracy in which way and how much a specific textbook deviates from the word range it claims to represent. The requirement is that we compare the vocabulary in the textbook against each one of the three predefined ranges.

The sequence of ranges and the conditions derived from the cognitive processes underlying language acquisition ask for some additional requirements regarding vocabulary distribution along textbooks. Firstly, new words must be gradually introduced, possibly within the adequate communicative context; secondly, new words should mainly pertain to the range the textbook belongs to; thirdly, the working vocabulary as a whole should include enough instances and offer abundant opportunities for practicing the new words previously introduced, be it in the same word range or in the previous ones. The result of such an analysis, implemented with computational tools, will clearly reveal if textbooks are the right instruments and if they can be considered suitable guides for teaching/learning vocabulary in a most efficient way.

III. A CASE STUDY: THE LEXICAL COMPONENT IN A TEXTBOOK

Our goal in this section is to analyse the vocabulary of a textbook and, as previously explained, to compare the results with the three word-based vocabulary ranges defined by Nation (1,000, 2,000, and 3,000 most frequent word families of English). This task requires that we first identify the vocabulary offered in the textbook. In a second stage, we will find out if the vocabulary matches the three aforementioned word ranges.

The computational tool used for counting and comparing vocabulary in the textbook will be RANGE (http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx), which classifies the vocabulary of any text into three frequency categories: the first 1,000, the second 1,000 and the third 1,000 most frequent word families of general English. Words not included within these first three categories appear as off-ranges. The classification of words as tokens, types and word families is also very relevant for our study. The identification of tokens vs. types allows for a contrast between the raw vocabulary input against the new words really introduced in a specific text. From the perspective of vocabulary acquisition we will later check if the textbook complies with the specific conditions governing knowledge acquisition, in particular (i) those regarding 'opportunities for repetition', which will depend on the frequency of occurrence of lexical items throughout the textbook, and (ii) the amount of activities favouring explicit or incidental learning.

The textbook analysed is *Valid Choice 2*, by Jane Lawrence and Alan Williams, published by Burlington Books (2006). The manual is adapted to the syllabus of the Spanish *Bachillerato*, Course 2, which could roughly be classified as an intermediate level. Following the official specifications and regulations, the methodological approach must adjust to the Communicative Method and to the principles underlying the *Common European Framework of Reference for Languages* (2001). The authors claim a particular emphasis on vocabulary learning as well. The book is structured in 6 main units, 10 pages each. Some specific sections are also included in the student's book: a section for 'exam preparation', a grammar appendix and a glossary. The glossary includes about 500 words (lemmas) only, which are defined as 'the most frequent words', with no further specifications. Such a glossary clearly contrasts against the 3,225 types -ca. 2,320 family words- identified in the manual (see next section).

III.1. Word counts and word ranges

All the units in the textbook reach 25,687 running words (tokens). Out of this total, 3,225 are distinct words (types). This figure amounts to about 2,320 word families. Regarding the word ranges defined by Nation (Nation, 2006), only 148 types belong to range 1, 630 to range 2 and 242 to range 3. The same types classified as word families would offer 113 for range 1; 434 for range 2 and 187 for range 3 (see Table 1).

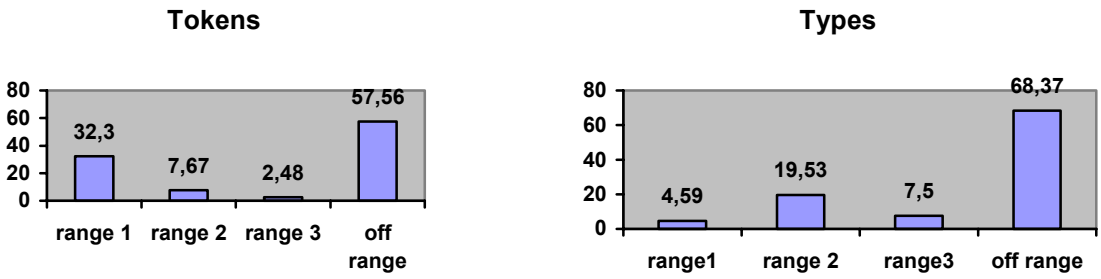
WORD range	TOKENS/%	TYPES/%	FAMILIES
(1) 1000	8296/ 32.30	148/ 4.59	113
(2) 2000	1970/ 7.67	630/ 19.53	434
(3) 3000	636/ 2.48	242/ 7.50	187
off lists	14785/ 57.56	2205/ 68.37	(not specified)
Total	25687	3225	734

Table 1. Tokens, types and word families for Ranges 1, 2 and 3 in *Valid Choice 2*.

The first striking figure refers to the amount of distinct words used in the textbook: *Valid Choice 2* contains 3,225 types (ca. 2,320 word families if we add those not specified in the range 'off lists'). The class hours during the academic year amount to 100. This fact implies that if students are to learn all the words included, they should learn 32 new types per hour, almost 100 per week, or 400 a month. In addition to that, the consolidation of the words already introduced in previous sessions should be also guaranteed. Such expectations exceed by far the most optimistic views on word acquisition. Ito (1995) concluded in an experimental study with Japanese students that they learned only 3 new words per day, that is, 20/22 per week. Our textbook lies too far away from expectations. You may argue that textbooks should not only offer the words to be acquired by the learners. Specific communicative events and situations require the use of low frequency contextual vocabulary, which must not necessarily be a primary learning target. In *Valid Choice 2*, out of the total 3,225 words (types) introduced, 1,345 occur only once, and 528 occur twice in the texts and exercises. Instances of words occurring once or twice give a total of 1,873. It could be assumed that instances of low occurrence (fewer than three occurrences) can hardly be considered candidates for memorization and could be excluded. The exclusion of words occurring once and twice would lower the amount of words for acquisition to 1,352, half of the total of types in the textbook. Still, learning 13.5 types per day (1,352 in 100 hours) is far from what experimental studies predict as adequate and within the acquisition potential of learners.

Regarding the three word ranges specified by Nation, 8,296 tokens are included within *range 1* (32.3% of the total). The figure seems reasonable in terms of percentage, but it only covers 148 types (4.59%) and 113 word families (Table 1). The unbalance between tokens and types is due to the high frequency of a few lexical items in *range 1*, which does not contradict *per se* the normal distribution of words in texts. The problem is though that 852 types of range 1 do not occur in the textbook. It is hard to assume that students have already fully consolidated those 852 high frequency items. It would have been more reasonable to reinforce their acquisition in level 2, at least with occasional occurrences.

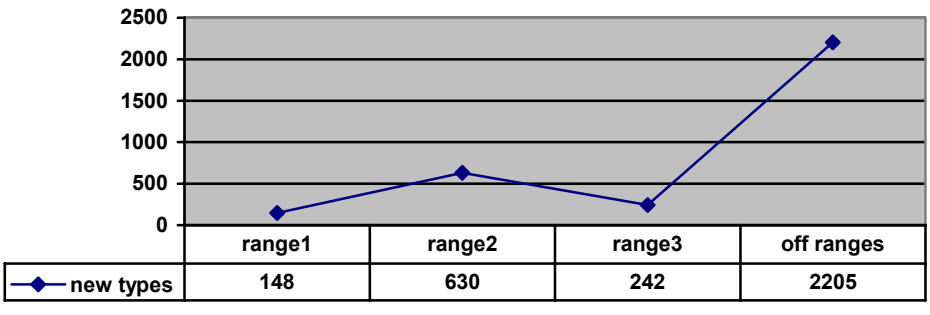
Valid Choice 2, a textbook for students who have already completed *Valid Choice 1* (designed for initial B1 level), should pay attention to the vocabulary of range 2 (second 1,000 most frequent words) and particularly to range 3 (third 1,000 most frequent words). Tokens in *range 2* in the textbook are actually 1,970 (7.67% of the total). They include 630 types (19.53%), and 434 word families (see Graphic 1).



Graphic 1. Types and tokens in *Valid Choice 2*.

These figures need some comments. The amount of tokens of range 2 is too low if compared to the total in the textbook, and more specifically if compared to the items included within range 1 (as the related percentages clearly show). The amount of types in range 2, however, is significantly higher: 630 (19.53%); this is also the case for word families. The relative lack of balance in the amount of tokens and types regarding the total of lexical items in the textbook implies serious negative consequences (as it was the case in range 1 vocabulary, but in the opposite direction). It means that the textbook introduces a reasonable amount of range 2 types (630/1,000), but their frequency is too low to favour effective consolidation or automatization, since students will find each new item only three times throughout the textbook (the average that results from dividing 1970 (tokens) by 630 (types)).

Range 3 is represented by 636 tokens in the textbook (2.48%). If the textbook is to reach the consolidation of level B1 (as prescribed by the Spanish official syllabus), this percentage lies far away from expectations. This level (defined for 'independent users' in the *Common European Framework)* requires a fluent communicative use of English in daily life, very much in line with the third 1,000 words included in range 3 plus the 2,000 words from the previous ranges 1 and 2. Accordingly, the types pertaining to range 3 in *Valid Choice 2* should equal at least the amount of words included in range 2; in any case, the vocabulary learned should follow a steadily ascending line from range 1 to range 3. In *Valid Choice 2*, the ascending curve for new vocabulary breaks off in range 2, falls down in range 3 and ascends abruptly in 'off ranges' (Graphic 2):



Graphic 2: New types along ranges.

In doing so, the book runs into a serious unbalance, which affects negatively the communicative potential of the vocabulary learned. From a pedagogical perspective, we

should expect that *Valid Choice 2* reinforces what has been learnt in *Valid Choice 1*, and should thus introduce new lexical items, which are proportionate to the learning potential of the students and to the ascending frequency line of general English. The new words should mostly appear first in range 2, and smoothly increase in range 3 (a higher level). A more advanced level (in that case 'off ranges' -B2?) is not the goal of this textbook and should consequently be poorly represented. This is not the case here: *Valid Choice 2* offers a strikingly high number of types above range 3: 2,205. The unbalance comes clearly into light in terms of percentage: the new items not included in range 1, 2 and 3 take 68.37% of the total of types in the book, against only 19.53% in range 2 and 7.50% in range 3 (apparently the closest to the goals of the manual). A sound distribution would ask just for the opposite: 68.37% of the new items (the highest figure) should belong to ranges 2 and 3, while the off-ranges interval should take lower percentages. Range 1 should be granted a moderate representation for consolidation purposes.

The conclusion is that *Valid Choice 2* is clearly unbalanced regarding

(i) the amount of vocabulary offered;

(ii) the distribution of vocabulary throughout the three ranges described by Nation (2006);

(iii) the frequency of vocabulary, which is too low and will not favour proceduralization and automatization through repetition;

(iv) the amount of words the students are expected to learn, which reach a level well above the more optimistic studies in the field.

III.2. Word frequency in presentation texts and activities

Textbooks are typically structured in two main sections: a first section with texts through which vocabulary and grammar relevant for the lesson are introduced in context, and a second section with activities, which aim at practicing the linguistic elements and grammar, selected as the goals of the unit and introduced in the first section. We will analyse only the distribution of vocabulary in each one of those sections.

Valid Choice 2 deviates significantly from pedagogically based expectations. The section with the texts should abound in new types, while the section with the activities should increase the amount of tokens in relation to types. The reason is obvious: presentation texts are specially selected to introduce new vocabulary, they are supposed to include repetition of words only occasionally. On the other hand, the section with activities is specifically designed to practice with words and grammatical structures as a means to consolidate acquisition. Table 2 illustrates quantities in the section with texts and in the section with activities:

SECTION WITH TEXTS ONLY:

WORD range	TOKENS/%	TYPES/%	FAMILIES
(1) 1000	2821/35.33	130/ 6.79	101
(2) 2000	568/ 7.11	350/18.28	283
(3) 3000	205/ 2.57	133/ 6.95	109
off ranges	4390/54.98	1302/67.99	(not specified)
Total	7984	1915	493

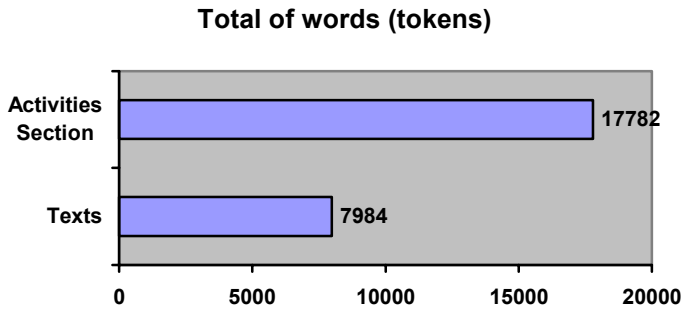
SECTION WITH ACTIVITIES ONLY:

WORD range	TOKENS/%	TYPES/%	FAMILIES
(1) 1000	5520/31.04	131/ 5.14	104
(2) 2000	1386/ 7.79	473/18.56	342
(3) 3000	438/ 2.46	180/ 7.06	144
off ranges	10438/58.70	1765/69.24	(not specified)
Total	17782	2549	590

Table 2. Tokens, types and word families in the text and activity sections from *Valid Choice 2*.

From the analysis of these data, several facts stand out:

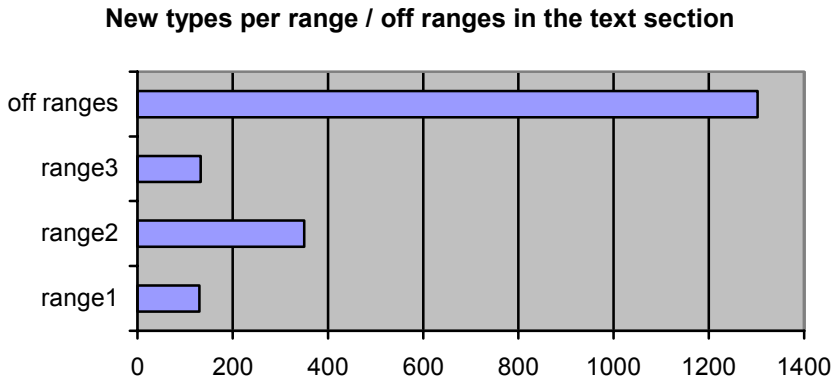
Fact 1: the total number of lexical items in the activity section only doubles that in the text section. Consequently, the opportunities for repetition are low: each token introduced in the first section will be repeated only twice on average. This is clearly shown in Graphic 3, which illustrates the relationship between the total number of words in each section.



Graphic 3. Total number of words in the text and activity sections from *Valid Choice 2*.

Fact 2: As can be seen in Table 2, the amount of tokens utside the word ranges is too high in both sections: it takes 54.98% and 58.70% in the first and second section respectively, that is, more than half of the words are above the 3,000 more frequent words threshold.

Fact 3: As Graphic 4 reveals, there is a strongly marked unbalance between the new words introduced in ranges 1, 2 and 3 and the rest of words outside these ranges.



Graphic 4. New types per range and off ranges in the text section from *Valid Choice 2*.

Range 1 counts with only 130 types, range 2 includes 350 and range 3 is represented with 133; 1,302 types fall outside these ranges (67.99%). Figures are similar in percentage for the activity section (131, 473, 180 and 1,765 respectively), reinforcing the 'functional' unbalance between both sections: the ideal proportion would ask for a significantly higher number of tokens in the activity section. Higher frequency of occurrence favours acquisition because it grants more opportunities for repetition and hence for automatization. However, tokens in the activity section total only 17,782 words (Graphic 3), just 2.2 times more than in the text section (7,984). The opportunities for repetition are very poor indeed.

The conclusion is necessarily negative regarding the lexical distribution in each one of the sections. The opportunities for repetitive practice in the activity section are very low and this fact distorts the functional expectations of the text and activity sections (introducing new material and practising respectively). The textbook does not offer teachers and students the expected and necessary opportunities for automatizing vocabulary acquisition. Moreover, (i) as indicated in section III.1., the amount of lexical items introduced exceeds by far the potential of learners for vocabulary acquisition and the rate of vocabulary learning, (ii) the lexical items introduced do not keep in line with the frequency lists; too many of them (1,302 out of a total of 1,915 in the text section) are not included in the three most frequent word ranges. This means that students will be primed to learn words of poor potential for communication in the level prescribed (B1).

III.3. Explicit vs. incidental vocabulary learning activities

Vocabulary knowledge is necessary for language fluency (Anderson & Freebody, 1981; Goulden et al., 1990; Laufer, 1998; Laufer & Nation, 2001; Read, 2000). As indicated in section II.1, explicit (conscious) and incidental (non-reflective and based on language use) learning activities are both important for vocabulary acquisition. Explicit learning is important because it attracts the attention of students and so it triggers the transfer of data from short-term to long-term memory; incidental learning is also relevant because it favours lexical consolidation and automatization, more slowly indeed but adding the advantage of contextualization and more realistic communicative contexts. As for lexical acquisition, a textbook may therefore offer explicit or incidental opportunities depending on the kind of activities included. Explicit vocabulary learning will be the object of activities in which the students' attention is directly drawn on to specific words or phrases by means of various strategies, such as the ones suggested by the following instructions:

Match the synonyms below.

Complete the sentences with a suitable adjective from the list above.

Incidental vocabulary learning will be triggered by activities in which students are involved in language use, such as reading, writing, speaking or listening, or whenever they must engage in exercises centred on the reception, interpretation, reshaping and transmission of meaning, as shown in the following instructions:

Skim the text and find out what Shakira's greatest challenge was.

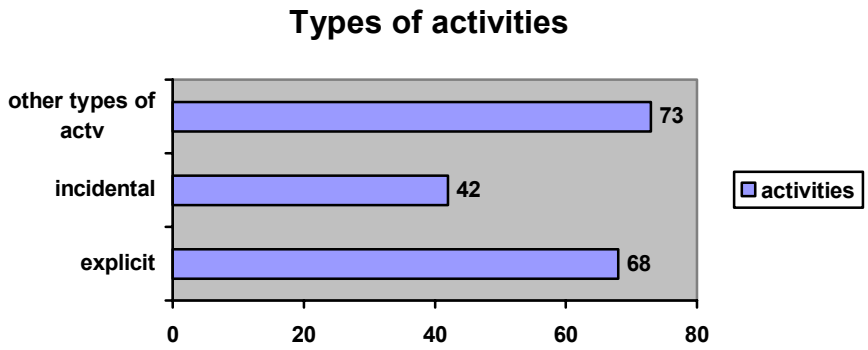
Are the sentences below true or false? Find evidence in the text to support your answers.

A careful analysis of the activities in *Valid Choice 2* offers the following results (Table 3):

Unit	Activities favouring explicit learning	Activities favouring incidental learning	Total of activities per unit
1	10	7	29
2	12	8	31
3	14	9	31
4	11	5	31
5	10	6	30
6	11	7	31
TOTAL	**68**	**42**	**183**

Table 3. Type of learning (explicit and implicit) and related lexical activities.

Graphic 5 below illustrates those figures:



Graphic 5. Number of activities per type of learning in *Valid 2*.

From a global point of view, the amount of explicit vocabulary learning activities is reasonably high, since it reaches 1/3 of all the activities in the coursebook. Activities that favour vocabulary incidental learning are also high: they cover 23% of all the activities. It must be added that the proportion of explicit and incidental vocabulary activities is homogeneously distributed all along the units, as the distribution of activities per unit shows. We must therefore conclude that, from the point of view of the amount of exercises devoted to vocabulary learning, *Valid Choice 2* is on the right track to reach the expected goals for vocabulary acquisition.

IV. CONCLUDING REMARKS

The analysis carried out in this paper does not allow for an optimistic conclusion regarding vocabulary presentation and distribution in the textbook examined. An overall evaluation leads to conclude that the manual is not in line with some of the basic and fundamental requirements for lexical acquisition, specifically the suitability of the words selected regarding the frequency list of general English and the lack of opportunities for

rehearsal and repetition of the lexical items the students are supposed to learn. It is true that textbooks are limited tools for language learning, since they are limited in size and in the communicative situations available. Still, the unbalance exhibited by *Valid Choice 2* in all the aspects analysed but one exceeds what may be considered reasonable. The efficiency of learning (the main goal all methods and manuals should strive for) would ask for a more comprehensive and solid understanding of the issues involved in vocabulary acquisition, as well as for a better control of the words to be learnt from the two following perspectives: (i) what is communicatively relevant, and (ii) some fundamental conditions based on the cognitive processes determined by human biological nature.

REFERENCES

Anderson, J. R. (2005). *Cognitive Psychology and its Implications*. (6th ed.). New York: Worth Publishers.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews.* Newark, DE: International Reading Association, pp. 77-117.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In K. Spende & J. Spence (Eds.), *The Psychology and Learning of Motivation* 2. New York: Academic Press, pp. 89-195.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Available on http://www.coe.int/T/DG4/Linguistic/CADRE_EN.asp

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

Goulden, R., Nation, P. & Read, J. (1990). How Large Can a Receptive Vocabulary Be? *Applied Linguistics*, vol. 11, 341-363.

Howatt, A.P.R. with Widdowson, H. G. (2004). *A History of English Language Teaching*, Oxford: Oxford University Press.

Ito, T. (1995). *The Acquisition of Basic Vocabulary by College-Age Japanese Students in an Intensive EAP Program*. Paper given at the *Annual Meeting of the Japan Association of College English Teachers*, Tokyo, September.

Kelly, L. G. (1969). *25 Centuries of Language Teaching*. NY: Newbury House Publishers.

Laufer, B. (1998). The Development of Passive and Active Vocabulary in a Second Language: Same or Different? *Applied Linguistics*, vol. 19:2, 225- 271.

Laufer, B. & Nation, P. (2001). Passive Vocabulary Size and Speed of Meaning Recognition. *EUROSLA Yearbook 1*, 7-28.

Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of Task-Induced Involvement. *Applied Linguistics*, vol. 22:1, 1-26.

Laufer, B. & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength and Computer Adaptiveness". *Language Learning,* vol. 54: 3, 339-436.

Lawrence, J. & Williams, A. (2006). *Valid Choice 2*. Burlington Books.

Nagy, W. E. & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of Reading Research* (vol. 3). Mahwah, NJ: Erlbaum, pp. 269-284.

Nation, I. S. P. (1993a). Using dictionaries to estimate vocabulary size: essential, but rarely followed, procedures. *Language Testing,* vol. 10:1, 27-40.

Nation, I. S. P. (1993b). Vocabulary size, growth and use. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon*. Amsterdam/Philadelphia: John Benjamins, pp. 115-134.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation I. S. P. (2006). How Large a Vocabulary is Needed for Reading and Listening? *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, vol. 63:1, 59-82.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Sánchez, A. (1997). *Los métodos en la enseñanza de idiomas. Evolución histórica y análisis didáctico*. Madrid: SGEL, S. A.

Sánchez, A. (2000). Language Teaching before and after 'digitilized Corpora'. Three main issues. *Cuadernos de Filología Inglesa*, vol. 9:1, 5-37.

Sánchez, A. (2009). *La enseñanza de idiomas en los últimos cien años. Métodos y enfoques.* Madrid: SGEL, S. A.

Sánchez, A. & Criado, R. (in press). Cognitive underpinnings of repetitive practice in the learning of EFL. *Proceedings of the 27th AESLA International Conference*.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.

Ullman, M. T. (2004). Contributions of Memory Circuits to Language: the Declarative / Procedural Model. *Cognition,* vol. 92, 231-270.