ABSTRACT
Despite the current plethora of corpora studies concerning language teaching, we find comparatively few regarding the analysis of didactic materials. The present study aims at comparing vocabulary input of a Primary Education coursebook against General English from the British National Corpora. Results are discussed both from a quantitative and a qualitative perspective. A considerable part of the vocabulary analysed falls out of the 2000 most frequent words in General English. Furthermore, the semantic fields that dominate the didactic units try to adapt to the functional language required by the children’s specific communicative context. Therefore, the study confirms that communication needs of children are afforded a greater weight than frequency criteria.

Keywords: Corpus, EFL, Frequency, Textbook, Vocabulary

I. INTRODUCTION

English language is constituted by around 54,000 word families (Webster, 1963). The amount of vocabulary known by an educated native speaker ranges between 17000 (Zechmeister et al., 1995) and 20000 (Nation and Waring, 1997) word families. It is quite complicated, though not impossible, that a L2 student learns that amount. The average L2 learner aims at getting to know the basic vocabulary to communicate with native speakers of that L2 – which involves the knowledge of a much smaller number of items. Francis and Kucera (1982) estimate that 2000 word families are enough for basic production and understanding. In turn, Laufer (1992) and Nation (2001) increase the minimum to 3000 words. Beyond disparities, when we claim that a student should know at least from 2000 to 3000 words we must define what knowing a word implies.

Linguists seem to disagree about what it means to know a word. Some of them (Richards, 1976; Nation, 1990; 2001) understand word knowledge as a taxonomy made up of different sub-knowledges regarding form, meaning, use and metalinguistic information. For others such as Faerch et al. (1984), Palmberg (1987) or Melka (1997), word knowledge develops along a continuum, with different degrees of familiarity. Laufer (1998) and Laufer and Paribakht (1998) conceive word knowledge mainly in passive and active terms, whereas authors such as Sánchez (2007) adopt the semantic content as the main parameter in their analysis. Thus, to Sánchez the word structure is organized in terms of the so called ‘lexical constellations’. Cognizant of the multidimensional nature of the word, Sánchez compares the different aspects of word meaning to the astronomical relationships between the planets and satellites. Similar to what happens in the universe, the different semantic features of a word are not organized independently ones from the others. They bear relationships that can be relatively more or less closer or direct.

On the other hand, some researchers identify word knowledge with Lexical Competence. In this regard, they have tried to reduce this competence to two or three dimensions. This is the case of Henriksen (1999) and Meara (1996). The former talks about three aspects on which word knowledge develops: partial-precise dimension, depth of knowledge and receptive-productive dimension. The latter set the knowledge of individual words aside by focusing on two vocabulary dimensions. According to this
author, only two dimensions constitute the basis for vocabulary knowledge: size or number of words, and organization or connections between words. The former should be given more importance in the first stages of the learning process, while the latter should be laid more importance as the learner’s level increases.

Since it is quite improbable that the average L2 learner can equal a native speaker in terms of vocabulary size, it is necessary to select which words are a priority in L2 vocabulary teaching. Nation and Gu claim that “in terms of usefulness, all words are not created equal” (Nation, 2007: 20). The concept of ‘usefulness’ in this context refers to communicative usefulness. Put it another way, the primary and ultimate goal of language learning is communication. Communication is possible in the greatest part thanks to vocabulary. In this regard, vocabulary learning should focus on words that provide students with the greatest benefit. The benefit offered by words will be inevitably linked to the communicative context of the user. Given the global character of the English language, the number of learner profiles has multiplied in the last decades. In addition to General and Academic English, nowadays we have to mention Business English, Medical English or Computer English. That is why it is necessary to offer the learners vocabulary input that provides them with the highest reward.

Many vocabulary researchers (Sinclair and Renouf, 1988; White, 1988) propose frequency as the main or one of the main criteria regarding vocabulary selection in coursebooks. Thus, Nation (2001) states that primary attention should be paid to the most frequent words in General English. In a narrow sense, low frequency words would not really deserve classroom time, and the rest of instruction after dealing with those frequent words should be mostly devoted to different learning and guessing strategies. Those statements are supported by studies which show that knowing the 1000 most frequent words leads to 78% understanding of an average text. That percentage increases up to 84% after having learned the second 1000 most frequent words. Another different issue is whether that percentage is enough for guessing the rest of words in that text. In this sense, researchers seem to disagree on the proportion needed to guess those words that are unknown by the learner. For Francis and Kucera (1982) around 80% is enough for understanding a written text. It implies that one out of five words in the text is unknown by the learner. Others (Li Na & Nation, 1985; Nation, 1990; Laufer, 1992) argue that only knowing at least 95% words in a text will a reader be able to discern the meaning of the unfamiliar items.

The Spanish Ministry of Education states that students should be able to communicate minimally in a foreign language after their compulsory academic years. Following Sinclair and Renouf, textbooks should primarily focus on the most frequent words in English, so that basic communication is guaranteed. Yet, frequency and functionality do not always have to coincide. That is, words as demanded by the communicative needs of students must engage in semantic fields that are contained in the communicative context of those learners. The present study pursues to compare the vocabulary input of one of the widely used textbooks in Spanish Primary Education against the 2000 most frequent words in General English.

From the Traditional Method to the Communicative one, vocabulary has always been present, albeit its relevance has waxed and waned along the history of language teaching. Those antecedents notwithstanding, vocabulary has been gaining positions since the mid-seventies, and nowadays it is one of the cornerstones in language teaching. O’Dell (1997) provides us with illustrative examples in an attempt to represent the evolution of vocabulary in the last thirty years. Five of the most well-known EFL textbooks show how vocabulary is remarkable by its absence at the beginning of the 70s in the Kernel Intermediate Coursebook (1971), until it is given pride of place in Willis’ and Lewis’ Lexical Syllabus (1993) by the mid-nineties. Vocabulary develops from being hardly mentioned to the centre of language learning.
Cognisant of this phenomenon, materials designers have pursued to remark the role of lexicon in their textbooks, lending considerable weight to vocabulary. Nonetheless, there is lack of agreement about the way vocabulary should be introduced. Nowadays, in great part due to the influence of the Communicative Method, coursebooks present vocabulary contextualized in dialogues, narrations, songs or other sorts of texts. The followers of this practice (Brown, Collins & Duguid, 1989; Watson & Olson, 1987; Nagy, 1997) hold that it is from context that L1 speakers learn vocabulary. Moreover, they remark that one of the major limitations of vocabulary lists is the reduced number of words that can be effectively covered. That is why they propose approaches such as the book flood (Elley, 1991; Scott et al., 1997) consisting of the implicit introduction of vocabulary by means of considerable number of readings.

However, not all researchers (McKeon, 1993; Schmitt & Schmitt, 1995; Ito Tae, 1995) share this opinion. Context is not the panacea of vocabulary learning, and has two substantial limitations. First, some learners react to unknown words in a text by merely dropping them. Second, in order to infer meaning from context it is necessary to know most words that are part of the text. As mentioned above, at least 95% of the words in a text must be known by the learner if he or she wants to infer the meaning of the unknown words. Those facts limit the use of context in low levels, and leads designers to combine short stories or dialogues with definition-based instruction and other tools such as or translation. In the light of those comments, an eclectic introduction of vocabulary seems to be the most appropriate approach. Yet, whether it is implicitly or explicitly presented, in context or out of it, vocabulary appears as one of the main interests in courses.

Coursebooks exert a great influence on the learning process and – when used adequately- they constitute a useful tool for language teachers. Materials presumably mirror the syllabus and the curricular aims – which at the same time regulate L2 teaching. The choice of an appropriate coursebook is one of the challenges that both teachers and institutions have to face every year. Given the current spate of textbooks available, they sometimes find themselves with dozens of proposals from “publisher representatives calling round and dazzling us with their new books” (Grant, 1987: 119). That is, despite the fact that materials design is a noble professional track, the publishing market is sometimes flooded with “books beautifully presented with jazzy covers and attractive artwork which distracts the eye and dulls the brain” (opt. cit.). Then, it is a task of teachers and institutions to go beyond the cover by choosing the materials that better cater for the learner’s needs - and it is the task of materials designers to cater for those needs.

The burgeoning interest in vocabulary shown by most current coursebooks can be appreciated in the wide number of activities devoted to vocabulary. Furthermore, current textbooks are organized in thematic units, each of them about a specific topic. Those topics are not random. They try to illustrate the possible communicative contexts where the learners may find themselves represented. Yet, hardly any comment is found on vocabulary choice. Having a look at the didactic guides behind the design of those courses we can be acquainted with the pedagogical bases of the project, as well as the main endeavours, among which we may highlight comprehension and production – both at the oral and written levels-, interest in other cultures or the use of new technologies in L2 communication.

However, we miss the reasons behind the choice of the specific words that conforms the vocabulary input of coursebooks. This seems to be just the tip of the iceberg. The Spanish documents for the regularization of FL teaching - Diseño Curricular Base, Ley de Educación 2003 and 2007- provide no explicit info on the specific words that the learner should know at the end of Primary and Secondary Education. Regarding the Foreign Language section of Curriculum de Enseñanzas mínimas para Edu Primaria (2006), the only reference to vocabulary states that one of the aims in this cycle is “el reconocimiento y uso del léxico previamente utilizado” (opt. 2006: 43093). Yet, in no
place of the document the specific vocabulary that children should have learnt is explicitly mentioned. Within the section of evaluation criteria, the document alludes to some semantic fields to be treated: “Se evalúa la capacidad de utilizar expresiones y frases para hablar en términos sencillos sobre su familia y otras personas, el tiempo atmosférico, la ropa, libros, juegos y sobre contenidos de las diferentes áreas” (opt. 2006: 43095). We have to admit that this is an attempt to highlight the vocabulary required by functions. Yet, once again, the paper fails to further specify the terms related to family, weather, clothes, books or games that should make up the coursebook units. As we can see, allusions to vocabulary are vague and inconclusive. The curriculum touches upon the topics but it does not get into the real question: which vocabulary, which words should feed each of those topics? It is here when corpora might have a role. The advent of modern corpora meant a major advance in Applied Linguistics. It allowed linguists to make far more detailed and objective studies of how language is actually used. O’Dell (1997) mentions two important applications of corpora to materials design: compilation of better word frequency lists – which allow more confident decisions on which frequent vocabulary to include – and, data on concordances, that is, how words are used in context. We have to bear in mind, though, that general corpora collect authentic pieces of language - which implies any kinds of words- whereas a textbook is constrained by the characteristics of the target audience, that is, level, cognitive abilities, functions and sociocultural environment. Yet, though shyly, the role of corpus in the classroom is afforded more and more weight. Hence, there have been serious attempts to develop teaching materials informed by the corpus (Carter, Hughes & McCarthy, 2000; McCarthy & O’Dell, 1999; 2001). They claim the important role of frequency in vocabulary choice, and base their designs on lists provided by different researchers. Nonetheless, we doubt that those works carried out by vocabulary researchers have really taken root in the classroom, since corpora are authentic samples and textbooks are not. In fact, McCarthy and Carter (2001: 338) hold that not always do students aim for authentic communication and “both teacher and learner may operate according to the theatre of the classroom”. Despite the differences between corpora and textbooks, the former may offer interesting and reliable quantitative and qualitative vocabulary data which can be used for many purposes including textbooks design. Therefore, nowadays there is no excuse to design but good-quality didactic materials. A different question is whether publishers really take research findings into consideration or just pay lip-service to them.

II. METHOD

II.1 Research questions

The present study pursues answering the following three questions:
Which is the frequency of the coursebook vocabulary in overall terms?
Which is the frequency of the coursebook vocabulary in each unit?
Which is the degree of representativeness in the textbook of the 2000 most frequent words in General English?

II.2 Instruments and procedure

II.2.1 Coursebook

The coursebook chosen for the analysis is entitled Bugs 3. It aims at young learners in their third year of Primary Education. There are two main reasons why this textbook was selected as the object of our analysis. First, it is widely used at numerous schools in
the Region of Murcia. It is very well known among teachers, students and parents due to its positive effect on the children’s learning process. The present course is built upon eight main units, plus an introduction. The set is completed with two special sections: Christmas and Easter. The units are delivered as follows: the first term comprehends the introduction and units 1, 2 and 3. The second part of the academic year develops units 4 from 6, whereas units 7 and 8 are displayed in the last term.

Bugs 3 is claimed to fall within the Communicative Method. Vocabulary is presented mostly, though not exclusively, in context by means of dialogues, stories or songs. Those are complemented with different activities where vocabulary learning is one of the aims. Units are built upon the topics of animals, clothes, the body, food, sports, routines and holidays.

II.2.2 Range

The coursebook analysis was carried out by means of the RANGE programme. Among other things, this computer program allows you to classify the vocabulary of any text into frequency categories. RANGE includes three different lists: list one and two (hereafter L1 and L2) contain the first and second most frequent words in General English, whereas list three (L3) presents words that are not found among the 2000 most frequent words but are usually found in upper secondary and university texts. The words that do not fit in any of the three lists appear in the “not found” category.

Furthermore, RANGE distinguishes among three different units of estimation: tokens, types and families. A token is defined as “every word form in a spoken or written text” (Nation, 2001: 7), so that each time a word occurs is counted. The number of tokens in a textbook may give us an idea about the amount of input the learners are exposed to in raw terms – which falls out of the purpose of the present study. A family consists of a headword, its inflected forms and its closely related derived forms. Adopting a family as counting unit presents the serious problem of deciding which forms should be included and which should not. What is more, we cast doubt on the idea that young learners know the word bedroom once they have learned bed. They might guess that bedroom is related to bed and even what it means, but we must distinguish between guessing ability and learning.

Regarding our study, we have adopted the type as counting unit. A type is any different form that appears in a text. It implies that words such as walk and walked are counted as two different items. At first sight, this may not be adequate, since a person that has learned the word walk possibly also knows walked, or at least can identify the latter with the former. However, we have to think that we deal with low levels and young learners. We can not assume that children knowing walk also may know walked or walking. Finally, though RANGE registers any kind of word, our analysis focuses exclusively on content words, that is, nouns, verbs, adjectives and adverbs.

III. RESULTS

III.1 Vocabulary Frequency: the whole book

After the coursebook was analysed by Range, we obtained four frequency lists: L1 containing items from the 1000 most frequent words in General English, L2 with that vocabulary pertaining to the second 1000 most frequent words, a third list (L3) with words of lower frequency but commonly present in upper secondary education and also university texts; finally a set of words which do not belong to any of those categories, labelled the “not found words” (hereafter NF). It is important to remember that our main aim is to observe the weight of the coursebook vocabulary within L1 and L2 against the
rest of words outside those lists. Thus, for convenience’s sake, we have grouped words into three categories. We have maintained L1 and L2 and have put L3 and NF words together into only one set, what we have called the “unusual words” (hereafter UNW).

Figure 1 shows how the outcomes point out a clear predominance of L1 with almost 56% (55.91%) words belonging to this frequency rank. This is followed by UNW words with 28.13% and finally by the least represented, L2 with just 15.95% presence in the course. Focusing on target vocabulary (those words that are especially highlighted to be learned), up to 40% (39.8%) partake the UNW category, closely followed by L1 with 33.4%. Again, L2 is the least remarkable, containing the lowest number of target words: 59 items; put it another way, only 11.3% of target vocabulary belongs to L2.

![Figure 1. Percentages of representation of frequency lists](image)

Regarding parts of speech, we find 227 nouns, 78 verbs, 54 adjectives and 17 adverbs. More than half the nouns, 72% verbs, almost 60% and the whole of adverbs are part of L1. However, the proportion amongst all of them differs from the one that is usually given in authentic texts. Proportionally, authentic texts contain 3 nouns every 2 verbs and 1 adjective. In this case, the proportion is dramatically inferior: less than 1 verb and less than 1 adjective every 3 nouns. Thus, presence of nouns is remarkably over the prototypical distribution.

**III.2 Vocabulary Frequency: individual units**

In addition to the overall analysis of the textbook, we should also afford insight into each didactic unit. We have omitted the especial sections of Introduction, Christmas and Easter, and just focused on the main course units from 1 to 8.
Figure 2. Development of the three frequency lists along the coursebook.

As with the results of the book as a whole, the predominance of L1 is manifested in every unit. Its presence oscillates between almost 80% in unit 5 down to 53.21% in unit 3. By contrast, not always the weight of UNW word list is higher than L2. If we organize the units according to the three terms of the academic year, we can see that the first term offers a higher presence of UNW words (from 17.15% in unit 1 to 31.2% in unit 3) than L2 words (ranging from 10.10% in unit 2 to 15.6% in unit 3). Yet, tables turn in the second term. Despite the light advantage of UNW list in unit 4, regarding units 5 and 6 L2 words are clearly superior in number with 11.76% and 12.62% versus 8.4% and 8.7% of UNW. Finally, the third term shows again the hegemony of UNW (20.87% and 15.03%) over L2 (14.78% and 10.46%).

Figure 2 shows the behaviour of each word list along the eight units. The most regular path is represented by L2, with soft upward and downward moves. Evolution of L1 and UNW is opposite, with marked dips of the former in units 3 and 7, which coincide with the UNW peaks.

We can also analyse didactic units from a qualitative perspective. A qualitative analysis answers the third question of our study regarding representativeness in the textbook of the 2000 most frequent words in General English. Figure 3 presents the classification of target vocabulary in each unit according to frequency. Thus, unit 1 combines school materials with pets; unit two focuses on clothes and family, whereas unit 3 mostly deals with body parts and wild animals; units 4, 5 and 6 are devoted to food, sports and everyday routines, respectively; summer holidays and weather are seen in unit 7, and finally, unit eight presents a traditional tale, where royal characters, cultural public places and several physical, mental and social features are boarded.

<table>
<thead>
<tr>
<th></th>
<th>L1</th>
<th>L2</th>
<th>UNW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Black, chair, dog, eye, fish, green, head, leg, rule, white</td>
<td>Bird, brown, desk, mouth, orange, wing, yellow</td>
<td>Ant, antennae, grasshopper, hamster, notebook, schoolbag, spider, tortoise</td>
</tr>
<tr>
<td>2</td>
<td>Blue, brother, dad, dress, game, go, make, mum, party, red, shoes, sister, trainers, wear</td>
<td>Antie, coat, hat, sock</td>
<td>Cousin, jeans, costume, lantern, pumpkin, t-shirt, uncle, witch</td>
</tr>
<tr>
<td>3</td>
<td>L1</td>
<td>L2</td>
<td>UNW</td>
</tr>
<tr>
<td>L1</td>
<td>L2</td>
<td>UNW</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>Big, body, eat, fish, great, live, long, sleep, small</td>
<td>Bird, ear, mouth, sea, tail, tooth</td>
<td>Tongue, crocodile, giraffe, gorilla, grasslands, jungle, leopard, lion, snake, tiger, whale, whisker, zebra</td>
<td></td>
</tr>
<tr>
<td>Egg, fish, five, four, glass, milk, one, six, three, two</td>
<td>Bread, cheese, chicken, chip, fruit, juice, knife, plate, vegetables</td>
<td>Sausages, spoon, cereal, ice cream, macaroni, rice, salad</td>
<td></td>
</tr>
<tr>
<td>Fifty, football, forty, jump, pass, play, run, score, thirty, throw, twenty, walk</td>
<td>Bike, goal, ride, swim</td>
<td>Tennis, basketball, rollerblade, skate, skateboard</td>
<td></td>
</tr>
<tr>
<td>Bed, dinner, get dressed, go, lunch, school</td>
<td>Breakfast, lunch, shower, stretch</td>
<td>Shampoo, soap, splash, munch, zip</td>
<td></td>
</tr>
<tr>
<td>Bear, cold, country, have, land, park, summer, sunny</td>
<td>Camera, cloudy, cream, hat, mountain, raining, snow</td>
<td>Beach, continent, lake, pole, seal, towel, fox, ocean, penguin, polar, reindeer, suitcase, sunglasses, swimming pool, swimsuit, trunks, whale</td>
<td></td>
</tr>
<tr>
<td>Beautiful, cat, English, father, give, help, king, man, marry, match, old, poor, speak, young</td>
<td>Boot, clever, coat, hat, queen, rich, river, theatre</td>
<td>Crown, museum, prince, princess, rabbit, donkey, handsome, zoo</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3 Target vocabulary from units 1 to 8 in *Bugs 3*

All those semantic fields are closely related to the communicative environment of young learners. As the main aim of a textbook is to fulfil the students’ needs as L2 users, the frequency criteria is placed in the background, and is taken over by functionality – which at the end of the day, represents the ultimate goal of linguistic communication. This functional character does not necessarily have to contrast with General English in all cases. As we can observe from the chart, a significant number of target words belong to the first 1000 most frequent items, and some of them – though in a more modest proportion- are among the second 1000 most frequent vocabulary.

Nonetheless, the proportion of the three lists varies from unit to unit, according to the semantic fields that predominate in each of them. Due to space limitations, we will comment on just four out of the eight units in the coursebook. We consider those four units most representative of children’s communicative context, and, at the same time, they are the ones that better represent that contrast between the functional and the frequent.

The selected units are unit 3 (wild animals), unit 4 (food), unit 5 (sports) and unit 7 (holidays). All target words in unit 4, except numbers one to six, are directly related to food. Thirteen of those twenty food words are not among the 1000 most frequent ones. Although words such as *sausages, spoon* or *ice cream* are classified as unusual vocabulary, they are highly functional in the learners’ communicative context. Regarding unit 5, all sports but *football* cannot be found within the most frequent words. Thus, *basketball, tennis, skateboard or rollerblade* - foci of that unit- belong to a low frequency rank in General English. A similar case is found in unit 3, with the whole set of wild animals - *crocodile, giraffe, gorilla, leopard, lion, snake, tiger, whale, zebra* – labelled as unusual.

Finally, unit devoted to summer holidays also offers a great number of its target words from the UNW category. In fact, central terms to summer holidays such as *beach, sunglasses, trunks* or *suitcase* fall out of the most frequent General English, but are
paramount to the semantic field of unit 7. Those examples show how the communicative context of children and therefore their linguistic functions do not usually coincide with those in General English corpora.

IV. DISCUSSION

Our results echo Donzelli’s (2007) and Vassiliu’s (2001) works. The former analysed the input of a textbook from the fourth year of Primary Education. The analysis of the latter encompasses a wider sample in quantitative and qualitative terms with several coursebooks from different levels. Similar to our case, both found a general predominance of L1 followed by almost one third of ‘unusual’ vocabulary. Again, the second 1000 most frequent words were the least represented in textbooks. From a qualitative perspective, we observe that as it happens with our analysis, target vocabulary is related to sports, or wild animals. Among the target words that were classified as unusual in Donzelli’s study we can highlight basketball, tennis, zebra or lion – terms that are also found in our analysis.

Those outcomes lead us to think that our study is not an isolated case and that there seems to be a significant presence of unusual words in didactic materials. Thus, we agree with Bell and Gower (1998: 117) that “the language presented in many coursebooks bears little relationship to real use and more to coursebook convention”. When Bell and Gower talk about coursebook convention they refer to nothing short of the attempt of coursebooks to cater for the communicative needs and functions of children- which as we have checked, do not always come to terms with frequency.

It is true that a large number of L1 words is found in our analysis. The surprising fact is the relatively high percentage of UNW words and the consequently low amount of L2 words, regarding frequency criteria. At first sight, an increase of L2 words in detriment of unusual words might be desirable. First, Spanish students of EFL in their third year of Primary Education are supposed to have been learning that language for six years, from their first course of kindergarten. After six years of instruction, a considerable proportion of L1 vocabulary is supposed to have been learned, or at least touched upon. Thus, a reduction of L1 words in pursuit of a greater presence of L2 words would not be excessively preposterous. Second, a reduction of the UNW list would be also more than justified if the main aim was to learn the most frequent words. In this regard, there is no immediate need to include so many unusual words, and on top of that, to label them as target.

If coursebooks adopted frequency as their main criterion, they would have to include words such as account, article, declare, population or faith at the stages of learning. All those terms are among the first 1000 most frequent words, and are found in semantic fields such as politics, law, journalism or religion. We cast doubt that they are the kind of vocabulary normally handled by children. Thus, those words and many others from L1 would be extremely difficult to comprehend and completely unpractical for young learners, as that kind of input is far from children’s context, both cognitively and socially.

In turn, we find ourselves at pains to think that the high number of unusual words in textbooks is the result of chance. A textbook must adapt to the communicative context of the target audience. Hence, the course must be built upon familiar topics for students in consonance with their functional needs. The selection of topics bears relationship to what Reda (2003) calls the “visitor’s wing” English - also known as the ‘basic’ area of the English language. It postulates that “the process of designing the EFL/EIL lexical syllabus does not involve exposing learners to natural language use but to a microcosm of this use” (opt. 264). The so called “wing” relies on functional English, and it adapts to the goals of the average learner. In fact, the principle of the visitor’s wing – which seems to be
followed by most coursebooks - is one of the main reasons for the high presence of unusual words.

In the light of our results, it is clear that “Qualitative criteria [...]are highly significant, particularly where applications of corpus insights to the language classroom are concerned” (McCarthy & Carter, 2001: 337). Therefore, when incorporating corpora to didactic materials we have to be aware of both the pedagogic and communicative context of learners. This includes the learnability and teachability of items regarding the learners’ environment and cognitive status. That is, difficulty in learning and teaching a word depends on external and internal factors to the word. Regarding the latter, we can find the semantic features of a word (Oxford & Scarcella, 1994; Laufer, 1997). The semantic features of a word refer to the different aspects that shape the meaning of that word. Among those features abstractness is especially relevant in the case of children. L2 young learners have not developed many abstract concepts in their L1, which makes it very difficult for them to learn words such as faith, freedom or hope. An adult does not face any problem understanding those terms, whereas it is quite difficult for children to come to terms with such a kind of words. Therefore, we can conclude that the learner’s cognitive and social features condition his/her communicative needs and consequently the vocabulary required for fulfilling those needs.

V. CONCLUSION

The present study compares the input from a Primary Education textbook against the 2000 most frequent words in General English. The outcomes show that there is a significant amount of vocabulary that falls out of the highest frequency levels. We discuss why this may be so. Primary Education materials target young learners groups. Thus, young learners’ environment is shaped by specific cognitive and social features. Frequency is considered one of the keys in vocabulary selection, but its role in materials design is submitted to the primary importance of functions.

McCarthy and Carter (2001) nicely reflect this idea by stating that “Corpora can afford considerable benefits for classroom teaching, but pedagogic processes should be informed by corpus, not driven or controlled by it” (opt. 338). Therefore, regarding vocabulary choice for didactic materials, we should make decisions beyond frequency, taking children’s special features and communicative needs as the main criteria for design.

References


Donzelli, G. (2007). Foreign language learners: words they hear and words they learn: A case study. ELIA, 7(1), 103-125

780


Laufer, 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 and M. McCarthy (Eds.) *Vocabulary: Description, acquisition and pedagogy* (pp. 141-155). Cambridge: CUP


Vassiliu, P. (2001). *Lexical input and uptake in the Greek low level EFL classroom*. PhD. University of Wales, Swansea

