Jaqueline Mora Guarín
University of La Rioja
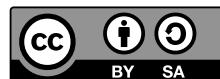Spain
jamora@unirioja.es

THE IMPLEMENTATION OF PROTOTYPICAL CONTENT WORD LISTS TO ANALYSE THE VOCABULARY INPUT CONTAINED IN ENGLISH LANGUAGE TEACHING TEXTBOOKS

1   INTRODUCTION[1]

Textbooks are key elements that involve the presentation of material, development activities such as learners' practice and communicative interaction, stimulation and ideas for classroom activities, and a reference source for learners on grammar, vocabulary, and pronunciation (Cunningsworth, 1995). In EFL contexts, textbooks are considered a resource or guide to language learning and teaching (Asher, 1994; Harmer, 1991). Also present is the idea of textbooks as crucial components of language education that provide the teacher with ideas for lessons and the complementary input learners need when learning a new language (Richards, 2001). Moreover, language learning textbooks provide the necessary input into classroom lessons through different activities, readings, and explanations (Tomlinson, 2012). Therefore, the central notions with regard to textbooks involve them being an essential resource for language teaching and learning and being containers of vocabulary input. In this respect, the present study aims to analyze the vocabulary contained in two ELT textbooks differing in age and course levels based on semantic prototypes through the elaboration of prototypical content word lists. To this end, we aimed to ascertain whether there is evidence of word prototypicality in each textbook and to determine whether there are differences and similarities regarding the type of prototypical content words found in each textbook.

2   PREVIOUS RESEARCH

Although few studies have examined prototypical words in the analysis of ELT textbooks, the current work aims to provide a general overview of the research conducted so far on ELT textbooks in instructional contexts.

Studies on vocabulary input in EFL textbooks initially follow two research tendencies: non-corpus and corpus-based studies (Jiménez & Mancebo, 2008). Researchers from the former set of studies focused on the rate at which new vocabulary is introduced in EFL coursebooks (Schofield, 1991), the design of vocabulary activities in school manuals by language learners (Martinez, 1999), the vocabulary related to social success (Ojeda Alba, 2004 in Jiménez Catalán & Ojeda Alba, 2009), and the adoption of the guidelines proposed by the European Framework of Reference (Mancebo, 2005). Some of these studies highlight the deficiencies or weaknesses of textbook writers, textbook designers, and vocabulary selection criteria. As for the latter, corpus-based studies have concentrated on the design of EFL syllabuses (Sinclair & Renouf, 1988; Willis & Willis, 1988; Cheng, 2011; Papajani, 2015); the learning and teaching of English instructional programmes (Stevens, 1991; Baleghizadeh & Saharkhiz, 2012); and the analysis of previous language teaching materials (Gavioli, 1997; Gray, 2016; Littlejohn, 2022), which has been a neglected research area since it has not been funded by educational projects (Leech, 1998).

Research on the vocabulary input contained in ELT textbooks for Primary and Secondary Education has encompassed studies about word frequency, word repetition, cultural vocabulary input, and very little on semantic prototypes.

Researchers on vocabulary input have centred their attention on word frequency as one of the main criteria as far as EFL materials are concerned (Sinclair & Renouf, 1988; White, 1988; Jiménez Catalán & Fitzpatrick, 2014). Furthermore, in general English, most of the attention should be paid to the most frequently used words since they have been regarded as a crucial dimension in the assessment of learners' lexical competence (Nation, 2001). To the best of our knowledge, studies on ELT textbooks regarding word frequency have been scarce in the fields of Primary and Secondary Education (Guerra & Jiménez Catalán, 2022; Criado & Sánchez, 2009; Rahmat & Coxhead, 2021; Larsson, 2017), although varied in aim and scope. Guerra and Jiménez (2022) focused on the analysis of four primary textbooks to identify the most frequent content words they contained and to determine whether the input provided to 6th graders was identical in terms of content words, finding that although there was a shared thematic structure in the distribution of topics among the four textbooks, there was an unequal number of units and additional material. In Secondary Education, Criado and Sánchez (2009) analysed the role frequency plays in language and vocabulary learning and showed that the frequency of vocabulary is clearly unbalanced since it seems to be too low, and there were few opportunities for the students to rehearse and repeat the lexical words they are supposed to learn in a textbook. Rahmat and Coxhead (2021) explored the corpora of three Indonesian EFL textbooks and determined that the Secondary Education materials contain a large number of high-frequency words, at more than 80% of the total, and up to 7% of the textbook comprises Indonesian words, which play roles in learning objectives and pre-vocabulary teaching activities. Larsson (2017) focused on investigating the vocabulary profiles in six

English language learning textbooks. The results indicated that with regard to word frequency, many of the tokens in all the textbooks belonged to the first 1,000 most frequent words (K1) followed by the second 1000 most frequent words at the K2 level within the general frequency word lists.

Analyses of word repetition suggest that the recurrence of words plays an essential role in assisting L2 learners to gain an in-depth understanding of both receptive and productive knowledge regarding form, meaning, and usage of the target vocabulary (Nation, 2013). Research in this respect has outlined that repetition has played an important role in learning vocabulary through ELT textbooks. For example, Jiménez Catalán and Mancebo (2008) indicated that the comparison of Primary and Secondary Education materials revealed that publishing houses do not agree regarding the frequency of word occurrence and the number of words learners are exposed to regarding coursebooks. The results showed many occurrences of the repetition of words in the school manuals. Norlund (2016) explored the frequency distribution of content words in the two most used series of textbooks in Sweden, with the results showing that publishers do not consider frequency in their choice of what words to include in textbooks, even though frequency lists and research on the importance of knowing the most frequent words are available.

Turning to cultural vocabulary input in ELT textbooks, some research has been carried out in the Spanish educational system of our region, the Autonomous Community of La Rioja. For instance, Canga and Cifone (2015, 2021) examined how intercultural competence was integrated into the vocabulary content of Primary and Secondary Education materials, finding that the total amount of words related to culture encountered in both textbooks is not balanced and that the target culture prevails over the source or international types of culture.

Regarding analyses of the vocabulary input of ELT textbooks by means of semantic prototypes, little work has been conducted to date, and we only found two recent studies on this linguistic and semantic field (Mora and Cifone Ponte, 2021a; Cifone Ponte and Mora, 2021b). These two studies mainly focused on examining the semantic prototypes in EFL textbooks in 6th-grade Spanish Primary Education to identify how multilingual cultures were represented in the vocabulary input of the materials. The results revealed that the target culture is pervasive, while immigrant learners' cultures in schools are scarcely included in the materials. Moreover, there were similarities and differences regarding the cultural aspects through semantic prototypes in the two EFL textbooks selected, and there were few elements related to cultural identity because most of the vocabulary contained in the materials is centred on using words from the target language (English).

The literature presented above mainly refers to studies that have dealt solely with word frequency, word repetition, and cultural aspects, given that research on prototypical content word lists in ELT materials has yet to be included because it has yet to be published.

The present study investigates the vocabulary input contained in two textbooks differing in course level (6th year of Primary Education and 12th year of Secondary Education). It has the following objectives: (i) to ascertain whether there is evidence of prototypicality in the vocabulary input contained in each textbook within the 50 most frequent content words, and if so, to determine whether there are differences and similarities between both textbooks in terms of semantic prototypes; and (ii) to find out whether the elaboration of content word lists might cast light on the types of prototypical words found in the selected ELT textbooks.

3   METHOD

3.1   Sample

The sample used in the present study consisted of two textbooks, one for Primary Education (*New Tiger 6*) for children between the ages of 11 and 13 and one for Secondary Education (*Key to Bachillerato 2, first edition)* for adolescents between the ages of 16 and 17. Table 1 gives more details of both textbooks.

Table 1: Description of the Selected ELT Textbooks

Course Level	Textbook	Authors	Publishing House	Year of Publication	Number of Units	CEFR Level
6th grade	*New Tiger 6*	Carol Read Mark Ormerod	MacMillan	2018	7	A1-A2
12th grade	*Key to Bachillerato 2*	Ben Wetz	Oxford	2014	6	B2

These two ELT textbooks were mainly selected due to the identification of prototypical words in the input received by the two groups of EFL informants differing in age and grade level (6th and 12th graders), whose output of lexical availability on prototypes was also analysed in a previous study (Mora, 2024). Additionally, these materials have been and are currently used as samples in several investigations within the framework of research projects conducted in the Autonomous Community of La Rioja in the north of Spain. Besides, not only has the Spanish Ministry of Education approved these textbooks, but also our regional Education Council in La Rioja, Spain. Moreover, the analyses of these coursebooks can be helpful for other teachers and researchers, given that they were published in the UK and are used in different countries to teach English. Furthermore, the selection of these two course levels stems from the fact that 6th grade is the final year of Spanish Primary Education, and 12th grade is the final year of compulsory Secondary Education. In this respect, comparisons in the vocabulary input of the materials will allow us to determine whether there are significant differences by grade level regarding the prototypical words contained in each grade's ELT textbook.

3.2   Data processing, edition, and lemmatization

We first converted the textbooks into .txt files, one per textbook, and submitted them to the computational software programme WordSmith Tools (version 7.0) for the corpus compilation. The lists of words were extracted in alphabetical order and by means of frequency (repetition or number of occurrences of the word through the text). The units of analysis considered were the *word* and the *lexical unit of meaning.* The former is defined as "any sequence of letters (and a limited number of other characteristics such as hyphen and apostrophe) bounded on either side by a space or punctuation mark" (Carter, 1998:4). We also followed Sinclair (2004), who defined a word as a *lexical item* or *lexical unit* to refer to word patterns (phrasal verbs, idioms) and not to individual words, as is the case of corpus studies. The latter focused on the combination of words such as phrasal verbs and idioms (e.g., go out, meet up, fall in love). Regarding lemmatization, we followed Read (2000) and Nation (2001), who defined it as grouping the base and inflected forms of a word under the same heading. In this respect, plural nouns and verb inflections were removed to count the word once (e.g., *dog-dogs: dog*, *examines-examined: examine*).

The two ELT textbooks' editing and lemmatization criteria followed the procedural steps stipulated by the Applied Linguistics Group from the University of La Rioja (GLAUR). Some of the criteria included (i) correction of spelling mistakes, if applicable as a result of the scanning of the textbooks; (ii) deletion of symbols and words written in different languages; (iii) compound words and lexical phrases consisting of two words, or more, were linked and considered one lexical unit and counted as one word (e.g., *CD-player, tidyup, windowshopping, hookon*); (vi) abbreviated words such as *TV* and *mum*, and acronyms such as *PE*, *CEO*, and *IT*, were included in the counting since they appear as lexical entries in dictionaries. Two dictionaries were used as a guide in this, the Cambridge Online Dictionary and WordReference Dictionary; (v) proper nouns referring to people, cities, and places were discarded, but we kept those referring to nationalities (e.g., *English*, *Spanish*, *French*); (vi) brand names (e.g., *Coca Cola*) were discarded; (vii) fictional characters and names were deleted (as seen primarily in 6th grade textbooks); (viii) titles of films, books and videogames (e.g., *Harry Potter*, *PS4*) were deleted; (ix) verbs were changed to their bare infinitives; (x) comparative and superlative adjectives were joined to their corresponding positive forms (e.g., *higher: high*).

As for the identification of prototypes in the two ELT textbooks, content words (nouns, verbs, adjectives, and adverbs) were used in our analysis. Thus, based on the lists of function words stated by Nation (2001), O'Shea (2013), and Ozturk (2019), prepositions, conjunctions, determiners, definite and indefinite articles, numbers, and auxiliary verbs, including contractions were ignored for this study, since, as opposed to function words, content words have meaning when used alone (Richards & Schmidt, 2002). Moreover, since a review of the literature suggests that prototypical word lists have not been created regarding the corpora in ELT textbooks, we created some preliminary criteria,

such as: (i) phrasal verbs were kept and linked and counted as one word (e.g., goout, meetup, workout); (ii) the word form *there be* was deleted; (iii) *have got/has got/ haven't got* word forms were removed; (iv) auxiliary verbs such as *do*, *get* and *have* were removed, unless they were used in the textbooks as verbs (e.g., do the shopping, get a new house, have a haircut); (v) the proper names of people, countries, and places were deleted (e.g., Phil, Scott, Jamaica, Wales, San Francisco).

Regarding the prototypical content words contained in the two ELT textbooks, we elaborated prototypical word lists following the semantic approach to the classification of content words proposed by Dixon (2005) and Viberg (1984). Dixon's classification consisted of *concrete* (people, volcano), *abstract* (time, year), and *speech act* (answer, question) nouns, *motion* (run, move), *stative* (believe, like), and *speaking* (talk, call) verbs, and adjectives of *value* (good, bad) and *similarity* (different, alike). Viberg's (1984) classification focused on the identification and typology of the *verbs of perception* (listen, look).

To the best of our knowledge, this study is the first attempt to elaborate prototypical word lists to analyse semantic prototypes in ELT textbooks. It focused exclusively on the 50 most frequent content words, with *frequency* referring to the number of times the words occur in a text. The higher the frequency of a word, the more prototypical it is within a text.

4   RESULTS

Before presenting the results regarding the two research questions set out in this study, Table 2 presents the distribution of the corpus and content words per course-level textbook for the sake of clarity.

Table 2: Distribution of Corpus and Content Words per Textbook

	Corpus (content + function words)	Tokens (content words) *Lemmatized	Types (content words) *Lemmatized
New Tiger 6	120,449	8,477	1,452
Key to Bachillerato 2	394,506	27,315	3,486

The results show differences in the number of words contained in terms of the general corpus (content and function words within the textbooks), tokens (total number of content words within the textbooks), and types (different content words within the textbooks). As seen in Table 2, the adolescent EFL learners' textbook (*Key to Bachillerato 2*) had the highest number of words, including content and function words, whereas the EFL children's material (*New Tiger 6*) had the lowest. As for tokens and types, the results showed that, as expected, the Secondary Education textbook *Key to Bachillerato 2* had more than the Primary Education Textbook *New Tiger 6*. Thus, the standardized token-type-ratios (TTRs)

indicated that the lexical variation between the two textbooks differed, at 37.74 for *New Tiger 6* and 45.50 for *Key to Bachillerato 2*. Moreover, within the number of content word types, the frequency of occurrence was up to a minimum of nine times, which was a low percentage in each textbook, at 16.25% for *New Tiger 6* and 18.15% for *Key to Bachillerato 2.* In contrast, the frequency of occurrence equal to one time had the highest percentage for both *New Tiger 6* (33.81%) and *Key to Bachillerato 2* (37.14%).

With reference to our first research question about the evidence of prototypicality in the two ELT textbooks, the results revealed that according to the frequency '*F*' (repetition or number of occurrences of the word through the text), there were some prototypical content words for each book, as highlighted in bold in Tables 3 and 4. The analysis of how the prototypes were represented in each textbook was based on the same semantic approaches that were applied to the content words.

Table 3: Prototypical Words Contained in New Tiger 6

New Tiger 6							
W	F	WC	Type of Category	W	F	WC	Type of Category
1. Listen	133	V	Perception	26. Make	41	V	Motion
2. Read	126	V	Motion	27. Place	40	N	Abstract
3. Like	103	V	Stative	28. Project	39	N	Abstract
4. People	103	N	Concrete	29. Answer	38	N/V	Speech Act/ Speaking
5. Go	102	V	Motion	30. Club	37	N	Concrete
6. Think	87	V	Stative	31. Favourite	37	ADJ	Value
7. Find	77	V	Motion	32. Take	36	V	Motion
8. Say	77	V	Speaking	33. Talk	36	V	Speaking
9. Day	63	N	Abstract	34. Friend	35	N	Concrete
10. See	59	V	Perception	35. Learn	35	V	Motion
11. Help	56	V	Motion	36. Year	34	N	Abstract
12. Story	56	N	Abstract	37. Name	33	N	Abstract
13. Play	55	V	Motion	38. Put	33	V	Motion
14. Good	53	ADJ	Value	39. Sport	32	N	Concrete
15. Country	51	N	Concrete	40. Volcano	32	N	Concrete
16. Time	51	N	Abstract	41. Call	31	V	Speaking
17. World	51	N	Concrete	42. Past	31	N	Abstract
18. Write	50	V	Motion	43. Ask	30	V	Speaking
19. Live	49	V	Stative	44. Different	30	ADJ	Similarity
20. Use	48	V	Motion	45. Home	30	N	Concrete
21. Tree	46	N	Concrete	46. Look	30	V	Perception
22. Know	43	V	Stative	47. Question	29	N	Speech Act
23. Want	43	V	Stative	48. Visit	29	V	Motion
24. Park	42	N	Concrete	49. Activity	28	N	Abstract
25. School	42	N	Concrete	50. Culture	28	N	Abstract

Note. W: Word, F: Frequency, WC: Word class, N: Noun, V: Verb, ADJ: Adjective

Table 3 shows that in the children's textbook *New Tiger 6*, the classification of content words consisted of different category subtypes. Namely, nouns (*concrete, abstract,* and *speech act),* verbs (*motion, stative, speaking,* and *perception*), and adjectives (*value* and *similarity*). Concerning concrete nouns, the highest number of words (24%) were referred to as real physical objects, things, or places (e.g., *people, country, school),* followed by abstract nouns (20%) related to ideas, qualities, or states (e.g., *time, project, culture*), and speech act nouns (4%) that serve a function in communication (e.g., *question, answer*). Regarding verbs, most of the words were highly related to motion verbs (24%) which express the meaning of action (e.g., *read, go, play*), proceeded by stative verbs (10%) that refer to a state of being (e.g., *think, know, want*) followed by speaking verbs (8%) related to those words used to report what someone has said (e.g., *say, ask, talk*), and verbs of perception (6%) that convey the experience of one of the physical senses (e.g., *listen, see, look*). Adjectives were the least represented, with a mere 6% in total, split into adjectives of value (4%), which refer to the quality added to the noun (e.g., *good, favourite*), and similarity (2%), which are related to the comparison of two things, states, or events (e.g., *different*). In this respect, semantic prototypes in the 6th graders' ELT textbook were mainly related to concrete nouns (24%) and motion verbs (24%).

The first five positions on the list show that the most frequent words in the textbook, with over 100 occurrences, were *listen* (133/1.57%), *read* (126/1.49%), *like* (1.03/1.22%), *people* (103/1.22%) and *go* (102/1.20%), and these can be considered the most prototypical in the 6th graders' textbook. The next five words on the top ten list also had rather high frequencies, and these were *think* (87/1.03%), *find* (77/0.91%), *say* (77/0.91%), *day* (63/0.74%) and *see* (59/0.70%). The findings thus showed that the top ten prototypical words in *New Tiger 6* referred to motion verbs (*read, go,* and *find*), perception verbs (*listen* and *see*), stative verbs (*like* and *think*), and verbs of saying (*say*). As for nouns, within the top ten most frequent only two stood out, with *people* as a concrete noun and *day* as an abstract noun referring to a unit of time. As seen in Table 3, the most prototypical word in the children's ELT textbook was *listen*, whereas the least was *culture*, at least within the 50 most frequent words in the text. Figure 1 shows the subcategory word types in *New Tiger 6*.
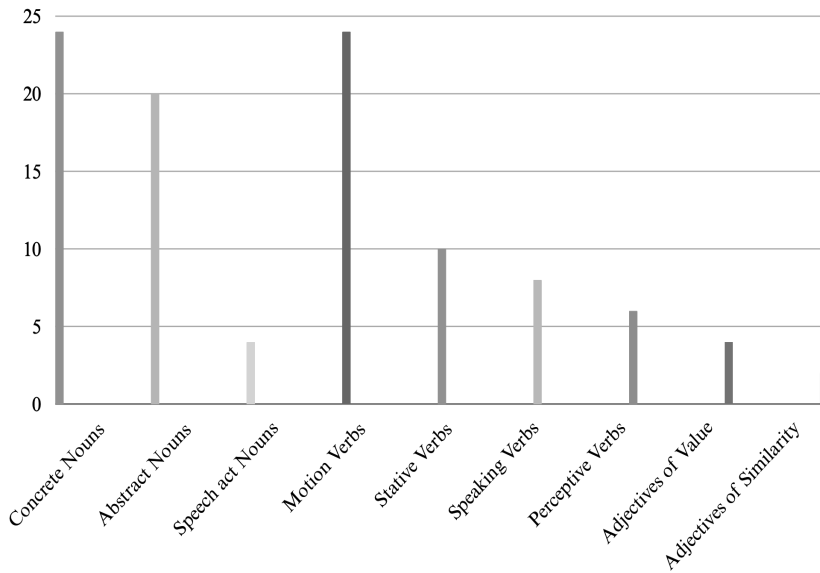
Figure 1: Prototypical Representation of Subcategory Word Types in New Tiger 6

Table 4: Prototypical Words Contained in Key to Bachillerato 2

Key to Bachillerato 2							
W	F	WC	Type of Category	W	F	WC	Type of Category
1. Notebook	355	N	Concrete	26. Look	111	V	Perception
2. Use	342	V	Motion	27. Ask	110	N	Speaking
3. Word	336	N	Abstract	28. Job	109	N	Concrete
4. People	313	N	Concrete	29. Choose	106	V	Motion
5. Sentence	263	N	Abstract	30. Listen	104	V	Perception
6. Question	260	N	Speech Act	31. Year	104	N	Abstract
7. Say	246	V	Speaking	32. See	96	V	Perception
8. Complete	238	V	Motion	33. Key	95	N	Concrete
9. Think	215	V	Stative	34. Exam	94	N	Abstract
10. Write	214	V	Motion	35. School	93	N	Concrete
11. Text	206	N	Abstract	36. Speak	93	V	Speaking
12. Answer	204	N/V	Speech Act/ Speaking	37. Thing	92	N	Concrete
13. Verb	204	N	Abstract	38. Want	91	V	Stative
14. Go	180	V	Motion	39. Take	89	V	Motion
15. Make	169	V	Motion	40. Find	88	V	Motion
16. Good	164	ADJ	Value	41. Paragraph	88	N	Abstract
17. Read	153	V	Motion	42. Walk	88	V	Motion
18. Like	140	V	Stative	43. Example	87	N	Abstract
19. Form	134	N	Abstract	44. Activity	86	N	Abstract

20. Correct	128	ADJ/V	Value	45. Unit	86	N	Abstract
21. Work	121	V	Motion	46. Know	85	V	Stative
22. Give	115	V	Motion	47. Language	80	N	Abstract
23. Past	113	N	Abstract	48. Need	80	V	Stative
24. Phrase	113	N	Abstract	49. Start	80	V	Motion
25. Time	112	N	Abstract	50. Change	79	V	Stative

Note. W: Word, F: Frequency, WC: Word class, N: Noun, V: Verb, ADJ: Adjective

The findings in Table 4 indicate that – similar to the children's textbook (A1/A2 CEFR levels) – the classification of the prototypical words in the adolescents' textbook *Key to Bachillerato 2* (B2 CEFR level) included nouns (concrete, abstract, and speech act), verbs (motion, stative, speaking, and perception), and adjectives (value). Unlike *New Tiger 6*, the most common nouns in *Key to Bachillerato 2* were abstract nouns (30%) (e.g., *word, sentence, text*), followed by concrete nouns (12%) (e.g., *notebook, people, job*), and nouns referring to speech acts (4%), such as *question* and *answer*. As in the children's material, motion verbs were also highly represented (26%) in the adolescents' textbook, such as *use, complete*, and *write*. Motion verbs were followed by stative verbs (e.g., *think, live, want*) and verbs of perception (e.g., *look, listen, see*), at 12% each. Speaking verbs accounted for 6% of the total, represented by words such as *say, ask,* and *speak*. As for adjectives, unlike *New Tiger 6*, *Key to Bachillerato 2* only included the *value* subtype (e.g., *good* and *correct*), with a frequency of 6%.

As can also be seen in Table 4, the highest positions in the word list for the 12th grade ELT textbook ranked from positions one to 31, since all these words had more than 100 occurrences. The most frequent prototypical word in this specific textbook (*Key to Bachillerato 2*) was the concrete noun *notebook* (355/1.30%), whilst the least frequent was the stative verb *change* (79/0.29%). Within the top ten most frequent words in the adolescents' textbook, the results showed that the most prototypical words – apart from *notebook* – were *use* (342/1.30%), *word* (336/1.23%), and *people* (313/1.15%). Other prototypical words appeared from the fifth position onwards in the word list, although at a lower frequency level. For instance, *sentence* (263/0.96%), *question* (260/0.95%), *say* (246/0.90%), *complete* (238/0.87%), *think* (215/0.79%), and *write* (214/0.78%). The findings showed that, as in the children's textbook, the top ten most prototypical words in the adolescents' material mainly corresponded to motion verbs (*use, complete,* and *write*). However, just one stative verb (*think*) and one speaking verb (*say*) had high frequencies, and no perceptive verbs were included in the top ten most frequent words in the prototypical word list. With regard to nouns, there were two concrete nouns (*notebook* and *people*), two abstract nouns (*word* and *sentence*), and a speech act noun (*question*). Therefore, these top ten words might be considered the most prototypical ones in *Key to Bachillerato 2*. Figure 2 shows the classification of the subtype categories used for the elaboration of the prototypical word lists in the adolescents' textbook.
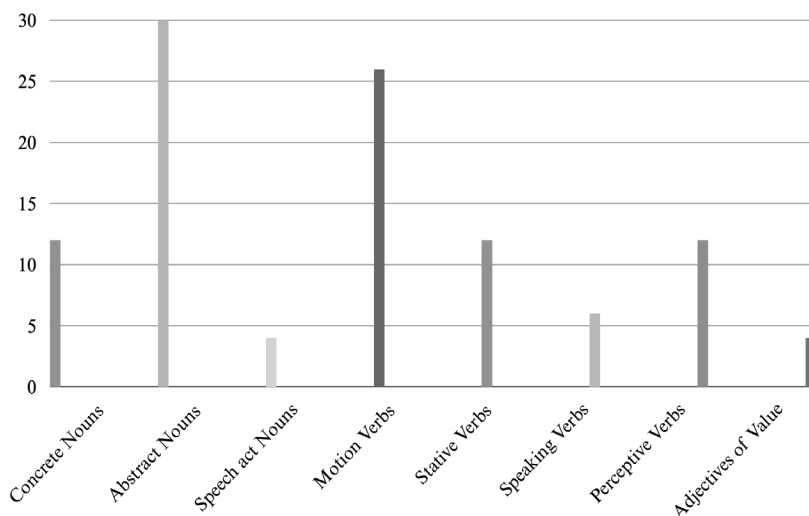
Figure 2: Prototypical Representation of Subcategory Word Types in Key to Bachillerato 2

Regarding our second research question, the results showed that despite the differences in terms of the frequency of the words contained within the 50 most frequent words per textbook, there were still some similarities among the prototypical words with regard to the type of content word subcategories. For instance, most of the prototypical content words were related to motion verbs in both the children's and adolescents' ELT textbooks. Moreover, the subcategories proposed for the elaboration of the prototypical content word lists in the ELT materials coincided with the classification of nouns, verbs, and adjectives. Another notable finding is that the word *people* appears in fourth places in both textbooks (Tables 3 and 4), although with a different number of occurrences. Within the most 50 frequent words, 25 prototypical words were found be the same in both ELT textbooks, despite the different course levels. Most of the shared prototypical words were verbs (64%), followed by nouns (32%) and a single adjective (4%). Table 5 shows the classification of these prototypical content words in the children's and adolescents' textbooks for ease of readability.

Table 5: The Prototypical Content Words that Appeared in Both ELT Textbooks

	Nouns		Verbs		Adjectives	
New Tiger 6 & Key to Bachillerato 2	Concrete	*People, school*	Motion	*Read, go, find, write, use, make, take*	Value	*Good*
	Abstract	*Time, past, year, activity*	Stative	*Like, think, know, want*		
	Speech Act	*Question, answer*	Speaking	*Say, ask*		
			Perception	*Listen, see, look*		

Note. Source: The 50 most frequent words per textbook

There were notable differences between both textbooks with regard to the most prototypical content words, as shown in Table 5. For instance, although the most frequent words were related to concrete nouns and motion verbs in the children's textbook *New Tiger 6*, the most representative prototype was a verb of perception (*listen*). In the case of the adolescents' textbook *Key to Bachillerato 2*, the highest percentage of words were related to abstract nouns and motion verbs, while the most prototypical content word was a concrete noun (*notebook*).

Furthermore, the results also revealed that among the most frequent concrete nouns in *New Tiger 6* there was evidence of subcategories such as Environment (*tree* and *volcano*), Geographic Places (*world* and *country*), and Town (*park, school,* and *club*). In contrast, in *Key to Bachillerato 2* the highest percentage of nouns were abstract ones, and the leading subcategory was Units of Language (*word, sentence, text, verb, paragraph,* and *phrase*).

5   DISCUSSION

Content words were the focus of this study with the aim of identifying prototypical words in the two EFL learners' ELT textbooks, given that they refer to notions, concepts, facts, and feelings (Nation, 2001; Gardner, 2013). Our results showed that when comparing the 6th and the 12th graders' ELT textbooks (*New Tiger 6* and *Key to Bachillerato 2*), content words, as expected, were more common in the older learners' B2 level material. As for the total number of words (tokens) and different words (types), our results indicated that there were fewer of these words in the young learners' textbook were than in the older learners' coursebook. These results were not surprising, since differences in factors such as age and course level have an impact on the learning language process. The data for the children's textbook *New Tiger 6* revealed that the total number of content and function

words was in line with the results presented by Guerra and Jiménez Catalán (2022). However, our data regarding content words exclusively (8,477 tokens and 1,452 types) did not agree with Guerra and Jimenez's results (20,350 tokens and 1,535 types). This difference derives from the fact that for our specific analysis of prototypes we eliminated all function words, including auxiliary verbs and their corresponding contractions (e.g., *be: aren't, isn't- do: don't, doesn't, didn't*), as suggested in the list of function words proposed by Nation (2001, p.430-431). Regarding *Key to Bachillerato 2*, such comparisons are irrelevant since, to the best of our knowledge, no research has been conducted concerning this 12th graders' textbook. However, the standard token-type ratio (STTR) in the list of content word types showed that, as expected, *Key to Bachillerato 2* had a higher lexical variation than *New Tiger 6*.

As stated in the section on methodology, the elaboration of prototypical word lists was carried out to identify patterns of prototypical words contained in the learners' ELT textbooks. These lists followed the semantic approaches put forward by Dixon's (2005) and Viberg's (1984) classification of words. As previously mentioned, the classification of the 50 most frequent words was conducted using these semantic approaches to determine whether there was evidence of semantic prototypes in the children's and adolescents' materials.

The data collected revealed that in the children's textbook (*New Tiger 6*), the most prototypical words corresponded to concrete nouns and motion verbs, as Dixon (2005) proposed in his classification of words. This may indicate that nouns and verbs are the most predominant words in the learners' materials (Guerra & Jiménez Catalán, 2022; Norlund, 2015; Hsieh et al., 2011). However, the results must be interpreted cautiously since this study only considered one of the many textbooks used in 6th primary education. The findings also showed that abstract nouns and speech act nouns, alongside stative, speaking, and perceptive verbs, appeared with low frequencies. Nevertheless, the most prototypical word, which was the most frequent in the textbook, was *listen*, a verb of perception. An interpretation of this result might be that this verb conveys the experience of a physical sense, and perception is more accessible for young children than mental states (Davis & Landau, 2020). Within the top ten most frequent words, most of the prototypical words were mainly verbs, with only two nouns. Most of the verbs belonged to the subcategory of motion verbs, followed by perceptive, stative, and speaking verbs. This tendency may suggest that verbs are the most frequent words and thus the most prototypical in the children's textbook, given that verbs constitute one of the largest classes of content words (Börjars & Burridge, 2001), which, in turn, are essential for vocabulary acquisition.

The results for the adolescents' textbook (*Key to Bachillerato 2*) showed that, unlike *New Tiger 6*, prototypical words were highly associated with abstract nouns followed by motion verbs. The findings for nouns within the top 50 most frequent content words in the 12th graders' material showed that abstract nouns outnumbered concrete nouns, unlike in the

6th graders' textbook. This may stem from the fact that textbook designers would have considered that the cognitive changes which take place during adolescence include a shift from concrete to more abstract and complex thinking (Lang et al., 2022). In addition, according to the Common European Framework of Reference (CEFR), learners who achieve a B2 upper intermediate level of performance are conditioned to understand the main ideas of complex texts on concrete and abstract topics, including some technical discussions. Regarding verbs, the results for this textbook were in line with those for the children's textbook, since the leading ones were motion verbs followed by stative, perceptive, and speaking verbs. As for adjectives, they were exclusively related to the value subcategory. The most prototypical word found in *Key to Bachillerato 2* was *notebook*, a concrete noun that might be considered a basic word within the category of School, as well as being the most used object by adolescent learners in this educational context. Concerning the top ten words within the 50 most frequent content words, the results indicated that, as in the children's textbook, in the adolescents' book the most representative subcategory was motion verbs. However, some differences were noticeable, since within the top ten words there were a mixture of nouns and verbs. However, unlike *New Tiger 6*, the results for *Key to Bachillerato 2* did not reveal any verbs of perception. This may suggest that although perceptive verbs related to physical senses were indeed present in the adolescent's textbook, they were not the most frequent, and thus not the most prototypical words as they were in the children's book, being within the top ten most frequent content words.

The findings with regard to the shared prototypical content words in both textbooks revealed that despite difference in age and course level, the children's and the adolescents' textbooks shared 25 basic prototypical words (Table 5). All of these basic words (e.g., *people, school, activity, answer, question, read, write, think, say, ask*) are included within the 2,000 most frequent words at introductory levels, and 5,000 most frequent 5000 words at advanced levels. This means that the vocabulary input provided by ELT textbooks includes these words so that learners can acquire a general comprehension of the language (Alsaif & Milton, 2012). However, caution should be taken since the sample of textbooks in this study is limited to only two titles. However, the basic words highlighted in this study have also been found within the top 50 content words in other Primary Education textbooks, with Jiménez Catalán and Mancebo (2008) noting items such as *people, school, read, write,* and *play.* This shows that apart from the two textbooks used in this investigation, these basic words are also present in other ELT textbooks, which is unsurprising since they belong to the 1,000 most frequent words (K1) (Nation, 2014).

6 CONCLUSIONS

This study is the first on the representation of word prototypicality in children's and adolescents' ELT textbooks, and thus its elaboration of prototypical word lists contributes

to the exploration of semantic prototypes in students' materials. The findings with regard to the 50 most frequent content words in both books support the conclusion that although the most frequent word in *New Tiger 6* (*listen*) differed from that in *Key to Bachillerato 2* (*notebook*), there was evidence of shared prototypes within both top ten most frequent words, such as *people, think,* and *say*. The results showed that for the children's textbook the prototypical words were mainly motion verbs and concrete nouns, while for the adolescents' textbook they were mainly abstract nouns followed by motion verbs. The results also showed that the top ten most frequent words in *New Tiger 6* included motion and stative verbs, verbs of perception and saying, followed by a small number of concrete and abstract nouns. In contrast, in *Key to Bachillerato 2* the most prototypical words were motion verbs, abstract and concrete nouns, followed by a few stative and speaking verbs, along with one single speech act noun. These findings should be valuable for researchers, EFL teachers, and textbook designers, given that the insights with regard to the prototypical words in ELT textbooks open new and first lines of investigation on vocabulary learning through the elaboration of prototypical word lists based on content word semantic approaches.

However, the findings should not be generalized since we only considered two textbooks, one each for learners from the 6th and 12th grades. We did not include a higher number of textbooks, given that in the selection criteria of the materials, we only considered those most and still used by EFL learners in the Autonomous Community of La Rioja, Spain. In further research, it could be useful to elaborate a more general and complete word list, including the classification of words by means of the semantic approaches proposed by Dixon (2005) and Viberg (1984) to ease the analysis of prototypes in ELT textbooks. As previously mentioned, ours is the first attempt to elaborate prototypical word lists based on the frequency of occurrences of the words contained in textbooks. However, for the specific purpose of our study, we exclusively focused on only two textbooks instead of a general word list that could be more useful for any textbook in FL research.

BIBLIOGRAPHY

ALSAIF, Abdullah/James MILTON (2012) Vocabulary input from school textbooks as a potential contributor to the small vocabulary uptake gained by English as a foreign language learner in Saudi Arabia. *Language Learning Journal* 40, 21–33. DOI: 10.1080/09571736.2012.658221.

ASHER, Ron (1994) *The Encyclopedia of Language and Linguistics*. *Volume 2*. Pergamon.

BALEGHIZADEH, Sasan/Arash SAHARKHIZ (2012) Linguistic Imperialism and EFL Textbooks: The Case of American English File. *e-Uluslararası Eğitim Araştırmaları Dergisi* 3/2, 36–49.

BÖRJAS, Kersti/Kate BURRIDGE (2001) *Introducing English Grammar*. Arnold Publishers.

CANGA ALONSO, Andrés/María Daniela CIFONE PONTE (2015) An Analysis of Cultural Vocabulary in ELT textbooks. *Odisea* 16, 83–96. https://doi.org/10.25115/odisea.v0i16.298.

CANGA ALONSO, Andrés/María Daniela CIFONE PONTE (2021) Cultural terms in EFL textbooks for young learners. *Encuentro: revista de investigación e innovación en la clase de idiomas* 29, 90–103.

CARTER, Ronald (1998) *Vocabulary: Applied linguistic perspectives*. Routledge.

CHENG, Vivian (2011) Infusing Creativity into Eastern Classrooms: Evaluations from Student Perspectives. *Thinking Skills and Creativity* 6, 67–87.

CIFONE PONTE, María Daniela/Jaqueline MORA (2021b) Immigrant learners' cultural identity in the vocabulary input of EFL textbooks through prototypical associations. *ELIA: Estudios de Lingüística Inglesa Aplicada* 21, 51–84.

CRIADO, Raquel/Aquilino SÁNCHEZ (2009) Vocabulary in EFL Textbooks. A Contrastive Analysis against Three Corpus-Based Word Ranges. A. Sánchez/P. Cantos (eds.), *A Survey on Corpus-based Research/Panorama de investigaciones basadas en corpus*. Murcia: Editum (Servicio de Publicaciones de la Universidad de Murcia), 862–875.

CUNNINGSWORTH, Alan (1995) *Choosing your coursebook*. Heinemann.

DAVIS, Emory/Barbara LANDAU (2020) Seeing and Believing: The Relationship between Perception and Mental Verbs in Acquisition. *Language Learning and Development* 17/1, 26–47. DOI: 10.1080/15475441.2020.1862660.

DIXON, Robert (2005) *A Semantic Approach to English Grammar* (2nd ed.). Oxford University Press.

GARDNER, Dee (2013) *Exploring vocabulary: Language in action.* Routledge.

GUERRA, María Isabel/Rosa María JIMÉNEZ CATALÁN (2022) Words and topics in ELT textbooks for young learners. *Language Teaching for Young Learners* 4/2, 264–288. https://doi.org/10.1075/ltyl.22002.alv.

HARMER, Jeremy (1991) *The Practice of English Language Teaching: New Edition*. Longman.

HSIEH, Ming YI/Fei-yu WANG/Sy-ying LEE (2011) A Corpus-Based Analysis Comparing Vocabulary Input from Storybooks and Textbooks. *The International Journal of Foreign Language Teaching* 1, 25–33.

JIMÉNEZ CATALÁN, Rosa María/Rocío MANCEBO (2008) Vocabulary input in EFL textbooks. *Revista española de lingüística aplicada* 21, 147–166.

JIMÉNEZ CATALÁN, Rosa María/Julieta OJEDA ALBA (2009) Girls' and boys' lexical availability in EFL. *ITL International Journal of Applied Linguistics*, *158*, 57-76.

LANG, Diana/Nick CONE/Tera JONES/Martha LALLY/Suzane VALENTINE (2022) Physical Development in Adolescence. D. Lang (eds.), *Individual and Family Development, Health, and Well-Being*. Iowa State University. Digital Press, 337–349.

LARSSON, Salome (2017) Vocabulary Profiles of English Language Learning Textbooks: A lexical analysis of textbooks used in EFL classrooms. Dissertation. Linnaeus University

LEECH, Geoffrey (1998) Learner Corpora: what they are and what can be done with them. S. Granger (ed.), *Learner of English on Computer*. Longman, xiv–xx.

LITTLEJOHN, Andrew (2022) The analysis and evaluation of language teaching materials. J. Norton/H.Buchanan (eds.), *The Routledge Handbook of Materials: Development for Language Teaching*. London. Routledge, 263–276.

MANCEBO, Rocío (2005) «¿Hasta qué punto los libros de texto de primera se adecuan a las directrices de enseñanza de vocabulario del marco común?» Paper presented at the commemoration of the European Day of Languages. University of Valladolid.

MARTINEZ, Begoña (1999) The influence of language learning theories on the design of vocabulary activities in L2 English course books. Small thesis, unpublished manuscript. University of La Rioja.

MORA, Jaqueline (2024) Semantic prototypes in EFL children's and adolescents' lexical output and vocabulary input in EFL textbooks. Doctoral dissertation. University of La Rioja.

MORA, Jaqueline/Maria Daniela CIFONE PONTE (2021a) Prototypical associations in EFL Textbooks: Representations of multilingual cultures in reading comprehension activities. *Cultural Diversity in Cross-Cultural Settings: A Global Approach*, 197–218.

NATION, Paul (2001) *Learning vocabulary in another language*. Cambridge University Press.

NATION, Paul (2014) How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language* 26, 1–16.

NATION, Paul (2013) *Learning vocabulary in another language (Second)*. Cambridge University Press.

NORDLUND, Marie (2015) Vocabulary in textbooks for young learners. *Educare* 1/1, 73–92.

NORDLUND, Marie (2016) EFL textbooks for young learners: A comparative analysis of vocabulary. *Education Inquiry* 7/1. https://doi.org/10.3402/edui.v7.27764.

O'SHEA, James (2018) Function Word Lists. Accessed on 11 July 2024 at: https://semanticsimilarity.wordpress.com/function-word-lists/.

OZTURK, Meral (2019) Function Words in Word Frequency Lists for EFL Learners. Accessed on 11 July 2024 at: https://meralozturkblog.wordpress.com/wp-content/uploads/2018/10/pdf-1.pdf

PAPAJANI, Josilda (2015) The Evaluation of the EFL Textbooks Used in the High Schools of Elbasan, Albania. *European Journal of Language and Literature* 1/7, 7–15. DOI: 10.26417/ejls.v1i1.p7-15.

RAHMAT, Yurieke/Averil COXHEAD (2021) Investigating vocabulary coverage and load in an Indonesian EFL textbook series. *Indonesian Journal of Applied Linguistics* 10/3, 804–814.

READ, John (2000) *Assessing vocabulary*. Cambridge University Press.

RICHARDS, Jack (2001) *The Role of Textbooks in a Language Program*. Cambridge University Press.

RICHARDS, Jack/Richard SCHMIDT (2002) *Longman: Dictionary of Language teaching and applied linguistics*. Pearson Education Limited.

SCHOFIELD, Phil (1991) Vocabulary rate in course books- living with an unstable lexical economy. *Proceedings of the 5th Symposium on the Description and or Comparison of English and Greek*. Aristotle University of Thessaloniki, 12–32.

SINCLAIR, John (2004) *Trust the Text: Language, Corpus, and Discourse*. Routledge.

SINCLAIR, John/Antoinette RENOUF (1988) A lexical syllabus for language learning. R. Carter/M. McCarthy (eds.), *Vocabulary and language teaching.* Longman, 140–160.

STEVENS, Vance (1991) Classroom concordancing: vocabulary materials derived from relevant authentic text. *English for Specific Purposes* 10, 35–46.

TOMLINSON, Brian (2012) Materials development for language learning and teaching. *Language Teaching* 45/2, 143–179. DOI:10.1017/S0261444811000528.

VIBERG, Åke (1984) The verbs of perception: a typological study. B. Butterworth/B. Comrie/Ö. Dahl (eds.), *Explanations for Language Universals*. Berlin/Boston: De Gruyter Mouton, 123–162. https://doi.org/10.1515/9783110868555.123.

WHITE, Lydia (1988) Island effects in second language acquisition. S. Flynn/W. O'Neil (eds.), *Linguistic theory in second language acquisition. Studies in Theoretical Psycholinguistics*. Springer, 144–172.

WILLIS, Jane/Dave WILLIS (1988) *Collins COBUILD Coursebook Student's Book 1*. Collins ELT.

POVZETEK

UPORABA SEZNAMOV PROTOTIPSKIH POLNOPOMENSKIH BESED ZA ANALIZO UVAJANJA BESEDIŠČA V UČBENIKIH ZA POUČEVANJE ANGLEŠKEGA JEZIKA

Učbeniki veljajo za ključno pedagoško komponento pri poučevanju in učenju tujih jezikov. Vendar so analize zastopanosti prototipskih vzorcev v besedišču, ki ga vsebujejo učbeniki za poučevanje angleškega jezika (ELT), redke, saj po nam znanih podatkih še ni bil izdelan noben seznam prototipskih besed. Namen te študije je predvsem razkriti, ali je prototipičnost besed zastopana v dveh učbenikih ELT za učence EFL, ki se razlikujeta po starosti in stopnji tečaja, pri čemer se osredotočamo na 50 najpogostejših besed v vsebini in nato pripravimo predhodne sezname prototipičnih besed na podlagi klasifikacije, uporabljene v semantičnih pristopih, ki sta jih predlagala Dixon (2005) in Viberg (1984). Ugotovitve so razkrile dokaze o prototipičnosti v obeh učbenikih, kljub razlikam v predvideni starosti in ravni tečaja. Rezultati so pokazali tudi, da glede na vrsto vsebinskih besed, ki jih najdemo v posameznem gradivu ELT, med knjigama obstajajo razlike in podobnosti, povezane s prototipskimi izrazi. Poleg tega so bile prototipske vsebinske besede, ki se nanašajo na konkretne in abstraktne samostalnike ter glagole zaznavanja, pogostejše kot druge vrste vsebinskih besed. Poleg tega so ugotovitve pokazale tudi, da so bili glagoli najbolj prototipične vsebinske besede, ki jih najdemo v obeh učbenikih ELT. Ta raziskava ima dve glavni pomembni izobraževalni posledici. Prva vključuje preučevanje pomenskih prototipov v učbenikih ELT za učence EFL, kar bi prineslo nova in bolj specifična spoznanja o vlogi, ki jo lahko imajo prototipne besede pri učenju besedišča v formalnem kontekstu. Drugi se nanaša na to, kako bi bilo mogoče prototipičnost besed v besedišču učbenikov ELT analizirati z uporabo seznamov prototipičnih besed v vsebini.

Ključne besede: semantični prototipi, vnos besedišča, učbeniki ELT, vsebinske besede, prototipni seznami besed

ABSTRACT

THE IMPLEMENTATION OF PROTOTYPICAL CONTENT WORD LISTS TO ANALYSE THE VOCABULARY INPUT CONTAINED IN ENGLISH LANGUAGE TEACHING TEXTBOOKS

Textbooks are considered a crucial pedagogical component in foreign language teaching and learning contexts. However, analyses of the representation of prototypical patterns in the vocabulary input contained in English language teaching (ELT) textbooks have been scarce since, to the best of our knowledge, no prototypical word lists have yet been produced. This study is primarily intended to reveal whether word prototypicality is represented in two EFL learners' ELT textbooks differing

in age and course level, focusing on the 50 most frequent content words and then elaborating preliminary prototypical word lists based on the classification used in the semantic approaches proposed by Dixon (2005) and Viberg (1984). The findings revealed evidence of prototypicality in both textbooks despite the differences in their intended ages and course levels. The results also showed that, according to the type of content words found in each ELT material, there were differences and similarities related to the prototypical terms between the two books. Moreover, prototypical content words referring to concrete and abstract nouns and perception verbs were more common than other types of content words. Furthermore, the findings also indicated that verbs were the most prototypical content words found in both ELT textbooks. This research has two main significant educational implications. The first involves studying semantic prototypes in the EFL learners' ELT textbooks, which would bring novel and more specific insights into the role prototypical words might play in vocabulary learning in formal contexts. The second refers to how word prototypicality in the vocabulary input of ELT textbooks might be analysed with the use of prototypical content word lists.