A child-oriented corpus-based study of semantically related verbs using sketch engine

Belén Labrador (iD)

Associate Prof. Dr., English Studies University of León, Spain, belen.labrador@unileon.es

ABSTRACT

This paper presents a corpus-based study on four pairs of high-frequency verbs: say and tell, do and make, see and look, and go and come, which are semantically related but do not show the same grammatical or collocational behavior, and therefore can easily be confused in their usage by Spanish young learners (YLs) of English. The study is child-oriented in two senses: First, the corpus used, a corpus of Children's Short Stories (CSS), was compiled with texts addressed to children, so they display specific linguistic features. Secondly, the research has been conducted with YLs in mind, so the choice of the pairs of common verbs studied is based on their difficulty for Spanish-speaking YLs and the data-driven learning (DDL) activities proposed are directed to YLs (aged 10-12). The methodology comprises the selection of verbs, their grammatical patterns, and an inventory of the strongest collocates for each category, based on word sketches and the typicality score (Log Dice) provided by Sketch Engine (SE) text analysis software. By identifying the most frequent semantically closely related verbs and their use in grammatical and phraseological structures, and by providing EFL student teachers with this information, awareness can be raised about a) how to find vocabulary especially worth teaching in primary education because they are high-frequency words, simple in meaning but difficult in usage, and b) how to teach these words, including direct and indirect corpus-aided language learning, like DDL, simplified and adapted for YLs, and the inclusion of authentic examples in the teaching materials.

Introduction

Learning vocabulary in a foreign language usually combines two main types of knowledge, each involving a particular direction: breadth of vocabulary refers to a surface-level knowledge of a wide range of words (on a horizontal axis), whereas depth of vocabulary implies a stronger knowledge of different aspects of the meanings and uses of a word (on a vertical axis). Depth of vocabulary is not achieved by a one-step procedure, nor is it a matter of all-or-nothing, i.e. you either know a word or you do not. The process of learning new words is gradual; we need to come across a word in different contexts to understand the different meanings and nuances of that word: "We need to see learning any particular word as being a cumulative process where knowledge is built up over a series of varied meetings with the word" (Nation, 2005, p. 48).

Knowing an L2 word, i.e., having a deep knowledge of that word, requires understanding its meanings and being able to spell, pronounce and translate it, but it also involves using it well, in correct grammatical patterns, with the right collocates, and in appropriate contexts (Thornbury, 2004). In addition, we need this varied repetition to recall the word better: "The more one engages with a word (deeper processing), the more likely the word will be remembered for later use" (Schmitt, 2000, p. 120).

The repeated exposure necessary to acquire this deep knowledge of words can be achieved spontaneously in the long term or it can be controlled and fostered by the teacher, who can look for multiple and varied examples where the target language occurs. However, coming up with these examples can be time-consuming and may not render all the expected results. Corpora can be used to find these examples, i.e., to 'arrange meetings' with the target vocabulary, thus facilitating learning. This study makes use of Sketch Engine (SE), a text analysis application that simplifies the compilation and exploration of corpora (Thomas, 2014), to retrieve an inventory of frequent collocates and patterns of core vocabulary, in this particular case, examples of some basic but commonly confused verbs found in children's stories. This type of information extracted from corpora can be used by teachers to provide young learners (YLs) with real examples from authentic materials. Examples supplied by data-extraction tools such as SE and selected by teachers are easy to remember by learners, as they trigger schemata activation, word associations, or situations in their minds that can help them to grasp the meaning of the concepts better, to learn about their phraseological and grammatical behavior and to integrate them into their mental lexicon.

This paper was prompted by the final aim of finding out ways for teacher trainers, teacher trainees, and novice teachers to assist YLs of English in the use of core vocabulary, thus enhancing their speaking and writing skills. The line of reasoning triggered by this overall purpose followed the path marked by these research questions: 1) which frequent verbs found in children-oriented texts can be easily confused with other frequent verbs that have a similar meaning? 2) how (in which co-text) are they used? i.e., what are their grammatical patterns and common collocates? The results of these two research questions can then be applied to English language lessons for YLs, in particular, for Spanish primary school children in the 3rd cycle (Years 5 and 6, with an age range of 10 to 12). The Spanish educational law includes, among the specific competencies for these years, the understanding and production of literary texts suitable for the learners' development level.

Regarding the first question, the choice of the word class of verbs was due to the fact that, among the core vocabulary, verbs tend to be slightly more complicated than other parts of speech for several reasons: a) their morphological variation (indicating tense, voice, number), b) the fact that they are the heads of verbal phrases and predicates, which makes them engage in more relationships with other constituents of the sentence such as subjects, direct objects, indirect objects, and prepositional complements, c) the fact that they are more abstract from a semantic point of view than, for example, concrete nouns, and d) the fact that they indicate actions, which are more dynamic and complex than objects, people, or places. Also, they are more easily confused in foreign language learning, as they tend to have one-to-many correspondences with their equivalents in the L1, in comparison with other word classes, which may have more straightforward, one-to-one correspondences.

The section on methodology describes how the selection of verbs and their study took place (to answer the first research question) and explains several decisions regarding the corpus and tools used for the research. The section dealing with the analysis and results provides the

answer to the second question by presenting an inventory of patterns, collocates, and examples for each verb based on their word sketches. Finally, the discussion offers some pedagogical implications, advocating the use of learner-centered, guided-discovery methods based on corpora (DDL) as an effective way of teaching vocabulary depth.

Semantically Closely Related Words

Different types of corpora have been used to analyze examples of near-synonyms or words that share some similarities from a semantic or functional perspective. For example, Kennedy (1991) compared the uses of the words *between* and *through*; Church, Gale, Hanks & Hindle (1991) of *strong* and *powerful*; Liu (2010) of *chief, main, major, primary*, and *principal*. AlAmro (2019) studied the usage and connotations of the English synonyms: *babble, blather, chatter, gibber, jabber*, and *prattle*. Alenizi (2021) analyzed the near-synonyms *attack, strike, offensive, aggression*, and *assault*. Wiliński (2017) identified distributional differences between the two semantically or functionally near-equivalent constructions 'on the brink of + noun' and 'on the verge of + noun', finding lexemes that exhibit a strong preference for either one or the other construction, using the distinctive-collexeme analysis (Gries and Stefanowitsch 2004). To name but a few examples from other languages, Janda (2009) analyzed the constructional profiles of several synonyms expressing sadness and happiness in Russian, Vandevoorde, Lefever, Plevoets, & De Sutter, (2017) examined several synonymous words in the semantic field of inchoativity in Dutch, and Almakrob, and Al-Ahdal (2020) investigated the co-occurrence of near-synonym pairs in the Holy Quran.

In addition, a considerable amount of translation and contrastive studies have focused on the comparison of similar words across languages, for instance, Xiao and McEnery (2006) undertook a cross-linguistic analysis of collocation, semantic prosody, and near-synonymy drawing data from English and Chinese. Labrador (2003) studied the synonymous series of English quantifiers *nobody, no one, somebody, someone, everybody, everyone*, and *anybody, anyone* and two corresponding paradigms of quantifiers in Spanish. Rojo López (2011) compared the words *crisis* and *recession* and their equivalents in Spanish. Gesuato (2007) focused on four pairs of English near-synonyms, which only have one corresponding Italian equivalent: a pair of nouns (*island* and *isle*), a pair of adjectives (*feeble* and *weak*), a pair of verbs (*to adore* and *to worship*) and a pair of adverbs (*gratefully* and *thankfully*).

The Use of Corpora in TEFL

There has been a growing interest in corpus-aided language learning over the last few decades (Huang, 2011). Sometimes this type of learning has been direct, encouraging students to act as researchers using corpora and undertaking guided-discovery tasks, as in the data-driven learning (DDL) approach (Johns, 1991) and the corpus-aided discovery learning (CADL) approach (Gavioli, 2000). Some other times the use has been indirect, through teaching materials and classroom activities based on the findings of corpus research (Boulton, 2010). Although the direct methods are usually applied in tertiary education (Jeaco, 2017), there have also been applications of DDL for children and teenagers (Braun, 2007, Crosthwaite, 2019), especially with "corpora designed specifically for pedagogic use with YLs – or 'pedagogic corpora' […] specifically designed to accommodate specific L1/L2 learners, levels, and needs"

(Crosthwaite, 2022, p. 7). The impact of the DDL approach is reflected in the literature: Pérez Paredes (2022) studied the uses and spread of data-driven learning (DDL) and corpora in language learning and teaching across five major CALL-related journals during the 2011–2015 period. Lusta et al. (2023) compiled a database of 89 studies discussing DDL interventions in language classrooms.

Although DDL seems to attract attention, several studies have revealed that teachers are not incorporating their knowledge about how to use corpora into their classroom practice (Pérez-Paredes, 2022). Some authors have pointed out barriers like the corpus interface being too technical (Vyatkina, 2020) and the lack of training (Farr & O'Keeffe, 2019, Bennett & Dhonnchadha, 2023). In addition, the need to guide the learners through the process of discovery, reflection, and awareness has been pointed out: "There is still a significant role to play for teachers of YLs in effectively preparing their students for DDL, and there should be no expectation that YLs will simply 'pick it up' themselves if left to their own devices" (Crosthwaite, 2022, p. 3). Similarly, Lusta et al. (2023) concluded that despite DDL's significant potential as a pedagogical tool, effective strategies such as tailored tasks, auxiliary guidance, supplemental support, and peer/group learning are necessary. Therefore, this study is intended for teachers, especially teacher trainers and trainees, so that they can use the ready-made materials offered here for this particular group of verbs, but also learn how to replicate this method of using SE to identify patterns of core vocabulary and then design customized tasks including these patterns in their teaching.

Methodology

The text analysis application SE (Kilgarriff, et al., 2010) was used to compile an *ad-hoc* corpus of children's short stories written in English (CSS) and to retrieve the data for this study: the word sketches of the verbs selected. The CSS is made up of 454 stories which amount to 517,321 tokens. They were extracted from children-oriented websites during the year 2016. Some of the stories come from well-known folk tales or fairy tales originally written by Andersen, the Brothers Grimm, Joseph Jacobs, Perrault, Beatrix Potter, and Aesop; some others are new stories written by contemporary writers like Nathan Oser, Daniel Henshaw, David Lambert, Clare O'Dea, Deirdre Sullivan, and Jim McCarthy.

The method used for the selection of verbs consisted in observing the list of the 25 most frequent verbs in the corpus and then selecting some pairs that were similar in meaning, hence prone to difficulty, among the candidate verbs. Table 1 below shows the list of verbs, and the rank of the verbs selected, which include: two speech verbs or verbs of communication, say (3rd) and tell (13th), two action verbs or verbs of creation and transformation, do (4th) and make (11th), two verbs of perception related to the sense of sight, see (7th), and look (10th), and two verbs of movement or motion, go (5th) and come (6th).

Table 1. *List of the 25 Most Frequent Verbs in the CSS Corpus and their Frequencies*

	Verb	Frequency
1	Be	17 958
2	Have	6 548
3	**Say**	3 904
4	**Do**	3 061
5	**Go**	2 718
6	**Come**	2 292
7	**See**	2 083
8	Get	1 408
9	Take	1 352
10	**Look**	1 264
11	**Make**	1 257
12	Know	1 087
13	**Tell**	1 064
14	Think	1 060
15	Find	1 007
16	Ask	950
17	Give	883
18	Want	721
19	Hear	705
20	Live	669
21	Eat	611
22	Run	609
23	Sit	545
24	Leave	545
25	Use	515

The frequency rates of these target words were checked against other word frequency databases, from large reference corpora and specific wordlists for YLs. All of these verbs appear in very high positions in all of these lists: the wordlist of the top 60,000 lemmas based on COCA, the Corpus of Contemporary American English (https://www.wordfrequency.info/samples.asp), the wordlist of ESL Kid Stuff (http://www.eslkidstuff.com/WordLists.htm) and Dolch's List of Basic Sight Words (http://www.eslkidstuff.com/WordLists.htm), a list of 220 words, prepared by E.W. Dolch, which makes up from 50 to 75 percent of the reading material encountered by students. These words are generally known as DolchWords, high-frequency words, or 'sight words'. Table 2 below shows the positions of the verbs in these wordlists, which is evidence of the importance of these verbs as core vocabulary in English to be targeted in the ELT primary classroom.

In addition, the selection of the verbs was made on the grounds of the difficulty that cross-linguistic divergence causes. For example, a single verb in Spanish can be translated as two verbs in English (in this case, *decir* as *tell* or *say* and *hacer* as *do* or *make*). Also, there are differences in some semantic features of verbs with very similar meanings, like *ver/ mirar* (*see/ look*) and *ir/ venir* (*go/come*). This causes more difficulty when one of these verbs can be translated using the other verb of the pair, instead of the equivalent, for instance, *come* in the expression 'I'm coming' is better rendered into Spanish as 'Ya voy' (literally, 'I already go'), that is, using the verb *ir* (*go*) instead of *venir* (*come*). Therefore, the criteria for selecting the

verbs under scrutiny were a) that they were salient in terms of frequency, b) that they were salient in terms of similarity in meaning with other verbs in the high-frequency lists, and c) that they were salient in the difficulty they can cause in English-Spanish transfer.

Table 2. *Position of the Verbs Selected in Word Frequency Databases.*

COCA		*ESL Kid Stuff*		Dolch's list	
Say - 26th	Tell - 92nd	Say - 134th	Tell - 150th	Say - 183rd	Tell - 141st
Make - 50th	Do - 15th	Make - 69th	Do - 52nd	Make - 114th	Do - 41st
See - 56th	Look - 81st	See - 71st	Look - 75th	See - 48th	Look - 26th
Go - 31st	Come - 61st	Go -79th	Come - 80th	Go - 35th	Come - 64th

After the selection was made, the usage of the verbs was analyzed with the information retrieved from SE. The word sketches, summaries of the words' grammatical and collocational behavior, were then observed, and the collocations that did not reach a minimum Dice Score of 8 were discarded. The typicality score (Log Dice) indicates how strong the collocation is; a low score means that the words in the collocation also frequently combine with other words or in other grammatical relations, so the collocation is not so strong. Log Dice is the association measure provided and recommended by SE when dealing with word sketches (https://www.sketchengine.eu/my_keywords/mi-score/) and has been used in previous studies (Hu and Yang, 2015; Labrador, 2022; Yuting, 2020), where Log Dice > 8 was used as a threshold (Labrador 2022, p. 4) or lower scores of 5 (Yuting, 2020, p. 18) and 6 (Hu and Yang, 2015, p.35).

The results in the following section present the collocations in descending order from higher to lower Dice Score, indicated in brackets in the tables. Some decisions were taken to solve some problems in the extraction process: SE retrieves patterns according to the order of the elements but a) does not distinguish between the direct object and indirect object and, in some cases, the classification is wrong, for example, there are a lot of cases of inversion, typical in a literary style such as that of stories, where the subject comes after the verb *say* and SE has classified those examples under the heading of 'objects of say'. Also, I have included other categories, like '+ Prepositional Phrase' to merge some instances found separately in the results retrieved from SE. I have discarded some cases because they were also wrongly classified, for example, the word *day* appeared among the strong subject collocates of *make* with examples like these: 'There she sat all through the summer days making rainbows.' In another example, the word *moon* appeared among the strong object collocates of *do* but in those cases, do acted as an auxiliary verb, not as the main verb in examples like 'Why does the moon stay awake at night?' Finally, a decision was made to exclude all pronominal subjects and objects, as they are less important from a lexical and phraseological perspective.

Results

The results have been presented in tables for each pair of verbs, where similar patterns are shown in parallel columns. Table 3 shows that *say* engages in more grammatical patterns than *tell*, as it is commonly found in coordination with other verbs (example 1), in inversion (example 2), and followed by prepositional phrases (example 3).

(1)   And it turned out that not just anybody could *go and say* hello to the Queen in her big house.

(2)   'Well, then, listen to me,' *said the fairy*.

(3)   The elephant always pretended not to hear, and sometimes used to *say in a very loud voice*, 'If there are any worms about, they had better get out of my way'.

Table 3. *Grammatical patterns and strong collocates of say and tell.*

SAY	TELL
VERB + AND + SAY: smile (12.2), laugh (11.1), go (9.1)	
SUBJECT + SAY: mother (9.9), woman (9.3), boy (9.3), Rabbit (9.3), father (9.2), man (9.2), people (8.9), chief (8.7), child (8.7), King (8.5), voice (8.5), brother (8.4), wife (8.4), fox (8.4), tree (8.2), Pig (8.2), butcher (8.2), lion (8.1)	SUBJECT + TELL: man (9.2), merchant (8.9), cat (8.8), father (8.6), witch (8.6), mom (8.5), Shirin (8.5), grandmother (8.5), Adem (8.5), Gerda (8.5), mother (8.4), Geraint (8.4), Princess (8.3), Prince (8.2)
SAY + SUBJECT: man (10.5), woman (10.0), Boots (9.3), lad (9.3), princess (9.3), cat (9.1), king (9.0), fairy (8.9), mother (8.7), Ania (8.7), Rabbit (8.4), soldier (8.4), Gerda (8.4), giant (8.4), Raven (8.3), emperor (8.3), dwarf (8.2), wind (8.2), wife (8.2), brother (8.1), prince (8.1), Halvor (8.0), Gareth (8.0)	
SAY + TO + INDIRECT OBJECT: man (10.4), son (9.9), father (9.9), wife (9.8), mother (9.8), steward (9.7), genius (9.7), brother (9.6), king (9.4), child (9.3)	TELL + INDIRECT OBJECT: king (9.5), mother (9.4), father (9.3), anyone (9.1), parent (9.1), people (9.0), husband (8.9), wife (8.8), servant (8.8), Enid (8.8), daughter (8.6), everyone (8.5), son (8.5), Geraint (8.5), man (8.4), boy (8.3), Arthur (8.2), lord (8.2), gnome (8.2), Ettarde (8.2), sister (8.0), brother (8.0), lady (8.0)
SAY + DIRECT OBJECT: good-bye (13.5), (no) more (10.7), word (8.7), anything (8.0)	TELL + DIRECT OBJECT: story (9.8), truth (9.6), lie(s) (8.7), tale (8.2)
MODIFIER + SAY, SAY + MODIFIER: then (10.1), again (9.4), all (9.3), always (9.0), proudly (8.9), so (8.7)	MODIFIER + TELL, TELL + MODIFIER: therefore (9.4), then (9.3), also (9.0), as (8.9), not (8.1)
SAY + WH-WORD: when (8.9), how (8.3)	TELL + WH-WORD: how (10.1), where (10.0), what (8.4)
SAY + PREPOSITIONAL PHRASE: with smile (12.7), in voice (11.9)	

Both *say* and *tell* are used with similar types of subjects, although there is a tendency for *say* to be used with direct speech and tell in other patterns, always followed by a noun phrase or a pronoun (examples 4): people, like *woman, boy, man, people, child*, (sometimes specifying family relations, e.g. *mother, father, brother, wife, mom, grandmother*) occupations, or positions, for instance, *chief, king, butcher, merchant, witch, princess, prince*), personified animals, such as *rabbit, fox, pig, lion, cat*, things or metonymical references, in particular, *tree, voice* in the case of *say* and proper nouns in the case of *tell*, as in examples 5.

(4)   His *mother said*, 'where did you go?'.
        Jack's *mother told* him to sell their only cow.

(5)   'Little Ania,' the *voice said*, 'why are you so unhappy?'.
        They exchanged stories and *Shirin told* Stephen all about life in Tehran.

Tell is followed by a wider range of indirect objects than *say*, whose indirect objects are always preceded by the preposition *to*, (example 6). In the case of *tell*, indirect objects include compounds like *anyone* and *everyone*, and proper nouns.

(6)   They had heard what their stepmother had *said to their father*.
        The little boy didn't *tell anyone* about this because who would have believed him?

In the case of the direct objects, there are very distinctive collocates for each verb: *say good-bye, (no) more, a word,* and *anything* and *tell a story, the truth, lie(s), a tale* (example 7):

(7)   Then the children *said goodbye* to the mermaid.
        We may not believe someone who often *tells lies*, even when he *tells the truth*.

With modifiers, the situation is similar: *then* is found around both verbs but *say* is either preceded or followed by *again, all, always, proudly, so,* whereas *tell* collocates with *therefore, also, as, not* (examples 8).

(8)   He told her what had happened, and *then said*, 'I love you more than anything else in the world'.
        Jack was happy and *as told* by his mother, he carried the cheese bar on his head.

As for wh-words, both verbs are frequently followed by *how*, and then *say* by *when* and *tell* by *what* and *where* (examples 9).

(9)   Before, madam, I answer your question, allow me to *say how* grateful I am for this meeting.
        They fell into a great storm, and when it was blown over, and it got still again, they couldn't *tell where* they were.

To sum up, *say* is more often found in coordination with other verbs (*smile/ laugh/ go* and *say*), in inversion (*said the fairy*), followed by prepositional phrases (*in a very loud voice*), and to introduce direct speech (*…said, 'where did you go?'*). In the case of the direct objects, some recurrent collocates are *say good-bye, (no) more, a word,* and *anything*. On the other hand, *tell* is followed by a wider range of indirect objects, including compounds like *anyone* and *everyone* and the collocates functioning as direct objects are: *tell a story, the truth, lie(s), a tale*.

Table 4. *Grammatical patterns and strong collocates of do and make*

DO	MAKE
	VERB + AND + MAKE: laugh (10.8), try (10.7)
SUBJECT + DO: lad (9.3), boy (9.0), prince (8.6), woman (8.5), people (8.5), father (8.4), king (8.4), Boots (8.3), Peter (8.3), monkey (8.2), thing (8.2), sister (8.1), child (8.1), brother (8.1) Ogalussa (8.0), folk (8.0)	SUBJECT + MAKE: hand (8.8), brother (8.5), Jack (8.4), king (8.4), son (8.2)
DO + OBJECT: nothing (but) (10.6), work (10.5), thing (10.5), something (10.3), anything (10.3), deed (10.0), duty (9.5), harm (9.0), everything (8.4), trick (8.2)	MAKE + OBJECT: fun (9.7), noise (9.6), haste (9.4), way (9.1), money (8.8), living (8.8), signs (8.8), fire (8.7), bread (8.6), soup (8.6), game (8.6), music (8.6), king (8.5), hole (8.5), attempt (8.3), ceremony (8.3), bed (8.3), effort (8.1), friends (8.0)
DO + MODIFIER: not (12.6), so (10.0), even (8.5), really (8.3), before (8.1), please (8.0)	MAKE + MODIFIER: altogether (9.3), then (8.9), always (8.9), well (8.4), all (8.4), also (8.1), only (8.0)
DO + WH-WORD: whatever (11.2), when (10.9), what (10.5)	
DO + OBJECT + ADJ: someone + good (12.3)	MAKE + OBJECT + OBJECT COMPLEMENT: someone + happy (12.3), suitable (11.6), rich (10.8), proud (10.4), strong (10.2), angry (10.2), sad (9.8), hungry (9.5), miserable (9.5), comfortable (9.5), invisible (9.5)
DO + ADJ: wrong (11.9), more (9.7)	MAKE + ADJ: merry (10.7), sure (10.5), ready (10.2)
	MAKE + PARTICLE: up (10.9), out (9.4), off (8.8)
	MAKE + PREPOSITIONAL PHRASE: up mind (13.8), of wood (10.4)

As can be seen in Table 4, the main difference between *do* and *make* is that *do* is followed by wh-words (example 10) and *make* occurs coordinated with other verbs (example 11), particles (example 12), and prepositional phrases (example 13).

(10)   I thanked the tailor for his counsel and said I would *do whatever* he advised.

(11)   They took some water into their mouths to *try and make* the same gurgling noise as the bird.

(12)   It's just a cock-and-bull story you have *made up* to trick me into releasing you.

(13)   He could not *make up his mind* to which of them to give his kingdom.

Do occurs with more varied, though similar types of subjects, most probably because it is often used as an auxiliary verb in negative sentences and as a substitute for predicates (examples 14); in the case of *make*, there is a case of an agent of a passive clause (example 15).

(14)   But *Boots didn't* mind that a bit.

'Your majesty, open the book.'; the *king did* so.

(15)   He was armed with a helmet and shield which no weapon *made by mortal hands* could pierce.

These two verbs collocate strongly with different direct objects; most of those that go with *do* are more general, including compounds with *-thing*: *nothing (but), work, thing, something, anything, deed, duty, harm, everything, trick,* whereas those following *make* are usually more

specific collocates: *fun, noise, haste, way, money, living, signs, fire, bread, soup, game, music, king, hole, attempt, ceremony, bed, effort, friends* (examples 16).

(16)   He had *done nothing but* look at her all the time at the ball.

The rest of the ducks always *made fun* of him.

The modifiers of each verb are also different: *do not/ so/ even/ really/ before/ please* and *make sure/ altogether/ then/ always/ well/ all/ also/ only* (examples 17).

(17)   I tried to do what I had *done before*.

He was *always making* mistakes.

Both verbs participate in a similar structure: someone + object + adjective, but the adjective in the constructions with *do* is only *good*, it is part of a fixed expression and it is acting like a direct object in the clause, while there is a varied list of adjectives in the constructions with *make* and they are object complements*: happy, suitable, rich, proud, strong, angry, sad, hungry, miserable, comfortable, invisible* (examples 18).

(18)   'If you're such a child, it will *do you good* to go and learn a lesson'.

She was very glad to see the hen, which would *make them rich* once more.

The adjectives or quantifiers that typically follow these two verbs are also different: *do wrong/ more/ merry, make sure/ ready* (examples 19).

(19)   But the robin could *do more* for her than she could dream of.

And from that day on, the little boy always did his best to *make sure* that no one felt sad and alone.

All in all, as far as their grammatical patterns are concerned, these verbs behave similarly except for the fact that *do* is followed by wh-words whereas *make* occurs in coordination with other verbs (*laugh/ try and make*), with particles and with prepositional phrases (*make up your mind*). Among the direct objects, *do* tends to collocate with more general words like *nothing (but), work, thing, something, anything, deed, duty, harm, everything, trick,* whereas *make* tends to collocate with more specific nouns: *fun, noise, haste, way, money, living, signs, fire, bread, soup, game, music, king, hole, attempt, ceremony, bed, effort,* and *friends*. Some other significant patterns including adjectives are: *do someone good, make someone happy/ suitable/ rich/ proud/ strong/ angry/ sad/ hungry/ miserable/ comfortable/ invisible.*

Table 5 shows that *look* engages in two patterns where *see* doesn't - it is followed by adjectives like *sad, different, beautiful, funny, ugly, more* and by particles like *around, over, up, down, out, upon,* (examples 20 and 21).

(10)   Her mother sat on her bed and *looked very sad*.

The poor duckling, who had crept out of his shell last of all, and *looked so ugly*, was bitten and pushed and made fun of, not only by the ducks, but by all the poultry.

(11)   The Monkey *looked down* at the Shark and listened.

Table 5. *Grammatical patterns and strong collocates of see and look*

SEE	LOOK
SEE + AND/ OR + VERB: hear (11.9), love (10.9)	
VERB + AND + SEE: go (11.2), wait (10.5), come (10.4)	VERB + AND + LOOK: look (12.2), stand (10.8), go (9.2)
SUBJECT + SEE: people (9.3), man (9.1), anyone (9.0), Geraint (8.8), hawk (8.7), prince (8.7), child (8.5), pig (8.5), Enyd (8.5), dwarf (8.3)	SUBJECT + LOOK: rose (8.9), eye (8.9), Peter (8.7), chief (8.7), maiden (8.7), everyone (8.7), everything (8.7), Adem (8.7), Gerda (8.6), queen (8.5), child (8.3), boy (8.2), wife (8.2)
SEE + OBJECT: anything (9.0), nothing (8.8), father (8.8), child (8.7), light (8.5), man (8.5), something (8.3), knight (8.2), elephant (8.2), girl (8.1), face (8.1), tree (8.0), bird (8.0)	
MODIFIER + SEE, SEE + MODIFIER: never (11.1), ever (10.8), just (9.3), again (9.2), not (9.0), first (8.9), soon (8.9), neither (8.8), last (8.5), even (8.4), only (8.3), before (8.3), distinctly (8.1)	MODIFIER + LOOK, LOOK + MODIFIER: back (9.5), closely (9.2), neither (9.2), forward (9.1), everywhere (8.9), about (8.9), up (8.7), as (8.7), just (8.7), earnestly (8.5), upwards (8.5), around (8.5), again (8.3), even (8.1), out (8.1)
SEE + WH-WORD: what (11.9), how (11.7), who (9.7), where (9.6), which (9.2)	LOOK + WH-WORD: how (8.3)
	LOOK + ADJ: sad (10.6), different (10.3), beautiful (10.2), funny (9.8), ugly (9.7), more (8.7)
	LOOK + PARTICLE: round (11.5), around (11.3), over (11.3), up (10.9), down (10.4), out (10.0), upon (9.8)
SEE + PREPOSITIONAL PHRASE: with + (adj, poss det) + eyes (12.0), in the distance (11.5)	LOOK + PREPOSITIONAL PHRASE: in seashell (12.9), after rock (12.2), into sky (11.9), like flame of fire (11.4)), for food (10.3), at sky (10.2), for flower (9.9), for boy (9.8), for something (9.8), at tree (9.5), at egg (9.5), at Gerda (9.5)

On the other hand, *see* is followed by objects (*anything, nothing, father, child, light, man, something, knight, elephant, girl, face, tree, bird*), as in examples 22 and *look* isn't. Both verbs collocate with other verbs coordinated with the conjunction *and*, and in the case of *see*, also with *or*. There is also repetition of the verb *look* for emphasis in this type of combination (examples 23).

(12) When he arrived, he *saw the children* playing in the garden.

(13) The Country Mouse one day invited his friend to *come and see* him at his home in the fields.

He *looked and looked* for the shoe but could not find it.

Both verbs are either preceded or followed by modifiers (examples 24).

(14) She could *just see* the moon and stars.

She *looked neither* to the right nor to the left.

The verb *see* tends to be followed by more wh-words than *look* (examples 25).

(15) Early in the morning, a peasant, who was passing by, *saw what* had happened.

'These are my old northern-lights,' said the Reindeer, '*look how* they gleam!'.

Finally, the pattern where the verb is followed by a prepositional phrase (examples 26) seems to be more relevant in the case of *look*.

> (16)   It was something we all *saw with our own eyes*.

> Once upon a time, there was a hungry fox that was *looking for something* to eat.

In summary, unlike *see, look* is followed by particles and by adjectives like *sad, different, beautiful, funny,* and *ugly*. In contrast, only *see* is followed by objects (*anything, nothing, father, child, light, man, something, knight, elephant, girl, face, tree, bird*). Some collocates in prepositional phrases are *see with our own eyes, see in the distance, look in/ into/ like/ at/ for something*.

Table 6. *Grammatical patterns and strong collocates of go and come.*

GO	COME
VERB + AND + GO: see (11.2), get (11.1), stand (10.7), find (10.4), turn (10.0), take (9.7), ask (9.6), sit (9.6), go (9.5), come (9.1)	VERB + AND + COME: see (10.4), take (10.0), go (9.1)
SUBJECT + GO: man (9.1), father (9.0), boy (8.9), thing (8.8), wife (8.7), sun (8.7), child (8.6), time (8.5), woman (8.3), knight (8.3), son (8.2), tortoise (8.2), people (8.0), gnome (8.0), week (8.0), Bruno (8.0)	SUBJECT + COME: time (9.5), man (9.4), night (9.4), day (9.1), morning (9.1), evening (8.7), woman (8.7), people (8.5), voice (8.1), friend (8.1), giant (8.0)
	COME + SUBJECT: Troll (9.2), child (8.1),
GO + OBJECT: home (12.0), way (10.4), bit (10.1), round (9.8), day (9.1), step (9.0)	COME + OBJECT: home (11.7), day (9.5), morning (8.8), way (8.1)
GO + INFINITIVE OBJECT: sleep (11.7), work (10.1), look (9.6), be (9.6), see (9.5), do (9.5), take (9.5), have (9.4), happen (8.9), kill (8.8), visit (8.8), play (8.7), live (8.7), make (8.5), eat (8.5), get (8.4)	COME + INFINITIVE OBJECT: see (11.9), visit (11.7), know (10.4), pass (10.3), help (10.1), meet (9.7), pick (9.7), try (9.3), eat (9.3), be (9.0)
GO + -ING: hunting (9.1)	COME + -ING: running (12.1), rushing (11.4), crashing (11.2)
MODIFIER + GO, GO + MODIFIER: back (11.4), away (11.0), again (9.9), far (9.7), then (9.7), straight (9.0), now (9.0), just (9.0), down (8.9), alone (8.6), all (8.6), not (8.6), together (8.5), inside (8.4), about (8.4), forth (8.2), well (8.2), up (8.2), there (8.2), thither (8.2), quickly (8.1), immediately (8.0)	MODIFIER + COME, COME + MODIFIER: soon (11.9), back (11.7), here (10.0), again (9.8), there (9.7), never (9.4), out (9.1), down (9.0), then (9.0), up (8.7), rapidly (8.5), forward (8.2), always (8.1), quickly (8.1), inside (8.0), all (8.0), not (8.0), close (8.0), hither (8.0)
GO + WH-WORD: when (10.8)	COME + WH-WORD: when (10.3)
GO + ADJ: first (11.0), outside (10.8), high (10.4)	COME + ADJ: true (12.2), near (11.7)
GO + PARTICLE: out (11.8), on (10.8), down (10.6), up (10.5), along (10.5), off (10.1), away (9.8), over (9.5), in (9.2), through (8.7)	COME + PARTICLE: out (11.7), down (10.6), up (10.6), along (10.2), over (9.8), in (9.7), on (9.4), off (8.8)
GO + PREPOSITIONAL PHRASE: up mountain (12.9), in search (12.4), on way (12.3), for walk (12.1), into room (11.7), in direction (11.5), on journey (11.4), into kitchen (11.2), into forest (11.1), to bed (11.1), into house (10.7), to school (10.2), into garden (10.2), into wood (10.1), to ball (9.8), to field (9.8), to palace (9.6), to house (9.6), to forest (9.5), to river (9.2), to camp (9.2), to home (9.1), to place (9.1), to flounder (9.0), to market (9.0), to man (8.9), to wood (8.8), to lake (8.8), to town (8.7), to king (8.5), to room (8.5), to castle (8.5), to tree (8.5), to beach (8.2)	COME + PREPOSITIONAL PHRASE: for corn (12.5), into land (11.6), from land (10.9), into room (10.7), to house (10.1), to life (9.8), to end (9.8), to castle (9.4), to town (9.4), to place (9.3), to palace (9.3), to tree (9.1), to lake (9.1), to ceremony (8.9), to edge (8.8), to wood (8.8)

Table 6 shows the grammatical patterns and strong collocates of the verbs of movement, *go* and *come*. In this case, the patterns are similar; the only difference lies in the fact that the verb *come* is sometimes found followed by the subject - example 27.

(17)   The third night there *came a Troll* that had nine heads.

All the other patterns are common to both verbs, although the types and variety of collocates may differ. The combination with other verbs in coordinated structures (examples 28) seems to be more important in the case of *go.*

(18)   He was scared, but also needed to *go and see* what had happened to the beast.

Then he *turned and went* up the stairs.

Come and see for yourself.

The subjects include references to animate beings, as in examples 29, (*man, father, boy, wife, child, woman, knight, son, tortoise, people, gnome, Bruno* in the case of *go* and *man, woman, people, voice, friend,* and *giant* in the case of *come*), inanimate beings, as in examples 30 (*thing* and *sun* only in the case of *go*) and time references, as in examples 31 (*time, week* with *go* and *time, night, day, morning,* and *evening* with *come*).

(19)   When *the children went* to bed that night the fairy music came to them in their

dreams.

While he was eating, *the giant came* home.

(20)   So, *things went* on happily for a while.

It was always dull and grey in that country, and when the *sun had gone* down there was no soft light.

(21)   As *time went* by, Jack lost weight and his health became very poor.

When *night came*, they were very far from home.

The most frequent objects found in the corpus with these verbs are similar: *home, day,* and *way* are the same for both verbs and there is a little more variety of strong collocates in this pattern in the case of *go* (examples 32).

(22)   Now we can't *go a step* farther.

They are *coming this way*!

Some more verbs are found as collocates of *go* in the infinitive form (examples 33), mainly because of the grammatical construction 'going to + inf' indicating future time, whereas some more verbs are found as collocates of *come* in the -ing form (examples 34).

(23)   Mummy why doesn't the moon *go to sleep*?

She was *going to have* a new brother or sister.

Then King Midas *went to visit* his daughter and the moment he hugged her, she turned into a gold statue.

All the people of the country *came to see* the great contest.

How did you *come to know* about him?

(24)   Puss in boots *went hunting*, captured a wild boar and brought it to the King.

> The next morning Marina *came running* into the kitchen, a big smile on her face, and a song on her lips.

There is a wide range of modifiers for both verbs, as examples 35 show:

(25)   The rabbit laughed and *went away*.

Everything *went well* for a week or a fortnight.

Come quickly and help me to save them.

Come inside and let me show you something.

In both cases, the only wh-word that often appears after these verbs is *when* (examples 36):

(26)   You can *go when* you like.

He told them to *come again when* the grain was ripe.

Other common patterns consist of the verb followed by an adjective (examples 37), the verb followed by a particle (examples 38), and the verb followed by a prepositional phrase (examples 39).

(27)   My uncle *went first,* and I followed him.

Clouds can *go higher* than dragons.

He thought that his wish had *come true*.

As they *came nearer,* they could see it was a dwarf.

(28)   She therefore *went out* in the garden.

But, one day when Thumbellina *went down* for her nap, a frog hopped through the open window.

They soon reached the island, and the monkey *came down* from the dolphin's back.

Mike's birthday was *coming up* in a week.

(29)   I am now *going in search* of him.

Then he *went on his way,* still very angry.

Once they *went for a walk* together and met an elephant.

Then she began to breathe a little, and little by little she *came back to life*.

And there the show *came to an end*.

In a nutshell, except for the fact that the verb *come* is sometimes followed by the subject, all the other patterns are common to both verbs, as well as some of the collocates, like *go/ come home, go/ come this way*, and they are also followed similar prepositions. Among the distinctive strong collocates that stand out in each case, we can find *go to sleep, go hunting, go outside, go high, go up the mountain, go in search of, go on one's way, go for a walk, go in the direction of, go on a journey, go to bed, go to school, come running, come true, come near, come to life,* and *come to an end*.

Discussion and Implications for TEFL to YLs

In the methodology section, I dealt with the first research question. A selection was made of four pairs of verbs typically used in children-oriented texts based on a) their frequency, b) their

similarity in meaning within each pair, c) their variety (two speech verbs: *say* and *tell*, two verbs of creation and transformation, *do* and *make*, two verbs of sight, *see* and *look*, and two verbs of movement, *go* and *come*) and d) their difficulty for YLs, as these verbs can easily be confused. The difficulty does not lie in the form or in the translation of the verbs, which is quite straightforward. The difficulty stems from the vocabulary depth, basically, from the correct use of these verbs with other words. Durrant et al (2021) reviewed 248 studies on linguistic features in L1 and/ or L2 writing using quantitative corpus linguistics methods, especially focusing on writing development and quality over time. In this large-scale research, they pointed out that the amount of vocabulary used does not measure the amount of vocabulary students know, as it is necessary to assess whether the vocabulary items have been used correctly. They consider formulaic language as a component of a writer's depth of vocabulary knowledge (words used in the context of other words and the appropriate use of collocations).

That is the reason why the section on results, dealing with the second research question, presented the grammatical patterns and strong collocates of these pairs of verbs in a juxtaposed manner. These patterns and collocates were found by analyzing the word sketches retrieved by SE with a Dice Score higher than 8. All this corpus-based information can then be applied to the teaching of English as a foreign language in primary school through DDL activities where YLs are provided with literal concordance lines from children-oriented corpora, or also modified/ adapted concordances (Wicher, 2019) and given the role of detectives – "every student a Sherlock Holmes" (Johns, 1997, p. 101). In all cases, reflection must be included in the process, as DDL goes hand in hand with reflective practice. Both of them enhance learner autonomy through awareness-raising activities designed to foster introspective and conscious learning. This is particularly important for teacher trainers and trainees who advocate a data-led, evidence-based, and practitioner-focused dialogic reflective practice (Walsh & Mann, 2015). Authentic language use and reflective learning are crucial in the design of teaching and learning tasks for YLs (Crosthwaite, 2022, Whyte & Schmid, 2018), although DDL for YLs should always be more teacher-led than for adults. Some activities are presented below to illustrate the use of DDL with the support provided, as examples to be used in an English language class in Year 5 or 6 at a Spanish Primary School.

Scanning-based activities

The students can be given a list of selected concordances, either in concordance-like format or edited as a list of sentences, as is the case shown below, to search for keywords or phrases. This list is only a short sample extract (two examples of each verb from those included in the results section). The concordances should be carefully selected by the teacher to include the intended target language (in the scope or depth that is appropriate to the learners' needs and level of proficiency). Also, the teacher needs to make sure that the students understand the relationship between these sentences - that they are not part of the same text; therefore, they do not have a connection with one another.

1. Then the children *said* goodbye to the mermaid.
2. The elephant always pretended not to hear, and sometimes used to *say* in a very loud voice, 'If there are any worms about, they had better get out of my way'.
3. We may not believe someone who often *tells* lies, even when he *tells* the truth.
4. Jack's mother *told* him to sell their only cow.
5. 'Your majesty, open the book.'; the king *did* so.
6. I tried to *do* what I had *done* before.
7. She was very glad to *see* the hen, which would *make* them rich once more.

8. And from that day on, the little boy always *did* his best to *make* sure that no one felt sad and alone.
9. It was something we all saw with our own eyes.
10. *Come* and *see* for yourself.
11. The poor duckling, who had crept out of his shell last of all, and *looked* so ugly, was bitten and pushed and *made* fun of.
12. He *looked* and *looked* for the shoe but could not find it.
13. Mummy why doesn't the moon *go* to sleep?
14. Then he *went* on his way, still very angry.
15. He thought that his wish had *come* true.
16. While he was eating, the giant *came* home.

Next, the students need to find the answers to some questions, 'hidden' somewhere in the concordances. The types of questions can be related to the form, as in a), but also related to both the form and meaning of the verbs and the collocations or phrases, as in b) and c):

a) identifying the keywords: the teacher can ask the learners to underline all the instances (including all forms) of the verbs: *say, tell, do, make, see, look, go, come*.

b) comprehension yes/ no and wh-questions about an element, e.g. Did the duckling *look ugly*? (Yes, it did), who doesn't *go to sleep*? (The moon), what did the children *say to* the mermaid? (They *said goodbye*), etc.

c) questions to elicit the use of strong collocates, e.g. What is the opposite of *telling lies*? (*Telling the truth*), which is the word in front of *home*? (*came*), which is the word in front of *sure*? (*make*), can you find an expression that means that something very good is finally real? (*come true*), etc.

Activities based on multimodal resources

Including multimodal resources to go with the concordance lines has been suggested before as useful support material that facilitates understanding of the language use and draws YLs' attention: "By generating appropriate accompanying tasks that require the same kind of noticing and pattern-sleuthing processes as found under a more 'typical' DDL approach […] one may greatly expand their repertoire of available DDL tools to include more multimodal resources, and – crucially – tools that YLs, in particular, can both easily use and find appealing" (Crosthwaite, 2022, p. 7). The form-based questions for the concordances above can be combined with other more content-based enigmas, and they can all be embedded in problem-solving tasks like scavenger hunts or escape rooms with narrative elements partly coming from these familiar classical stories and partly new. In order to adapt the level of difficulty to different age and competence groups, some audiovisual support, like pictures, videos, songs, animal sounds, and so on, can be added next to the concordance lines. Different types of tasks can be based on this combination of concordances and multimodal resources: matching concordances and images/ sounds, finding factual mismatches and correcting them (e.g., a picture of a hen instead of a duckling in line 11 of the list above), modeling concordances to provide new captions for pictures, finding the solution to enigmas in games specifically designed to include the selected concordances, in applications like *BreakouEdu* or game-based learning platforms such as *Kahoot*.

Story-related activities

Vocabulary learning should not be approached as rote learning of lists of isolated words, and it can be argued that concordance lines are isolated chunks of language. However, when we are using a specialized corpus, like CSS, (restricted to a particular genre as graphic as children's short stories), many of the sentences are so vivid that they can easily spark the imagination and be considered micro-stories, which can trigger schemata activation. It is not difficult to picture mental images like 'The elephant always pretended not to hear, and sometimes used to *say in a very loud voice*, 'If there are any worms about, they had better get out of my way' (example 3), 'Once upon a time, there was a hungry fox that was *looking for something* to eat' (example 26), 'Clouds can *go higher* than dragons' (example 37). In addition, in many cases, it is easy to trace the sentence back to a familiar story, e.g., '*Jack's mother told* him to sell their only cow' (example 4), 'The Country Mouse one day invited his friend to *come and see* him at his home in the fields' (example 23), 'Then King Midas *went to visit* his daughter and the moment he hugged her, she turned into a gold statue' (example 33).

The co-text provided by concordance lines is not as broad as the context supplied by a whole story, or a whole text in general; nevertheless, this co-text or "contextual sentence" (Syrha, 2021, p. 115) is meaningful as well, and since it is small, it can better help learners to focus on some particular language items and to become aware of the words that tend to collocate with those lexical units. It is particularly helpful at an intermediate stage in the gradual process of literacy development where images (mental as well as physical) and phrases or short sentences combine to create meaningful stories: "For primary-age learners, having access to multimodal forms of data for DDL is very important, given that such learners progress in their literacy development from pictures with word captions, to picture books, then on to chapter books" (Crosthwaite, 2022, p. 6).

Taking these and similar concordances as prompts for stories, several story-related activities can be designed around them, for instance, a) reading/ telling a story to them and every time they hear a particular pattern or collocation, they have to stand up or make a gesture, b) giving each of the students a concordance and ask them to tell/ write a short story that begins/ includes/ ends with that concordance, c) giving them an incomplete comic strip or cartoon with only a panel where a concordance appears and ask them to complete the rest of the strip, d) turning stories with some selected concordances into jigsaws that they need to puzzle out so that they can sequence short paragraphs, sentences or panels to form short narratives.

There is a combination of implicit and explicit teaching in all these activities given above as examples. YLs need to pay attention to the uses of keywords and find out the rules or trends, but these activities also allow for incidental learning of other grammatical and lexical aspects involved. In the case of fairy tales, and the use of the verbs that have been the object of study in this paper, the literary style usually involves fronting, as in *said he, there came, off they went, out came a chicken.* The coordination with other verbs is also an outstanding syntactic feature. Other common grammatical structures include *say + to +* indirect object as opposed to *tell +* indirect object, *go* and *come* followed by an infinitive, all these verbs, except for *make* followed by a *wh*-phrase, the verb *look* followed by an adjective, like *sad/ different/ beautiful*, and the verb *make* followed by a noun phrase and then by an adjective, e.g. *make* someone *happy/ rich/ proud/ angry/ sad*.

Regarding the typical vocabulary found in stories, which can be more or less specialized, in the surroundings of the key verbs studied, there are nouns related to animate beings, such as *child*, *daughter, grandmother, mermaid, troll, emperor, princess*, *knight, gnome, dwarf, witch, fox, tortoise,* time references, for instance, *day, night, morning, evening*, place references, for example, *home, room, house*, *land, palace, castle, tree, sky, lake, field, garden*, adjectives like *happy, sure*, frequency adverbs like *always, never*, manner adverbs like

quickly, closely, discourse markers like *therefore, so*, compounds like *nothing, everywhere*, etc. Finally, from a phraseological point of view, the results have shown some strong collocations of these verbs with noun phrases and prepositional phrases, (more or less fixed, i.e. invariable expressions or expressions with some slight variations such as different articles, determiners, or adjectives, the noun in the singular or plural form, etc.) e.g. *say goodbye/ with smile/ in voice/ words, tell story/ truth/ lie/ tale, do nothing (but)/ work/ something/ duty/ harm/ everything/ trick, make fun/ noise/ haste/ way/ money/ friends/ sure/ up your mind, made of wood, see with eyes/ in the distance/ look after/ into/ like/ for/ at, go up mountain/ in search/ on way/ for walk/ into/ to, come for/ into/ to/ from*.

By increasing the exposure to these grammatical constructions and phraseological combinations in which these basic verbs engage and prompting YLs to observe authentic examples of usage, their production skills can be highly reinforced, since not only do they learn new vocabulary but also, more importantly, how to use it in complete utterances. 'Due to the incremental nature of vocabulary acquisition, repeated exposures are necessary to consolidate a new word in the learner's mind' (Schmitt & Carter 2000, p. 4).

Conclusions

This paper has supported the idea that DDL can be used with YLs and that specialized corpora of children-oriented texts are well-suited for the implementation of this method. A corpus-based study of the word sketches of some basic semantically closely related verbs has been presented to illustrate how particular language items can be explored and how they can be used for teaching English as a foreign language to speakers of other languages, in particular, to Spanish YLs in the last stage of Primary Education. Understanding the behavior of these high-frequency and easily confused verbs can contribute to reinforcing vocabulary learning and fostering the use of idiomatic collocations of core vocabulary. The intended teaching of keywords and their patterns and collocates, along with reflective practice, complements all the incidental learning achieved by the use of communicative teaching methods. Promoting deeper and explicit knowledge of the use of basic vocabulary among language teacher trainees and raising their awareness of the usefulness of incorporating DDL activities in communicative lessons, rather than using deductive grammar exercises, can better equip them to guide their YLs through the learning process of the foreign language.

Disclosure Statement

References

AlAmro, M. (2019). A Corpus-based study on English synonyms: Babble, blather, chatter, gibber, jabber and prattle. *International Journal of Social Sciences and Educational Studies, 6*(1), 122-133. https://doi.org/10.23918/ijsses.v6i1p122

Alenizi, A. M. (2021). A Corpus-based analysis of near synonyms: The case of attack, strike, offensive, aggression and assault. *TESOL International Journal, 16*(3.2), 167-187.

Almakrob, A. Y., & Al-Ahdal, A. A. M. H. (2020). An investigation of the near-synonyms in the Quran: A collocational analysis. *The Asian ESP Journal, 16*(5.1), 327-341.

Bennett, C. & Dhonnchadha, E.U. (2023). Becoming corpus literate: An in-service EFL teacher education framework for integrating corpora into EFL teaching. *Applied Corpus Linguistics, 3*(1), 100048. https://doi.org/10.1016/j.acorp.2023.100048

Boulton, A. 2010. Data-driven learning: taking the computer out of the equation. *Language Learning, 60*(3), 534–572. https://doi.org/10.1111/j.1467-9922.2010.00566.x

Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL, 19*(3), 307-328. https://doi.org/10.1017/S0958344007000535

Church, K., Gale, W., Hanks P., & Hindle, D. (1991). Using statistics in lexical analysis. In U. Zernik (Ed.), *Lexical acquisition* (pp. 115-64). Erlbaum. https://doi.org/10.4324/9781315785387-8

Crosthwaite, P. (ed.) (2019). Data-driven learning for the next generation. *Corpora and DDL for Pre-tertiary learners*. Routledge. https://doi.org/10.4324/9780429425899

Crosthwaite, P. (2022). DDL for Younger Learners. In R. Jablonkai & E. Csomay (Eds.), *Routledge handbook of corpora and English language teaching and learning* (pp. 377-389). Routledge. https://doi.org/10.4324/9781003002901-30

Durrant, P., Brenchley, M., & McCallum, L. (2021). *Understanding development and proficiency in writing: Quantitative corpus linguistics approaches*. Cambridge University Press. https://doi.org/10.1017/9781108770101

Farr, F., & O'Keeffe, A. (2019). Using corpora to analyse language. *Routledge handbook of English language teacher education*, 268-282. https://doi.org/10.4324/9781315659824-22

Gavioli, L. (2000). The learner as researcher: introducing corpus concordancing in the classroom. In G. Aston (Ed.), *Learning with corpora* (pp. 108-137). CLUEB.

Gesuato, S. (2007). How (dis)similar? Telling the difference between near-synonyms in a foreign language. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom. Selected papers from the Sixth International Conference on Teaching and Language Corpora (TaLC 6). University of Granada, Spain, 4-7 July, 2004* (pp. 175-190). Brill. https://doi.org/10.1163/9789401203906_012

Gries, S., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics, 9*(1), 97-129. https://doi.org/10.1075/ijcl.9.1.06gri

Hu, C. & Yang, B. (2015). Using Sketch Engine to Investigate Synonymous Verbs. *International Journal of English Linguistics, 5*(4), 29-41. http://dx.doi.org/10.5539/ijel.v5n4p29

Huang, L. S. (2011). Corpus-aided language learning. *ELT Journal*, *65*(4), 481-484. https://doi.org/10.1093/elt/ccr031

Janda, L. A. (2009). What constructional profiles reveal about synonymy: A case study of Russian words for sadness and happiness. *Cognitive Linguistics, 20*(2), 367-393. https://doi.org/10.1515/COGL.2009.018

Jeaco, S. (2017). Helping language learners get started with concordancing. *TESOL International Journal, 12*(1), 91-110.

Johns, T. F. (1991). Should you be persuaded: Two examples of data-driven learning. *English Language Research Journal 4*, 1–16.

Johns, T. F. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A.Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115). Longman. https://doi.org/10.4324/9781315842677-9

Kennedy, G. (1991). Between and through: The company they keep and the functions they serve. In K. Aijmer and B. Altenberg (Eds.), *English corpus linguistics* (pp. 95-111). Longman.

Kilgarriff, A., Kovár, V., Krek, S., Srdanovic, I., & Tiberius, C. (2010). A quantitative evaluation of word sketches. In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the XIV Euralex international. Leeuwarden, Netherlands: Fryske Academy congress* (pp. 372–379). Euralex.

Labrador, B. (2003). Typicality in the Use of English and Spanish Indefinite Quantifiers: A Functional Approach. *Babel, 49*(1), 41-58. https://doi.org/10.1075/babel.49.1.04lab

Labrador, B. (2022). Word sketches of descriptive modifiers in children's short stories for teacher training in teaching English as a foreign language. *Linguistics and Education, 69*, 101036.

Liu, D. (2010). Is it a chief, main, major, primary, or principal concern?: A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics*, *15*(1), 56-87. https://doi.org/10.1075/ijcl.15.1.03liu

Lusta, A,, Demirel, O., & Mohammadzadeh, B. (2023). Language corpus and data-driven (DDL) in language classrooms: A systematic review. *Heliyon, 9*(12), e22731. https://doi.org/10.1016/j.heliyon.2023.e22731

Nation, P. (2005). Teaching vocabulary. *The Asian EFL Journal, 7*(3), 47-54.

Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning 35*, 36-61. https://doi.org/10.1080/09588221.2019.1667832

Rojo-López, A. M. (2011). Distinguishing near-synonyms and translation equivalents in metaphorical terms: Crisis vs. recession in English and Spanish. *Review of Cognitive Linguistics, 9*(1), 280-314. https://doi.org/10.1075/rcl.9.1.13roj

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

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal, 9*(1), 4-9. https://doi.org/10.1002/j.1949-3533.2000.tb00220.x

Syrha Goveas, M. (2021). Contextual understanding of vocabulary - A heuristic approach aided by modern technology. *The Asian EFL Journal, 28*(3.1), 110-118.

Thomas, J. (2014). *Discovering English with the Sketch Engine*. Laptop Languages. https://doi.org/10.14705/rpnet.2014.000246

Thornbury, S. (2004). *How to teach vocabulary*. Pearson.

Vandevoorde, L., Lefever, E., Plevoets, K., & De Sutter, G. (2017). A corpus-based study of semantic differences in translation. The case of inchoativity in Dutch. *Target. International Journal of Translation Studies, 29*(3), 388-415. https://doi.org/10.1075/target.15009.van

Walsh, S., & Mann, S. (2015). Doing reflective practice: a data-led way forward. *ELT Journal, 69*(4), 351-362. https://doi.org/10.1093/elt/ccv018

Wicher, O. (2019). Data-driven learning in the secondary classroom: A critical evaluation from the perspective of foreign language didactics. In P. Crosthwaite (Ed.), *Data Driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners* (pp. 31-46). Routledge. https://doi.org/10.4324/9780429425899-3

Wiliński, J. (2017). On the brink of-noun vs. on the verge of-noun: a distinctive-collexeme analysis. *Research in Language, 15*(4), 425-443. https://doi.org/10.1515/rela-2017-0024

Whyte, S., & Schmid, E. C. (2018). Classroom technology for young learners. In S. Garton & F. Copland (Eds.), *Routledge handbook of teaching English to young learners* (pp.338-355). Routledge. https://doi.org/10.4324/9781315623672-22

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103–129. https://doi.org/10.1093/applin/ami045

Yuting, X. (2020). Using Sketch Engine to Investigate Synonymous Verbs Complete and Finish in English. *International Journal Advances in Social Science and Humanities*, 11-25.

Vyatkina, N. (2020). Corpora as open educational resources for language teaching. Foreign Language Annuals, 53(2), 359-370. https://doi.org/10.1111/flan.12464

Copyrights