

Stretched verb collocations with give: their use and translation into Spanish using the BNC and CREA corpora

SILVIA MOLINA-PLAZA

*Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Navales,
Arco de la Victoria s/n, 28040 Madrid, Spain
(email: Silvia.Molina@upm.es)*

EDUARDO DE GREGORIO-GODEO

*Universidad de Castilla-La Mancha, Departamento de Filología Moderna, Facultad
de Letras, Avda. Camilo José Cela s/n 13071 Ciudad Real, Spain
(email: Eduardo.Gregorio@uclm.es)*

Abstract

Within the context of on-going research,¹ this paper explores the pedagogical implications of contrastive analyses of multiword units in English and Spanish based on electronic corpora as a CALL resource. The main tenets of collocations from a contrastive perspective – and the points of contact and departure between both languages – are discussed prior to examining the commonest types of verb + noun combinations as a significant case of so-called ‘de-lexicalized’, ‘light’, ‘empty’, ‘thin’, ‘stretched’ or ‘support verbs’. A qualitatively and quantitatively-oriented case study is accordingly conducted, determining the weight of *dar* in support verb constructions within the *Corpus de Referencia del Español Actual* (CREA) and of the English equivalent stretched verb constructions with *give* within the British National Corpus (BNC). Based on the empirical data obtained in this way, this paper provides relevant insights for more accurate translations, helping to enhance the collocational competence of L2 students, who tend to avoid constructions including empty verbs like *give* in favour of full-verb forms. The detailed findings in this paper come to shed light on the potential of CALL resources for improving the collocational usage of foreign-language learners, as quantitative and qualitative comparisons of collocations based on electronic corpora serve to highlight the similarities and, more importantly, the lexical and typological differences between both languages, thereby substantiating the invaluable role that corpus analysis may play for language teaching in general and for collocational knowledge and proficiency in particular.

Keywords: Collocations, stretched verbs, translation, phraseology, CALL

¹ BFF 2003-02540 and PAI 07-0018-0804 aiming to produce an electronic dictionary of collocations and idioms for Spanish-speaking EFL learners: *Diccionario de unidades fraseológicas inglés/español* [DUFIE].

1 Introduction: CALL and phraseology at the crossroads

Computer-assisted language learning (CALL) is concerned with the use of computer technology in second language learning, thereby attempting to provide computer-based resources to improve language learning and teaching practices. But to what extent may CALL contribute to students' learning of phraseological units, that is, lexico-grammatical clusters or 'chunks' incorporating everyday experience? By and large, phraseology is concerned with the study of those chunks which, be they collocations or idioms, constitute some crucial cognitive, textual and pragmatic tools to be mastered by the language learner. As Sokolik (2001: 487) underlines in her overview of CALL, "corpus linguistics and concordancing can help provide the data and tools that students and instructors need to make sense out of usage". CALL applications and phraseology may thus provide invaluable resources for the student's knowledge of multiword units such as de-lexicalized verbs which often function differently from *stretched* or *support* verbs. Prior to undertaking a detailed case study exploring this phenomenon and its implications for language learning, some remarks will be made on CALL, phraseology and their interface. As described in more detail below, this paper will thus examine the potential of CALL resources for improving foreign language learners' phraseological knowledge.

1.1 CALL

No-one would dispute nowadays that multimedia and the Internet have an incredible potential to make a huge impact on language teaching. For the moment, however, both technologies present challenges which are every bit as great as the opportunities they open up. Experienced teachers know that they are very demanding in terms of time: a good CD-ROM can take hours to explore. Likewise, the riches of the Internet, especially since the advent of the Worldwide Web, can lead users into hours of 'surfing' with very little concrete material to show at the end of it. CALL has thus become a new, but well-established, scholarly domain researching the pedagogical possibilities provided by computers and the Internet for increasing language learners' communicative skills (Warschauer & Kern, 2000; Warschauer, 2001; Blake, 2001; Davies, 2002; Godwin-Jones, 2005; Oster, Ruiz Madrid & Sanz Gil, 2006). CALL may be broadly defined as "any process in which a learner uses a computer and, as a result, improves his or her language" (Beatty, 2003: 248). Evolving from more technical approaches to and definitions of the term CALL, such as Levy's (1997), to the emphasis on its pedagogical implications (Egbert, 2005), the notion of CALL has come to be "used to describe the introduction of computers into the field of L2 learning and teaching" (Ruiz-Madrid, 2007: 64). As Sokolik (2001: 486) stresses in her examination of the forms and functions of computer technology in language learning, "it is clear that computers are providing instructors and students alike with a new battery of tools with which language can be learned more effectively". Work by Hanson-Smith (2000), Warschauer and Kern (2000), Chapelle (2001), Godwin-Jones (2000), Jeong-Bae (2004), Szendeffy (2005), Smith and Baber (2005) or Kern (2006), just to quote a few examples, evidences the potential of computers for language learning and teaching having shaped CALL as a vibrant academic field. In particular, concordancing, and other corpus-related activities based on

electronic corpora – which is closely related to the case study conducted in this contribution – “lends itself to a range of classroom activities [...] particularly as large-scale corpora become generally available, together with the techniques for analyzing them through corpus linguistics” (Johnson & Johnson, 1999: 83), thereby substantiating the potential of corpora for classroom activities in language learning and teaching.

Bearing this context in mind, it is obvious that CALL education is growing fast as the learning of languages is linked with technology competence. This paper is focused on project-based learning (Beckett & Millers, 2006) exploring how collocational competence may be integrated into English language programs in both the Technical University of Madrid and the University of Castilla-la Mancha, Spain, by developing a multiword unit dictionary which will enhance the collocational competence of students. We also focus on a specific type of collocation, *give + noun*, in the two languages, in order to ascertain the points of contact and departure.

1.2 *Phraseology: definitions, contrastive analysis and use*

Phraseology has now come of age. For a long time, both linguists and psychologists have paid attention to multi-word units as syntagmatic patterns (Firth, 1951: 190–215; Cowie, 2004: 37–52) and tried to analyze how they are represented in the lexicon. They have noticed that language is acquired in cohesive lexico-grammatical clusters or ‘chunks’ which capture everyday experiences and constitute crucial cognitive, textual and pragmatic tools to be mastered. Many authors have highlighted the enormous difficulty in clearly defining and delimiting this complex of features that interact in various, often untidy, ways and represent a broad continuum between non-compositional – or idiomatic – and compositional groups of words (Moon, 1998: 6). This author calls them fixed expressions and idioms (FEIs).

These FEIs or chunks are usually examples of formulaic language (Nattinger & de Carrico, 1992; Wray, 2002), in which word strings occurring together tend to convey holistic meanings that are either more than the sum of the individual parts or else diverge significantly from a literal, or word-for-word meaning, and operate as a single semantic unit. Gries (2008: 4 and ff.) identifies a set of parameters worthy of attention that are typically implicated in phraseological studies:

- i. the *nature* of the elements involved in the phraseologism;
- ii. the *number* of elements involved in a phraseologism;
- iii. the *number of times* which an expression must be observed before it counts as a phraseologism;
- iv. the permissible *distance* between the elements involved in a phraseologism;
- v. the degree of *lexical and syntactic flexibility* of the elements involved;
- vi. the role that *semantic unity* and semantic non-compositionality/non-predictability play in the definition.

He then defines what a *phraseologism* is (*op. cit.*, 2008: 6), a definition accepted in this study: “the co-occurrence of a form or lemma of a lexical item and one or more additional linguistic elements of various kinds which functions as one semantic unit in a clause or sentence and whose frequency of co-occurrence is larger than expected on the basis of chance”.

1.3 Bridging the gap between CALL and phraseology

An exploration of the literature on the subject reveals that work in CALL has focused on various topics related to lexicography such as on-line dictionaries and the introduction of their use as part of CALL (Campoy, 2004: 47–72); use of audio files in computer mediated dictionaries (Sobkowiak, 1999: 246 ff.); dictionary usage guidance in paper dictionaries versus online dictionaries (López & Campoy, 2003); user typology (de Schryver, 2003: 151); or simultaneous look-up in dictionaries with the same information (Luzón, 1999), to quote a few relevant examples. Given the emphasis of this article on phraseology, it is to be noted that, in addition to grammar and vocabulary practice computer programs, the use of concordance programs has been discussed as another means for vocabulary and grammar practice where “the students can view many examples of usage and compare them to their own writing without having to search manually through many pages of text” (Hanson-Smith, 2004: 111) as well as on-line concordancers which help students use search engines to find typical collocations and grammatical or rhetorical items on the internet (Mills, 2000). However, no work has been done so far on specific types of collocation that support verb constructions, which are frequent in oral and written genres. Our examples in English and Spanish reveal that these flexible collocations are widely used in discourse simply because they are adaptable to a wide range of situations. However, we have noticed that non-native speakers tend to underuse these de-lexicalized collocations. Therefore, CALL is a useful means of paying attention to them, enhancing the students’ production and understanding of these phrasemes as well as improving their communicative competence in general.

1.4 Scope, purposes and methodology

The studies in L2 raise an awareness that a significant proportion of the language that we produce is made up of collocations and idioms, labelled as formulaic sequences or phrasal units by other authors (Sinclair, 1997: 82). If this is the case for native speakers, a logical consequence is that learners of English will also find these formulaic sequences very important and useful. Keeping this context in mind, it is fairly obvious that teachers should try to increase the acquisition of these sequences by learners, and design teaching materials highlighting the similarities and differences between English and Spanish. In section 2 of this article, through a detailed case study exploring phraseological units from a contrastive perspective, we report on current research (DUFIE, see footnote 1) on multiword units taken from a contrastive approach in English and Spanish. Section 3 focuses on a subgroup of these collocations known as *de-lexicalized*, *light*, *empty*, *thin*, *stretched* or *support* verbs (Jespersen, 1942; Mel’čuk, 1993; Allerton, 2002). In particular, the overall aim and the subsequent methodology of this paper is two-fold:

Firstly, to present on-going research (DUFIE) into contrastive multiword units in English and Spanish. The main tenets of collocations from a contrastive perspective – and the points of contact and departure between both languages in this specific type of collocation – are accordingly examined through a case study which considers the pedagogical implications of phraseology usage within a broader CALL.

Secondly, to analyze these collocational structures in the subsequent sections, with a focus on the commonest types of verb + noun combinations, and especially on what is known as de-lexicalized, light, empty, thin, stretched or support verbs (Jespersen, 1942; Mel'čuk, 1993; Allerton, 2002), such as *give one's approval for*. The analysis in the case study may also serve as an example of the CALL implications of collocational usage by foreign language learners.

More specifically, we want to determine the weight of *dar* in the most widely accessed on-line Spanish corpus, that is, the *Corpus de Referencia del Español Actual* (CREA) and of its English equivalent support verb constructions (i.e., *give access*, etc.) in the British National Corpus (BNC). Both corpora have over a 100 million words, are linguistically representative and easy to access. The aim of this study based on empirical data is to provide relevant insights for more accurate translations and to enhance the collocational competence of L2 students (Bahns, 1993), who tend to avoid these constructions in favour of full verb forms (i.e., *to access*, *account*, etc). Examples from our teaching experience show that the translation of collocations is difficult for non-native speakers. They also prove that many collocation translations are idiosyncratic in the sense that they are unpredictable by syntactic or semantic features. Taking a broader pedagogical perspective drawing upon the implications of corpora for language learning and teaching, we aim to demonstrate that combined and integrative use of phraseology and CALL may thus provide most helpful insights not only for contrastive collocational analyses across languages but, more importantly, for foreign language learners and teachers who need to acquire collocational knowledge and competence. Arguably, quantitative and qualitative comparisons of collocations based on electronic corpora are greatly needed to highlight the similarities and, more importantly, the lexical and typological differences between both languages for a more fruitful use by English as a Second Language (ESL) learners. With a view to facilitating lexical and cultural acquisition in the L2, all these findings are to be reflected in an electronic translation-oriented electronic dictionary of collocations and idioms for advanced learners of English which may be employed as an invaluable CALL instrument.

2 Background to the study: the DUFIE project

This project started from a strong conviction that the collocational needs of students learning English are not adequately met by the range of bilingual or monolingual collocation dictionaries available. Collocations (e.g., *horse riding*), proverbs (e.g., *you can lead a horse to water but you can't make it drink*) and idioms (e.g., *eat like a horse*) are grouped indiscriminately along with other phraseological units. From a translation perspective, three facts are to be taken into account for translating multiword units as accurately as possible:

- a) Literal translation of the constituents should be avoided because phraseological units only rarely have the same form in the TL, leading to an 'unnatural' equivalent at best (e.g., *a severe winter* > **“un invierno severo”*). The adjectives *inclemente* or *crudo* are better options here, with three tokens and seven respectively in the Spanish Royal Academy Corpus, CREA (*Corpus de Referencia del Español Actual*). Literal translation may also lead to a

mistranslation at worst, which is totally meaningless for a non-native speaker (e.g., *no querer alhajas con dientes* > *‘‘not to want jewels with teeth’’). The latter Spanish expression means ‘‘to reject a valuable present such as a jewel (*alhaja*) because it is not a real gift in fact’’.

- b) The starting point for the rendering of the phraseological unit should be its meaning, rather than considering it as a unit of form. Moreover, as phraseological units can be polysemous, it is the meaning in context that must be rendered. Thus, the meaning of *estar chungo* is different in the sentence *María está chungo* (i.e., ‘‘María is in a bad way, out of sorts’’) and the sentence *La cosa está chungo* (i.e., ‘‘Things aren’t too good’’).
- c) The meaning of the SL phraseological unit can be rendered in one of three ways:
 - by an equivalent phraseological unit in the TL, if one exists: *al pan, pan y al vino, vino* > ‘‘to call a spade a spade’’;
 - by a single word (*in order to* > ‘‘para’’);
 - by a paraphrase (*nobody took a blind bit of notice* > ‘‘nadie le hizo ni pizca de caso’’).

The problems that learners face while translating or trying to produce acceptable collocational patterns in their speech, writing, and translation show a certain regularity. We were concerned by the similarities in this research but, above all, by the diversity in the meaning, structure and use of set phrases in the two languages, bearing in mind that these set phrases constitute a major aspect of any language and become a frequent difficulty in translation courses, a fact which has been proven empirically by Tirkkonen-Condit (2002) in her study of translation errors hampering the target texts.

The aforementioned arguments substantiate the need to produce a better phraseological dictionary for advanced learners of English, to improve their performance in translation in particular and enhance their collocation skills in general. Selecting the most significant FEIs proved a difficult task, as we were struck by the significant diatopic, lexical and even grammatical variation in their use. These variations have been included whenever possible.

So, the aim of this lexicographical project is to include updated general English collocations, idioms and multiword units with their translation equivalents into Spanish. If there is not an equivalent multiword unit, we look for a functional equivalent. At the first stage, the most significant 10,000 multiword units from the Bank of English and the BNC were included, explaining relevant cultural information with usage notes when necessary. We have included terminological collocations as well, as many disciplines or technical domains create their own set phrases and multiword terms. At the second stage, the most relevant English multiword units have been translated into Spanish, including examples of real usage following Cowie’s taxonomy (Cowie & Mackin, 1993: xii–xiii).

Figure 1 includes just a few examples of a DUFIE entry with the word *give* with its 59 most frequent collocates. One example of real usage from the BNC, the Bank of English or the Internet is included in each multiword unit. Hyperlinks with other entries are marked in yellow. Collocations or idioms that take on different meanings in different contexts are indicated in translation separated by slashes offering different alternatives as in example 4. Last but not least, the reader gets information on

GIVE v.

1. **give advice; give a chance** *dar, ofrecer, proporcionar un consejo; dar, ofrecer la oportunidad.*
2. **give a bash on; give colour; give credit** *dar un porrazo; a) dar color, b) hacer que algo parezca creíble o verosímil; dar crédito a alguien.* Occasionally someone may tell you that refusing credit can be the basis for legal action on the grounds that a reputation has been unfairly damaged by a refusal to give credit in a business where giving credit is normal.
3. **give chase** *perseguir algo/ alguien, salir en persecución, ir tras/ en pos de algo/alguien.* Ben loved to run along the public footpath through the ripe fields, but at times, although not to venture into the field itself, he would often spot a rabbit and give chase.
4. **not give a damn/ hoot** *importar un bledo/ un pepino/ un pimiento/ un pito (col.)* Jordan, from his expression, didn't give a damn about either Leila or the group.
5. **give an example** *dar un ejemplo.* Can you give an example when you say that the law is designed?
6. **give evidence for//against** *prestar declaración a favor// en contra de.* They admitted manslaughter and gave evidence against Allen.
7. **give a fillip** *estimular algo.* I am sure that he also welcomes today's news that the Abbey National, Halifax and Nationwide building societies are reducing their lending rates to, as they put it, give a fillip to the housing market.
8. **good to give** *es bueno dar (...)* It is good to give credit where credit is due.
9. **give a growl** *dar un gruñido/ gruñir.* Sid gave a growl and informed that as I knew nothing about modern techniques.
10. **give sb a dirty LOOK** *echarle a alguien una mirada asesina.* At this Mum turned round and gave her a dirty look before she moved along with Dad, followed by the rest of us.

Fig. 1. Examples of multiword units with *give* as a key word.

whether the multiword unit is typical of British or American English – *blow your own trumpet* (BrE)/*blow your own horn* (AmE) – whether it is more common in oral English, and register indicators (colloquial, slang, etc.).

This dictionary focuses on lexical collocations, which do not normally comprise prepositions, infinitives or clauses as grammatical clusters. The prototypical structure of lexical collocations is usually formed by substantives, adjectives, verbs and adverbs. The electronic dictionary will help to identify these specific subgroups of collocations such as support verb constructions made up by verb + noun, as in section 3 of the paper.

3 Case study: contrastive analysis of support verb constructions with *give/dar* in two comparable corpora using the BNC and the CREA

A light verb, such as *give*, *take*, or *make*, is combined with a wide range of complements from different syntactic categories to form a new predicate called a light verb construction or support verb construction (LVC/SVC). From a semantic perspective, these constructions consist of a predicative noun and a support verb. The fundamental idea of a SVC is the realization of the arguments (i.e., the semantic actants) of the predicative noun as syntactic actants of the support verb. Prototypical support verb constructions are semi-compositional structures consisting of a semantically transparently used noun and a verb that is semantically reduced and adapted to the construction.

These expressions form a cline of idiomaticity from ‘clearly idiomatic expressions’ (e.g., *take time*, *have a look*) to ‘relatively idiomatic expressions’ such as *have a chance*, *take a walk*, *make a statement*, where the meaning of individual words is retained up to a certain extent and there are expressions that retain the core meaning of these verbs at the other extreme (Biber, Johansson, Leech, Conrad & Finegan, 1999: 1027): e.g., *you can take a snack in your pocket*, *he made a sandwich*. These support verbs (e.g., *do*, *make*, *get*, *give*, *do*, *take* and *have*) are, according to corpus research, some of the most frequently used words in the English language. Some people also call them ‘empty’ verbs because they usually have very little meaning. They are also particularly important for speakers of Latin languages because students avoid or do not even feel the need to use them as there is normally a one word translation between their L1 and English. However, when speaking, learners can sound odd or strange to the native speaker as their choice of language is different.

We decided to analyze support verb constructions with *give*, which is halfway between full lexical verbs and auxiliary verbs, in the BNC, and wanted to compare them with their Spanish homologues in the CREA, as the literature on this topic points out that in languages such as French, Italian, Spanish and English, support verb constructions are semi-productive (Wierzbicka, 1982; Alba-Salas, 2002; Kearns, 2002). Both corpora are considered comparable in corpora studies (Hornero, Luzón & Murillo, 2006) as they are both over 100 million words, despite the fact that there are differences in the way they have been compiled.

Hence, the data in this paper consist of twenty-four support verb collocations in English with *give* and their Spanish translation, *dar*, in most cases attested in bilingual dictionaries and bilingual translated texts from the EU website. We have chosen noun collocates with a high frequency in both languages. Our references in Spanish are Ueda

(1989), who selected the most common nominal collocates (2,727) and the dictionary *Diccionario de Uso del Español (Dictionary of Spanish Usage)* by Moliner (2000), as it indicates the verb collocates that usually go with the de-lexicalized base. In English the raw frequencies of the infinitives comprise occurrences in all the BNC subcorpora and in Spanish we have followed the same criterion with the CREA. We have excluded tense and number variation but we have included lexical variation to check whether the support verb construction with *give/dar* was more usual or not.

3.1 English

Twenty-four noun phrases combining with *give* in the infinitive form were selected using concordances. It is imperative to notice that most nouns are abstract (e.g., *gives a speech*) in comparison with common collocations in the oral register with other support verb constructions such as *have a baby* etc. However, there are examples such as *give someone a present, give money, give a book* which have literal meaning and the noun is concrete. Light verb usages of *give* constitute a continuum of meaning from literal uses at the core to figurative ones located at the periphery such as *give a push, give a kick* (see Figure 2).

Other *give* collocates have been barred such as *give somebody a present, give up one's seat* because its core meaning – ‘pass good news to owner’ – is rare in this use. There are also other core meanings of *give* such as ‘allow/permit/grant freedom to act’ as in *give somebody access, give somebody admission, give somebody authority* which are not so common either according to Allerton (2002: 180) and, consequently, have been excluded from the analysis as well.

Attention is paid instead to the commonest core meaning – perform a service for others – as exemplified by *give somebody aid, give somebody an answer, give somebody an appointment*, etc. In most cases, *give* as a thin verb is used for action that is beneficial to the speaker, performing a service for them, with some exceptions: *give somebody a beating* or *give somebody one's attention* is not necessarily of benefit to the recipient. The nouns that collocate with *give* in our data are in line with its sub-meanings: what is given is usually a right or permission to act (i.e., *give access, admission*, etc.). The thing given can also be information (*give an answer, give assurance*) or a different kind of general service (*give aid, applause, attention*, etc.), or a particular one, according to Allerton (2002: 181) such as *give an appointment, a bath*, etc. This also seems to be the case with the Spanish counterparts. See graph 1 for the five most common uses in our data.

In our corpus-based approach, the collocational patterning of *give* collocates with words with pleasant connotations such as *advice, assistance, hand*. If the starting point is the lexical item, one may find a range of verbs which collocate with the noun. Such is the case with *attention*, which not only collocates with *give** but also with *pay**. Concordancing these two verbs shows that *give* is associated with informative texts and *pay* has a more positive semantic prosody, as in the following examples:

- (1) *There are two aspects of its definition to which we must first give attention.*
- (2) *After visitors have departed, nurses should pay attention to a patient's non-verbal behaviour as well as to what is said.*

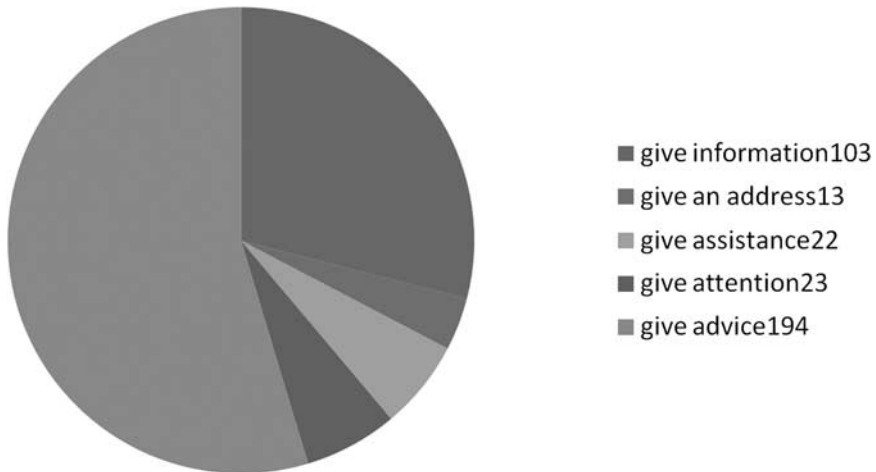
Support verb constructions with <i>give</i>	Number of tokens in the BNC	Example
1) <i>give advice</i>	194	Staff are available in all electricity shops to give advice to anyone in difficulty.
2) <i>give information</i>	103	They will also be able to give information about local voluntary agencies.
3) <i>give an answer</i>	31	You have one minute to give an answer .
4) <i>give/ pay sm. attention (to smt./ sb.)</i>	23/235	There are two aspects of its definition to which we must first give attention .
5) <i>give/ render assistance to sb</i>	22/5	As a result, Chairman Roger Annan has resigned although he will continue to give assistance to the Executive on special projects.
6) <i>give a hand to sb</i>	19	He called to his wife, 'Will you give a hand , Nessie?'
7) <i>give an address to sb</i>	13	He borrowed some benches from the old school room at Stalling Busk for the congregation to sit on, engaged the Rector of Spennithorne to give an address , and finally put his plans into
		action in August 1956.
8) <i>give/ offer sb aid</i>	11/0	It isn't enough to give aid .

Fig 2. Support verb constructions with give in the BNC.

9) <i>give evidence for/ against</i>	9/19	I would expect someone who claimed to believe in ghosts to give some evidence for their existence, (...).
10) <i>give chase</i>	9	The police car was speeding up to give chase when the co-driver warned his partner to ease off.
11) <i>give a lift</i>	7	Having checked that they are heading for Prague, I give a lift to the two young men at the head of the queue.
12) <i>give an injection</i>	5	If this does not help, the doctor may be able to give an injection to dry up secretions and make the noise less upsetting.
13) <i>give a look</i>	2	I'll give a look for it before you go
14) <i>give a prize</i>	2	Provide paints and give a prize for the best-decorated egg.
15) <i>give authorization</i>	2	The inspectorate told the company it would be unlikely to give authorization under those conditions.
16) <i>give news</i>	2	While as early as Elizabethan times crime chap-books were published to give news about recent crimes, (...)
17) <i>give a smile</i>	1	She finds it warming that the check-out girls give her a smile.
18) <i>give a solution/ solutions</i>	1/1	We want to rush in and give solutions, or reassurance when dealing with pain and confusion.

Fig. 2. Continued.

QUANTITATIVE FINDINGS



Graph 1.

It is also important to point out that *give* tends to appear interpersonally marked with modal verbs such as *must* or *need to*, which provide subjective meaning with a deontic value as in:

- (3) *Although it is often claimed that the camera does not lie, when using old photographs for historical purposes pupils need to give attention to the processes involved in choosing a subject and taking a photograph.*

Another point of interest is that *pay* appears in imperative forms:

- (4) *Pay attention to contrasting colours, [...].*

After examining the highest collocation in our corpus, *give attention*, the collocational patterning of *give information* was examined. Its 103 examples in the BNC show that this support construction is frequently encountered and accessed in English. From the ideational standpoint, this collocation is used to provide objective and fact-driven data as in the following example: *Europe wants shops to give information on the efficiency of all white goods, televisions and hi-fis, computers, ovens, lights and other household appliances down to the smallest toaster.* It is fairly obvious, however, that language is multifunctional and the interpersonal function is also prevalent in many instances of this collocation. Lexical markers such as adjectives and adverbs are often stance markers evaluating the content of utterances as in: *So there's a kind of critical period here that we have got a lot of information about.*

Give an answer is the third highest collocation. The key concept is 'answer' and falls under the *suasion* function (Wilkins, 1976: 46), specifically used to persuade, suggest, advise, advocate, etc., that is, affecting the behavior of others. A recommended course of action is put forward to solve an existing problem, as in the following news report:

- (5) *A Northern Ireland nurse has condemned local maternity hospitals for failing to give information useful to expectant mums.*

Last but not least, it is to be noted that there is a dispersion of results in the remaining collocations with *give* indicating different types of verbal processes. *Give recommendations* is found in the Introductions as many as three times in the present simple active:

- (6) *The manual's sections take each aspect of a hotel operation in turn and give recommendations for action.*

Interpersonality is lacking in the three examples and this SVC is normally used cataphorically as in the former example. However, many of these *give* collocates may be related to the interpersonal function and tend to have positive semantic prosodies and occupy Rheme position, e.g., *Provide paints and give a prize for the best-decorated egg.*

As a conclusion to the analysis of the BNC data, our findings of the lexicogrammatical patterns of *give + noun* reveal three additional phenomena:

- a) Texts strongly favour full lexical verbs to verbalize phenomena and processes in the BNC instead of support verb constructions, as the raw frequency of tokens of *assist/give assistance* attest: *assist* (2512) vs. *give assistance* (22). The Spanish support verb constructions with *dar* clearly show this preference as well: *gruñir* (49) versus *dar un gruñido* (2).
- b) *Give* as a thin verb in our data suggests action that is beneficial in the sense of facilitating action by other people, informing or performing a service for them.
- c) Lexical variation should be taken into account: there are 23 instances of *give attention* in our corpus versus 235 of *pay attention*. Corpus analysts should therefore analyze not only the verbs but also the nouns as a starting point in order to get the full range of verb collocates.

By and large, the work herein presented may be regarded as a step towards a better understanding of collocational links of support verb constructions for the purpose of learning and Natural Language Processing (NLP). As stated above, phrases with these types of verbs are fairly frequent in English and pose serious problems for learners and translators alike. It must be borne in mind that the support verb cannot be translated by the default equivalent of the 'heavy' sense of the verb. It is incorrect to translate *give evidence* into Spanish as * "dar declaración", as the appropriate collocation is "prestar declaración". Point 3.2 discusses how *give + noun* collocates behave in Spanish.

3.2 Theoretical underpinnings and quantitative/qualitative analysis of light verb constructions in the Spanish data

Koike (2001) has studied lexical collocations in general, but has also paid attention to support verb constructions which he terms general functional verbs (2001: 69), and sheds light on the *dar + noun* support construction, which is highly productive in his frequency lists. He comments that the verb *dar* ["to give"] is transitive and loses its original semantic meaning when it is used in de-lexicalized verb constructions such as *dar una bofeta* ["to slap somebody in the face"]. In very general terms, he also states that *dar* can be used with either concrete or abstract nouns such *dar un regalo*

[“give a present”] or *dar información* [“give information”] respectively. *Dar* admits synonymy in certain constructions (e.g., *dar/pegar una paliza*; *dar/emitir una opinión*; *dar/causar/producir vergüenza*) and diatopic variation (e.g., *dar una bofetada* in Peninsular Spanish versus *fletar una bofetada* in Chile and Peru).

Furthermore, Koike notices five other important facts about *dar* as a support verb construction:

- a) Verbs related morphologically to nouns which appear in these support verb collocations do not show a tendency to form collocations themselves. Thus, the verb *aconsejar* does not collocate as the noun *consejo* [“advice”] does with the verb *dar* in the delexicalized verb constructions *dar un consejo* [“give advice”].
- b) Action, motion and sound nouns are compatible with *dar* such as (i) nouns related to blows: *dar(se) un golpe* > “to hit/ bang” (your head, etc.), *dar una torta* > “to hit somebody”; and (ii) nouns related to sounds: *dar un aullido* > “to give a howl”. There are also other animal sounds using *dar + noun* construction not commented upon here for brevity’s sake (see Koike, 2001: 111 for a full account).
- c) *dar + se*, the impersonal pronoun, has mostly negative semantic prosodies: *darse un atracón* > *to gorge or stuff oneself*, *darse un batacazo* > *to fall over and bang your arm, leg, etc.*, *darse una paliza* > *to work one’s butt off (AmE)/ to slog one’s guts out (BrE)* or *darse un susto* > *give him/her a fright*. Nevertheless, some are neutral: *darse una ducha* > *to take/have a shower*; *darse maquillaje* > *to put one’s make-up on*.
- d) Some constructions have a causative value arising from the noun phrase: *dar alegría* > *make sb happy*, *dar angustia* > *cause great anguish or distress*.
- e) Koike (2001: 85) points out that the support verb construction *dar + substantive* is the most common light verb collocation, followed by *tener* (to have + noun). The former is a common lexical collocation with different processes:
 - Processes of communication > *dar respuesta*, *dar consejos*.
 - Processes of cognition and decision-making > *dar la aprobación*, *dar ayuda*.
 - Processes of change of state > *dar un premio*.
 - Dynamic durative verbs (taking place over a period of time) such as activities performed by animate agents: *dar articulación*.
 - State of emotion or attitude verbs: *dar un gruñido*.

We have studied the equivalent phraseological units whenever possible in the Spanish corpus with *dar*. There are sometimes several possible translations into Spanish as stated above (*give information* > *dar/proporcionar información*), but we have worked mainly with the direct translation to see the quantitative differences between languages (see Figure 3). Ten tokens of *dar + noun* are shown in box 2. No tokens have been found of *dar un gruñido* [“give a growl”] or *dar articulación* [“give articulation to smt.”], exactly the same as in English. Their past tense forms are also practically negligible: 2 and 0 tokens in that order. Graph 2 shows the 5 most common occurrences of SVC *dar + noun*.

Our way of looking at collocation is to start with the verb and then move to the noun, which raises the question of a word’s semantic prosody. *Dar* collocates with

Support verb constructions with <i>dar</i>	Number of tokens	Example
<i>dar respuesta</i>	462	Dirigimos esta carta a "La Vanguardia" para dar respuesta a la carta publicada el 19 de mayo.
<i>dar información</i>	108	Ni la policía ni el juzgado accedieron a dar información .
<i>dar/ ofrecer consejos</i>	63/3	Puede que sea demasiado joven para dar consejos a dos directores consagrados.
<i>dar problemas a algn.</i>	12	La aorta dilatada no suele dar problemas serios casi nunca.
<i>dar/ ofrecer ayuda</i>	9/14	La aventura era excitante. La subida se hacía cada vez más difícil y los excursionistas respondían mucho mejor de lo que el guía esperaba. La solidaridad se hacía presente cada vez más en el grupo, presto a dar ayuda a quien la necesitara.

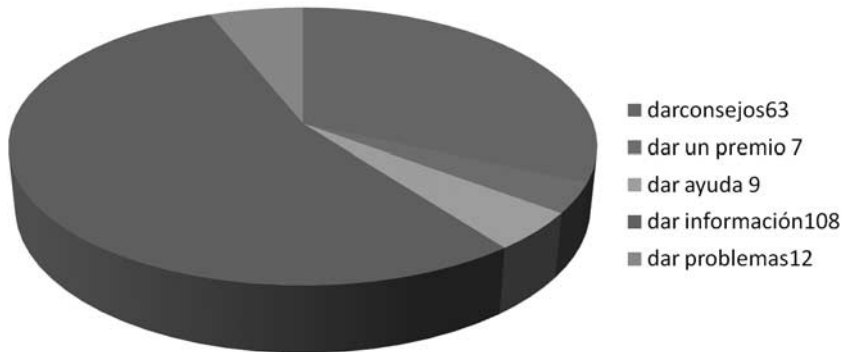
Fig. 3. Support verb constructions with *dar* in CREA.

words with both unpleasant connotations e.g., *dar un gruñido* [“to give a growl”] and positive semantic prosody such as *consejo* [“advice”] or *premio* [“award”]. At a later stage we decided to study the noun to check with concordances with which verbs an item frequently collocates. This second procedure also sheds light on a range of verbs which collocate with the nouns such as *dar/ofrecer consejos* [“give advice”].

There are 462 tokens in total for *respuesta* [“answer”] which coupled with *dar* proposes a solution to different problems. It is the most significant collocation in the Spanish corpus as can be seen in graph 2, followed by *dar información*, *dar consejos*,

dar problemas and *dar ayuda*. The analysis will focus on the three most common collocates, so that the quantitative results of the most common collocates in both languages will be discussed.

QUANTITATIVE FINDINGS



Graph 2.

Looking at the 35 tokens in the technical texts genre, the active voice patterning of *dar respuesta* has a cataphoric function indicating the content of the answer as in the example below:

- (7) *Los empresarios, agricultores y campesinos, deben **dar respuesta** a la demanda de los consumidores sobre***

However, when this construction occurs in the past tense, it has an intertextual function as it makes reference to a previous solution or lack of it as in the following example:

- (8) *¿No le hablaron? EVA. Sí, señor, pero no **dio respuesta**.*

When we looked at this noun in the plural, unlike the tokens in the singular form, we realized that it patterned quite differently. Whereas the singular form does not collocate with adjectives, the plural shows a tendency to collocate with evaluative adjectives such as *immediate*, *clear*, *satisfactory* (cf. Thompson & Hunston, 2000: 6):

- (9) *Las explicaciones de coyuntura no bastan ya para **dar respuestas satisfactorias**.*

Dar información is also a highly productive collocation in our corpus. Its occurrence is similar to English and is used overwhelmingly in the news genre (61 tokens out of 108). This finding is not very surprising as *información* is admittedly a keyword in the press genre. The remaining 47 tokens for *dar información* are scattered in different text types, predominantly in those with a strong ideational component: Politics and Economics (40 tokens), Science and Technology (12 tokens), etc.

Dar consejos contributes a notion of ‘abstract transfer’ as *give advice* in English, while in *dar un gruñido*, *dar* incorporates a nuance of ‘emission’. It is scattered in different genres and is related to the interpersonal function in oral texts, for example:

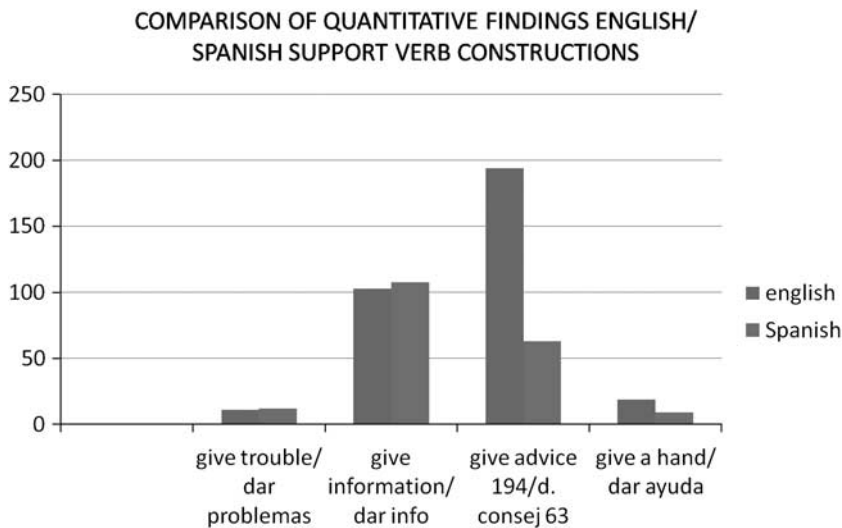
- (10) *Bueno, a mí no me gusta dar consejos a nadie, y menos en público, ¿no?*

It should also be noted that *dar consejos* falls into the category of ‘proposing/recommending a solution’ to different problems as well as *dar/ofrecer ayuda* [‘give help’], which also appears in CREA with 9 and 14 tokens respectively.

Finally, seven SVC constructions with support verbs other than *dar* have been found. The figures of the full lexical verb have been included because they clearly indicate how Spanish speakers and writers strongly favour the use of the full lexical verbs with the notable exception of the first collocate, *hacer público* in 1 and *soltar una carcajada* in 5, which is preferred instead of the full lexical form:

1. *Hacer público* (163)/*publicar* (2510) [‘give an airing to’]
2. *prestarle ayuda* (13)/*ayudar* (6442) [‘give help’]
3. *prestar atención* (434)/*atender* (5150) [‘give/pay attention’]
4. *poner una inyección* (12)/*inyectar* (299) [‘to give an injection’]
5. *soltar una carcajada* (17)/*carcajearse* (11) [‘to give a laugh’]
6. *echar/lanzar una mirada a alguien* [‘to give a look’].

A final examination of graph 3, comparing the quantitative results, shows a similar use of the collocations *give trouble*, *give information* and *give a hand* in both languages. The only marked quantitative difference is *give advice*, with 194 tokens in English versus 63 tokens in Spanish. This difference may be due to the different compilation techniques in both corpora. Despite the fact that both have a similar number of words, the CREA corpus is strongly biased in favour of written texts, with just 10% of oral texts. As this particular collocation is more related to oral speech in both languages, this difference would be easily balanced if the oral component in both corpora were the same.



Graph 3.

4 Conclusions

From a strictly phraseological viewpoint, the present paper has tried to offer an overview of the synchronic usage of multiword units in corpus data in English and

Spanish, focusing on support verb constructions with *give+noun* collocations and showing that they are inherent to language use. The collocational patterns which we have studied are related to what is typically said, rather than what can be said, but they admit lexical, diatopical, register and tense variation. From a quantitative point of view, we expected to find more support verb constructions with *give* in English, as it is a frequent de-lexicalized verb according to the *Collins Cobuild English Grammar* (Sinclair, 1990: 147). Nonetheless, its figures are fairly low with the exception of *give advice* and *give information*. The Spanish data show that the use of the *dar + noun* construction is significantly higher than in English. On the other hand, from a qualitative point of view, this bilingual comparison exercise of support verb collocations renders interesting insights about certain coincident phenomena in both languages. First of all, these support verb constructions are related mainly to abstract nouns, although it is also feasible to find some constructions with concrete nouns in Spanish. Secondly, the use of these clusters tends to portray a factual view of reality (*give information/dar información*). Thirdly, nouns play an important role in both English and Spanish collocations because they select the verb and its syntactic demands. This verb selection materializes in fixed expressions like *hacer una excursión* [“to go on an outing”] since we use the verb *dar* (literally, “to give”) with the noun *paseo* in the support verb collocation *dar un paseo* [“to go for a walk”].

Considering the overall CALL implications of a study like this, it seems to be clear that English as a second language (ESL) teachers should draw attention to these contrasts if they want their students to use collocations as native speakers do. Obviously, this requires explicit training in the use of collocations, through the use of corpora in class, and of de-lexicalized collocations in particular – in accordance with to *CEF* standards – to raise both an awareness of ideational, interpersonal and textual functions, and also of authorial and distributional factors at play in each genre. An electronic dictionary like DUFIE may be an invaluable resource in this respect and, more importantly, the analysis herein conducted may provide insights into the benefits and possibilities of CALL applications like this for language learning education in general, and collocational knowledge and proficiency in particular. McCarthy (1990: 12) stresses that “the relationship of *collocation* is fundamental in the study of vocabulary”, and thereby is a major area of concern for learners of EFL, which converts bilingual dictionaries of multi-word expressions into an invaluable resource for both learners of EFL and translators. Assuming that “all fluent and appropriate language use requires collocational knowledge” (Nation, 2001: 318), electronic dictionaries like this may become a fundamental instrument for CALL-oriented educational practices.

The pedagogical implications of a study like this are clear. Indeed, the study herein conducted is just an example of the possibilities for foreign language education based on CALL resources. As it is, it is extremely important that students grasp not only the conventional grammar but also these support verb constructions (obviously with *give* but also with other verbs such as *make*, *take*, etc.) in connection with syntax, semantics, pragmatics and each respective culture. Furthermore, awareness of the collocational points of contact and divergence should be raised among teachers and students in order to facilitate acquisition by underscoring the potential lexical, genre and register differences. Our teaching experience over the years has proved that

students learn English more naturally if encouraged to use and build collocation repertoires in the classroom. An electronic dictionary like this may be most useful in this respect. As substantiated by the case study herein presented, the use of corpora in class offers great potential for language teaching in general and for collocational knowledge in particular, which seems to be consistent with current trends in corpus linguistics applications of language teaching and learning (cf. Gabrielatos, 2005; Aijmer, 2009).

Apart from the examples studied in this paper, the DUFIE dictionary also offers standard and typical phraseology of other types, simultaneously available to language learners and translators with an on-line dictionary as a reference database identifying and translating multiword units in both English and Spanish. Although our primary data has been selected from corpora, the Internet is also playing an increasingly crucial role in sorting out the most recurrent and widely used grammatical and lexical collocations.

References

- Aijmer, K. (ed.) (2009) *Corpora and language teaching*. Amsterdam: John Benjamins.
- Alba-Salas, J. (2002) *Light verb constructions in Romance: a syntactic analysis*. Ph.D. thesis, Cornell University.
- Allerton, D. J. (2002) *Stretched verb constructions in English*. London/New York: Routledge.
- Bahns, J. (1993) Lexical collocations: a contrastive view. *ELT Journal*, 47(1): 56–63.
- Bank of English. <http://www.collins.co.uk/books.aspx?group=153>
- Beatty, K. (2003) Computers in the language classroom. In: Nunan, D. (ed.) *Practical English language teaching*. Boston: McGraw Hill, 247–266.
- Beckett, G. H. and Millers, P. C. (2006) *Project-based second and foreign language education: past, present and future*. Greenwich, CT: Information Age Publishing.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman grammar of spoken and written English*. London: Longman.
- Blake, R. (2001) What language professionals need to know about technology. *ADFL Bulletin*, 32(3) [*Chairing the foreign language and literature department, part 2*]: 93–99.
- BNC (British National Corpus, Sampler, XML version) (2005) [Distributed by Oxford University Computing Services on behalf of the BNC Consortium]. <http://www.natcorp.ox.ac.uk/>
- Campoy, M. C. (2004) Computer mediated lexicography: an insight into online dictionaries. In: Campoy, M. C. and Safont, P. (eds.) *Computer-mediated lexicography in the foreign language learning context*. Castelló de la Plana: Universitat Jaume I, 47–72.
- Chapelle, C. A. (2001) *Computer applications in second language acquisition: foundations for teaching, testing and research*. Cambridge: Cambridge University Press.
- Cowie, A. and Mackin, R. (1993) *Oxford dictionary of phrasal verbs*. Oxford: Oxford University Press.
- Cowie, A. (2004) Phraseology – The Hornby legacy. In: Williams, G. and Vessier, S. (eds.) *Proceedings of the eleventh EURALEX International Congress*, Vol. 1. Lorient: Université de Bretagne-Sud, 37–52.
- CREA (*Corpus de referencia del español actual*). <http://www.rae.es>.
- Davies, G. D. (2002) ICT and modern foreign languages: learning opportunities and training needs. *International Journal of English Studies*, 2(1) monograph issue.
- de Schryver, M. (2003) Lexicographers' dreams in the electronic-dictionary. *International Journal of Lexicography*, 16(2): 143–199.

- Egbert, J. L. (2005) Conducting research on CALL. In: Egbert, J. L. and Petrie, G. M. (eds.) *CALL research perspectives*. Mahwah, NJ: Lawrence Erlbaum, 3–8.
- Gabrielatos, C. (2005) Corpora and language teaching: just a fling, or wedding bells? *TESL-EJ*, 8(4) [A1]: 1–37.
- Godwin-Jones, R. (2000) Emerging technologies: Literacies and technology tools/trends. *Language Learning & Technology*, 4(2): 11–18.
- Godwin-Jones, R. (2005) Messaging, gaming, peer-to-peer sharing: language learning strategies & tools for the millennial generation. *Language Learning and Teaching*, 9(1): 17–22.
- Gries, S. (2008) Phraseology and linguistic theory: a brief survey. In: Granger, S. and Meunier, F. (eds.) *Phraseology. An interdisciplinary perspective*. Amsterdam/Philadelphia: John Benjamins, 3–26.
- Hanson-Smith, E. (ed.) (2000) *Technology-enhanced learning environments*. Alexandria, VA: TESOL.
- Hanson-Smith, E. (2004) Computer-assisted language learning. In: Carter, R. and Nunan, D. (eds.) *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press, 107–113.
- Hornero, A. M., Luzón, M. J. and Murillo, S. (eds.) (2006) *Corpus linguistics. Applications for the study of English*. Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien: Peter Lang.
- Jeong-Bae, S. (ed.) (2004) *Computer-assisted language learning: concepts, contexts and practices*. New York: APACALL.
- Jespersen, O. (1942) *A modern English grammar on historical principles, VI*. Copenhagen: Ejnar Munsgaard/London: Allen & Unwin.
- Johnson, K. and Johnson, H. (1999) *Encyclopedic dictionary of applied linguistics*. Oxford: Blackwell.
- Kearns, K. (2002) *Light verbs in English*. <http://www.ling.canterbury.ac.nz/kate/lightverbs.pdf>
- Kern, R. (2006) Perspectives on technology in learning and teaching languages. *TESOL-Quarterly*, 40: 183–210.
- Koike, K. (2001) *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Alcalá de Henares: Universidad de Alcalá.
- Levy, M. (1997) *Computer-assisted language learning: context and conceptualization*. Oxford: Clarendon Press.
- López, M. and Campoy, M. C. (2003) User guides in computer-mediated dictionaries. In: Posteguillo, S., Ortells, E., Prado, J. R., Bolaños, A. and Alcina, A. (eds.) *Internet in linguistics, translation and literary studies*. Castelló de la Plana: Universitat Jaume I Publicacions, 439–452.
- Luzón, M. J. (1999) Dictionary websites: transforming and adding value to the dictionary. In: Campoy, M. C and Safont, P. (eds.) *Computer-mediated lexicography in the foreign language learning context*. Castelló de la Plana: Universitat Jaume I, 73–94.
- McCarthy, M. (1990) *Vocabulary*. Oxford: Oxford University Press.
- Mel'čuk, I. (1993) La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère. *Études de linguistique appliquée*, 92: 82–113.
- Mills, D. (2000) Enthusiasm, experience and collaboration: technology in the IEI at UIUC. In: Hanson-Smith, E. (ed.) *Technology-enhanced learning environments*. Alexandria, VA: TESOL.
- Moliner, M. (2000) *Diccionario de uso del español*. Madrid: Gredos.
- Moon, R. (1998) *Fixed expressions and idioms in English: a corpus-based approach*. Oxford: Oxford University Press.
- Nation, I. S. P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

- Nattinger, J. and de Carrico, J. (1992) *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Oster, U., Ruiz Madrid, M. N. and Sanz Gil, M. (eds.) (2006) *Towards the Integration of the ICT in Language Learning and Teaching: Reflection and Experience*. Castelló de la Plana: Servei de Publicacions de la Universitat Jaume I.
- Ruiz-Madrid, N. (2007) Integration of the ICT in language learning. In: Usó-Juan, E. and Ruiz-Madrid, N. (eds.) *Pedagogical reflections on learning languages in instructed settings*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 62–93.
- Sinclair, J. (1990) (Editor-in-Chief) *Collins COBUILD English Grammar*. London: Collins.
- Sinclair, J. (1997) Corpus evidence in language description. In: Wichmann, A., Fligelstone, S., McEnery, T. and Knowles, G. (eds.) *Teaching and Language Corpora*. London/New York: Longman, 27–39.
- Sobkowiak, W. (1999) *Pronunciation in EFL machine-readable dictionaries*. Poznan: Motivex.
- Smith, D. and Baber, E. (eds.) (2005) *Teaching English with information technology*. London: Modern English Publishing Ltd.
- Sokolik, M. (2001) Computers in language teaching. In: Celce-Murcia, M. (ed.) *Teaching English as a foreign or second language*. Boston, MA: Heinle & Heinle – Thomson Learning, 477–488.
- Szendeffy, J. de (2005) *A practical guide to using computers in language teaching*. Ann Arbor: University of Michigan Press.
- Thompson, G. and Hunston, S. (2000) Evaluation: an introduction. In: Hunston, S. and Thompson, G. (eds.) *Evaluation in text. Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Tirkkonen-Condit, S. (2002) Process research: state of the art and where to go next? In: Tirkkonen-Condit, S. (ed.) *Across languages and cultures* 3(1): 5–19 [Joensuu: University of Joensuu, Savonlinna School of Translation Studies].
- Ueda, H. (1989) *Estudio cuantitativo del léxico español*. Tokyo: Publicaciones del Dpto. de Idiomas Extranjeros, Univ. de Tokio.
- Warschauer, M. (2001) On-line communication. In: Carter, R. and Nunan, D. (eds.) *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press, 207–212.
- Warschauer, M. and Kern, R. (2000) *Network-based language teaching: computers and practice*. New York: Cambridge University Press.
- Wierzbicka, A. (1982) Why can you have a drink when you can't *have an eat? *Language*, 58(4): 753–799.
- Wilkins, D. A. (1976) *Notional syllabuses: A taxonomy and its relevance to foreign language curriculum*. Oxford: Oxford University Press.
- Wray, A. (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.