
Remarks on the Elaboration of an English–Spanish Word-combination Dictionary

Eduardo de Gregorio-Godeo, *Universidad de Castilla-La Mancha, Ciudad Real, Spain* (eduardo.gregorio@uclm.es)

Abstract: Through a focus on word-combination analysis in bilingual dictionaries, this article presents the epistemological and methodological background for an ongoing research and editorial project aiming to produce an English–Spanish dictionary of multi-word combinations. After a discussion of the treatment of word combinations as a phenomenon in different types of dictionaries, the lexicographic process guiding the elaboration of the dictionary is described. In addition to the preliminary plan and the fieldwork undertaken for this dictionary-making project, a detailed account is given of the principles determining the inclusion of entries and their presentation within the dictionary. The typology of word combinations included in the dictionary and the structure of entries is accordingly dealt with prior to making some remarks regarding the final presentation-and-revision stage currently being carried out. Because this article is based on current research intending to compile an English–Spanish dictionary of multi-word expressions, emphasis is constantly laid on a usage context where the source language is English and the target language Spanish. By considering the potential users of the English–Spanish word-combination dictionary examined here, some concluding remarks are made with regard to the educational implications of this kind of dictionary primarily aimed at intermediate- to advanced-level Spanish-speaking EFL learners.

Keywords: MULTI-WORD EXPRESSIONS, MULTI-WORD COMBINATIONS, COLLOCATIONS, IDIOMATIC EXPRESSIONS, ENGLISH–SPANISH BILINGUAL DICTIONARIES, LEXICOGRAPHY, PHRASEOLOGICAL UNITS

Opsomming: Opmerkings oor die bewerking van 'n Engels–Spaanse woordeboek van woordverbindings. Deur op die ontleding van woordverbindings in tweetalige woordeboeke te fokus, gee hierdie artikel die epistemologiese en metodologiese agtergrond vir 'n voortgesette navorsings- en redigeringsprojek met as doel die totstandbrenging van 'n Engels–Spaanse woordeboek van meerwoordige verbindinge. Na 'n bespreking van die behandeling van woordverbindings as 'n verskynsel in verskillende woordeboeksoorte, word die leksikografiese proses beskryf wat die bewerking van die woordeboek rig. Benewens die voorlopige plan en die veldwerk gedoen vir hierdie woordeboeksamestellingsprojek, word 'n uitvoerige verslag gegee van die beginsels wat die insluiting van inskrywings en hul aanbieding in die woordeboek bepaal. Die tipologie van die woordverbindings ingesluit in die woordeboek en die struktuur van die inskrywings word dus ook behandel voor daar 'n aantal opmerkings gemaak word oor die finale aanbieding- en hersieningsstadium wat tans voltooi word. Omdat hierdie artikel gebaseer is op huidige navorsing met die doel om 'n Engels–Spaanse woordeboek van meerwoordige uitdruk-

kings saam te stel, word die klem gedurig geplaas op 'n gebruikerskonteks waar die brontaal Engels en die doeltaal Spaans is. Deur inagneming van die potensiële gebruikers van die Engels-Spaanse woordeboek van woordverbindings wat hier bespreek word, word 'n aantal slot-opmerkings gemaak met betrekking tot die opvoedkundige implikasies van hierdie soort woordeboek wat primêr gerig is op middelbare- tot gevorderdevlak- Spaanssprekende EVT-aanleerders.

Sleutelwoorde: MEERWOORDIGE UITDRUKKINGS, MEERWOORDIGE VERBINDINGS, KOLLOKASIES, IDIOMATIESE UITDRUKKINGS, ENGELS-SPAANSE TWEETALIGE WOORDEBOEKE, LEKSIKOGRAFIE, FRASEOLOGIESE EENHEDE

1. Introduction

Taking word-combination analysis in bilingual dictionaries as a case in point, this article presents the epistemological and methodological background underlying a research and editorial project aiming to produce an English-Spanish dictionary of word combinations. After a discussion of the treatment of word combinations as a phenomenon in different types of dictionaries, the lexicographic process guiding the elaboration of the dictionary is described. An account is accordingly made of the principles determining the inclusion of entries and their presentation within the dictionary. Given that this article is based on current research aiming to build an English-Spanish dictionary of multi-word expressions, emphasis will be laid at all times on a usage context where the source language is English and the target language is Spanish.¹ After the potential users of the English-Spanish word-combination dictionary here presented have been considered, some final remarks are made about the educational implications of such a dictionary primarily aimed at intermediate- to advanced-level Spanish-speaking EFL learners.

2. Word combinations and lexicographic analyses: a focus on collocations and idioms

Within the domain of lexicology, phraseology is progressively passing through a certain process of segregation as a separate branch of linguistics. While lexicology has to do with words and their meanings, phraseology focuses on such collocations of words (i.e. 'phraseologisms', 'phraseological units', 'collocations' or 'idioms'), where the meaning of the whole is different from the sum of the literal meanings of the words, comprising a phraseological unit. 'Phraseological units', which are the object of study of phraseology, happen to be stable word-groups with partially or fully transferred meanings. According to Gläser (1998: 125), a 'phraseological unit' may be conceived of as a lexicalized, reproducible, billexemic or polylexemic word group in common use, which has relative syntactic and semantic stability, may be idiomatized, may carry connotations, and may have an emphatic or intensifying function in a text.

As major types of phraseological units, 'collocation' and 'idiom' are familiar terms for users of English monolingual and bilingual dictionaries having achieved a certain linguistic proficiency. Such linguistic phenomena are often conceived of as 'multi-word combinations' (Ilson 2002: 333) whose meaning is more than the sum of the meaning of their components. As Benson, Benson and Ilson maintain in their approach to this lexical phenomenon, 'word combinations' come into being when "certain words regularly combine with certain other words or grammatical constructions" (1997: ix). To a great extent, when exploring the word combinations of a language, both collocations and idiomatic expressions are at some point examined.²

The notion of 'collocation' refers "to the tendency for certain words to occur together. The term itself comes from the verb *collocate*, meaning 'to go together'" (Finch 2000: 152).³ It is common practice for monolingual dictionaries to include information about units 'above' the word level, such as "units including more than one complete word, i.e. compounds and idioms like *black-bird, bank on, give up, night owl, hammer and tongs, at all, kick the bucket*" (Ilson 2002: 333). As substantiated by Ilson's investigation of lexicographic practices, most dictionaries generally incorporate syntagmatic information about the use of items in forming sentences, which may include notes on complementation and, which is the case in point here, information on "collocation with specific words or types of words ([e.g.] *fond of vs fondness for; [or] the association of capsize with boats or ships*)" (2002: 335).

According to Bussmann, collocations are characteristic word combinations "which have developed an *idiomatic* semantic relation based on their frequent co-occurrence" (1996: 81; emphasis added). Consequently, collocations are semantically and syntactically close to such lexico-grammatical units as 'idioms', that is, "a sequence of words which is semantically and often syntactically restricted, so that they function as a single unit. From a semantic viewpoint, the meanings of the individual words cannot be summed to produce the meaning of the idiomatic expression as a whole" (Crystal 2003: 225-226).

Idiomatic expressions are semantically related to collocations to such an extent that, as Crystal underlines, "an alternative terminology refers to idioms as 'habitual collocations'" (2003: 226). In fact, collocations and idioms may be taken to partake of the same stock of prefabricated units which, following the Russian tradition of phraseology, various authors have labelled 'word combinations' and also 'phraseological units' or 'phrasal lexemes' to delineate the "ready-made memorized combinations in written and spoken language" (Cowie 1998: 1), comprising both "'word-like' units, which function syntactically at or below the level of the simple sentence, and 'sentence-like' units, which function pragmatically as sayings, catchphrases, and conversational formulae" (Cowie, 1998: 4).⁴ Studies on the phraseology of English like that of Altenberg (1998: 120 *et passim*) likewise acknowledge idioms and collocations to be types of word-combinations. Indeed, Cowie himself highlights that "collocations of words in familiar literal senses are at one end of a broad spectrum

of word combinations in English. At the other are idioms: combinations whose constant re-use in a fixed form has led to a radical change of meaning" (1988: 131).

3. A project in context: English–Spanish word-combination lexicographic analysis

As Cowie asserted at the end of the last decade, there is now a "wider recognition of the crucial part that ready-made memorized word combinations "play in first- and second-language acquisition and adult language production [...] native-like proficiency in a language depends crucially on a stock of prefabricated units, 'prefabs', varying in complexity and internal stability (1998:1). Textbooks for EFL learners often examine collocations and idioms as fundamental aspects of vocabulary learning. This is also the trend in specific vocabulary-learning manuals for students of EFL (e.g. Thomas 1991; Redman 1997; McCarthy 2001, 2002). There are various monolingual dictionaries of idioms for EFL learners. A few examples of a long list that would be impossible to include exhaustively here are the *Longman Dictionary of Idioms* (Long 1979); the *Oxford Learner's Dictionary of English Idioms* (Warren 1994); and the *Collins COBUILD Dictionary of Idioms* (1995). Monolingual dictionaries of phrasal verbs as a special category of idiomatic expressions are also widely available on the market, for instance, the *Collins COBUILD Dictionary of Phrasal Verbs* (1989), including a workbook by Goodale (1993); the *Oxford Dictionary of Phrasal Verbs* (Cowie 1993) or the *Longman Phrasal Verbs Dictionary* (2000).⁵ There are likewise various monolingual dictionaries of collocations, including *Selected English Collocations* (Kozłowska 1993); *A Dictionary of English Collocations* (Kjellmer 1994); *The BBI Dictionary of English Word Combinations* (Benson, Benson and Ilson 1997); and the *Oxford Collocations Dictionary* (2004).

Current monolingual dictionaries for learners of English similarly incorporate information on collocations and idioms. Again, an exhaustive list could not possibly be included here, some well-known examples being the *Longman Language Activator* (1993); the *Longman Dictionary of Contemporary English* (1995); the *Oxford Advanced Learner's Dictionary* (Hornby 2000); and the *Collins COBUILD New Student's Dictionary* (2002). However, bilingual sources of English word combinations for Spanish learners of English in the form of user-friendly dictionaries are rather limited. Ordinary English–Spanish/Spanish–English dictionaries often include information on idioms and, to a lesser extent, collocations. Representative cases are the *Gran diccionario español–inglés, English–Spanish* (García-Pelayo y Gross 1993); the *Cambridge Word Selector, Inglés–Español* (1995); the *Diccionario Oxford avanzado para estudiantes de inglés, español–inglés inglés–español* (1996); *El Diccionario Oxford español–inglés, inglés–español* (2001); the *Diccionario de inglés contemporáneo para estudiantes: English–Spanish, español–inglés* (Sánchez Benedito and Gámez Gámez 2001); and the *Collins Dictionary: Español–Inglés, English–Spanish* (2003).

Although there exist English–Spanish dictionaries of phrasal verbs as a fundamental subcategory of idiomatic expressions like that of Khalaili and Marina (1984), comprehensive bilingual dictionaries of idiomatic expressions and collocations are almost non-existent. Thus, exceptions like the phraseological dictionaries of Carbonell Basset (1971, 1995, 1996) tend to lay a strong emphasis on proverbs and other idiomatic expressions, so that, in addition to disregarding a great deal of English idioms which are not proverbs, they do not on the whole take collocations into account. As substantiated by this overview of the literature on English–Spanish word-combination analysis, a thorough lexicographic examination of English–Spanish word combinations needs undertaking. Such a dictionary would be particularly useful for EFL learners whose mother language is Spanish. In fact, it is somewhat surprising that such a project should not have been contemplated to date, considering that word-combination dictionaries have already been produced which explore English phraseology from the viewpoint of languages with considerably fewer speakers than Spanish, for example, Polish (Osuchowska 2001).

4. Steps towards an English–Spanish dictionary of word combinations

Bearing in mind the above-mentioned lack of lexicographic resources providing translations of English word combinations into Spanish, we proceed to describe the approach followed for the elaboration of an English–Spanish dictionary of such phraseological units. Admittedly, this procedural account may serve as a basis for further lexicographic work exploring English multi-word combinations from the viewpoint of other target languages. The following subsections present the lexicographic process undertaken for the research and editing process leading to an English–Spanish dictionary of word combinations.

4.1 Lexicographic process

The project is consistent with the basic principles of contemporary lexicographic practices, thereby comprising all four stages of successful lexicographic work as mentioned by Hartmann (2001: 14–20), namely 'planning', 'fieldwork', 'description' and 'presentation.' The last stage includes final revision as well.

4.1.1 Preliminary plan

Hartmann stresses that "to be successful, lexicography as dictionary making requires careful planning and implementation of the compilation process on the basis of market research and the specification of the potential users' reference needs to be met" (2001: 20). As discussed above, an examination of the current availability of bilingual English–Spanish dictionaries gives evidence of

a significant lack of specific dictionaries dealing with word combinations. Therefore, a dictionary of this kind including collocations and idioms seemed to be a must, and, given the specificity of word combinations in every language, it seemed that the project would be beneficial for EFL learners as well as translators.⁶ After some preliminary work had been carried out through some local university- and regionally-funded research, the positive attitude of the Spanish Department of Education in granting funding on a national competitive basis confirmed the expected positive results of the project.⁷

4.1.2 Fieldwork

Further to the initial planning stage of the dictionary-making process, the first step in the research is to produce a thorough database of word combinations in English that might be used as a basis or 'macrostructure' for the alphabetically-ordered dictionary 'entries' in English, including Spanish equivalents, usage examples and other information within the English-Spanish dictionary of word combinations.⁸ The macrostructure of lexical entries has thus been completed on the basis of existing English monolingual dictionaries of multi-word expressions including *The BBI Dictionary of English Word Combinations* (Benson, Benson and Ilson 1997) or the *Oxford Collocations Dictionary* (2004), plus various other current English-Spanish dictionaries such as *El Diccionario Oxford español-inglés, inglés-español* (2001), and general-use monolingual English dictionaries like the *Oxford Advanced Learner's Dictionary* (Hornby 2000). Specialised dictionaries of English idioms, such as the *Collins COBUILD Dictionary of Idioms* (1995), have been consulted as well to build a macrostructure of word combinations in English consisting of collocations and idioms. Rather than come up with a new repertoire of English multi-word expressions defined in English, the main focus of the project was to provide Spanish equivalents for word-combination compilations in English, fieldwork therefore mainly relying on 'secondary sources'.⁹

The BNC (British National Corpus) (2001) has been adopted as a key source of information for the usage examples of the entries in the dictionary. Examples are easily retrieved through SARA, a PC-based concordance programme available on the commercial CD-based version of the BNC. The BNC has been chosen on grounds of its relevance for its present world-wide English-related lexicographic analyses: "the BNC has been used for the dictionaries of Oxford University Press, Longman and Chambers, the three publishers who contributed to its compilation" (Leech 2002: 91).¹⁰ However, when usage examples are absent in the BNC, other English-language corpora are consulted, for example, the *Collins COBUILD English Collocations on CD-ROM* (1995). The Internet is drawn upon as a last resource for usage contexts when no examples are found within existing language corpora.

Again, equivalents in Spanish for the word combinations in the macrostructure of the dictionary are provided on the basis of existing bilingual Eng-

lish–Spanish general-use dictionaries. A wide range of such bilingual dictionaries have been consulted in this regard. The list includes, but is not limited to, dictionaries like the *Diccionario Oxford avanzado para estudiantes de inglés, español–inglés inglés–español* (1996); the *Gran diccionario español–inglés, english–spanish* (García-Pelayo y Gross 1993); *El Diccionario Oxford español–inglés, inglés–español* (2001); the *Diccionario de inglés contemporáneo para estudiantes: English–Spanish, español–inglés* (Sánchez Benedito and Gámez Gámez 2001); the *Collins Dictionary: Español–Inglés, English–Spanish* (2003); or the *Cambridge Word Selector, Inglés–Español* (1995). When necessary, the CREA (Corpus de referencia del español actual, i.e. Reference Corpus of Contemporary Spanish) has been consistently used to test the appropriateness and idiomaticity of the equivalents in Spanish.

4.1.3 Description

4.1.3.1 Word combinations included in the dictionary

For the purposes of this dictionary, we have followed Benson, Benson and Ilson's (1986: 252–254) overall taxonomy of 'lexical combinations', namely: (a) 'free combinations' of words, which are those whose "components are the freest in regard to combining with other lexical items"; (b) 'idioms', which are "relatively frozen expressions whose meanings do not reflect the meanings of their component parts"; and (c) 'collocations', which are "fixed"/"recurrent" word combinations, that is, "loosely fixed combinations" between free word combinations and idioms.¹¹ Nonetheless, free combinations of words have not been considered in the dictionary, since their meaning may be found by examining the meaning of their individual constituent words in general dictionaries.¹²

4.1.3.1.1 Collocations

Benson, Benson and Ilson (1997) distinguish between 'grammatical collocations' and 'lexical collocations'.¹³ A 'grammatical collocation' is "a phrase consisting of a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or clause" (1997: xv); for instance, *decide on* (*decidirse por*) in *decide on a boat*. In contrast, 'lexical collocations' "do not contain prepositions, infinitives, or clauses. Typical lexical collocations consist of nouns, adjectives, verbs and adverbs" (1997: xxx); for example, *warmest regards* (*saludos afectuosos*) in *I send warmest regards*. The dictionary is mainly concerned with lexical collocations, which Spanish-speaking users of English might easily convert into wrong collocations (e.g. **deserted children/abandoned children* [*niños abandonados*]). Moreover, although there is a large number of English collocations following the Verb + Noun structure (*abandon hope*), this lexicographic repertoire does not include all possible free word combinations. In their discussion of English word combinations, Benson, Benson and Ilson (1997: xxx) make

this point clear by considering the construction *condemn murder* (*condenar delitos*), which is a free word combination in English, the verb *condemn* combining with an unlimited number of nouns (e.g. *condemn abortion, abuse of power*) and *murder* similarly combining with hundreds of verbs (e.g. *accept, acclaim, advocate*); however, *commit murder* (*cometer un crimen*) is a collocation. Following Benson, Benson and Ilson's (1997: xxx-xxxiii) taxonomy of collocations, the main types of lexical collocations included in the dictionary are the following:

- (a) **Verb + Noun / Pronoun / Prepositional Phrase**
Most of these collocations consist of a verb of action or activation plus a noun or a pronoun (e.g. *come to a conclusion, make an appointment, set a record* [*llegar a una conclusión, concertar una cita, fijar un récord*]). In many cases, the collocation incorporates an idea of eradication (e.g. *override a veto* [*anular el veto*]).
- (b) **Adjective + Noun**
Typical examples include *strong/weak tea* [*té fuerte/flojo*], which may not be replaced by *might/feeble tea*. Sometimes it is possible for various adjectives to collocate with a single noun. In these cases, only the most frequent lexical collocations are included. Furthermore, it is necessary to bear in mind that in English many nouns have an adjectival function when they are placed before another noun (e.g. *placement test* [*test de nivel*]). Such collocations appear in the dictionary entry of the second noun. Nevertheless, if the meaning of the second noun within the word combination is different from its meaning when used independently, the word combination is included in the dictionary on the basis of the second word (e.g. *acceptance speech* [*discurso de ingreso en una institución pública*]).
- (c) **Noun + Verb**
These are collocations where the verb reflects characteristic actions of the noun, be it a person or a thing (e.g. *bombs explode/go off* [*las bombas explotan*]). Combinations which are easily predictable are not considered in the dictionary (e.g. *dancers dance, teachers teach, writers write* [*los bailarines bailan, los profesores enseñan, los escritores escriben*]).
- (d) **Lexical collocations indicating the 'unit' commonly associated with a noun**
Typically, their structure in English is *noun₁ of noun₂*. Such collocations project the meaning that an individual belongs to a larger group (e.g. *a pack of wolves, a swarm of bees* [*una banda de ladrones, un enjambre de abejas*]), or the specific and concrete character of a unit with regard to larger elements (e.g. *a bit of advice, an act of violence* [*un consejo, un acto violento*]).
- (e) **Adverb + Adjective**
Some examples of this type of collocation are *utterly abhorrent; patent-*

ly/totally/utterly absurd [totalmente aborrecible; completamente, absolutamente absurdo].

(f) Verb + Adverb

Here typical examples could be *appreciate deeply, greatly, keenly, sincerely, very much; argue bitterly, heatedly, passionately, strenuously, vehemently* [agradecer sumamente, muchísimo; discutir apasionadamente, acaloradamente, vigorosamente].

4.1.3.1.2 Idiomatic expressions

The dictionary takes account of a wide range of idiomatic expressions of contemporary British and American English. Idiomatic expressions fall within different categories such as: (a) traditional idioms (e.g. *spill the beans* [irse de la lengua]); (b) new phrases (e.g. *it's all gone pear-shaped* [ha salido fatal]); (c) metaphorical phrases (e.g. *face the music* [afrentar las consecuencias]); (d) two-word phrases (e.g. *wild card* [comodín]); and (e) various other similes (e.g. *like two peas in a pod* [ser como dos gotas de agua]).¹⁴ The dictionary does not incorporate 'phrasal verbs' as a characteristic type of multi-word combination because there are various dictionaries of phrasal verbs, both monolingual and bilingual, on the market.¹⁵

4.1.3.2 Structure of entries

The dictionary is articulated in accordance with the typical procedure of alphabetical order, which "is based on the written form of the lexically relevant units rather than on their meaning" and adopts a 'semasiological' approach in entries, that is, "going from name to notion" (Ilson 2002: 291). The basic principles guiding the organisation of entries within the dictionary are as follows:¹⁶

- (a) Key headwords, compounds included, are alphabetically ordered in the dictionary.¹⁷ Single-word compounds precede those written as two words. Homographs follow this order: adjective, adverb, noun, verb. For instance, *tease* (n.) precedes *tease* (v.). Determiners, prepositions and pronouns are not often headwords.

TEASE I n.

[person who teases] a terrible **tease** *un bromista*.

TEASE II v.

1. **to tease a person cruelly** *burlarse o reírse de una persona*. By then she had acquired a distinctive Geordie accent and she was upset when her friends at school teased her about her rounded vowels and up and down, sing-song voice.

2. **to tease an animal** *provocar a un animal*. He would meddle with fishing nets, pull up anchors and tow boats, tease dogs and tow swimmers.

3. **to tease a fabric; wool** *cardar un tejido; lana*. As a test of his strength Utnapishtim challenged him to stay awake for six days and seven nights —; But while Gilgamesh sat there resting on his haunches, a mist of sleep like soft wool teased from the fleece drifted over him (...).

- (b) Entries contain at least one context of usage. The key headword is written in small capital letters. English collocations are written in bold and Spanish equivalents appear in italics. Regular font face has been adopted for usage examples, the word combination in question being underlined. Here is an example:

MACHETE n.

to brandish, wield a machete *blandir un machete*. We can't take time to talk to the stylist before they wield the machete.

- (c) Cross-references within the dictionary are highlighted in yellow. For example, in the following entry *penny* is highlighted in yellow because the same idiom may be found in the entry for *penny* within the dictionary:

TEN n.

1. **to be two/ ten a penny** *ser baratísimo / no valer nada*. Uncritical testimonials to the postmodern's attractions are ten a penny, and conservative denunciations thereof not much scarcer.

2. **ten to one** (inform.) *diez a uno / te apuesto lo que quieras*. The great bulk of those who thronged Emmett place last night wanted to see the film, and they outnumbered protestors by a minimum of ten to one.

- (d) When a word collocates with others, the dictionary pinpoints series of collocations in an alphabetical order (e.g. **ABILITY: to demonstrate, display, exhibit, show**). This helps not only to save space, but also to display synonyms and near-synonyms. So, in the series for *ability*, *demonstrate ability* and *display ability* are treated as synonyms. However, collocations which are not synonyms are separated by a semicolon (;) (e.g. **domestic; physical; sexual ABUSE**). Synonyms are thus grouped together and separated by commas within the series of collocations. As shown in the example below, **to operate, run, use, work; shut down a machine**, *to operate, run, use* and *work* are treated as synonyms, and are likewise separated from *shut down* (which is not a synonym) by a semicolon:

MACHINE n.

1. **to operate, run, use, work; shut down a machine** *usar, utilizar; apagar una máquina*. A knowledge of how to set up and operate a machine could reduce the likelihood of a breakdown, so minimising the time spent on repairs.

- (e) Collocations whose meaning may be difficult to grasp by the student are often defined between brackets on the left of the collocation:

MARKET n.

[...]

5. **a bear ("falling"); bull ("rising") market** *mercado/bolsa a la baja; en alza*. Spicer points out that the sector has thrived in a bull market but the advertising boom may be coming to an end.

Information about register, dialects and other social factors is also indicated between brackets, e.g. formal (form.), American English (IAm), British English (IBr); etc:

MARINE n.

1. **a mercantile (IBr), merchant (IAm) marine** *marino mercante*. It is true that certain measures of nationalisation had been undertaken in the first months of Soviet government — for example, the Merchant Marine had been taken over in January 1918 (...).

- (f) As the following example of an entry from the dictionary instantiates, idiomatic expressions are written in blue at the end of entries:

MAT n.

1. **to weave a mat** *sacudir una alfombra/un felpudo*. The Lele are subsistence cultivators, growing maize, ground-nuts, and raffia palms from whose fronds they weave mats which are used as a special currency — a rare case of money really growing on trees.

2. **a bath; exercise; prayer; welcome mat** *alfombra de baño; para hacer ejercicio/colchoneta; para arrodillarse y rezar; de bienvenida*. For elderly people, grab rails at strategic places (including the bathroom/toilet) and use of a non-slip bath mat are devices likely to prevent accidents.

3. **to go to the mat (for sb)** (IAm) *hacer todo lo que puedes para solucionar un problema difícil, para ayudar a algn que tiene menos poder que tú*.

- (g) Examples have been taken from the BNC and, to a lesser extent, from the *Collins COBUILD English Collocations on CD-ROM* (1995). Although this is the case on very few occasions, the dictionary sometimes includes collocations without usage examples as no examples have been found in the language corpora utilized. However, an equivalent is always provided in Spanish, if the word combination is believed to be problematic for Spanish speakers:

AIR CONDITIONER n.

[...]

3. **to run; turn on an air conditioner; turn off an air conditioner** *poner/encender; apagar el aire acondicionado*.

- (h) Alternative translations of a collocation or idiom are separated by a slash (/):

MAD adj.

[...]

5. mad cow disease *enfermedad de las vacas locas/encefalopatía espongiiforme bovina*. Excuses abound: world markets have collapsed, diet-conscious Europeans are eating less red meat, some people in Britain fear it will give them mad cow disease.

- (i) Usage notes are highlighted in light blue, and range from pragmatic information to differences between British and American English, through other grammatical questions. They are also used to make reference to 'false friends', namely "terms in two languages which are phonologically and graphologically similar (cognates), but have subtly different meanings" (Malmkjær 2002: 82). Here is an example:

TAPE n.

[...]

9. audiotape; videotape *cinta de audio; de vídeo*. The screen replaces the car windscreen, and on it is shown a road scene that has been recorded on videotape, and transferred to disk.

Nota de uso: se pueden encontrar los compuestos juntos (videotape) o separados (video tape).

- (j) The use of a long underscore (____) in an idiomatic phrase indicates that various nouns, adjectives or verbs may be inserted in the gap:

MAN n.

[...]

33. our man in _____ *nuestro hombre en _____*. I don't know why our man in Madrid came to see me; just idle curiosity, perhaps — so few people have ever met me, so many seem anxious to do so.

4.1.4 Presentation and Revision

At this stage of the editing project, the dictionary on the whole is close to completion. The final stage of the project still needs to be undertaken. This will entail formatting, printing and proofreading the whole dictionary several times. A thorough revision of the dictionary will accordingly have to be carried out prior to its final publication.

5. Final remarks

Through the presentation of a research project attempting to produce an English–Spanish lexicographic repertoire of word-combinations using examples from the British National Corpus, this article has tried to shed light on the vital role of corpus-based language analysis for dictionary-making practices, chiefly as regards "bilingual dictionaries that have English as their source language"

(Benson 1985: 61). Focusing on collocations and idioms as fundamental multi-word-combination categories in English, this article has tried to disentangle the lexicographic mechanisms contributing to developing systematically-organised lexical repertoires of equivalents in Spanish.

This project is intended to compensate for the shortage of bilingual dictionaries providing instant and user-friendly access to Spanish translations of English word combinations. The fact is, as McCarthy stresses, that "the relationship of *collocation* is fundamental in the study of vocabulary" (1990: 12), and therefore is a major area of concern for learners of EFL. In a similar way, "idioms are a great source of difficulty for foreign learners" (Taylor 1990: 49), so that bilingual dictionaries of multi-word expressions are invaluable resources for both Spanish learners of EFL and translators. Dictionaries like this may thus help to facilitate and increase Spanish speakers' cultural awareness of English, since "all fluent and appropriate language use requires collocational knowledge" (Nation 2001: 318).

By way of conclusion, it could be said that an exercise like this may serve as a model for similar research and editing projects aiming to build word-combination bilingual dictionaries dealing with other languages. The production of a Spanish–English counterpart of this lexicographic repertoire is now the logical continuation of the present project.

Endnotes

1. According to Richards and Schmidt's definition of these terms when applied to bilingual dictionaries, the 'source language' is "the language out of which a translation is made" (2002: 496) and the 'target language' is "the language into which a translation is made" (2002: 539).
2. There is some degree of terminological confusion regarding word combinations and their subclasses. As Mel'čuk (1998: 23-24) evidences, the terms 'collocation' and 'idiom' — or 'idiomatic expression' — have sometimes been used to refer not only to a subclass of the word combinations of a language, but as synonyms of notions like 'word combination' or 'fixed (frozen) phrases'. His use of the term 'set phrases' (1998: 23ff), or 'phrasemes' as a synonym of word combinations to refer to phrases which are *not free*, is quite illuminating in this respect. Together with the term 'word combination', the lexicographic literature on phraseology recurrently draws upon the terms 'word-combination' (e.g. Howarth 1996), 'multi-word combination' (e.g. Ilson 2002) and 'multi-word expression' (e.g. Hartmann and James 1998) to refer to "a phrase consisting of two or more words functioning as a single lexeme. The constituents are relatively stable (fixed expression), and, if used idiomatically, their combined meaning is more or other than the sum of the parts, e.g. *fly-by-night*, *face the music*" (Hartmann and James 1998: 97).
3. The notion of 'collocation' was first introduced by Firth (1957) in his semantic theory to evoke the phenomenon of word combinations giving rise to semantic and idiomatic relations based on their co-occurrence.
4. See Cowie (1998) and Mel'čuk (1988) for detailed and complex terminological discussions outside the scope of this contribution.

5. Most of them are being constantly updated with new editions on the market.
6. This dictionary might accordingly be included within the category of dictionaries for the foreign learner, which are "essentially general-purpose dictionaries, but tailored to the needs of a specific group of users" (Jackson 1988: 174).
7. Preliminary contacts with a number of publishing firms similarly encouraged us to pursue such a project.
8. A dictionary 'macrostructure' is "a succession of articles, [that is,] entries, so ordered that any article may be found through an explicitly stable search procedure, an algorithm" (Ilson 2002: 291). An 'entry' is "the basic reference unit in a dictionary" (Hartmann and James 1998: 50). Entries consist of (a) a 'lemma' allowing compilers to locate and users to find entries in the word-list, and (b) a 'comment'. Comments may focus on 'formal' aspects (e.g. spelling, pronunciation, grammar) and semantic information (e.g. definition, usage, equivalents in the L2). As Hartmann and James add, "in case of multiple meanings of the lemma, the entry is subdivided into (usually numbered or otherwise marked) sections called 'sub-entries' or 'sub-senses', each of which provides the same basic information categories" (1998: 50).
9. During the fieldwork stage of dictionary making, 'primary' sources have to do with original material, whereas 'secondary' sources concern derivative material, chiefly other dictionaries (Svensén 1993).
10. As Summers states, "the British National Corpus is a collaborative initiative carried out by Oxford University Press, Longman, Chambers Harrap, Oxford University Computer Services, Lancaster University's Unit for Computer Research in the English Language (headed by Professor Geoffrey Leech), and the British Library. The project received funding from the UK Department of Trade and Industry and the Science and Engineering Research Council, and was supported by grants from the British Academy and the British Library" (1996: 266).
11. Benson, Benson and Ilson (1986: 252-254) also distinguish intermediate categories like 'transitional collocations', which fall between idioms and collocations, and 'compounds', that is, lexical elements of more than one word. Such categories are not contemplated here to avoid excessive complexity. The terms 'idiomatic expression' instead of 'idiom', and 'restricted collocations' instead of 'collocation', are sometimes found in the literature on word combinations.
12. A 'free combination of words' (Cowie 1981: 226) is a combination of words following only the general rules of syntax: the elements are not bound specifically to each other and so they occur with other lexical items freely; that is, the meanings of the words combine compositionally and may be substituted by synonyms (e.g. *run a business*, where *run* may be substituted by *manage*, in the same way as *a business* may be substituted by *a bank*).
13. The authors admit that terms like 'recurrent combination' or 'fixed combination' are sometimes employed instead of 'collocation' (1997: 15).
14. Cf. Long (1979: ix-x).
15. A 'phrasal verb' is "a unit in English which is formed from a verb with the addition of a preposition or adverb which can variously precede or follow an object: e.g. *take up*" (Matthews 1997: 279).
16. For further examples of entries in the dictionary, a final appendix has been included at the end of this article with the first entries of the letter *m* in the dictionary.
17. A 'headword' is the typographically canonical form of a word or a phrase which is chosen for the lemma, the position in the dictionary structure where the entry starts (Hartmann and James 1998: 68).

References

- Altenberg, B.** 1998. On the Phraseology of Spoken English: The Evidence of Recurrent Word-Combinations. Cowie, A.P. (Ed.). 1998: 101-122.
- Benson, M.** 1985. Collocations and Idioms. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 61-68. ELT Documents 120. Oxford: Pergamon Press.
- Benson, M., E. Benson and R. Ilson.** 1986. *Lexicographic Description of English*. Amsterdam/Philadelphia: John Benjamins.
- Benson, M., E. Benson and R. Ilson.** 1997. *The BBI Dictionary of English Word Combinations*. Second Edition. Amsterdam/Philadelphia: John Benjamins.
- BNC.** 1995–. *The British National Corpus*. Oxford: Oxford University Computing Services.
- Bussmann, H. (Ed.).** 1996. *Routledge Dictionary of Language and Linguistics*. London/New York: Routledge.
- Cambridge Word Selector, Inglés–Español: Diccionario Temático del Inglés Contemporáneo.* 1995. Cambridge: Cambridge University Press.
- Carbonell Basset, D.** 1971. *Diccionario de modismos inglés–español*. Madrid: Dos Continentes.
- Carbonell Basset, D.** 1995. *Diccionario fraseológico inglés–castellano, castellano–inglés, frases, expresiones, modismos, dichos, locuciones, idiomatismos, refranes, etc. A Phraseological Dictionary, English–Spanish, Spanish–English, Phrases, Expressions, Locutions, Clichés, Idioms, Sayings, Saws, Proverbs, etc.* Barcelona: Ediciones Serbal.
- Carbonell Basset, D.** 1996. *Dictionary of Proverbs: Sayings, Saws, Adages, English and Spanish — Diccionario de refranes: proverbios, adagios, castellano e inglés*. Barcelona: Ediciones del Serbal.
- Collins COBUILD Dictionary of Phrasal Verbs.* Sinclair, J. and R. Moon (Eds.). 1989. Glasgow: HarperCollins.
- Collins COBUILD Dictionary of Idioms.* 1995. Sinclair, J. and R. Moon (Eds.). Glasgow: HarperCollins.
- Collins COBUILD English Collocations on CD-ROM.* 1995. London: HarperCollins.
- Collins COBUILD New Student's Dictionary.* Sinclair, J. (Ed.). 2002. London: HarperCollins.
- Collins Dictionary: Español–Inglés, English–Spanish.* 2003. Barcelona: Grijalbo / London: HarperCollins.
- Corpus de referencia del español actual (CREA). Real Academia Española, Banco de datos [online]. <http://www.rae.es> [26 March 2010].
- Cowie, A.P.** 1981. The Treatment of Collocations and Idioms in Learners' Dictionaries. *Applied Linguistics* 2(3): 223-235.
- Cowie, A.P.** 1988. Stable and Creative Aspects of Vocabulary Use. Carter, R. and M. McCarthy (Eds.). 1988. *Vocabulary and Language Teaching*: 126-139. London/New York: Longman.
- Cowie, A.P.** 1993. *Oxford Dictionary of Phrasal Verbs*. Oxford: Oxford University Press.
- Cowie, A.P.** 1998. Introduction. Cowie, A.P. (Ed.). 1998: 1-20.
- Cowie, A.P. (Ed.).** 1998. *Phraseology. Theory, Analysis, and Applications*. Oxford: Clarendon Press.
- Crystal, D.** 2003. *A Dictionary of Linguistics and Phonetics*. Fifth Edition. Oxford: Blackwell.
- Diccionario Oxford avanzado para estudiantes de inglés, español–inglés inglés–español.* 1996. Oxford: Oxford University Press.
- El Diccionario Oxford español–inglés, inglés–español — The Oxford Spanish Dictionary Spanish–English, English–Spanish.* 2001. Oxford: Oxford University Press.
- Finch, G.** 2000. *Linguistic Terms and Concepts*. London: Macmillan.

- Firth, J.R.** 1957. Modes of Meaning. Firth, J.R. (Ed.). *Papers on Linguistics, 1934–1951*: 190-215. London/New York: Oxford University Press.
- García-Pelayo y Gross, R.** 1993. *Gran diccionario español-inglés, English-Spanish*. Barcelona: Larousse.
- Gläser, R.** 1998. The Stylistic Potential of Phraseological Units in the Light of Genre Analysis. Cowie, A.P. (Ed.). 1998: 125-143.
- Goodale, M.** 1993. *Collins COBUILD Dictionary of Phrasal Verbs Workbook*. Glasgow: HarperCollins.
- Hartmann, R.R.K.** 2001. *Teaching and Researching Lexicography*. Applied Linguistics in Action Series. Harlow: Longman.
- Hartmann, R.R.K. and G. James.** 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Hornby, A.S.** 2000. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
- Howarth, P.A.** 1996. *Phraseology in English Academic Writing: Some Implications for Language Learning and Dictionary Making*. Lexicographica Series Maior 75. Tübingen: Max Niemeyer.
- Ilsón, R.F.** 2002. Lexicography. Malmkjær, K. and J.M. Anderson (Eds.). 2002: 331-339.
- Jackson, H.** 1988. *Words and their Meaning*. London: Longman.
- Khalaili, K. and B. Marina.** 1984. *English Phrasal Verbs in Spanish*. London: Edward Arnold.
- Kjellmer, G.** 1994. *A Dictionary of English Collocations*. Oxford: Clarendon Press.
- Kozłowska, C.D.** 1993. *Selected English Collocations*. Warsaw: Wydawn Nauk PWN.
- Leech, G.N.** 2002. Corpora. Malmkjær, K. and J.M. Anderson (Eds.). 2002: 84-93.
- Long, T.H. (Ed.).** 1979. *Longman Dictionary of Idioms*. Harlow: Longman.
- Longman Dictionary of Contemporary English*. Summers, D. (Ed.). 1995. London: Longman.
- Longman Dictionary of Phrasal Verbs*. Courtney, R. (Ed.). 2000. Harlow: Longman.
- Longman Language Activator: The World's First Production*. Summers, D. (Ed.). 1993. London: Longman.
- Malmkjær, K.** 2002. Contrastive Linguistics and Cross-Linguistic Studies. Malmkjær, K. and J.M. Anderson (Eds.). 2002: 82-84.
- Malmkjær, K. and J.M. Anderson (Eds.).** 2002. *The Linguistics Encyclopedia*. Second Edition. London/New York: Routledge.
- Matthews, P.H.** 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- McCarthy, M.** 1990. *Vocabulary*. Oxford: Oxford University Press.
- McCarthy, M.** 2001. *English Vocabulary in Use: Upper Intermediate*. Cambridge: Cambridge University Press.
- McCarthy, M.** 2002. *English Vocabulary in Use: Advanced*. Cambridge: Cambridge University Press.
- Mel'čuk, I.** 1988. Collocations and Lexical Functions. Cowie, A.P. (Ed.). 1988: 23-53.
- Nation, I.S.P.** 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Osuchowska, D.** 2001. *English at Work: An English-Polish Dictionary of Selective Collocations*. Warsaw: PWN.
- Oxford Collocations Dictionary: For Students of English*. Crowther, J. 2004. Oxford: Oxford University Press.
- Redman, S.** 1997. *English Vocabulary in Use: Pre-Intermediate and Intermediate*. Cambridge: Cambridge University Press.

- Richards, J.C. and R. Schmidt.** 2002. *Longman Dictionary of Language Teaching and Applied Linguistics*. Third Edition. Harlow: Longman.
- Sánchez Benedito, F. and F. Gámez Gámez.** 2001. *Diccionario de inglés contemporáneo para estudiantes: English–Spanish, español–inglés*. Madrid: Pearson.
- Summers, D.** 1996. Computer Lexicography: The Importance of Representativeness in Relation to Frequency. Thomas, J. and M. Short (Eds.). 1996. *Using Corpora for Language Research*: 260-266. Harlow: Longman.
- Svensén, B.** 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- Taylor, L.** 1990. *Teaching and Learning Vocabulary*. New York: Prentice Hall.
- Thomas, B.J.** 1991. *Advanced Vocabulary and Idiom*. London: Nelson.
- Warren, H. (Ed.).** 1994. *Oxford Learner's Dictionary of English Idioms*. Oxford: Oxford University Press.

Appendix: Examples of entries from the letter M in the dictionary

MA'AM

1. Wham bam thank you ma'am (oral) *encuentro sexual que no se vuelve a repetir*. He was, it appears, neither rapacious nor peremptory, no 'for kicks', no 'wham bam, thank you, mam!' — he tried to make sure both of them enjoyed the time; and he was fun.

2. wham bam *algo que ocurre de repente sin preaviso o preparación y que termina rápidamente*. The problem arises when you fall from a screen above and land on 'monkey features': wham bam, every single life lost! — smacked bums all round and general loss of brownie points.

MACE n.

[staff used as a symbol of authority] **a ceremonial mace** *maza ceremonial*. The Liberal Democrats complained that too many schools have outside lavatories; that the Tories were profligate in setting up a chair of maritime history at a local university; and that they had spent too much on a ceremonial mace.

MACHETE n.

to brandish, wield a machete *blandir un machete*. We can't take time to talk to the stylist before they wield the machete.

MACHINE n.

1. to operate, run, use, work; shut down a machine *usar, utilizar; apagar una máquina*. A knowledge of how to set up and operate a machine could reduce the likelihood of a breakdown so minimising the time spent on repairs.

2. an adding, calculating machine *una máquina calculadora*. In this social service, therefore, the institutional element consists not in the great organisation over which John Boyd-Carpenter presides, with its large and efficient staff and its famous calculating machine at Newcastle.

3. an answering; washing machine *un contestador; una lavadora*. They could not afford a washing machine.

4. a cash, money access (IAm); cigarette; slot (IBr), vending machine *un cajero; un máquina expendedora de cigarrillos; máquina expendedora de bebidas*. Film fans can put down 20fr in a slot machine outside the Cinema Museum for a current programme.

5. a composing, linotype, typesetting; copy, copying, duplicating; fax machine *linotipia; fotocopiadora; fax*. The fax machine started chuntering away.

6. a heart-lung; X-ray machine *máquina de circulación extracorpórea; de rayos X*. I'm sorry, sir, one of the guards replied in English, but we have to check your case with the X-ray machine.

7. an earth-moving; milking; milling; sanding; sewing machine *excavadora; ordeñadora; fresadora; pulidora; máquina de coser*. You can bring your own sewing machine and basic sewing equipment

8. a mincing machine (IBr; IAm meat grinder) *picadora de carne*. Miss Paula McCloskey, 26, also injured her other leg when she became entangled in a mincing machine at the M U P factory in Richill, Co Armagh, in June, 1990, the High Court in Belfast was told yesterday.

9. a fruit (IBr), **slot** (IAm); **pinball machine** (IBr pintable) *tragaperras; flipper millón*. There ain't no music playing or nothing like there usually is, just the sound of people talking and the clunking of the fruit machine.

Nota de uso: En IBr se encuentra a veces *pintable* en vez de *pinball machine*.