

Designing an English-Spanish Dictionary of Word Combinations Including Usage Examples from Corpora

Eduardo de Gregorio-Godeo¹

1. Introduction

As maintained by Bussman, a ‘corpus’ may be characterised as “a finite set of concrete linguistic utterances that serves as an empirical basis for linguistic research” (1996: 106). A language corpus is “a collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting-point of linguistic description or as a means of verifying hypotheses about language” (Crystal, 1997: 95). Corpus linguistics may thus be described as “an approach to investigating language structure and use through the analysis of large databases of real language examples stored on computer” (Richards and Schmidt, 2002: 127). By exploring the case word-combination analysis in bilingual dictionaries, this paper presents an ongoing project aiming to produce an English-Spanish dictionary of word combinations. After discussing the role of corpus linguistics for lexicographic analyses, and the approach to word combinations in dictionaries, the lexicographic process followed for the elaboration of the dictionary is herein described. An account is accordingly made of the principles determining the inclusion of entries and their presentation within the dictionary.

2. Corpus Linguistics and its Applications for Lexicographic Analysis

Today corpus linguistics has become a vibrant field of linguistic research. As stressed by McEnery and Wilson, “from being a marginalized approach used largely in English linguistics and specifically in studies of English grammar, corpus linguistics has started to widen its scope [...] corpus linguistics is a lively subject, with corpus-based approaches being taken to many research questions in linguistics” (2001: 1). Corpus-based linguistic analyses have become a fundamental tool for Applied Linguistics on the whole, so that corpora are currently drawn upon across a wide range of domains of linguistic enquiry, so that “corpora, and the study of corpora, have revolutionised the study of language, and the applications of language, over the last few years” (Hunston, 2002: 1). As McEnery and Wilson’s (2001: 104-128) thorough overview reveals, corpus linguistics is fruitfully being employed in numerous language-related disciplines.

Corpora have become a fundamental tool for lexicography, a domain of linguistic enquiry which deals with “the writing of dictionaries, or the scholarly study of this activity” (Trask, 1997: 131) in general, and, in particular, with “the principles that underlie the process of compiling and editing a dictionary” (Jackson, 1988: 247). In fact, “modern dictionaries are now usually based upon huge corpora of English, from which words, forms, spellings, meanings and grammatical behaviour are extracted, thus allowing lexicographers to appeal directly to the observed facts of language” (Trask, 1999: 166).

¹ Universidad de Castilla-La Mancha (Spain), Departamento de Filología Moderna
e-mail: Eduardo.Gregorio@uclm.es

3. Word Combinations in Lexicography: The Case of English-Spanish Lexicographic Analyses

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). By and large, both ‘collocations’ and ‘idioms’ are often conceived of as ‘multi-word combinations’ (Ilson, 2002: 333) whose meaning is more than the sum of the meaning of their components.

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). 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 *blackbird, bank on, give up, night owl, hammer and tongs, at all, kick the bucket*” (Ilson, 2002: 333). Collocations are characteristic word combinations “which have developed an *idiomatic* semantic relation based on their frequent co-occurrence” (Bussmann, 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: 225–26).

Idiomatic expressions are semantically related to collocations to such an extent that, as Crystal underlines, “an alternative terminology refers to idioms as ‘habitual collocations’” (*ibid*, 226). In point of 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 as ‘word combinations’ – and also as ‘phraseological units’ or ‘phrasal lexemes’ – to delineate the “ready-made memorized combinations in written and spoken language” (Cowie, 1998: 1), which comprise 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). Studies on the phraseology of English like Altenberg’s (1998: 120 and *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).

It has to be admitted that 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 Khalaili and Marina’s (1984), comprehensive bilingual dictionaries of idiomatic expressions and collocations are almost non-existent. Thus, exceptions like Carbonell Basset’s (1971, 1995, 1996) phraseological dictionaries 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 take into account collocations on the whole. Therefore, a thorough lexicographic examination of English-Spanish word combinations needs undertaking. Such a dictionary would be particularly useful for learners of English as a Foreign Language whose mother language is Spanish.

4. Producing an English-Spanish Dictionary of Word Combinations

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 echoed by Hartman (2001: 14-20), namely ‘planning’, ‘fieldwork’, ‘description’ and ‘presentation.’ The last stage includes final revision as well.

4.1.1 Preliminary Plan

Hartman 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 market of bilingual English-Spanish dictionaries evidences 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 learners of English as a foreign language and translators alike.

4.1.2 Fieldwork

Further to the initial planning stage of the dictionary-making process, the first step in the research involved producing 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. 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.

Given that, 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 has mainly relied on 'secondary sources', that is, existing dictionaries.

The BNC (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 English-related lexicographic analyses all over the world nowadays: "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). In any case, when usage examples are not found on the BNC, other English-language corpora are consulted, by way of 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 English-Spanish general-use dictionaries. A wide range of such bilingual dictionaries have been consulted to this regard. The list includes, but it 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).

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: (i) 'free combinations' of words, which are those whose "components are the freest in regard to combining with other lexical items"; (ii) 'idioms', which are "relatively frozen expressions whose meanings do not reflect the meanings of their component parts"; and (iii) 'collocations', which are "fixed"/"recurrent" word combinations, that is, "loosely fixed combinations" between free word combinations and idioms. 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.

4.1.3.1.1 Collocations

Benson, Benson and Ilson (1997) make a distinction between 'grammatical collocations' and 'lexical collocations'. 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" (Benson, Benson and Ilson, 1997: xv); for instance, *decide on* (*decidirse por*) in *decide on a boat*. In contrast, 'lexical collocations'

“do not contain prepositions, infinitive or clauses. Typical lexical collocations consist of nouns, adjectives, verbs and adverbs” (*ibid*, 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 convert into wrong collocations very easily (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 (*ibid*, 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 (*ibid*, xxx-xxxiii) taxonomy of collocations, the main types of lexical collocations included in the dictionary are as follows:

(i) 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*]).

(ii) Adjective + Noun

Typical examples include *strong/weak tea* [*té fuert/flojo*], which may not be replaced by *might/feeble tea*. Sometimes it is possible for various adjectives to collocate with a single noun. In those 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*]).

(iii) 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*]).

(iv) Lexical collocations indicating the ‘unit’ commonly associated to 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*]).

(v) Adverb + Adjective

Some examples of this type of collocation are *utterly abhorrent; patently/ totally/ utterly absurd* [*totalmente aborrecible; completamente, absolutamente absurdo*].

(vi) Verb + Adverb

Here typical examples could be *to 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: (i) traditional idioms (e.g. *Spill the beans* [*irse de la lengua*]); (ii) new phrases (e.g. *it's all gone pear-shaped* [*ha salido fatal*]); (iii) metaphorical phrases (e.g. *face the music* [*afrentar las consecuencias*]); (iv) two-word phrases (e.g. *wild card* [*comodín*]); and (v) various other similes (e.g. *like to peas in a pod* [*ser como dos gotas de agua*]). 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, in the market.

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 follows²:

(i) Key headwords – compounds included – are alphabetically ordered in the dictionary. 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 For further examples of entries in the dictionary, a final appendix has been included at the end of this paper with the first entries of letter *m* in the dictionary.

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 (...).

(ii) 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.

(iii) 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.

(iv) 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 in the dictionary, 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 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, in **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.

(v) 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 (...).

(vi) As the following example of 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*.

(vii) 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 includes sometimes collocations without usage examples as no examples have been found in the language corpora utilized. At any rate, an equivalent is provided in Spanish at all times, since 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*.

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

MAD adj.

[...]

5. mad cow disease *enfermedad de las vacas locas/ encelofatía 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.

(ix) 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 video tape, and transferred to disk.

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

(x) 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. Conclusions

Focusing on collocations and idioms as fundamental multi-word-combination categories in English, this paper has disentangled the lexicographic mechanisms contributing to developing systematically-organised lexical repertoires of equivalents in Spanish. In particular, 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 paper has 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). The project means to compensate for the shortage of bilingual dictionaries providing instant and user-friendly access to Spanish translations of English word combinations. As it is, McCarthy stresses that “the relationship of *collocation* is fundamental in the study of vocabulary” (1990: 12), and thereby 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), which converts bilingual dictionaries of multi-word expressions into an invaluable resource 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).

References

- Altenberg, B. (1998) On the Phraseology of Spoken English: The Evidence of Recurrent Word-Combinations, in A.P. Cowie (ed.) *Phraseology. Theory, Analysis, and Applications*, pp. 101–122. Oxford: Clarendon Press.
- Benson, M. (1985) Collocations and Idioms, in R. Ilson (ed.) *Dictionaries, Lexicography and Language Learning (ELT Documents 120)*, pp. 61–68. Oxford: Pergamon Press.
- Benson, M., E. Benson and R. Ilson (1986) *Lexicographic Description of English*. Amsterdam: John Benjamins Publishing Company.
- Benson, M., E. Benson and R. Ilson (1997) [1986] *The BBI Dictionary of English Word Combinations (2nd edition)*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- BNC [The British National Corpus] (2001) Oxford: Oxford University Computing Services.
- Bussmann, H. (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 English Collocations on CD-ROM (1995) London: HarperCollins. *Collins Dictionary: Español-Inglés, English-Spanish (2003)* Barcelona/London: Grijalbo/HarperCollins.
- Cowie, A.P. (1988) Stable and Creative Aspects of Vocabulary Use, in R. Carter and M. McCarthy (eds) *Vocabulary and Language Teaching*, pp. 126–39. London/New York: Longman.
- Cowie, A.P. (1998) Introduction, in A.P. Cowie (ed.) *Phraseology. Theory, Analysis, and Applications*, pp. 1–20. Oxford: Clarendon Press.
- Crystal, D. (1997) [1978] *A Dictionary of Linguistics and Phonetics (4th edition)*. Oxford: Blackwell.
- Crystal, D. (2003) [1978] *A Dictionary of Linguistics and Phonetics (5th 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*. Houndmill, Basingstoke: Macmillan.

- García-Pelayo y Gross, R. (1993) *Gran diccionario español-inglés, English-Spanish*. Barcelona: Larousse.
- Hartman, R.R.K. (2001) *Teaching and Researching Lexicography*. Harlow: Longman.
- Hornby, A.S. (2000) *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
- Hunston, S. (2000) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Ilson, R.F. (2002) [1991] *Lexicography*, in K. Malmkjær and J.M. Anderson (eds) *The Linguistics Encyclopedia* (2nd edition), pp. 331–39. London/New York: Routledge.
- Jackson, H. (1988) *Words and their Meaning*. London: Longman.
- Khalaili, K. and B. Marina (1984) *English Phrasal Verbs in Spanish*. London: Edward Arnold.
- Leech, G.N. (2002) *Corpora*, in K. Malmkjær and J.M. Anderson (eds) *The Linguistics Encyclopedia* (2nd edition), pp. 84–93. London/New York: Routledge.
- Malmkjær, K. (2002) *Contrastive Linguistics and Cross-Linguistic Studies*, in K. Malmkjær and J.M. Anderson (eds) *The Linguistics Encyclopedia* (2nd edition), pp. 82–84. London/New York: Routledge.
- McCarthy, M. (1990) *Vocabulary*. Oxford: Oxford University Press.
- McEnery, T. and A. Wilson (2001) [1996] *Corpus Linguistics* (2nd edition). Edinburgh: Edinburgh University Press.
- Nation, I.S.P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Oxford Collocations Dictionary: for Students of English* (2004) Oxford: Oxford University Press.
- Richards, J.C. and R. Schmidt (2002) [1985] *Longman Dictionary of Language Teaching and Applied Linguistics* (3rd edition). Harlow: Longman.
- Sánchez Benedito, F. and F. Gámez Gámez (2001) *Diccionario de inglés contemporáneo para estudiantes : English-Spanish, español-inglés*. Madrid: Pearson.
- Taylor, L. (1990) *Teaching and Learning Vocabulary*. New York: Prentice Hall.
- Trask, R.L. (1997) *A Student's Dictionary of Language and Linguistics*. London: Arnold.
- Trask, R.L. (1999) *Key Concepts in Language and Linguistics*. London/New York: Routledge.

Appendix: Examples of entries in the dictionary

M

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 cant 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.

The official Collins English-Spanish Dictionary online. Over 100,000 Spanish translations of English words and phrases. With wide-ranging coverage of contemporary Spanish and English, and a wealth of examples and idioms, this authoritative dictionary is ideal for everyone using Spanish and English today. First published in 1971, the Collins Unabridged Spanish Dictionary pioneered a new approach to bilingual dictionary publishing which continues to inspire millions of language users today. The Collins Spanish online dictionary offers you: Comprehensive and up-to-date coverage of today's language. Over 420,000 translations of current Spanish and English. Thousands of useful phrases, idioms and examples. Audio and video pronunciations. Images for hundreds of entries. Notable English language corpora include the following: With a computer, we can now search millions of words in seconds. The search word or phrase is often referred to as the 'node' and concordance lines are usually presented with the node word/phrase in the centre of the line with seven or eight words presented at either side. These are known as Key-Word-in-Context displays (or KWIC concordances). From Corpus to Classroom: Language Use and Language Teaching. Cambridge University Press, 2007). - Corpus data do not only provide illustrative examples, but are a theoretical resource. - Corpus data give essential information for a number of applied areas, like language teaching and language technology (machine translation, speech synthesis etc.). As an English native speaker, I go eeeeeugh. This is blazingly wrong. (We might say "I cannot recall his face.") Today, corpora composed of texts containing hundreds of millions of words are available to the compilers of dictionaries. This is done by means of an Internet crawler that has been specially designed for this purpose by the Dan-. <http://lexikos.journals.ac.za>. Once a number of useful word lists have been found by the Internet crawler, these lists are copied and pasted into a so-called lemma loader (see Figure 1), another tool developed by Ordbogen.com and conceived by Professor Emeritus Henning Bergenholtz from the Centre for Lexicography in Aarhus. The corpus used in this study consisted of Anglicisms collected from the press by the Antena Chilena de Neologismos for a period of seven years (2006-2012). This procedure was done both manually and semi-automatically and data was stored in the database of the Observatori de Neologismos (OBNEO, 2003). Anglicized forms like "pen drive" or "pendrive" do not appear in regular English dictionaries (Cambridge Dictionary, 2013; The Free Dictionary, 2013; Merriam Webster's Dictionary, 2013; WordReference, 2013), although internet definitions can be found for this portable data-storage device. The word "retail", for example, appeared 49 times in the corpus, but was typographically marked (in italics) only three times. Previous studies based on Chilean press also report. Category. 3 Phraseology in legal English-Spanish dictionaries: description When searching for legal equivalents, apart from the wide range of electronic tools available (such as Google or discussion forums), translators continue to use both monolingual and bilingual dictionaries as the first port of call. Monolingual legal dictionaries provide definitions of legal concepts that form a conceptual network of a legal system in one language, whereas bilingual legal dictionaries provide target language equivalents of source language legal concepts (Biel 2008: 27-28).