Collocation dictionary as an elaborate pedagogical tool for Spanish as a foreign language

I. Introduction

After examining English monolingual collocational dictionaries, Nuccorini (2003: 387) concludes that a dictionary of collocations has to be "appropriate to individual needs." The author takes into consideration that very often the intended addressees of a dictionary cannot coincide with the final users, and even the same individual can use the dictionary according to different roles, namely, as a language teacher, translator, etc.

From our viewpoint – trying to specify Nuccorini’s claim – a dictionary of collocations has to satisfy both advanced learners of a foreign language and native speakers. The former need to learn collocations in L2 because either it is not possible to translate them word by word (to pay attention (Eng.), presto ‘to lend’ atención (Sp.), fare ‘to make’ attenzione (It.), kiinnittää ‘to fix’ huomattava (Finn.), die Aufmerksamkeit schecken ‘to give (as a present)’ (Ger.) or there are no symmetrical correspondences (a piece of news (Eng.) vs. una noticia (Sp.); to queue (Eng.) vs. hacer cola (Sp.).) The latter, the native speaker – but also the learner – could need to have quick access to all the ways of expressing a particular meaning through different perspectives according to the participants of the situation designated by the lexical unit (LU). If the LU were analysis (person X’s Analysis of phenomenon Y), the combinations could be: X carries out an A of Y; Y

Begoña Sanromán Vilas* and Margarita Alonso Ramos**
*University of Helsinki, **Universidade da Coruña

Examples of analysis are taken from Mel’čuk (1998: 39). Only one of the verbs used (to carry out) was found in the Oxford Collocations Dictionary, and another one (to undergo) in The BBI Dictionary of English Word Combinations.

The LF used for encoding the "light" or "support" verbs is Oper. We use here the term "empty" verb in a broad sense, the reader can find a more detailed description of those verbs in Alonso Ramos (2004: 85-96).
COLOCATION DICTIONARY AS AN ELABORATE PEDAGOGICAL TOOL

Although our idea of combining a dictionary plus exercises is not a new one – the Oxford Collocations Dictionary (OCD), for example has done it earlier (see also Alonso Ramos 2006b) – DIce, following the idea developed in LAF, is provided in the lexical co-occurrence section, with a description of collocations by means of a gloss (por amor ‘como una muestra de A’ (for love ‘as an act showing feeling’); cobrar cariño ‘empezar a sentir C’ (to receive affection ‘to start to feel A’). Glosses will be proved to be especially useful in the elaboration of pedagogical material. Learners, as in the case with the OCD, will not be forced to use a general monolingual dictionary of Spanish to understand and be able to use the collocation: each group of collocates will be preceded by its gloss or paraphrase.

This paper is organized into four sections. After this introduction, section 2 provides a description of DIce as presented on the web. In section 3 we propose a set of exercises to check as well as to improve the collocational competence of Spanish learners. Section 4 summarizes the paper and draws some conclusions.

2. The Diccionario de colocaciones del español (DIce)

In this section we will describe DIce following the demonstration available in Internet at <http://www.dicesp.com> and <http://dicesp.cesga.es>, where ten emotion nouns are presented. DIce gives the information mainly through the base of the collocation, which is the headword (or lemma). In this sense, DIce is oriented towards language production. So if the user needs to know which verbs combine with admiración ‘admiration’ to express the meaning ‘to feel’, he will find the following collocates: sentir, tener, rendir, deber, dispensar, tratar or profesor (‘to feel, to have’). However, access through collocates is also possible. The user can select one of the previous verbs, for instance rendir, and check which other bases, as well as the combinations of the previous ones, are part of the collocation.

3. Cooccurrence and collocation

As mentioned above, collocations are a combination of words that are used together more than expected by chance, either repeatedly or in a particular context. Collocation consists of collocating words in a particular order and, optionally, to have specific paraphrases. Collocations provide a way to express emotions, and they are used in different sentences.

4. The Diccionario de emociones del español (DIEM)

The Diccionario de emociones del español (DIEM) is a dictionary that contains expressions relating to emotions. The dictionary contains over 1,000 expressions, divided into six categories: admiration, love, respect, fear, surprise, and sadness.

5. The Diccionario de emociones del español (DIEM)

The Diccionario de emociones del español (DIEM) is a dictionary that contains expressions relating to emotions. The dictionary contains over 1,000 expressions, divided into six categories: admiration, love, respect, fear, surprise, and sadness.

6. The Diccionario de emociones del español (DIEM)

The Diccionario de emociones del español (DIEM) is a dictionary that contains expressions relating to emotions. The dictionary contains over 1,000 expressions, divided into six categories: admiration, love, respect, fear, surprise, and sadness.
The purpose of the gloss is to provide hints to the user about how to select the suitable collocate among those presented. From our viewpoint, a gloss is "a brief indication of the meaning of the collocate in connection with the base" (Alonso Ramos 2005: 22). In other words, we can also say that a gloss is a translation of the LF into a natural metalanguage. This semantic information is not found in English collocational dictionaries such as the English BBI Dictionary of English Word Combinations (BBI), the LTP Dictionary of Selected Collocations (LTP) or OCD. These dictionaries order collocations into word classes and within these classes collocates are grouped together according to semantic similarity.

The particular case of OCD is contradictory. Although the intended users of this dictionary are students of English, there is no semantic information that helps the learners to choose a particular collocate. In the introduction, the authors say that the order of the groups of collocates "tries to be as intuitive as possible" (Crowther, Dignen, and Lea 2002: 3). However, a learner may have difficulties wondering what the difference between a great idea and a big idea is, taking into consideration the fact that both adjectives can be seen as synonyms, but in OCD they are presented in different groups. We consider that the use of intuition does not help much if the user is a non-native speaker; for instance, how could he find out that to take a chance does not have exactly the same meaning as to have a chance, but to have a shower means the same as to take a shower? Similarly, Klotz
(2003: 60), in his review of this dictionary, offers some examples of collocates that are grouped together but they are not synonyms (e.g., biting and brilliant as collocates for satire, and delicately and lightly for scented).

The coverage of the lexical co-occurrence in DICE is very extensive. Two randomly chosen LUs presented in the web demonstration, admiración 1 and sospecha 1, contain 92 and 137 related LUs, respectively. BBI has only 21 for admiración and 34 for sospecha. OCD contains 29 for admiración and 57 for sospecha. LTP includes 31 collocates for admiración and 73 for sospecha. Even if a strict comparison between DICE, on the one side, and BBI, OCD, and LTP, on the other, is not possible because we are dealing not only with different languages but also with different concepts of collocation, this can give a rough idea of the exhaustivity DICE is trying to achieve.

Additionally, DICE's demonstration on the web includes an exercise module that contains production- and comprehension-oriented questions. The program gives the percentage of correct answers.

3. Elaboration of pedagogic material from DICE

After the presentation of DICE, we will concentrate on its pedagogic application. As we commented in the introduction, the idea of combining dictionary and exercises is not a new one. OCD, for example, has already included some exercises as “study pages” in the middle of the dictionary. One of the authors of LTP has edited a book, Teaching Collocations (Lewn 2000), where six out of the eleven chapters are written by classroom teachers describing how they introduce collocations and create activities and exercises based on them. The importance of collocations in the teaching and learning of a foreign language has been emphasized by many authors as an essential knowledge required in fluency and nativelike selection (see Nation 2001: 317-318). Within the application field, we should mention two books of English vocabulary exercises, The Words You Need (Rudzka, Channell, Putseys, and Ostyn 1981) and More Words You Need (Rudzka, Channell, Putseys, and Ostyn 1985), including several collocational activities. In the chapter dedicated to emotions (“Do You Love Me?”), Rudzka et al. 1981: 118-137), for instance, there are exercises which focus on collocates such as (1):

1. Find appropriate adjectives to fill in the blanks.
   1. a(n) ... stranger; 2. a(n) ... feeling; 3. a(n) ... grasp, etc. (Rudzka et al. 1981: 133.)

Conversely, some activities ask for bases of the collocations (2)

2. Add appropriate nouns to the following adjectives:
   1. (an) confident; 2. (an) innovative; 3. (an) positive ... (Rudzka et al. 1981: 132.)

One characteristic of the exercises in Rudzka et al. 1981: 132-133) which we share is that they do not only concentrate on syntagmatic relations, but also on the paradigmatic ones. So there are also activities where learners must look for synonyms or make nouns from verbs, and vice versa. However, we noticed that Rudzka et al. make fine-grained distinctions about the meaning of terms belonging to the same semantic fields - using componential analysis - but when they deal with collocations, no semantic support is given to the learner.

In the following, we will describe some of the exercises we propose to teach the use of collocations from the perspective of SFL. Because of the semantic information included in DICE by means of glosses, learners will be able to find the answers to the exercises without using a general dictionary. As DICE is mainly oriented towards production, but can also be used as a comprehension oriented dictionary, we will divide the exercises into two groups depending on the production or comprehension orientation. However, more attention will be paid here to the production-oriented exercises and to verbal collocations.

The first set of exercises will be focused on production activities. One type of exercises is a multiple-choice test. In this case, the learner is provided with the gloss of the collocation and is asked to search for the collocate (3):

3. Rodea la respuesta que mejor se corresponda con la frase ‘causar pánico en varios personas’: ['Circle the answer that best corresponds to the gloss ‘to cause panic in several people’].
   a. espantar el pánico; b. reproducir el pánico; c. sembrar el pánico; d. extender el pánico.

Type (3) exercises are suitable when collocates are specific to one particular base. However, we can create exercises that can be reproduced in several bases or
semantic fields. Exercises (4) and (5) deal with a particular type of verb + noun collocation:

(4) Transforma las siguientes oraciones de manera que se pueda sustituir el verbo en negrita por una expresión equivalente: un verbo de apoyo seguido del nombre que aparece entre paréntesis [Transform the following sentences replacing the verb in bold with an equivalent expression: a light verb followed by the noun in brackets]:

a. **Firma este documento, por favor.** (FIRMA)
   - a. Pedro le _________ a los avances. (Pedro started to feel fear)
   - b. No deberías _________ el rincón por tan poco cosa. (you continue feeling bitter)
   - c. De repente, le _________ el respeto, (I stop feeling respect).
   - [a. tomó, cogió; b. guardar; c. perdi]

b. **Castigarán a los culpables.** (CASTIGO)
   - c. No te preocupes: su vida no _________ (PELIGRO)

(5) Encuentra sinónimos para los verbos que aparecen en cursiva [Find synonyms for the verbs in italics]:

a. **Sentí una gran alegría,** que todavía ahora no podría explicar.
   - a. Vamos a _________ A una persona que está aquí.
   - b. Sufrí una grave dolencia que le mantuvo retenido durante un tiempo.
   - c. Le dio un puntapié a un objeto duro que rodeó unos pasos delante suyo.

In these examples the verb (in this case, the collocate) is a semantically empty (or quasi-empty) verb that can be combined with the noun to verbalize the meaning of the noun. These collocations are codified in terms of LFs as Oper. In particular, in exercise 4 we combine paradigmatic and syntagmatic co-occurrence, which is not separated in real texts. Thus, the learner has to relate the full verb firmar "to sign" to the collocation sacar poder una firma "to put/ write a signature". Both types of information would be available in DICE under the entry of firma "signature". This exercise is also useful because the user has to be aware of the different structural patterns of firmar (X firma el documento Y) and sacar poder (X sacar poder una firma en el documento Y). Exercise (5) offers collocates and asks for synonyms of these collocates. One should note here that when the collocate is an empty meaning, as in (4) and (5), it is impossible to make a gloss or paraphrase. For this reason, our preliminary decision is to introduce a generic collocate into DICE for each semantic field (Alonso Ramos 2006a). In (5) we have offered sentir "to feel", sufrir "to suffer" and dar "to give" that co-occur with feelings, diseases and nouns designating blows, respectively. Other collocates that could replace sentir in (5a) would be tener, experimentar, o llevarse. In (5b), the learner could use tener or padecer, instead of sufrir, and for dar in (5a) there are several synonyms, such as propinar, asustar, averecer, salir, sacudir, hacer, etc.

As said above, DICE contains only emotion nouns. However, we have used nouns like firmar or castigo to propose a model of exercises following DICE's methodology.

Another type of exercise allows practice with phasal verbs, that is to say, the learner must search for collocate verbs denoting "to start to" (LF Incep.), "to continue" (LF Cont.) or "to finish" (LF Fin.). Exercise (6) can serve as an example:

(6) Rellena los huecos valiéndote de las glosas ante paréntesis [Fill in the gaps according to the glosses within brackets]:

   a. Pedro le _________ MIEDO a los avances. (Pedro started to feel fear)
   b. No deberías _________ RECOR por tan poco cosa. (you continue feeling bitter)
   c. De repente, le _________ el respeto. (I stop feeling respect).
   
   [a. tomó; cogió; b. guardar; c. perdi]

Exercise (7) tries to draw attention to the difference between two verbal collocates, namely, to semantically empty verbs and realization verbs. The former type was treated in exercises (4) and (5); the latter one is a value of the LF Real., that is, a collocate with the meaning 'to do what you are supposed to do with the referent of the LU'. For instance, a realization verb for promesa 'promise' is cumplir 'to fulfil' and an anti-realization verb is romper 'to break', while the semantically empty verb is hacer 'to make'. With these notions in mind, we can ask students to do (7):

(7) Completa los huecos con el verbo adecuado [Fill in the gaps with the appropriate verb]:

   Si Pedro le _________ a Ana que irá al cine el sábado, Pedro le _________ promesa a Ana. Si el sábado Pedro va al cine, Pedro _________ su promesa; pero si no va, Pedro _________ su promesa.

As we can observe, (7) also combines paradigmatic (prometer 'to promise') and syntagmatic (hacer una promesa 'to make a promise') co-occurrence. Similar exercises can be done to distinguish collocations such as hacer una invitación 'to make an invitation' and aceptar una invitación 'to accept an invitation', tender una trampa 'to set a trap' and caer en una trampa 'to fall into a trap', tomar una tentación 'to feel temptation' and caer en una tentación 'to yield to temptation'.

A relatively frequent problem for Spanish learners is the use of pronominal verbs. Within the semantic field of emotions, the pronominal/non-pronominal verb contrast is a productive phenomenon that can be exploited. So, on the one hand, many pronominal verbs, for instance alegrarse 'to be happy', are equivalent to a collocation whose collocate is a semantically empty verb, sentir alegría, which takes the experiencer of the emotion as its grammatical subject (LF Oper.). On the other hand, non-pronominal verbs, such as alegrar 'to make happy', correspond to the same type of collocation, causa alegría, but here the grammatical subject is
the cause of the emotion (LF Opese). Thus, both verbs and collocations express the same meaning showing an opposite diathesis, as can be seen in exercise (8):

(8) Construye oraciones equivalentes siguiendo estos pasos: 1) transforma los verbos plenos en colocabiones verbo + nombre de emoción, o viceversa; 2) invierte las funciones de sujeto y objeto preposicional de la oración original [Construct equivalent sentences following these two procedures: 1) transform full verbs into verb + emotion noun collocations, or vice versa; 2) reverse the subject and prepositional object functions in the original sentence]

a. Sentía vergüenza de su cuerpo. ≡ [Su cuerpo le avergonzaba.]
b. La entrevista ha provocado el enfado de la familia real británica.
c. Los comentarios del profesor preocuparon a María.
d. Pedro se asombró con la noticia.

From the point of view of the comprehension activity, we can propose exercises where learners have to choose the most suitable gloss for one particular collocation, such as (9):

(9) Si leemos en el periódico “el gobierno de EEUU tiene fuertes sospechas de que frán ha desarrollado programas de armas bacteriológicas”, entenderemos que las sospechas son fuertes porque:

a. se refieren a un hecho considerado muy muro;
b. el gobierno está muy seguro de que el hecho es cierto;
c. duran mucho tiempo;
d. son compartidas por varias personas.

The correct answer is (9b). In Spanish, when the person who suspects is very sure about the fact, the suspicion (sospecha) is strong (fuerte). If the suspicion refers to a fact which is considered negative, as is the case in (9a), then the suspicion is serious (grave) or terrible (terrible). For a lasting suspicion (9c), the collocate is permanent (permanente) and if the suspicion is shared by many people, then it is extended (extendida) or generalized (generalizada). Another exercise could deal with the difference within a polysemic collocate such as (10):

(10) Busca otro coloquial que exprese el mismo significado que dar in the following oraciones [Search another collocate which expresses the same meaning as dar (lit. ‘to give’) in the following sentences]:

a. Le di a CONPASIÓN del perro y se lo llevó.12
b. Los aviones siempre le dieron MIEDO.

4. Summary and Conclusions

In this paper we have introduced DICE as a useful tool in SFL. We have focused on the semantic information it contains, showing in particular that DICE provides each collocate with a gloss explaining its meaning in basic Spanish within a particular collocation. The presence of this semantic information is the main feature which distinguishes our approach from other existing collocation dictionaries. In contrast to the common practice of simply grouping collocates by part of speech and vague semantic similarity, our approach to provide a gloss makes possible the right selection between collocates which have the same part of speech, but different meanings.

Furthermore, the correspondence between the gloss and the LF opens an interesting avenue to continue further research on the use of LFs as an interlingua. Let us note the important role that LFs can play in different systems of natural language processing, such as a text generation or machine translation, and more particularly, in the line of this article, in the computer assisted language learning systems.

DICE has been conceived as an electronic lexical database. This allows us to provide more information and to implement a flexible means of access, free of space limitations. The electronic medium and the structured information, which is systematised by LFs, facilitate the integration of the dictionary and the exercise module into an on-line language learning environment. Note that the research line started with DICE can be continued along several different strands. Thus, the data in DICE already contain a considerable corpus of collocates which can be used for both theoretical and computational aims; for example, to train a machine-

---

12 Example extracted from DUE.
learning program for automatic recognition of collocations (Wanner et al. 2005). A further line to develop is the possibility of automatic generation of exercises based on the data included.

References

Dictionaries


Other Literature


Begoña Sanromán Vilas and Margarita Alonso Ramos 295


