

John Benjamins Publishing Company



This is a contribution from *Spanish Word Formation and Lexical Creation*.

Edited by José Luis Cifuentes Honrubia and Susana Rodríguez Rosique.

© 2011. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

PART IV

Applications

Lexical collocations and the learning of Spanish as a foreign language

State of the art and future projects

Marta Higuera García

Instituto Cervantes

“La formación de palabras en español constituye, más que un tema, un ámbito temático”¹
(Almela 1999)

This chapter deals with the definition of *collocation* in books used for the teaching of Spanish as a foreign language (SFL). Different definitions of this concept are revised, and its characteristics are deeply discussed in order to distinguish them from free combinations and compounds. Throughout history, the semantic approach has prevailed over the statistic approach. On the one hand, not every recurring combination forms a collocation, and only certain kinds of combinations are considered collocations, as has been pointed out in the two typologies proposed by Corpas (1996) and Koike (2001). On the other hand, the idea of arbitrariness of collocations should be replaced by the concept of lexical selection. This concept determines which lexical classes select predicates, as was done in the *Redes* dictionary.

Keywords: collocations; compounds; dictionaries; Teaching of Spanish as a Foreign Languages

1. Introduction

From a very broad perspective, which considers word formation to be a thematic field, there is room for a reflection such as the one we are going to propose, which starts with the phraseology and the context of teaching vocabulary to non-native speakers and which attempts to base its proposals on perspectives of theoretical Linguistics and Psycholinguistics.

The aim of this paper is, firstly, to summarize the different contributions of Phraseology and Foreign Language Teaching (FLT) about collocations, pointing out

1. [Word formation in Spanish constitutes, rather than a topic, a thematic field].

differences and similarities with free combinations and compounds. We would quickly like to mention that we will be going over the various positions found and that, rather than answers, we will pose questions, since there is quite little consensus regarding the subject of collocations and the discipline that should approach them. Noting the diverging points and those in common between different opinions, as well as defending their importance both in acquisition and in the teaching of vocabulary and deriving didactic consequences are also priority goals in this paper.

Lastly, after this state of the art, we aim to arouse curiosity and motivate future teachers of SFL to continue investigating this topic. For this purpose, we will list new lines of research which, from our point of view, should be approached in the near future, both in lexicographic and didactic projects, before stating that collocations are fully incorporated in SFL teaching.

2. The controversial concept of collocation and its characteristics

Already a decade ago,² Alonso Ramos (1994: 9) highlighted the polysemous character of this term: in some works – mainly those didactically orientated, we might add – it was used to discuss probable or usual combinations of two words; while in other cases – in works on linguistics –, it is applied to constrained combinations in which a lexeme, the base, required the presence of another, the collocative, like in examples such as *actividad febril* [frenetic activity], *lucha encarnizada* [fierce struggle] or *esfuerzo ímprobo* [tremendous effort].

From the point of view of phraseology, an indisputable reference is the work by Corpas (1996), inspired in the definition by Haensch et al. (1982: 251): “...entendemos por colocación aquella propiedad de las lenguas por la que los hablantes tienden a producir ciertas combinaciones de palabras entre una gran cantidad de combinaciones teóricamente posibles”;³ Corpas broadens this definition as follows⁴ (1996: 66):

2. Even in the 80's the term, which as we will see allows various interpretations, still had a very limited use and was barely found outside the field of lexicography. However, in the 90's it began to receive the attention of various studies, as we will see below. Nevertheless, it is rare to find papers in Spanish applying this concept to the didactics of foreign languages (Higuera 2004a). The term has begun to be stabilized as of 1992 with the works of Írsula (1992); Aguilar-Amat (1993); Alonso Ramos (1993); Corpas (1996); Castillo (1998); Koike (2001); Bosque (2001b) and Penadés (2001).

3. [...we understand by collocation the property of languages by which speakers tend to produce certain combinations of words among a great number of theoretically possible combinations].

4. [We also use collocation to refer to the resulting combinations, that is, to the phraseological units formed by lexical units in a syntactic relation which, in themselves, do not constitute

También denominaremos colocación a las combinaciones así resultantes, es decir, a las unidades fraseológicas formadas por dos unidades léxicas en relación sintáctica, que no constituyen por sí mismas, actos de habla ni enunciados; y que, debido a su fijación en la norma, presentan restricciones de combinación establecidas por el uso, generalmente de base semántica: el colocado autónomo semánticamente (la base) no sólo determina la elección del colocativo, sino que, además, selecciona en éste una acepción especial, frecuentemente de carácter abstracto o figurativo.

We must thank Hausmann (1989: 1010) for the precision regarding the different status between the two collocated words;⁵ according to this author, one of them determines the selection of the other. Therefore, he distinguishes between the base, or the word that determines which words it can combine with, and the collocative, which is the determined element.

No less clarifying is the definition by Koike (2001: 76) in his monograph on lexical collocations, where he defines them as “a lexical-semantic link conditioned by syntax”. This author points out three formal and three semantic characteristics. Starting with formal characteristics, we come across frequent co-occurrence, which is one of the most important ones, but it is not limited to collocations, nor does it guarantee that a collocation exists. Sometimes one or various words appear interspersed with others, a concept which has been labeled “collocational span”.⁶ Secondly, combination constraints: collocations are preferable or habitualized combinations that present certain combination constraints, imposed through traditional use. It is

acts of speech or sentences; and which, due to their fixation in the norm, present combination constraints established through use, generally of a semantic base: the semantically autonomous collocated component (the base) not only determines the selection of the collocative, but also selects a special nuance, often of an abstract or figurative character].

5. For this author, the base is the noun in collocations of the noun + adjective, noun + verb and verb + noun kind; it is the verb in the verb + adjective kind; in the case of adjective + adverb, the base is the adjective; and in the noun + preposition + noun pattern, it is one of the two nouns. The base is semantically autonomous, but the collocative is not (Hausmann 1979 and Alonso Ramos 2002a); thus, for example, in order for the adjective *solo* [alone] to mean ‘sin leche’ [without milk] it must co-occur with *café* [coffee] and, for this reason, the collocation is normally described in the entry of the base in works such as the *Diccionario combinatorio del español* (DICE) where, starting from an entry such as *error*, the user can find out how to intensify that word – see Alonso Ramos (2002b) and Sanromán (2003) –, but not the *Redes* dictionary, which could be considered a dictionary of collocatives – see Bosque (2001a); Bosque (2004) and Higuera (2004b) –.

6. The words considered are four before and four after. In order to discover the co-occurrence of two lexical units A and B, we must calculate the frequency of A, the frequency of B, the co-occurrence of A and B, and in what sense they are used. Koike (2001: 27) believes that we can talk of collocations when the percentage of co-occurrence of two lexical units is above 20%.

a case of lexical predilection, with a lower degree of fixation than idioms, since it is possible to replace one of the two words from the collocation with a synonym: *superar/vencer/salvar una dificultad* [to overcome/to cope with/to resolve difficulties]. The last characteristic is formal compositionality. Collocations are compositional and variable both combinationally and morphologically, which allows certain formal flexibilities: component substitution: *violar/transgredir/vulnerar las normas* [to break/violate/disregard the rules]; adjective modification: *daba una explicación satisfactoria* [he gave a satisfactory explanation]; pronominalization, nominalization: *repicar la campana* and *repique de campanas* [to ring out bells and ringing of bells]; and passive transformation: *sea condonada la deuda* [let the debt be cancelled]. Nevertheless, collocations are not always compositional from a semantic point of view; some of them, due to semantic specialization of one of its components, are not so easy to interpret for someone unfamiliar with a given language and culture.

Next, Koike (2001: 28–29) goes on to summarize the three semantic characteristics. Firstly, he talks about the link between two lexemes, which could explain the change in grammatical category that affects some collocations: *luchar encarnizadamente; lucha encarnizada* [to battle fiercely, 'a fierce battle]. Secondly, he mentions the typicality of the relation: collocations express a typical relation between components, and this is perhaps one of the novelties of Koike's characterization. So, we can include as collocations *tocar la guitarra* [to play the guitar] and *rasguear la guitarra* [to strum the guitar] but not *comprar una guitarra* [to buy a guitar] or *guardar la guitarra* [to put away a guitar], since the noun *guitarra* can only establish a typical relation as a musical instrument. This typicality of the relation is normally present in the lexicographical definition (and this may be the reason why collocation and "entourage" have occasionally been mixed up). Lastly, Koike talks about semantic precision. As opposed to idioms, collocations present, for this author, a semantic precision or an unmistakable concept; especially those formed by noun-verb, which are used as definers, that is, they define a simple lexical unit in senses found in entries of the dictionary: for example, in one sense of *rumorear* [to rumour] we find the collocation *circular un rumor* [to spread a rumour].

From a linguistic and phraseological perspective – we are not going to go into the relation between both, the marginality of the latter and its recent boom thanks to cognitive linguistics (see Salvador 1995) –, works by other authors allow us to add other characteristics: they are a psychological unit for the speakers of a language (Benson et al. 1986), they are a case of lexical constraint in which there is directionality, since one lexeme selects the other (Alonso Ramos 1994: 16); they present arbitrary fixation in the norm, caused by repeated use (Hausman 1989; Corpas 1996: 53 and Zuluaga 2002: 106), and they are also characterized by syntactic regularity (Zuluaga 2002: 11) and by being partially compositional, because the base conserves its meaning, but the collocative adopts a special meaning which it only has in the presence of the other

element (Corpas 1996; Castillo Carballo 1998 & Koike 2001). For some authors they are a lexical unit (Lewis 1993; Corpas 1996; Bogaards 2001; Gómez Molina 2004 and Higuera 1997), while for others it is a relation of meaning (McCarthy 1990 & Schmitt 2000), an aspect we will go on later.

3. The *continuum* of idiomaticity and fixation

Despite the fact that the characteristics of collocations seemed to be clear in phraseology, however, it is clear that works on the SLT field, such as those by Lewis (1993 1997 2000) dealt with a much broader concept of collocations: their most notable characteristics were institutionalization and frequency of co-occurrence, but not so much, lexical constraint and directionality or typicality, perhaps for numerical reasons, since there are so few collocations that present all of these features. Consequently, many of the examples used in books for learning foreign languages under the label of collocations would not be considered so by most phraseologists. This broad perspective of the concept of collocation – proposed by Hausmann (1989), who claimed that whenever an idea is expressed with different lexemes in another language it should appear in a collocations dictionary – is what allows us to state that collocations are worked on in activities about topics, such as housework (*limpiar el polvo* [to dust], *hacer/estirar/arreglar las camas* [to make/do/make up the bed], *planchar la ropa* [to iron clothes], *fregar los platos* [to wash the dishes], etc.) or sports or hobbies done in our spare time [*montar a caballo* [horse riding], *jugar al tenis* [to play tennis], *hacer senderismo* [to go trekking], *ir de compras* [to go shopping], etc.].

In order to bring together both opinions and especially to list the kind of lexical units which language teachers should pay attention to, we proposed (Higuera 2004a, 2006) a not discrete classification of lexical units (in line with the work of Ruiz Gurillo⁷ 1998). In these papers we stated that free combinations, prototypical collocations (described above) and those that do not fulfill each and every characteristic (which we will call non-prototypical collocations⁸ or syntagmatic combinations) form a blurred group, with intersections, common areas and intermediate categories. In spite of the fact that prototypical collocations have been the most studied in phraseology, we hold

7. Ruiz Gurillo (1997; 1998) has also talked about a *continuum* in which collocations would take up an intermediate position between idioms, on one hand, and simply frequent or free combinations, on the other. She defines them as units with a certain degree of fixation, that do not present idiomaticity and that are subject to regular formation processes.

8. These only present arbitrary fixation in the norm, syntactic regularity, semantic transparency, co-occurrence frequency and institutionalization; while free combinations have syntactic regularity and semantic transparency.

that there are more non-prototypical ones and they can be more useful for the learner of a foreign language, since they are usually composed of frequent words and might not correspond to the same lexemes when the student translates them to his or her language.

We cannot refrain from mentioning the great number of papers published with the aim of distinguishing collocations from free combinations, compounds (lexical and syntagmatic) and from idioms (González Rey 1998; Martínez 2000; Ruiz Gurillo 2001; Higuera 2004). In this *continuum* of idiomaticity and fixation formed by phraseological units, collocations would take up the intermediate position between idioms, on the one hand, and free combinations or of simply frequent words, on the other.⁹ The consequence of adopting a not discrete view of these phenomena is that the concepts cannot be clearly defined in absolute terms, but rather through the intersection of properties that tend to coincide statistically or probabilistically; that is, they are conceived as parts of a *continuum*, because linguistic categories are complex (characterized by more than one property or feature) and asymmetrical (each unit from a group should not be defined by all the properties or features, but only some of them); therefore, it does not make sense to think that all idioms, for example, share the same features, as pointed out by Ruiz Gurillo (1997:71), just as collocations appear not to.

It is not always easy to distinguish collocations from free combinations of words, a fact that is not at all trivial, which poses one of the main difficulties that lexicographers have to find an answer to. For that reason, we turn to the example proposed by McCarthy (1990), in which he points out the significant differences we can find between the following examples: *hablar de deporte* [to talk about sport], *del trabajo* [about work], *de los precios* [about prices]... as opposed to *hablar de negocios* [to talk business]. In the first group we are looking at free combinations of the verb *hablar* [to talk] with different prepositional objects that refer to the huge number of subjects human beings can talk about. However, the second example has a different character: its meaning is transparent, but it refers to a state of affairs that does not merely refer to talking, but to 'starting a commercial relation, suggesting an idea to do business', etc. We are looking at a collocation because it is transparent and, furthermore, it is the formula any native speaker would use in that specific situation we have just described: to propose an idea, to make an offer..., therefore, it is a lexicalized and frequent unit.

9. Zuluaga (2002: 98) compares collocations with the dawn, free combinations with the day and phraseological units with the night and states that although it is true that we cannot establish rigid boundaries between day and night and less so between the dawn and the night, that does not mean we cannot have a clear concept, or precise intuition, of what day, night and dawn are.

Koike (2001: 30) also provides some criteria to distinguish collocations from free combinations: the latter are not produced with a stable frequency of co-occurrence, since they are subject to the speaker's choice; they have a lower degree of combination constraints (that is, collocations have constraints imposed by linguistic tradition but free combinations do not); they have greater combinatory, morphological and syntactic flexibility than collocations; they do not express a typical relation and, lastly, collocations are easy to memorize and are psychologically salient, as opposed to free combinations.

On the other hand, in our opinion, it is important to distinguish collocations from compounds¹⁰ and idioms. We will only briefly mention the distinction from compounds – due to restrictions of space –. The first thing we note is the different classification made of these elements in the *Gramática descriptiva del español* (Bosque & Demonte 1999) and in the *NGLE* (Real Academia Española 2009). A decade ago, that work divided them into: prototypical lexical compounds and syntagmatic compounds. According to Val, the former are a concatenation of two words to build a new lexical unit, as in *pelirrojo* [redhead], and the latter are atypical compounds or syntagmatic compounds, in which there is a reinterpretation of a syntactic construction, forming a new lexical unit, characterized by its lack of compositionality (see also Bustos Gisbert 1986: 69–177): *ojo de buey* [porthole], *tocino de cielo* [sweet made with egg yolk and sugar], *oro negro*¹¹ [black gold].

Nevertheless, the grammar of the RAE divides compounds in three groups: proper or univerbal compounds (*agridulce* [sweet and sour]); syntagmatic compounds, formed through juxtaposition of words that preserve their graphic and accentual independence, with or without a hyphen (*árabe-israelí* [Arab-Israeli], *teórico-práctico* [theoretical-practical], *problema clave* [key problem], *decreto ley* [order in council]), and syntactic compounds or nominal idioms, which are interpreted as lexical pieces, not as morphological units: *ojo de buey*, *mesa redonda* [round table], *media naranja* [better half]. This discrepancy in nomenclature proves that the subject requires further consideration, that it should be approached in the future from different points of view.

The characteristics compounds and collocations have in common – bearing in mind that neither of these groups is homogenous and, therefore, we can always find exceptions – would be the following: they are a stable combination of lexemes, which we can find isolated in other contexts outside that combination (Piera & Varela 1999: 4371); they are learned as a whole and are shared by a community, that is, they are not produced but reproduced (what Coseriu labelled “repeated discourse”)

10. The interesting paper by Alonso Ramos (2008) shows an in depth insight into the subject.

11. These examples are interpreted by Corpas as nominal idioms and this interpretation prevails in the *Nueva gramática* by the RAE.

(Ferrando 2002); they belong to certain diastatic and diaphasic variants (compounds: *matasanos* [derogatory term for doctor], *picapleitos* [derogatory term for lawyer] and collocations: *pillar un catarro*, *coger un catarro* [to catch/get a cold]); they are formed by the lexical categories that constitute the open classes of words (N, V, Adj., Adv., and Prep.) (Piera & Varela 1999: 4371); there is a syntactical relation between the two lexemes; they share structures¹² and, lastly, they allow rhetorical figures, such as metaphors or metonymy (compounds: *hombre rana* [frogman], *ciudad dormitorio* [dormitory town] and collocations: *profundo dolor* [terrible pain], *profunda aversion* [deep aversion]...).

The main difference between compounds and collocations is that only the former are equivalent to a grammatical category,¹³ while collocations are closer to phrases. As regards semantic aspects, both lexical compounds and collocations are compositional: therefore, the cases in which they can be mistaken are those in which the former do not have sufficient cohesion to be written in one word. Val (1999: 4760) studies the characteristics of prototypical lexical compounds in depth (see also Bustos Gisbert 1986) and we now compare them with collocations in order to find further differences: they present a phonological amalgam of their components with a single main accent in the second component, and they have morphological unity of the whole. There are also other differences (Higuera 1997: 40–41): they behave differently when faced with variation of gender and number; in order to form the plural of a collocation, both words must be in the plural: *chica morena*/**chicas morena*/**chica morenas* [brown-haired girl]. However, as it has been pointed out in different grammars of Spanish, in the formation of the plural of noun compounds the degree of cohesion reached by the components is relevant. When these elements have a high cohesion, they do allow the plural, and form it with the last component; if their cohesion is not complete, either the first or both components form it, hence the coexistence of: *guardias civiles* and *guardiaciviles* [civil guards]. As regards gender, in collocations where there are nouns and adjectives they must all agree in gender, while in compounds one can be feminine and the other masculine, since what matters is the gender of the resulting compound: *la chica morena*/**la chica moreno* as opposed to: *coche (m.) + cama (f.)* → *el coche cama* [sleeping car]. Also, as pointed out by Hernanz and Brucart (1987: 157), the second

12. Lexical compounds can be: nominal (N+N; V+N; V+V; V+Adv.; Adj.+N; N+Adj.); adjectival (Adj.+Adj.; N+Adj.); and verbal (N+V; Adv.+V). Syntagmatic compounds, according to Val, can be of various kinds: with a prepositional phrase (N+Prep.); nominal phrase (N+Adj. or Adj.+N); verbal phrase (VP+N); adjectives (Adj.+Adv.) and other types. Regarding the type of collocations, the classifications of Corpas and Koike are explained further on.

13. Nominal compounds are equivalent to nouns (*sacacorchos* [corkscrew], *bonobús* [bus pass], *telaraña* [cobweb]); adjectival compounds equal adjectives (*labiodental* [labiodental], *sadomasoquista* [somasochist], and verbal compounds are equivalent to verbs (*malcriar* [to spoil], *fotografiar* [to photograph]).

nominal element of a compound cannot be modified by a specifying element or a complement of its own: **coche cama comodísima* [comfortable sleeping car], **pez dos espadas* [two *swordsfish].

An even more complex subject is distinguishing collocations from syntagmatic compounds (in Val's proposal, or syntactical compounds in the *NGLE*). Ferrando (2002:100) also finds a series of differences regarding fixation: compounds present greater fixation and stability (collocations allow variations) and the way of meaning. On the one hand, compounds are more idiomatic (we find they have an unpredictable, added meaning, but in collocations the base preserves its meaning and the collocative presents a specialized meaning) and, furthermore, compounds play a denominational or designational role: syntagmatic compounds are equivalent to a single concept (*oso hormiguero* [anteater]), they refer to a single reality, they act as a single unit of meaning; this does not occur with collocations. This author also underlines three other differences: in collocations, one element selects the other; if we have the lexeme *odio* [hate] and we wish to intensify it, there is the adjective *mortal* [mortal]; if we start with the noun *miedo* [fear], we would probably choose *atroz* [terrible], while in the elements of a compound there is no selection constraint; in the case of the adjective, they differ in the manner of meaning: in N + Adj. collocations, the adjective has an intensifying function (*voluntad férrea* [strong will], *frío glacial* [ice cold], in the syntagmatic compound it specifies (*león marino* [sea lion], *barco pesquero* [fishing boat], *avión comercial* [commercial plane]). Lastly, in collocations there is no obligatory co-presence of its components, sometimes one component can be replaced by a synonym (*hambre canina/voraz* [great/extreme hunger]), but in compounds this is not always the case (*caja fuerte*, pero no **caja robusta* [safe] but not **[robust box]*, *centro comercial* pero no **núcleo comercial* ([shopping centre] but not **[commercial nucleus]*).

The conclusions reached by Ferrando (2002:106) are that collocations and syntagmatic compounds are adjacent phenomena, and that the difference between them is a matter of degree, since compounds present greater fixation and idiomaticity. Thanks to use, a free phrase can become a collocation and this, in turn, can become a syntagmatic compound.

We cannot consider the debate to be over, but we can observe that the *NGLE* incorporates syntagmatic compounds within idioms, as was also proposed by Ferrando, when he stated that they must be studied from phraseology, given that they are halfway between idioms and collocations, which are also included in the content of this discipline. From a didactic point of view, what we have looked at in this section is a minor problem that phraseology or lexicology should solve, since neither foreign language teachers nor students require this metalanguage; they need only learn to separate blocks of words or lexical units that have a unitary meaning and can be contrasted with other lexical units in their mother tongues. This is the strategy that Lewis (1993) labelled "pedagogical chunking" and is one of the main principles of the Lexical Approach for FLT.

4. Brief historical revision

There is no doubt that collocations are an enormously controversial phenomenon – and therefore, interesting – which have been studied from different perspectives and with various objectives that we will try to classify in this section, aiming to derive from this historical revision some didactic implications for their teaching.

4.1 Statistical approach versus semantic approach

Corpas (2001: 49) speaks of two basic approaches in collocation studies: statistical and semantic.¹⁴ The first was born in neo-Firthian surroundings when collocation was identified with combination, so that Sinclair (1991: 70) understood by collocation: “the co-occurrence of two or more words that are found in a short length of text”. Although his works have provided a very important methodology of study (because they used objective criteria),¹⁵ they were severely criticized due to their limitations derived from the direct application of statistics.

The semantic trend, however, starts off with combinations, but applies another series of criteria to identify collocations, and allows us to see two aspects. On the one hand, we find authors that hold that the base (the semantically independent element) selects its collocative (the element whose meaning is restricted by the base). This theory refers to the works by Hausmann, and has had a huge influence on lexicography and phraseology (Corpas, Alonso Ramos, Zuluaga, Koike, etc.). Collocations, therefore, present internal semantic links of various kinds: specialization and delexicalization or metaphorization of collocated elements; and they reflect a typical and true relation held between these elements in the real world which provide semantic precision

14. However, González Rey (2002:157–164) divides the consolidation of the concept of collocation within the language in general in three phases: broad conception of the concept, typical in the combinatorial method of English, since for its followers a collocation is any phrase formed by two or more words that co-appear with certain frequency in discourse (for authors such as Cowie (1981), Benson (1989) and Sinclair (1991)). These works were reproduced in dictionaries containing multilexical constructions of the English language, aimed specifically at learners of that language. The limited conception of this term corresponds to the combinatorial method of French and is defined by Melčuk, and is also the position adopted by the German-French author Hausmann, who, as we have already mentioned, created a theory applicable to lexicography, didactics and translation. In Spanish combinatory, the tradition of Hausmann has been inherited and reinterpreted through Corpas, so it follows that limited conception of the term.

15. The aim of his studies was to create dictionaries based on textual corpora and they culminated in the publication in 1987 of the *Collins COBUILD English Language Dictionary*.

(that is, collocations express an unequivocal concept for native speakers in a quick and economical way). Also within this line are the works by Alonso Ramos (1993) on lexical functions in the Meaning-Text Theory, that list the internal semantic links of collocations. On the other hand, there are the works by Bosque, who talks of lexical selection of predicates towards their arguments, as we will see below.

This dichotomy between the statistical approach and the semantic approach has been resolved clearly in favour of the latter. Proof of the fact that not any recurring combination forms a collocation is that there is a closed typology of lexical collocations. We will now go on to revise the classifications of Corpas and Koike to emphasize this idea.

Corpas' classification (1996:66–76) from her *Manual de fraseología* takes into account different criteria: grammatical category, syntactic relation between collocated elements and lexical functions of the Meaning-Text Theory (see Alonso Ramos 1993). According to the first two criteria mentioned, she distinguishes six types of collocations¹⁶

16. Some specifications on the different types:

- a. Noun (subject) + verb. In this kind of collocation the verb denotes an action typical of the person or thing designated by the noun.
- b. Verb + noun (object). The verbs in these collocations lead to three groups: collocations that share collocatives and a base belonging to the same semantic field (*desempeñar un cargo/ un papel/ una función...* [to carry out a job/a role/ a function]; collocations whose bases present a very limited combinatorial pattern (*conciliar el sueño/*atraer el sueño* [to get to sleep/to attract sleep]; *acariciar una idea/*tocar una idea* [to caress an idea/*to touch an idea])...; and intermediate cases in which there is a delexicalized verb (normally polysemous verbs such as *dar* [give], *tomar* [take], *hacer* [do/make] or *poner* [put], which carries out rather auxiliary functions and a generally deverbal noun, which provides an essential semantic load.
- c. Adjective + noun. The adjective is the collocative in these collocations: *fuentes fidedignas* [reliable source]; *importancia capital* [prime importance]; *relación estrecha* [close relationship]; *error garrafal* [terrible mistake]... From a semantic point of view, these collocations are characterized by the fact that the adjective normally intensifies the base, in a positive or negative sense, on a given scale.
- d. Noun + preposition + noun. This kind of collocations indicates either the unit within which there is a smaller entity or the group an individual belongs to. The first noun is the collocative, while the second is the base.
- e. Verb + adverb. Adverbs that belong to this type of collocations are adverbs of mode and intensity.
- f. Adjective + adverb. This group includes collocations that contain a participle, acting as an adjective, followed by an adverb, which is also normally of mode or intensity.

N.+ ver.	Ver. + n.	Adj. + n.	N. + prep. + n.	Ver. + adv.	Adj. + adv.
<i>Zarpar un barco</i>	<i>Desempeñar un cargo/</i>	<i>Oído/vista/ olfato fino</i>	<i>Rebanada de pan</i>	<i>Negar rotundamente</i>	<i>Profundamente dormido</i>
<i>Estallar una guerra</i>	<i>una función/ un papel</i>	<i>Dinero negro</i>	<i>Rebaño de ovejas</i>		<i>Rematadamente loco</i>
	<i>Conciliar el sueño</i>	<i>Visita relámpago</i>			
	<i>Tomar una decisión</i>				

Figure 1. Corpa's classification (1996)

	Types	Subtypes	Example
Collocations with simple lexical units	A. Noun + verb	A1.Noun (subject) + verb	<i>Rumiar [la vaca]</i>
		A2.Verb + noun (direct object)	<i>Contraer matrimonio</i>
		A3.Verb + preposition + noun	<i>Andar con bromas</i>
	B. Noun + adjective	B1.Attributive	<i>Lluvia torrencial</i>
		B2.Predicative	<i>La lucha fue encarnizada</i>
	C. Noun + <i>de</i> + noun		<i>Rebanada de pan</i>
	D. Verb + adverb		<i>Comer opíparamente</i>
E. Adverb + adjective/ participle		<i>Diametralmente opuesto</i>	
F. Verb + adjective		<i>Resultar ileso</i>	

	Types	Example
Complex collocations	A. Verb + nominal idiom	<i>Dar el santo y seña</i>
	B. Verbal idiom + noun	<i>Llevar a cabo un proyecto</i>
	C. Noun + adjectival idiom	<i>Tener una salud de hierro</i>
	D. Verb + adverbial idiom	<i>Reírse a mandíbula batiente</i>
	E. Adverbial idiom + adjective	<i>Fuerte como un roble</i>

Figure 2. Koike's classification (2001)

Secondly, we will take a look at Koike's classification (2001), since he introduces the novelty of distinguishing between collocations of simple lexical units and complex

collocations, as shown in Figure 2; the latter are formed by adding a lexical unit to another phraseological unit. As regards simple collocations, he suggests reducing the six types of collocations of Corpas to five, because he believes that the first two can be grouped into one – noun + verb –, but with three subgroups: noun (subject) + verb; verb + noun (direct object), and verb + preposition + noun. However, he adds a group composed of verb + adjective (for example, *salir indemne, resultar ileso* [to come off unscathed]), which means his proposal also contains six types.

4.2 Arbitrariness versus lexical selection

Towards the end of the 80's the work by Benson (1989:3) provides an insight into the arbitrariness of collocations when he reflects on the structure of a dictionary of collocations, since then, the lexicographers, similarly to foreign language teachers, cannot lose sight of what the real needs of the users of that dictionary are.

Thus, we can say that collocations should be defined not just as 'recurrent word combinations', but as 'arbitrary recurrent word combinations'. Hausmann assigns collocations to *langue*, i. e. the very system of the language. The arbitrary (as opposed to free) nature of collocations can be demonstrated when they are juxtaposed with corresponding collocations in other languages

(Benson 1985:11). (...)

The arbitrary nature of collocations can be easily demonstrated within English itself. One says *make an estimate* – but not **make an estimation*; *make an effort* – but not **make an exertion*; *commit treason* – but not **commit a treachery*...

That same year, and also from a lexicographic perspective, Hausmann (1989: 1010 and sections) dwells on the usefulness and need of a dictionary of collocations and insists on their arbitrariness, as would later be done by Corpas (1996), Zuluaga (2002) and Ferrando (1996).

C'est que l'idiosyncrasie de la collocation ne se relève définitivement que dans l'optique d'une autre langue qui combine, pour exprimer le même fait, des mots différents. On peut supposer qu'il suffit de contraster la combinaison avec celle, correspondante, de l'ensemble des langues de la terre, pour que n'importe quelle combinaison libre se transforme en collocation. Vu cet état de choses, on a intérêt à favoriser les solutions pragmatiques, à essayer de prendre en compte certaines grandes langues de comparaison, à éviter des cas trop évidents du genre *valise rouge* et à se résigner à ne pas être exhaustif dans la périphérie.

Nevertheless, Bosque's proposal and the publication of the *Redes* dictionary have shown that combinations of lexemes, which lead to collocations, are not arbitrary and they have been defined, described and classified in this lexicographic work.

Redes is not, strictly speaking, a dictionary of collocations, since it does not focus on this unit,¹⁷ but on lexical and/or semantic constraints that are found in the relations between predicates and arguments (Bosque 2004). For this author, as mentioned earlier (Bosque 2001b), collocations are a particular phenomenon of frequent co-appearitions and lexical selection, a broader notion that can be used to explain this and other phenomena of language. This vision of this linguistic phenomenon is very original, since it inverts the direction of the selection; that is, it suggests that it is the predicate that selects arguments and not the base that determines the collocative,¹⁸ as was thought until then.

Redes turns collocatives into entry-heads, so, for example, an adverb like *tajantemente* [outright] would be the entry-head and we might say it would be predicated in verbs like *rechazar* [to reject], *oponerse* [to oppose], *cerrar* [to close], *zanjar* [to settle], *separar* [to separate], etc.; that is, in each case, a group of words is given with which to combine the entry-head (which is always a predicate or selective element: N, V, Adj., Adv., Prep.), grouped in lexical classes (Bosque 2001a: 32). This work offers information that native Spanish speakers know as users of the language, but that does not appear in any other dictionary.

Therefore, Bosque understands that collocatives constitute predicates, and bases become their arguments; the combinatory dictionary directed by this author, *Redes*, studies, for example, entries such as *enérgicamente* [strongly], because the semantic classes¹⁹ this predicate forms can be characterized, but it does not study adverbs like *lentamente* [slowly], since they are applied to an excessively wide group of verbs.²⁰

17. As regards collocations, it does not consider them part of phraseology and recommends great care with the criteria of frequency, since they often offer information on the world and not on language (Bosque 2001b: 25–28). Regarding the criteria of preference, he also recommends caution, since often the most frequent collocation is interpreted as the only one, for example we speak of *desear ardentement* [to desperately desire], but it is no less true that we can apply this adverb to other verbs such as: *anhelar* [to long for], *defender* [to defend], *crear* [to create], *vivir* [to live], *amar* [to love], *besar* [to kiss], etc. Lastly, it is also possible that many of the preferable combinations are expressions of common places, favoured by the habits and beliefs of a community (love is fire/passion, etc.). The *Nueva Gramática de la Lengua Española* (2009:57) insists that lexical pieces are of interest to morphology and phraseology, and that lexicology should be the one in charge of collocations.

18. This idea that collocatives select a series of bases was already proposed in the concept of collocational field by Hausmann (1989), and Corpas also mentioned that instead of talking about pairs such as “collocative-base”, there were paradigms formed by related elements (*prestar ayuda/ atención/ asistencia/ colaboración...*[to lend a hand]).

19. Some examples of lexical classes would be: verbs of movement, of influence, of affection, of perception, of prepositional attitude, etc. Given that lexical classes are frequently repeated, *Redes* is full of cross references. Bosque (2001a).

20. It also does not include all nouns, verbs, idioms, etc. but only those that restrict their arguments according to semantic criteria.

In the following section we illustrate this opposition between arbitrariness and lexical selection, with two compared examples: on the one hand, a quote from Zuluaga and, on the other, the entry of the verb *cometer* [to commit], from the *Redes* dictionary, summarized in Figure 3.

4.3 Didactic implications

As we can see in the entry in *Redes*, the lexical classes that can go with this verb have been described, proving that it is not a case of arbitrary combinations but that the lexical meaning of that verb selects certain lexical classes, since in the very definition of the verb²¹ we can already infer that the actions committed are considered negative, which rules out combinations such as **cometer un acierto* [**to commit success*]. It is possible that in another language this meaning may be included in a single lexeme or there may be different verbs for committing positive actions and negative actions. Therefore, the first implication for SFL teachers is the need to become familiar with combinatory dictionaries which help understand this linguistic phenomenon.

This theoretical progress, which has shed its light on Spanish lexicography, has still not had the repercussion it deserves in FLT, and our aim in this article is to highlight this fact, since methodological changes for the teaching of collocation should be derived in the short-term. On the other hand, most of the definitions given by experts in vocabulary teaching of English as a Second Language (Gairns & Redman 1986; McCarthy 1990; Nation 1990, 2001; and Lewis 1993 1997 2000) and also by us in works prior to the publication of *Redes*, have defined collocations as the frequent co-occurrence of two words and have insisted on their arbitrary nature. A second didactic implication, as pointed out by Walker (2008: 307), would be that once the supposed arbitrariness of collocations is questioned, activities on collocations should be designed in a way that would help the student to discover why a word is combined with another.

Thirdly, perhaps another important didactic implication as regards teaching vocabulary is including word combinatory when explaining new vocabulary in the classroom; that is, providing the lexical classes that select a given word. Years of use of synonyms as the main form of presentation or explanation of unknown vocabulary, probably influenced by the importance of synonymic definitions in dictionaries, could be leading students away from the natural course of learning vocabulary, which

21. The first sense of the word in the DRAE says: 1. tr. “Caer, incurrir en culpa, yerro, falta, etc.” [To become guilty, make a mistake or fault]. As we can see, it does not indicate what can be “committed” in Spanish.

probably implies the comprehension of the extension²² of each predicate, that is, the entities to which it can be applied.

Foreign language teachers should always explain lexical units from their combinatory,²³ because knowing how a word combines is one of the dimensions of what knowing a word implies, and it is a vital piece of information – although not the only one – for the correct use of words in the syntagmatic axis (Higuera 2006a).

A teacher aware of the progress in psycholinguistics, who knows that words are stored in our lexicon creating networks with very diverse relations, and who expects his students to easily acquire this new vocabulary,²⁴ should attempt an explanation that could help create networks between words, insisting on combinatory, that is, the lexical classes that each predicate selects, as is shown in the analytical entries in *Redes*.

22. *Redes* is an onomasiological dictionary, that helps us codify, and therefore also takes charge, in a rigorous way, of the outlines (“*entourage*”) of words, and tries to describe the extension of the concepts (the entities to which they can be applied), as opposed to other dictionaries that focus solely on intensional information (the series of properties that distinguish it from other concepts). Thanks to the direction P > A (from the predicate to the arguments), *Redes* allows us to adequately relate the intension of the concepts with their extension. The extensional information is provided by the lists of examples and the intensional information, by lexical classes. (Bosque 2004: LXXXVIII and sections).

23. Over the last three decades a considerable number of combinatory dictionaries have been published for English and French (in Higuera 2004b, a revision of these works can be consulted), to which we can add the two titanic projects on the combinatory of Spanish: the DICE and *Redes*. These are two complementary approaches: for Alonso Ramos (2002a: 70) collocations are a codification phenomenon, hence the fact that she holds that collocations should be treated in the lexicographical entry of the base (as is done in the *Dictionnaire explicatif et combinatoire du français contemporain*, DEC, and the *Diccionario de colocaciones del español*, DICE: starting from an entry like *error* [error] the user can try to find out how to intensify that word). From a decoding perspective, however, the only way to rigorously deal with collocatives is that adopted by Bosque in *Redes*. In *Redes*, both *plantear* [to pose] and *problema* [problem] are entry-heads, while in the DICE only key words are defined (in the case of the collocation *plantear un problema* [to pose a problem] only *problema* appeared); to us it seems logical that a foreign student should need to know the combinatory capacity of the two words and not just of one of them. In the *Práctico* dictionary (Bosque 2006) both of them appear as entry-heads.

24. Recent studies on psycholinguistics make it possible to state that lexical acquisition is not only a sum of words that the student memorizes and stores one after another, but is a qualitative, gradual, multidimensional, progressive and dynamic process, where the learner weaves a complex network of relations – phonetic, graphic, semantic, morphological, syntactical, encyclopedic or personal – between lexical units, which associates and fixes new information with what he or she already knows. This intelligent storing of lexical units in which our lexical competence lies is what we know as lexicon.

Traditional explanation	Explanation bearing in mind the combinatory of words
<i>Cometer</i> is to carry out, for example <i>cometer un error</i> [to make a mistake]	<p><i>Cometer</i> is a verb meaning to carry out acts that are considered negative and can be combined with the following word classes:</p> <p>Nouns denoting errors: <i>error, fallo, equivocación, lapsus...</i></p> <p>Nouns denoting irregularity: <i>irregularidad, imperfección</i></p> <p>Nouns denoting senseless, imprudent, foolish or careless acts: <i>locura, disparate, barbaridad, tontería, estupidez...</i></p> <p>Nouns denoting results of breaking laws or rules: <i>falta, crimen, infracción, robo, asesinato...</i></p>

Figure 3. Two explanations of the verb *cometer*

This explanation, without doubt, contrasts with the explanation given by Zuluaga, that insists on arbitrariness:²⁵

La fijación que presentan las colocaciones está basada en el uso repetido y es arbitraria, no existe ninguna regla semántica que explique por qué decimos *cometer un error* en lugar de *hacer un error*. En *cometer un error* se realizan tanto el sistema como la norma, sin embargo, en *hacer un error*, sólo se realiza el sistema, es una combinación posible según las reglas del sistema, pero no está ratificada por la comunidad de hablantes, luego no se actualiza en la norma.

(Zuluaga 2002: 106)

Fourthly, we should be sensitive to the need to relate physical meanings of words with figured meanings, thus offering an integrating and non-accumulative approach to vocabulary. As explained by Bosque (2003: 23), predicates from column A have an entry in the *Redes* dictionary, and their physical uses (column B) are briefly explained, since they refer to extralinguistic (designational) information and the reader can add words to that list without knowledge of the language; however, *Redes* carefully studies column C, because we must have a knowledge of the language in order to complete it and understand it. Nevertheless, the elements in this column can be classified, as shown in the analytical entries.

25. [The fixation presented by collocations is based on repeated use and it is arbitrary, there is no semantic rule to explain why we can say *cometer un error* instead of *hacer un error*. In *cometer un error* both the system and the norm are followed, however, in *hacer un error*, only the system is carried out, it is a possible combination according to the system, but it is not confirmed by the speaking community, meaning the norm is not actualized] (Zuluaga 2002: 106).

A. Predicate	B. Physical uses	C. Figured extensions
Atesorar	Libros	Recuerdos, vivencias
Canalizar	Agua, río	Ayuda, demanda, sentimientos
Congelar	Agua, nariz	Negociación, imagen
Planear	Águila, avión	Sombra, duda, sospecha

Figure 4. Relationship between physical and figured meanings in *Redes*

5. Collocations in the acquisition and learning of vocabulary

Some powerful reasons to insist on teaching collocations when it comes to teaching vocabulary in a foreign language could be those mentioned below.

The collocational aspect is vital in the process of language acquisition, since the lack of knowledge of combinatory constraints on words in a language separates the productions of a native speaker from those of a non-native speaker. Castillo Carballo (2001) even goes further as to state that collocations are more complicated than idioms themselves, since the latter can be substituted by a free combination of words that expresses the same idea, but this operation is not possible with collocations, given the semantic precision that characterizes them. Through collocations we can determine the degree of knowledge the speaker possesses and, therefore, their teaching should not be limited to intermediate and high levels, but should be approached from the beginning of the process of learning a language.

Precisely because collocations are transparent and composed of familiar words, they go by unnoticed by the student.²⁶ If we draw the student's attention to this phenomenon, we are training him or her in the strategy of segmenting the input in units of meaning bigger than words, and making it easier for later incidental or unplanned learning of collocations. This way, we are paying attention to both types of learning, direct and incidental.

Furthermore, collocations are a principle that helps to organize and store vocabulary in the lexicon. Aitchison (1987:74–85) proved, through experiments consisting in asking a series of informers to say the first word that sprang to their minds when presented with a given lexical unit, that there are certain vocabulary

26. Hausmann(1989: 2013) underlines this fact: collocations are understood by learners, but they cannot reproduce them automatically, since they are in a way unpredictable and, therefore, they have to learn them: "Il doit l'apprendre, parce que les langues, dans la totalité des combinaciones logiquement possibles, font un choix idiosyncratique. La collocation est une unité, non de la parole, mais de la langue".

structuring principles, or different types of relations between lexical units, which she listed according to the order of frequency: coordination, collocations, superordination and, lastly, synonyms.

On the other hand, teaching multiverb complex lexical units consists in applying the idiom principle²⁷ (Sinclair 1991), and accepting that not every time we speak do we creatively construct our sentences word by word, but that we also resort to more complex segments that we have memorized, including collocations. In this sense, in the teaching of second languages it is used the term “lexical units”, more from a point of view of storage units than from a linguistic perspective.²⁸ Following this proposal we found the definition by Gómez Molina (2004: 497) of lexical unit:²⁹ “...la unidad de significado en el lexicón mental, que sirve como elemento vehiculador de la cultura del español y puede estar formada por una o más palabras (*cabeza, paraguas, dinero negro, a la chita callando, tomar el pelo, no hay más cera que la que arde, etc.*)”. The proposal we have made in specialized publications on didactics follows this line,³⁰ it is framed within the FLT and aims to clarify for teachers the type of units they should bear in mind in order to facilitate vocabulary learning for students; therefore, the term ‘lexical unit’ is referred to a processing unit or unit of meaning, with all the types of meanings we know exist.

Nevertheless, for other authors, for example McCarthy (1990: 6 and sections) they are not a lexical unit, but a relation of meaning, that accounts for syntagmatic or

27. The idiom principle, holds that a speaker has a wide number of semi-built blocks that he/she can combine when speaking, and is opposed to the open choice principle, which states that vocabulary is understood as a group of individual words that can be used very freely and are only limited by grammatical rules. This contribution by Sinclair means a step forward in the collocationist view on language.

28. From a linguistic perspective, lexical units are those that possess a unitary meaning and a referential unit and, therefore, are included in the dictionary (Bustos Gisbert 1986: 66). According to this concept of lexical unit we could not say that collocations are lexical units, since they are used to refer to frequent combinations of certain words, but do not always have a referential unit. What worldly reality does *saludar atentamente* [to wave politely] or *enemigo acérrimo* [worst enemy] refer to, for example?

29. [...the unit of meaning in the mental lexicon that acts as a vehicle element of the Spanish culture and can be formed by one or more words (*cabeza, paraguas, dinero negro, a la chita callando, tomar el pelo, no hay más cera que la que arde, etc.*)].

30. In spite of the fact that the units we are looking at in these works – namely: compounds, collocations, syntagmatic compounds, idiomatic expressions and institutionalized expressions or routine formulas – are of a very varied nature, in fact some are equivalent to a grammatical category, others to a phrase and others to a sentence. This heterogeneity can be justified from the moment we accept that there are many dimensions of meaning: collocational, referential, pragmatic, discursive, etc.

combinatory relations.³¹ Schmitt (2000:76 and sections) agrees with this approach, and explains the phenomenon of collocations as a property, a tendency, that appears in different degrees from free combinations to idiomatic ones.

In addition, we would like to underline the fact that knowing the collocation of words helps to distinguish synonyms (or words that share semes, like *mirar* – *ver* [to look, to see] and *coger* – *tomar* [to get, to take]) and to promote the creation of correct associations between words that tend to appear together in a specific language.

Also, we remind that teaching collocations is consequent with qualitative learning (Lewis 1993), that is, learning more about the words we know, as opposed to quantitative learning, which only focuses on the number of new words taught; lastly, it is a linguistic content that lends itself to autonomous learning, since the student can count on a series of lexicographic works that will help him or her with both codification and decoding of messages.

6. Future projects and lines of investigation

In spite of the rare presence of collocations in linguistics³² and of the minimal consensus regarding their nature and characteristics, paradoxically, the practical applications of this concept do not cease to bear their fruit in lexicography and FLT.

We will firstly comment on three ongoing projects. The first is the completion of the combinatory dictionary DICE, an absolute essential since it is the only one that applies the theory of lexical functions of the Meaning Text Theory to Spanish vocabulary. Since this dictionary is characterized by the use of a broad conception of collocation, not based exclusively on frequency of co-aparition, consulting it can help SFL students greatly, because in each entry all combination possibilities are given.

31. Selection (or “rektion” in some terminological systems) is also a relation established between two units. Some verbs govern or select certain prepositions, and the lexical meaning also restricts the combination possibilities of words. *NGLE* (2009:13).

32. In the *NGLE* (2009) collocations are timidly mentioned, as can be expected, since it is not a typical grammatical content, unless adjacent cases are studied with morphological units. The following is said of collocations: on page 855 it is said that we consider the following combinations to be collocations and not idioms: *enemigo acérrimo*, *esfuerzo ímprobo*, *error garrafal*. It insists that they are more productive than idioms, as a differentiating characteristic: the highest level of productivity is shown in syntactical units and the lowest in idioms, which do not allow the substitution of their components by a synonym, but which collocations do allow because they are more productive: *labor ímproba*, *tarea ímproba*, *trabajo ímprobo*. There is even an attempt to define ‘collocation’: frequent and preferential constrained lexical combinations, semantically restricted (p. 855 and 2654). On page 748 there is an insistence on the fact that collocations are fully integrated in syntax, since they allow coordination: *una ímproba y penosa tarea*, while idioms do not allow this characteristic: **caja fuerte y segura*.

For example, as pointed out by Sanromán (2003: 19), even combinations such as *gran silencio* and *silencio absoluto* [great silence and absolute silence] are both interpreted as collocations, because both express a 'high degree' and it is a case of constrained lexical concurrence. The contribution of lexical functions is undeniable, since they allow us to assign different values to a base, in order to obtain a more or less extensive number of collocations in each case, and to illustrate the complexity of the phenomenon of collocations, but at the same time this fact greatly hinders the publishing of the complete work, given the thoroughness of the treatment of the entry-heads (bearing in mind, for example, that the entry of the noun *esperanza* [hope] takes up ten pages).

Secondly, the logical evolution of the series of lexicographical publications from SM publishers points towards the publication of a third dictionary, the heir to *Redes*, aimed at SFL learners. This collocation dictionary project, in which we have had the pleasure of collaborating, will soon appear in the near future and its orientation is similar to that of the *Oxford Collocations Dictionary for Learners of English* (2001).

Thirdly, there is the project *Coloca-Te*³³ – Collocations and new technologies –, which aims to develop an active virtual learning environment to learn collocations in Spanish that will allow users to learn this content autonomously. We are also working on this project, which is directed by the researcher Margarita Alonso Ramos, and framed within the philosophy of lexical functions of the Meaning Text Theory. Other experts from other fields also participate: a team from the University of La Coruña in charge of the creation of contents in a learning environment based on the net, and a team from the Universtiy of Pompeu Fabra, in charge of the automatic processing of collocations in a learning environment based on the net. In the first phase we are working on the analysis of a corpus of learners of Spanish CEDEL2³⁴ in order to detect collocational errors, take note of them (using the Kwnotator tool) and classify them in a typology that will allow a later study. The first results³⁵ are the following:

- Of the total of collocations labelled, 61% of collocations are correct, 39% incorrect.
- Of those incorrect, 62% reveal lexical mistakes, 33% grammatical mistakes and 5% both.
- 54% affect the collocative 20% the base and 26% the whole collocation.
- 52% of mistakes are due to extension.
- The most frequent grammatical mistake is related to the verb complementation.

33. It is the I+D (innovation and development) project of the Spanish Ministry for Science and Innovation FFI2008-06479-C02-01/FILO. (Alonso Ramos 2010).

34. The information on the corpus can be found on: <http://www.uam.es/proyectosinv/woslac/cedel2.htm>.

35. Data presented in the SEDL congress 2009.

There is also a series of projects we consider necessary, from our point of view, and that have not begun yet, at least that we know of. One, for example, would be the creation of a guide to help both teachers and students to use combinatory dictionaries, since this requires certain abilities and metalinguistic knowledge, as they are dictionaries without definitions. This guide could be accompanied by activities that could be useful for independent learners and teachers wanting to approach this subject in their lessons, and could be inspired by already published works associated to the BBI dictionary for English.³⁶

Collocations should find an important space both in general books for the teaching of SFL and in specific books on autonomous vocabulary learning, and in one and the other we found that they are timidly making an appearance,³⁷ although rarely is the term collocation explicitly mentioned. This is much more frequent in manuals aimed at learning English as a foreign language, where we can even find books entirely dedicated to learning collocations.³⁸

A job which is still to be done is a thorough revision of the combinatory of vocabulary as a determining element when deciding the level at which a word is taught, which was taken into account when elaborating the *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español* (PCIC) (2006). Obviously, limitations of space meant we could not do this systematically enough (if we add the size of any monolingual dictionary and the information in a combinatory dictionary – for example *Práctico* – and we compare it with the space dedicated to general and specific notions of this project, the need to increase these inventories becomes clear). They should also be periodically updated using new linguistic corpus and data coming from studies on lexical availability of foreign students. Furthermore, we must consider that in this work only vocabulary from mainland Spanish variants were taken into account and other variants from American Spanish should be added.

As a result of this revision of collocation contents in the PCIC, it is possible that this knowledge might be more present in certification exams of linguistic competence, such as the DELE. Here, we can also see a field of elaboration of evaluation items, that allow the evaluation of collocation competence in learners of SFL on different levels.

Our intention was to offer a panoramic view of the progress made and of what is still to be done. As a conclusion, we could say that, curiously, the practical uses

36. Benson, M – Benson, E. – Ilson, R. & Young, R (1991): *Using the BBI: a workbook with exercises*, Benjamins.

37. Higuera, M. (2008): *Vocabulario. De las palabras al texto. A1*, Madrid: SM., and Higuera, M. (2009): *Vocabulario. De las palabras al texto. A2*, Madrid: SM.

38. McCarthy, M. & O'Del, F. (2005): *English Collocations in use*, Cambridge: Cambridge University Press.

of the concept of collocation are increasing in didactics and lexicography aimed at non-native speakers, while a theoretical consensus has not been entirely reached regarding this concept, despite it still being necessary to get to know more precisely the place that collocations should take up in Linguistics or in Phraseology.

References

- Aguilar-Amat Castillo, A. 1993. "En torno a la combinatoria del léxico: los conceptos de colocación e idiomatismo." In *Actas del XI Congreso de Lenguajes Formales*, In Carlos Martín Vide (ed.), 267–272, Barcelona: PPU.
- Aitchison, J. 1987. *Words in the Mind: an Introduction to the Mental Lexicon*. Oxford: Basil Blackwell.
- Almela Pérez, R. 1999. *Procedimientos de formación de palabras en español*. Barcelona: Ariel.
- Alonso Ramos, M. 1993. *Las funciones léxicas en el modelo lexicográfico de I. Melčuk*. Tesis doctoral, Madrid: UNED.
- Alonso Ramos, M. 1994. "Hacia una definición del concepto de colocación: de J.R. Firth a I.A. Melčuk". *Revista de Lexicografía* 1: 9–28.
- Alonso Ramos, M. 2002a. "Colocaciones y contorno de la definición lexicográfica. *Lingüística Española Actual* XXIV: 63–96.
- Alonso Ramos, M. 2002b. "Presentación del Diccionario de colocaciones y marcadores del español: estructura y objetivos." *IV Congreso de Lingüística General*, Cádiz, University of Cádiz, 47–61.
- Alonso Ramos, M. 2008. "Papel de los diccionarios de colocaciones en la enseñanza de español como L2." In *Proceedings of the Thirteenth EURALEX internacional Congress*, Bernal, E. & Decesaris, J. (eds.), 1215–1230. Barcelona: IULA, Pompeu Fabra University.
- Alonso Ramos, M. 2008. "Delimitando la intersección entre composición y fraseología." *Lingüística Española Actual* XXXI-2: 5–37.
- Alonso Ramos, M. 2010. "Towards a Motivated Annotation Schema of Collocation Errors in Learner Corpora." IREC 2010, in <http://www.lrec-conf.org/proceedings/lrec2010/summaries/751.html>.
- Benson, M. 1989. "The structure of the collocational dictionary." *International Journal of Lexicography* 2/1: 1–14.
- Boers, F. & Lindstromberg, S. (eds.) 2008. *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Berlin: Mouton de Gruyter.
- Bogaards, P. 2001. "Lexical units and the learning of foreign language vocabulary." *Studies on Second Language Acquisition* 23: 321–343.
- Bosque, I. & Demonte, V. (dirs.), 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa.
- Bosque, I. 2001a. "Bases para un diccionario de restricciones léxicas." *Moenia* 7: 11–52.
- Bosque, I. 2001b. "Sobre el concepto de 'colocación' y sus límites." *Lingüística española actual* XXIII/1: 9–40.
- Bosque, I. 2003. "La direccionalidad en los diccionarios combinatorios y el problema de la selección léxica". *Xarxa Temàtica en Gramàtica Teòrica*, Barcelona: (<http://seneca.uab.es/nlologiacatalana/investigacio/xarxes/bosque.pdf>).
- Bosque, I. 2004. *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM.

- Bustos Gisbert, E. de 1986. *La composición nominal en español*. Salamanca: University of Salamanca.
- Casares, J. 1992. [1950]. *Introducción a la lexicografía moderna*. Madrid: CSIC.
- Castillo Carballo, M.A. 1998. "El término colocación en la lingüística actual." *Lingüística española actual* XX/1, Madrid: Arco libros, 41–54.
- Castillo Carballo, M.A. 2001. "Colocaciones léxicas y variación lingüística: implicaciones didácticas." *Lingüística Española Actual* XXIII/1: 108–133.
- Corpas Pastor, G. 1996. *Manual de fraseología española*. Madrid: Gredos.
- Corpas Pastor, G. 2001. "Apuntes para el estudio de la colocación." *Lingüística Española Actual* XXIII/1: 41–56.
- Fernando, C. 1996. *Idioms and Idiomaticity*. Oxford: Oxford University Press.
- Ferrando Aramo, V. 2002. "Colocaciones y compuestos sintagmáticos: dos fenómenos léxicos colindantes." In *Léxico y gramática*, Alexandre Veiga et al. (eds.), 99–107. Lugo: Tris Tram.
- Gairns, R. & Redman, S. 1986. *Working with Words*. Cambridge: Cambridge University Press.
- Gómez Molina, J.R. 2004. "La subcompetencia léxico-semántica." In *Vademécum para la formación de profesores. Enseñar español como segunda lengua (L2)/ lengua extranjera (LE)*, J. Sánchez Lobato e I. Santos (eds.), 491–510. Madrid: SGEL.
- González Rey, M.I. 1998. "Estudio de la idiomatidad en las unidades fraseológicas." In *Estudios de fraseología y fraseografía del español actual*, Gerd Wotjak (ed.), 56–73. Madrid: Iberoamericana.
- Haensch, G. et al. 1982. *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Madrid: Gredos.
- Hatch, E. & Brown, C. 1995. *Vocabulary, Semantics and Language Education*. Cambridge: Cambridge University Press.
- Hausmann, F.J. 1979. "Un dictionnaire des collocations est-il possible?" *Travaux de Linguistique et de Litterature* 17/1: 187–195.
- Hausmann, F.J. 1989. "Le dictionnaire de collocations." In *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie. Enciclopédie internationale de lexicographie*, vol. 1, Franz Josef Hausmann et al. (eds.), 1010–1019. Berlin-New York: Walter de Gruyter.
- Hernández, M. Ll. & Brucart, J.M. 1987. *La sintaxis*. Madrid: Síntesis.
- Higuera García, M. 1997. "Las unidades léxicas y la enseñanza del léxico a extranjeros." *REALE* 8: 35–49.
- Higuera García, M. 2004a. *La enseñanza aprendizaje de las colocaciones en el desarrollo de la competencia léxica en el español como lengua extranjera*. Doctoral dissertation, University Complutense of Madrid.
- Higuera García, M. 2004b. "Necesidad de un diccionario de colocaciones para aprendientes de ELE." *Actas del XV congreso Internacional de ASELE*, in http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/15/15_0478.pdf.
- Higuera García, M. 2006. *Las colocaciones y la enseñanza del léxico*. Madrid: Arco Libros.
- Instituto Cervantes 2006. *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid: Biblioteca nueva.
- Írsula Peña, J. 1992. "Colocaciones sustantivo-verbo." In *Estudios de lexicografía y metalexigrafía del español actual*, Gerd Wotjak (ed.), 159–167. Tübingen: Max Niemeyer Verlag.
- Koike, K. 2001. *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Madrid: University of Alcalá & University of Takushoku.
- Lewis, M. 1993. *The Lexical Approach*. London: Language Teaching Publications.

- Lewis, M. 1997. *Implementing the Lexical Approach*. London: Language Teaching Publications.
- Lewis, M. 2000. *Teaching collocation. Further Developments in the Lexical Approach*. London: Language Teaching Publications.
- McCarthy, M. 1990. *Vocabulary*. Oxford: Oxford University Press.
- Martínez Marín, J. 2000. "Las unidades léxicas complejas en español: aspectos teóricos y descriptivos". *Revista de investigación lingüística* 2-III: 315–338.
- Martín García, J. 1999. *El diccionario en la enseñanza de español*. Madrid: Arco Libros.
- Nation, I.S.P. 1990. *Teaching and Learning Vocabulary*. Nueva York: Newbury House.
- Nation, I.S.P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, J.R. & DeCarrico, J.S. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Penadés Martínez, I. 2001. "¿Colocaciones o locuciones verbales." *Lingüística Española Actual* XXIII/1: 57–88.
- Piera, C. & Varela, S. 1999. "Relaciones entre morfología y sintaxis." In *Gramática descriptiva de la lengua*, Ignacio Bosque & Violeta Demonte (dirs.), 4367–4426. Madrid: Espasa.
- RAE 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Ruiz Gurillo, L. 1997. *Aspectos de fraseología teórica española*. Cuadernos de Filología, Anejo XXIV. Valencia: University of Valencia.
- Ruiz Gurillo, L. 1998. "Una clasificación no discreta de las unidades fraseológicas del español." In *Estudios de fraseología y fraseografía del español actual*, Gerd Wotjak (ed.), 13–38. Madrid: Iberoamericana.
- Ruiz Gurillo, L. 2002. "Compuestos, colocaciones, locuciones: intento de delimitación." In *Léxico y gramática*, Alexandre Veiga et al. (eds.), 327–339. Lugo: Tris Tam.
- Salvador, V. 1995. "De la fraseología a la lingüística aplicada." *Caplletra: Revista internacional de filología*: 11–30.
- Sanromán Vilas, B. 2003. *Semántica, sintaxis y combinatoria léxica de los nombres de emoción en español*. Doctoral dissertation, University of Helsinki, (<http://ethesis.helsinki.fi/julkaisut/hum/romaa/vk/sanroman/>).
- Schmitt, N. & McCarthy, M. 1997. *Vocabulary. Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Val Álvaro, J.F. 1999. "La composición." In *Gramática descriptiva de la lengua*, Ignacio Bosque & Violeta Demonte (dirs.), 4756–4841. Madrid: Espasa.
- Walker, C. 2008. "Factors which influence the process of collocation." In *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*, Boers & Lindstromberg (eds.). Berlin: Mouton de Gruyter.
- Zuluaga, A. 2002. "Los 'enlaces frecuentes' de María Moliner. Observaciones sobre las llamadas colocaciones". *Lingüística Española Actual* XXIV: 97–114.

