Collocation Chains: How to Deal with them?

Margarita Alonso Ramos(1), Leo Wanner (2)

(1) Universidade da Coruña
Campus da Zapateira s/n, 15071 CORUÑA
lxalonso@udc.es
(2) ICREA and Technology Department
Pompeu Fabra University, 08003 Barcelona
leo.wanner@upf.edu

Abstract

Collocation chains are pairs of collocations in which one element participates in both collocations. Such chains have hardly been considered so far in MTT and in other theoretical frameworks. However, some of them raise interesting questions with respect to their representation in the dictionary and treatment in NLP – e.g., in automatic text generation. In this paper, we give an overview of the different types of collocation chains and provide a preliminary suggestion for their treatment. For illustration, we use examples from Spanish.

Keywords

Collocation, lexical function, Explanatory Combinatorial Dictionary, automatic text generation, semantics, syntax, lexicology, collocation chain

1 Introduction

Speaking about collocations, we usually think of the occurrence of two lexical units (LUs) between which a restricted lexical co-occurrence holds – as, for example [*to*] TAKE and *WALK* in *John took a walk along the Hudson River*. However, in real corpora, we quite often come across what we call *collocation chains*: the occurrence of at least two collocations where one element is shared by both collocations. In what follows, we restrict the discussion to chains of two collocations. Consider, for instance, [*to*] *commit a violation of the law*. The noun VIOLATION is the base of the collocation *commit* [*a*] *violation*, and, at the same time, the collocate of the collocation *violation* [*of the*] *law*. Not all collocation chains are of the same kind. Some of them (as, e.g., [*to*] *take drastic measures*) can be treated as sequences of isolated collocations and thus do not need any new consideration with respect to their encoding in the dictionary and representation / mapping in a synthesis-oriented model (such as the Meaning-Text Model, MTM). Others require an extension of the standard collocation

representation as well as of synthesis rules; cf., e.g., *follow the rules rigidly*. The adverb *rigidly* modifies the verbal collocation *follow the rules* as a whole (rather than *follow* only).

In this paper, we take a closer look at collocation chains, focusing on Spanish material and giving occasionally evidence from German and English. To the best of our knowledge, collocations chains have not been studied in MTT yet, and only rather superficially outside MTT; cf., e.g., (Koike, 2001:147-149, 2004; Muñiz, 2004). Our goal is threefold. First, to provide the first sketch of a "collocation chain typology". Second, to determine how collocation chains are to be represented in a dictionary in the tradition of the Explanatory and Combinatorial Lexicology (Mel'čuk et al., 1995), using *lexical functions*, LFs, (Wanner, 1996). Third, to propose a way to represent the different kinds of collocation chains in the semantic (Sem) and deep-syntactic (DSynt) structures of an MTM and to encode the Sem-DSynt transition rules in a generation framework.

These three issues are dealt with in three separate sections. In Section 2, we present a preliminary typology of collocation chains. In Section 3, we address the problem of the representation of collocation chains in an Explanatory Combinatorial Dictionary. Section 4 elaborates on the representation of collocation chains in SemSs and DSyntSs and on the transition rules that map SemSs onto DSyntSs.

2 Towards a Typology of Collocation Chains

From a somewhat more formal view, a collocation chain is a sequence of three elements E_1, E_2, E_3, with collocational relations defined either between E_1 and E_2 and E_3 individually: E_1-E_2, E_1-E_3, E_2-E_3, or between E_1 and a complex unit consisting of E_2 and E_3: $E_1 - [E_2$-$E_3]$. An element can be a base (B), a collocate (C), or both ($|C/B|$). The following five basic cases are theoretically possible:[1]

1. CCB: two collocates share the base (as, e.g., *tomar*$_C$ *medidas*$_B$ *drásticas*$_C$ 'to take drastic measures');

2. CBB: two bases share the collocate (as, e.g., *las abejas*$_B$ *liban*$_C$ *el néctar*$_B$ 'the bees suck the nectar');

3. C $|B/C|$ B: one element is simultaneously collocate and base of two different collocations (as, e.g., *cometer*$_C$ *una violación*$_{|B/C|}$ *de la ley*$_B$ 'to commit a violation of the law');

4. C[CB]: a collocation is the base of a collocate (as, e.g., *[llamar*$_C$ *la atención*$_B]$ *poderosamente*$_C$ 'to draw the attention intensely')

5. B[CB]: a collocation is the collocate of a base (as, e.g., *[coger*$_C$ *el volante*$_B]$ *del coche*$_B$ 'to take the wheel of the car').[2]

Obviously a first rough distinction can be made between the first three cases and the last two: in the last two, one of the elements of the chain consists itself of a collocation. In what follows, we call this element *complex element*. Each of the first three cases potentially

---

[1] The order in which the elements are given does not reflect any dependencies or default order within sentences. To make the role of each element in the examples explicit, we annotate it with the corresponding index.

[2] Note that in this example (and in similar examples of this pattern in the course of the article), the complex unit (*coger el volante*) is a quasi-synonym of *conducir* 'drive', i.e., it is the value of the LF Real$_1$(*coche*).

contains several subcases which vary with respect to the syntactic dependencies between the elements. In order to assess which cases are problematic, we need to examine more closely these dependencies.

2.1 Syntactic Dependency between individual elements of collocations

Roughly speaking, we can distinguish configurations where between collocate(s) and base(s) a dependency relation holds and those where it does not.

2.1.1 Collocate and base are syntactically related

This is the most frequent case. It occurs in both patterns CCB and CBB. For CCB, the possible syntactic configurations are $C_1 \rightarrow B \rightarrow C_2$ and $C_1 \leftarrow B \rightarrow C_2$. The elements can be of different parts of speech; cf., e.g.,

> Base: ⁻; Collocates: Verb, Adj: *tomar medidas drásticas* 'take drastic measures'; *einen ausgedehnten Spaziergang machen* lit.'take a stretched-out walk'
>
> Base: ⁻; Collocates: Verb, Verb: *callar el miedo sentido* lit. 'to calm the felt fear'
>
> Base: ⁻; Collocates: Adj, Adj: *medidas drásticas pero justas* 'drastic but fair measures'

For the CBB pattern, we find, for instance, among the possible syntactic configurations, $B_1 \leftarrow C \rightarrow B_2$, which is exemplified, e.g., by *las abejas liban el néctar*; *el conductor conduce el coche* 'the driver drives the car', etc., where ABEJAS and CONDUCTOR are B_1, LIBAR and CO DUCIR are C and NÉCTAR and COCHE are B_2.

2.1.2 Collocate and Base are syntactically unrelated

The cases with no relation between the base and collocate are less frequent, but, as we will see later, much more interesting from the viewpoint of their treatment. In what follows, we restrict ourselves to the analysis of two configurations of pattern 2 (CCB) and one configuration of pattern 3 (C |B/C| B), namely $C_1 \rightarrow C_2 \rightarrow B$, $C_1 \leftarrow C_2 \rightarrow B$ and $C_1 \rightarrow B_1/C_2 \rightarrow B_2$

The configuration $C_1 \rightarrow C_2 \rightarrow B$ is instantiated by, e.g., *sufrir [una] oleada de atentados* 'to suffer a spate of terrorist attacks',[3] *[el] estado de alarma cunde* '[the] state of alarm spreads', as well as by *tener una sensación de alivio* 'to have a feeling of relief' and *tener un ataque de sueño* 'to have a fit of sleep' (with SUFRIR, CUNDIR, and TE ER as C_1, OLEADA, ESTADO, SE SACIÓ and ATAQUE as C_2, and ATENTADO, ALARMA, ALIVIO and SUEÑO as B). Note that SUFRIR and OLEADA and CUNDIR and ESTADO do not form collocations. The case of *tener una sensación de alivio* and *tener un ataque de sueño* is somewhat different. TENER and SE SACIÓ , and TENER and ATAQUE form a collocation. That is, this case can also be interpreted as the C |B/C| B-pattern (see below).

The nouns OLEADA, ATAQUE and ESTADO have the semantic feature called by Apresjan et al. (1989) *crystal*. They are also known as *light* (Bosque, 2001; Koike, 2003) or *transparent*

---

[3] The orientation of the dependency between OLEADA and ATENTADO is not fully agreed upon in MTT. We assume that OLEADA governs ATENTADO.

nouns (Fillmore et al. 2002, 2003). As Fillmore et al. (2003:244) put it: "a transparent noun is one which can appear as the first noun in $\square_1$-*of*-$\square_2$ constructions in contexts where the governing verb actually selects $\square_2$ rather than $\square_1$, the syntactic head". They suggest that transparent nouns can denote types, parts, portions, aggregates, borders, classifiers, and quantifiers. In the LF-terminology, the transparent noun C_2 in a $C_1 \rightarrow C_2 \rightarrow B$ dependency chain is most often Mult(B), Sing(B), Figur(B) or Gener(B) – if the generic noun is abstract enough, as, e.g., ESTADO 'state'. Henceforth, we refer to these LFs as *transparent LFs*.

The configuration $C_1 \leftarrow C_2 \rightarrow B$ as, e.g. *hacer solemnemente una promesa* 'to solemnly make a promise' is another example in which C_1 and B are syntactically not directly related. It is thus to be handled along the same lines as $C_1 \rightarrow C_2 \rightarrow B$.

The $C_1 \rightarrow B_1 / C_2 \rightarrow B_2$ (called *enchained collocation*) pattern as manifested in *cometer una violación de la ley*, *die Urherberschaft einer Tat leugnen* 'deny the authorship of an act', *die Lösung eines Problems finden* 'find the solution of a problem', etc. is very productive (and thus very common). Since B_1 (VIOLACIÓN, URHERBERSCHAFT, etc.) functions at the same time as collocate C_2, we have to label it in the DSyntS as an LF (rather than as a regular LU).

2.2 Syntactic Dependency between complex and individual elements of a collocation

The last two patterns (4) and (5) in our list of collocation chains share the property that one of their elements is complex, i.e. a collocation. In pattern (4), the complex element is the base, and in pattern (5), the complex element is the collocate.

Pattern (4) (C[CB]) as illustrated above by *llamar la atención poderosamente* is rather rare; it corresponds to the phenomenon Muñiz (2004:24) calls *integration* of collocations. Two subcases can be distinguished: $C_1 \rightarrow [C_2 B]$ and $C_1 \leftarrow [C_2 B]$. The above example illustrates the pattern $C_1 \leftarrow [C_2 B]$ (with *llamar la atención* as [$C_2 B$] and PODEROSAMENTE as C_1). C_1 indeed refers to $C_2 B$ as a combination, not to C_2:*llamar poderosamente*, or B: *atención poderosamente*. Further examples for this pattern are: *izar la bandera a media asta* 'to raise the flag at half-mast' and *abrir la puerta de par en par* 'to open wide the door'.

The representation of pattern 5 (B[CB]) does not require any extraordinary strategy. Again, there are two dependency configurations: $B_1 \rightarrow [C B_2]$ and $B_1 \leftarrow [C B_2]$; let us restrict ourselves to the second of them. It is illustrated by [*coger*$_C$ *el volante*$_{B2}$] *del coche*$_{B1}$, [*dar*$_C$ *rienda*$_{B2}$ *suelta*] *al llanto*$_{B1}$ 'to give free rein to crying', [*tomar*$_C$ *las riendas*$_{B2}$] *del negocio* $_{B1}$ lit. 'to take the reins of the business', etc. This configuration is more productive for symptomatic expressions of emotions, such as *quedarse sin habla de miedo* 'stay speechless with fear', *los dientes castañetean de miedo* 'the teeth are chattering with fear', *quedarse helado del susto* 'stay frozen with shock'.

3   Representation of Collocation Chains in Dictionary

We presuppose that the reader is familiar with the notion of LF and the representation of collocations in terms of LFs in the dictionary. The different types of collocation chains call for the distinction of three approaches to representation: (a) the standard way used in ECL-dictionaries; (b) in terms of separate lexical entries for collocates, (c) in terms of embedded lexical zones for collocates in the entries for the bases.

For patterns 1 (CCB), 2 (BBC) and 5 (B[CB]), the lexicographic description is standard. Each collocation is described in isolation in terms of LFs in the entry for its base (= keyword of the LF), no matter whether the value is an individual lexeme, an idiom, a collocation or a free phrase. We dispense with giving examples here; for extensive illustrative material, see the ECL-dictionaries.

In the case of pattern 3 ($C_1|B_1/C_2|B_2$), the description of $_{B1}/C_2B_2$ (in our example, *violación de la ley*) is standard. The description of $C_1B_{1/C2}$ (*cometer violación*) depends on whether $B_{1/C2}$ receives its own lexical entry or not. Until recently, in ECL, a collocate did not systematically receive an entry. However, the decision on opening a lexical entry for a collocate depends on several factors (Alonso Ramos, 2003). One of the factors is the capacity of the LU in question to select its own collocates. This is the case for the noun VIOLACIÓN: there is no better place than its lexical entry to state that it selects COMETER as its support verb. Additionally, VIOLACIÓN occurs in a number of other collocations as base.

The case of pattern 4 ($C_2[C_1B]$) is similar to the previous case. It is to be decided whether an entry for C_1 (*llamar* in our example) is to be opened or its collocate (PODEROSAMENTE) should be described in the entry of the base B (ATENCIÓN). For example, we could open an entry for LLAMAR with the particular meaning it has in co-occurrence with ATENCIÓN. In this entry, the adverb PODEROSAMENTE would be the value of Magn. However, if this adverb can modifiy LLAMAR only when it is combined with ATENCIÓN, it seems more appropriate to describe the relation between the adverb and the collocation in the entry for the noun. Thus, for the entry ATENCIÓN, we would open a lexical zone for the collocation in such way that we can assign information to some of the values of the LF, but not to all.

```
Caus₂Func₁:
        llamar [ART ~ a X]
                Magn = poderosamente
        suscitar [ART ~ a X]
```

4 Collocation Chains in Generation

In this section, we take a more systematic look at the encoding of collocation chains at the Sem- and DSynt-levels and their treatment during the transition between these two levels.
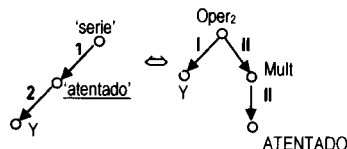
4.1 Representation of Collocation Chains

We present the SemSs and DSyntSs of collocation chains from some of the examples from above. The examples show that SemS of the chains we look at in the paper is fairly simple; it is the DSyntS (and thus also the Sem-DSynt transition), which deserves some attention.
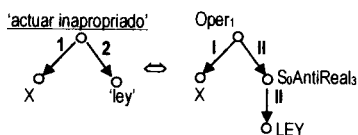
1. *tomar medidas drásticas*
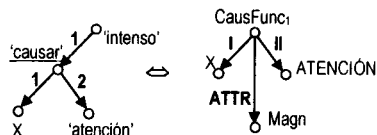   'take drastic measures'

'intenso'   Oper₁
  1   I II
'medida'   ⇔   X   MEDIDA
  X   ATTR
                     Magn
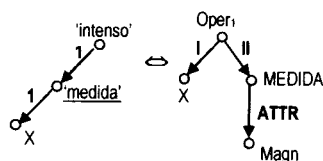
2. *sufrir una oleada de atentados*
   'suffer a spate of attacks'

'serie'   Oper₂
  1   I   II
  2 'atentado' ⇔ Y   Mult
  Y   II
                       ATENTADO

3. *cometer una violación de la ley*
   'comit a violation of the law'

'actuar inapropriado'   Oper₁
   1   2   I   II
  X   'ley'   ⇔   X   S₀AntiReal₃
                            II
                            LEY

4. *llamar atención poderosamente*
   'call powerfully for attention'

                        'intenso'   CausFunc₁
'causar' 1   I   II
   1   2   ⇔   X   ATENCIÓN
  X   'atención'   ATTR
                            Magn

These structures call for the following remarks:

(i) In accordance with Wanner and Alonso (2005), we interpret LF-labels at the DSynt-level as deep LUs, i.e., we do not use in the structures the functional notation f(L). This does not mean that the information concerning the functional interpretation of an LF is lost. Each LF receives an entry where we specify its government pattern (GP); the GP allows for the recovery of the keyword.

(ii) The case illustrated in (1) is simple since the representation of the chain at the DSynt-level is equivalent of the representation of two LFs in isolation.

(iii) At the DSynt-level, where we consider LFs as deep LUs, cases (2) and (3) are analogous: in both cases, an LF has as its DSyntA II another LF. The difference lies in their functional interpretation. The standard Oper-representation in DSyntSs requires the keyword to be actant II of the Oper₁-node. However, in (2), actant II is the Mult-node, which is not the keyword of Oper₂: the keyword is ATENTADO. In (3), although actant II of the Oper-node is also an LF, the LF-node acts as its keyword.

(iv) Case (3) is not trivial because it contains an LF-label (namely, S₀AntiReal₃) as keyword of the Oper-node. This means that we need to know the LU that functions as value of S₀AntiReal₃ in order to ensure that this LU possesses an Oper.

(v) (4) is the same as (3) with respect to the use of an LF-label (CausFunc₁) as keyword of another LF-labelled node (Magn). However, (4) requires a different description in the dictionary and thus a different treatment during the DSynt-SSynt transition.

4.2 Transition between Sem- and DSynt-Levels of Representation

The Sem-DSynt transition between the structures of the type shown in (1) has already been discussed, e.g., in (Iordanskaja et al., 1996; Kahane and Mel'čuk, 1999; Alonso Ramos,
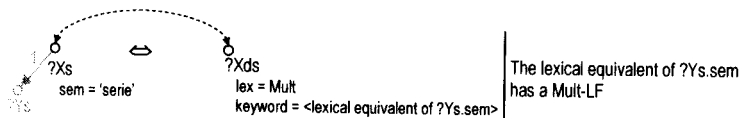
2007). Therefore, we focus on the transitions between structures illustrated in (2 – 4). Each transition rule is elementary in the sense that it covers the smallest fragment possible of a SemS that is being "transduced" into its DSynt-equivalent: single nodes and binary relations between nodes. As a result, each transition grammar consists of a set of *nodal* and *sagittal* rules (Kahane and Mel'čuk, 1999), with any of the rules being applicable to a multitude of structures. Let us introduce the types of rules shared by all our cases.
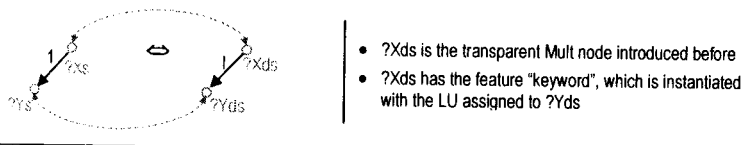
4.2.1 Regular Nodal and Sagittal Rules

Two types of regular nodal rules are pertinent for our task: lexical-semantic rules and lexical-functional rules. Furthermore, we need some sagittal rules of which we illustrate one. The lexical semantic rule cited below maps a communicatively dominant predicative semanteme (denoted here by the variable ?Xs) onto the corresponding deep LU that is realized as a noun in an LF-construction (assigned to the variable ?Xds). We assume that the choice of an Oper-construction is motivated essentially by the communicative structure, CommS, captured in the rule in question by the feature *aspectual focus_event*.[4]

?Xs ⟺ ?Xds
lex = \<nominal LU of 's'\>

| ?Xs is the communicatively dominant node |
| ?Xs has the comm. feature 'aspectual focus event' |
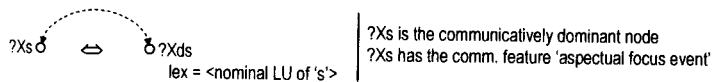
The following rule illustrates a "regular" lexical-functional rule which maps the semanteme 'serie' 'series' onto the LF-label Mult.[5] The keyword of Mult is the lexeme corresponding to the first argument of 'serie' (?Xs-1→?Ys). Both the lexeme and the relation are introduced by other elementary rules. Therefore, both appear in the context (showed in grey). Obviously, the condition to be fulfilled is that this LU has a value of Mult.

?Xs   ?Xds
sem = 'serie'   lex = Mult
         keyword = \<lexical equivalent of ?Ys.sem\>

The lexical equivalent of ?Ys.sem
has a Mult-LF

The following saggital rule below maps the semantic relation '1' onto the DSynt-relation 'I'. The relation '1' holds between a quantifier semanteme (such as 'series') and its argument. The relation 'I' holds between the LF label of a transparent LFs such as Mult, Sing, etc. and the corresponding keyword.

1 ?Xs   I ?Xds
?Ys   ?Yds

- ?Xds is the transparent Mult node introduced before
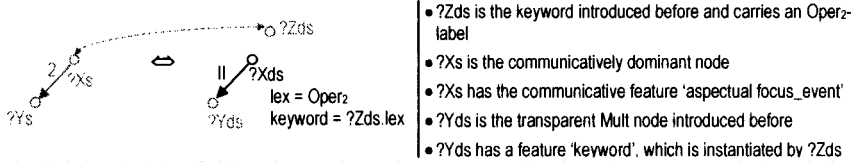- ?Xds has the feature "keyword", which is instantiated with the LU assigned to ?Yds

---

[4] This feature is intended to capture the different communicative aspectual perspectives on a semantic configuration. For instance, a support verb construction such as *take a walk* expresses a bound event, while the verb *to walk* expresses a process.

[5] To facilitate the readability, we have chosen a very specific rule. In practice, this rule is substituted by a more generic rule that handles the mapping of semantic configurations onto LF-labels.

In order to carry out the mapping between the SemSs and DSyntSs in (2–4, Section 4.1), we need, in addition to the regular rules sketched above, some idiosyncratic rules that take care of collocation chains.
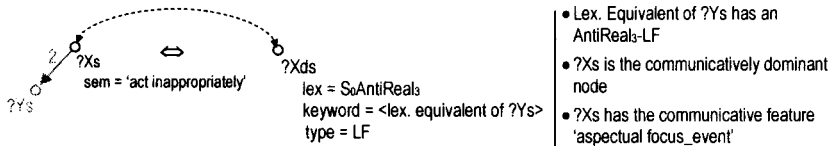
4.2.2 The Case of Transparent LFs

This case corresponds to the mapping of the structures in (2). To realize the collocation chain, we need a special $Oper_2$-rule. Any Oper-rule needs to consider the CommS. However, in this case, it must also ensure the availability of a transparent LF already introduced into the DSynS. The rule specifies that the second actant of the newly introduced $Oper_2$-node is a transparent LF and both LFs share the same keyword (?Zds.lex). $?Xs-2\rightarrow?Ys$ is the relation between the main node and its second actant.

- ?Zds is the keyword introduced before and carries an $Oper_2$-label
- ?Xs is the communicatively dominant node
- ?Xs has the communicative feature 'aspectual focus_event'
- ?Yds is the transparent Mult node introduced before
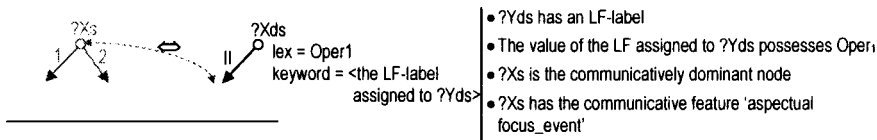- ?Yds has a feature 'keyword', which is instantiated by ?Zds

The left hand side of this rule consists of a context only since ?Xs , ?Ys and the relation between them are mapped by other rules.

4.2.3 The Case of Enchained LFs

In the case of enchained LFs of the type *cometer violación de la ley* (cf. (3) in Section 4.1), two rules similar to those already introduced above are essential: one for the introduction of the $S_0AntiReal_3$-node and one for the introduction of the $Oper_1$-node. A preliminary $S_0AntiReal_3$-rule looks as follows:[6]

- Lex. Equivalent of ?Ys has an $AntiReal_3$-LF
- ?Xs is the communicatively dominant node
- ?Xs has the communicative feature 'aspectual focus_event'

The semanteme 'act inappropriately' corresponds at the DSyntS-side to $AntiReal_3$. The nominalization enforced by the communicative features is encoded by S_0. The $Oper_1$-rule is somewhat different from the Oper-rule in Section 4.2.2:

- ?Yds has an LF-label
- The value of the LF assigned to ?Yds possesses $Oper_1$
- ?Xs is the communicatively dominant node
- ?Xs has the communicative feature 'aspectual focus_event'

[6] The rule is fairly idiosyncratic and implies that each verbal LF-rule must be duplicated in order to account for nominalization of the value of the verbal LF. A dynamic composition of LF-labels (in the spirit of Kahane and Polguère, 2001) during the transition would be more appropriate. However, this strategy has the disadvantage that it implied a modification of already introduced DSynt-nodes.
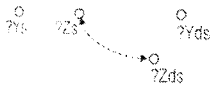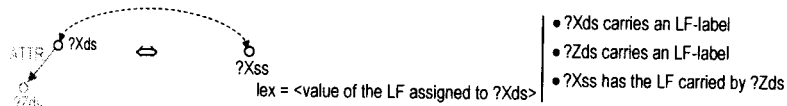
The conditions require in the dictionary an $Oper_1$-value (here, COMETER) for the value of the keyword-LF assigned to ?Yds (i.e., VIOLACIÓN).

4.2.4 The Case of Complex Elements

Chains containing complex elements of the kind *llamar la atención poderosamente* can be dealt with during the Sem-DSynt transition along the same lines as discussed in 4.2.3. However, they call for attention during the DSynt-SSynt transition because the collocate (such as PODEROSAMENTE) does not co-occur with all values of the LF applied to ATENCIÓN (in our case, $Caus_2Func_1$). In order to ensure the realization of Magn, we have to select a $Caus_2Func_1$-value which has a Magn-collocate. This is the tribute we must pay to account for having a collocation as a base. The following rule handles the $Caus_2Func_1$-lexicalization:

- ?Xds carries an LF-label
- ?Zds carries an LF-label
- ?Xss has the LF carried by ?Zds

lex = <value of the LF assigned to ?Xds>

The rule for the lexicalization of Magn ensures that a value is chosen that co-occurs with the already determined surface lexeme of $Caus_2Func_1$. We omit the discussion of the Magn-rule.

5 Summary and Conclusions

We presented some examples of collocation chains, showing that they cannot always be dealt with as isolated collocations. Dependency relations between the bases and collocates play a particular role in the complexity of a collocation chain. Furthermore, the double role of LFs, as deep LUs and as functions, which comes to bear in collocation chains raises again the question of the nature of LFs in an MTM and thus, subsequently, the nature of collocates as proper LUs. Finally, despite the preliminary nature of our SemDSynt-transition rules and of the discussion of the representation of collocation chains in the dictionary, we hope to have shown that they need to be taken into account in a comprehensive synthesis model.

Bibliography

Alonso Ramos, M. 2003. La nature des collocatifs: leur statut en tant qu'unités lexicales. In Grossmann F. and Tutin, A (eds.), *Les collocations : analyse et traitement, Travaux et Recherches en Linguistique appliquée*, Amsterdam: Editions De Werelt, 45-60.

Alonso Ramos, M. 2007. Towards the Synthesis of Support Verb Constructions: Distribution of Syntactic Actants between the Verb and the Noun. In Wanner, L. (ed.), *Selected Lexical and Grammatical Issues in Meaning Text Theory. In Honour of Igor Mel'čuk*, Amsterdam: John Benjamins, 93-133.

Apresjan, Yu. D. et al. 1989. Lingvističeskoe obespečenie sistemy ÈTAP-2. Moscow: Nauka.

Bosque, I. 2001. On the Weight of Light Predicates. In Herschensohn, J. Mallén, E., & Zagona, K. (eds.). *Features and Interfaces in Romance. Essays in honour of Heles Contreras*, Amsterdam: John Benjamins, 23–38.

Fillmore, C.J., Baker, C.F. and Sato, H. 2002. Seeing Arguments through Transparent Structures. *Proceedings of LREC*, Las Palmas, 787-791.

Fillmore, C.J, Johnson, C and Petruck, M. 2003, Background to FrameNet, *International Journal of Lexicography* 13.3: 235-250.

Iordanskaja, L., Kim, M. and Polguère, A. 1996. Some Procedural Problems in the Implementation of Lexical Functions for Text Generation". In Wanner. L. (ed.), 279-297.

Kahane, S. and I. Mel'čuk. 1999. Synthèse des phrases à extraction en français contemporain. *T.A.L.* 40.2: 25-85.

Kahane, S. and Polguère, A. 2001, Formal foundation of lexical functions. *Proceedings* of the *Workshop COLLOCATION: Computational Extraction, Analysis and Exploitation at the 39th Annual Meeting of the ACL*, Toulouse, 8-15.

Koike, K. 2001. *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*, Alcalá de Henares / Tokio: Universidad de Alcalá / Univ. Takushoku.

Koike, K. 2003. Sustantivos "ligeros", *Moenia*, 9: 9-20.

Koike, K. 2004. Colocaciones complejas en el español actual. In Almela Pérez, R., Ramón Trives, E., Wotjak, G., (eds.). *Fraseología contrastiva*. Murcia: Universidad de Murcia / Universität Leipzig.

Mel'čuk, I., Clas, A. and Polguère, A. 1995, *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-neuve: Duculot.

Muñiz, E. 2004. *El concepto de 'colocación' en español*, PhD thesis, Santiago de Compostela: Universidade de Santiago de Compostela.

Wanner, L. ed. 1996, *Lexical Functions in Lexicography and Natural Language*, Amsterdam: John Benjamins.

Wanner, L. and Alonso, M. 2005. What Type of Entity is a Lexical Function? In Apresjan, J. and Iomdin, L. (eds.), *East West Encounter: Second International Conference on Meaning-Text Theory*, Moscow: Slavic Culture Language Publishing House, 518-528.

Acknowledgements