

# **A Practical Example for the Use of AGFL in the Automatic Syntactic Analysis of Spanish**

M<sup>a</sup>. Paula Santalla del Río  
Computational Linguistics Department  
Royal Spanish Academy  
Felipe IV, 4, Madrid, 28014, Spain

Jan Cloeren  
Computational Linguistics Department  
Royal Spanish Academy  
Felipe IV, 4, Madrid, 28014, Spain

## **Abstract**

With the information derived from a corpus based verbal database developed at the University of Santiago de Compostela, we propose a preliminary way of codification of verb valency and subcategorization in the lexicon of an AGFL grammar as a safe way to face syntactic clause analysis with a more pure and simple AGFL syntax. An example is given with the verb “abandonar” about how to extract and simplify the information obtained from the database, and how to code and process it with an AGFL grammar.

## **1 Theoretical Principles**

Because of the initial state of the line of research which will be described in the remainder of this paper, this exposition is conceived as a preliminary proposal for a description and formalization of a certain type of syntactic sequences in Spanish. The syntactic descriptive frame in which this proposal will be developed is a constitutive and functional one. Constitutive because it makes explicit the constitutive relations established between elements of a sequence and the whole sequence, by means of its progressive analysis in its components until a minimum level of units (not syntactic units any more) is reached. Functional because it makes also explicit the functional relations between components and successive segments which result from analysis. Functions are in the center of description because it is assumed that are precisely functions (internal functions) which define units.

Besides categories and functions, as theoretical principles, mechanisms of construction, of expansion are recognized: coordination, subordination, interdependency. We will not develop here these concepts (for a first approach to our point of view, see Santalla, 1994, and Jiménez Juliá, 1993), but will concentrate on syntactic units and functions.

Among syntactic units in Spanish, the **clause** will be in the centre of this exposition. A clause is defined as the syntactic unit constructed around a function that we call **PREDICATE**, which is always assigned to a verb.

The first step in the internal description of the clause, consequently, concerns the external (in the sense of syntagmatic) description of the verb, its syntagmatic behaviour in the clause. Studying this behaviour for the verb, immediately becomes studying it for each verbal lexeme, not for the category itself, because it is the **valency** of each verb which determines how many and which argumental (not free) functions can appear around each concrete lexeme. Summarizing: the internal study of the clause drive us directly to the concept of verbal valency, which can be defined as:

“las peculiaridades combinatorias que presenta cada esquema verbal” (García-Miguel, 1995: 19)

But, what do we consider pertinent information for determination of valency? The point of view supported here can be found in García-Miguel (1995:31-34): valency is a complex of information which can be broke down in the following parts:

- *Quantitative valency*: **number of arguments** (obligatory or facultative) a verb requires in order to constitute a complete sense clause.
- *Qualitative valency*, which includes different types of linguistic information:
  1. *Syntactic*: **syntactic functions** assigned to each argument, and **syntactic categories** selected by the verb for each one of them.
  2. *Semantic*: **semantic functions**, semantic roles linked to the arguments, and **semantic categories** selected by the verb for each one of them.

About quantitative valency: looking for some kind of generalization, we have constructed a hierarchy of minimal sets of characteristics which represent the probability of a clause functional segment of being an argument with respect to any verb (a more detailed study of criteria which have been proposed to distinguish arguments and circumstantials can be found in García-Miguel, 1995:23-31):

TABLE I

+ obligatory or, otherwise, latent + lexically restricted by the verb +/- central <sup>1</sup>	- obligatory or, otherwise, latent + lexically restricted by the verb +/- central	- obligatory or, otherwise, latent - lexically restricted by the verb +/- central
+ central: Yo voy al cine -central Lo obligué <i>a venir</i>	+ central: Está escribiendo <i>una carta</i> - central: La abandonó <i>a su suerte</i>	+ central: <i>Le</i> abrí la puerta <i>a María</i> - central: Lee libros <i>todos los días</i>

<sup>1</sup>Which, referred to a possible argument, means that the segment in question is grammatically related to the verb: by concordance or cliticization (in Spanish, subject, direct object or indirect object).

We will assign quantitative valency to each verb taking into account the clause functional segments which comply with the conditions imposed on the first group on the left<sup>2</sup>. Nevertheless we will not relegate clause functional segments which comply with the center group of our hierarchy out of our description: they are included (see below 3) in what describes qualitative valency. The same holds for possible arguments which conform with the group of characteristics on the right and have a positive value for the last one (centrality)<sup>3</sup>. We are aware that, in this way, we are separating the unique concept of valency, but it proves to be more economic for a description which is expected to support formalization. And, on the other side, the information coded is exactly the same (richer in the case of central not valential constituents), and its reorganization can be systematically carried on.

About functions: possible valential and argumental functions are: subject, direct object, indirect object, prepositional complement (prepositional object, adverbial, modal) and predicative<sup>4</sup>.

About syntactic categories and semantic features: with the purpose of simplifying the formal syntactic rule of the predicate and the lexical one for the verb, they are coded together. The catalogue of values is not closed: as the research goes on, new values are added and included in the description and, consequently in the grammar. Anyway, we pretend to restrict values for semantic categories to animation, determination and countability<sup>5</sup>. The results of its combination with values which will prove to be necessary for syntactic categories is more difficult to predict at the actual initial state of the research.

About semantic functions: neither this information, nor its correspondence with syntactic functions, which should be present in a valency dictionary, can be coded in advance, because they are not linked to a formal, material mark: their number and categorization depend on the lexical character of the verb and they are recognizable only once it has obtained a concrete realization in a clause. From the practical point of view, this means that, for the purpose of this research, which is the identification of verbal syntactic and clausal schemes, this information can not be previously used, but subsequently derived.

---

<sup>2</sup>This is not the usual value assigned to quantitative valency, which usually includes also clause functional segments which comply with conditions defined for the centre group of our hierarchy (see García-Miguel, 1995:24).

<sup>3</sup>This affects in Spanish to some indirect objects.

<sup>4</sup>In this state of the research we have left out this argument, because its relation seems to be different from the one established by the rest of the possible arguments in Spanish.

<sup>5</sup>Actually, we are working only with the first one, because determination and countability are much more difficult to systemize. Anyway, we include them because we have collected this information in the database (when we were working manually with concrete examples), and we do not want to disregard this perhaps useful information in advance.

## 2 The database

With the theoretical and implied methodological principles described above, a database with information about verbs and their valency (in both senses, qualitative and quantitative) has been constructed in the University of Santiago de Compostela. Information about verbs and their construction has been derived from a *corpus*<sup>6</sup> of one million and a half words which includes oral language and text from spanish novel, essay, newspaper and drama after 1980.

For each occurrence of a verb in the corpus a record is collected with detailed information about the syntactic configuration of the linguistic sequence in such occurrence. Each record contains general information about the *clause* in which the verb is found (if it is independent or not, mood, polarity, verb form, order of constituents...), and, which is our main interest, about *arguments* with which it is combined in the concrete case under study (quantitative valency) and their syntactic functions and form (qualitative valency or *verbal scheme and subscheme*). The criteria for selection of argumental constituents have been quite extensive. In this way, we pretend to have guaranties that, at least, for what is more frequent, we have collected all the possible syntactic peculiarities of each verb<sup>7</sup>. Possible arguments of a verb, its quantitative valency, are considered to function as *subject, direct object, indirect object, prepositional complement (prepositional object, locative adverbial or modal adverbial) or predicative*. Alphanumerical keys refer to the form of these constituents: the syntactic categories they assume. Some syntactically relevant semantic features are also specified: *determination, animation and countability*. Processing all the records collected for each verb we obtain this type of summary information<sup>8</sup>:

ABANDONAR “to leave, abandon, give up”

[197 casos; 9 esquemas; 18 subesquemas]

Act.	S	(F= 2; 1,02% s/verbo)
	San	(F= 2; 100,00% s/esq.)
Act.	SD	(F=172; 87,31% s/verbo)
	San Dan	(F= 31; 18,02% s/esq.)
	San Din	(F= 126; 73,26% s/esq)
	Sin Dan	(F= 10; 5,81% s/esq.)
	Sin Din	(F= 5; 2,91% s/esq.)

<sup>6</sup>ARTHUS, which stands for *Archivo de textos hispánicos de la Universidad de Santiago de Compostela*.

<sup>7</sup>Even when, for indirect object, prepositional complement and predicative, the argumental character can be doubtful.

<sup>8</sup>Keys to this table: Act., active voice, Pron., middle voice, Pas., passive voice, S, subject, D, direct object, AD, prepositional complement (adverbial), SP, prepositional complement (prepositional object), A, agent complement, an, animate, in, inanimate, cf, infinitive clause.

Act.	SD AD	(F= 4; 2,03% s/verbo)
	San Dan ADin (en )	(F= 1; 25,00% s/esq.)
	San Din ADin (en )	(F= 2; 50,00% s/esq.)
	San Din ADin (sobre )	(F= 1; 25,00% s/esq.)
Act.	SD SP	(F= 2; 1,02% s/verbo)
	San Dan SPin (a )	(F= 1; 50,00% s/esq.)
	San Din SPin (a )	(F= 1; 50,00% s/esq.)
Pron.	S	(F= 3; 1,52% s/verbo)
	San	(F= 1; 33,33% s/esq.)
	Sin	(F= 2; 66,67% s/esq.)
Pron.	S SP	(F= 8; 4,06% s/verbo)
	San SPcf (a )	(F= 1; 12,50% s/esq.)
	San SPin (a )	(F= 7; 87,50% s/esq.)
Pas.	S	(F= 3; 1,52% s/verbo)
	San	(F= 1; 33,33% s/esq)
	Sin	(F= 2; 66,67% s/esq.)
Pas.	S A	(F= 2; 1,02% s/verbo)
	San Aan	(F= 2; 100,00% s/esq.)
Pas.	S SP	(F= 1; 0,51% s/verbo)
	Sin SPin (a )	(F= 1; 100,00% s/esq.)

It is clear, at this point, that one of the immediate applications of such a syntactic database is the elaboration of a dictionary of verbal construction. With this aim, each occurrence of a verb, and consequently, its construction in this concrete case, is assigned to a concrete meaning. Possible meanings are derived from various dictionaries for Spanish<sup>9</sup>. Combining this assignment with the one previously described we obtain this type of summary information:

---

<sup>9</sup>Real Academia Española, *Diccionario de la lengua española*, Espasa-Calpe, Madrid, 1992. Alvar Ezquerro, Manuel, *Diccionario actual de la lengua española*, Bibliograf, Barcelona, 1990. Moliner, María, *Diccionario de uso del español*, Gredos, Madrid, 1966-67.

ABANDONAR<sup>10</sup> "to leave, abandon, give up"

1. [SPD(aU)] Dejar desamparado algo o a alguien(D), dejar algo o alguien (D) a merced de algo(U) (÷ DALE 1, DRAE 1, DUE 1) [24%].
2. [SPD] Soltar algo(D), aflojar la tensión de algo(D) (÷ DUE 2) [1%].
  - 2b. [SPD] Renunciar a una idea, proyecto, etc.(D) (÷ DALE 2, DRAE 2, DUE 2) [3%].
  - 2c. [SPD] Dejar algo ya iniciado(D) (÷ DALE 2, DRAE 2, DUE 2) [7%].
  - 2d. [SPD(A)] Dejar de utilizar algo(D), prescindir de algo(D), dejar algo(D) en algún sitio(A) (DALE 3, DUE 2) [23%].
  - 2e. [SPD] Dejar el asegurado(S) por cuenta del asegurador [...] las cosas aseguradas(D) a fin de obtener el pago del seguro (÷ DALE 5, DRAE6) [0%].
3. [SPD] Apartarse de algo o alguien(D), marcharse de un lugar(D) (÷ DRAE 3, DUEp 3) [37%].
4. [S seP] Cesar en un esfuerzo, reclinarsse con dejadez (÷ DALE 4, DRAE 4, DUE 1) [1%].
  - 4b. [S seP] Desmoronarse, caer de ánimo (÷ DALE 9, DRAE 10, DUEp 3) [0%].
  - 4c. [S seP] Descuidar alguien(S) sus obligaciones, su aseo (÷ DALE 7, DRAE 8/9, DUEp 2) [0%].
5. [S seP aU] Darse a algo(U), dejarse dominar por algo(U) (÷ DALE 6, DRAE7, DUEp 3) [3%].
6. [S seP aU] Entregarse a algo o alguien(U), confiarse en algo o alguien(U) (÷ DALE 8, DRAE 5, DUEp 3) [1%].

### 3 Coding this information in AGFL format

As it is well known here, a formal grammar written in the AGFL formalism is mainly divided into two parts: the *metagrammar*, in which affixes are described, and the *main grammar rules*, which account for syntagmatic combinations of items. A third part, formally equal to the grammar rules, can be methodologically distinguished, the *lexicon*, in which the terminal symbols of the main grammar rules are identified. Because we have collected all the information referred above, our first interest when we have to describe the clause in Spanish is the coding of this verbal information in the lexicon. With this aim, our grammar presents (in its actual state) two types of lexical rules to describe verbs:

**verb (VALENCY, TYPE, SSUBCAT, DOSUBCAT, IOSUBCAT):**  
"concrete verb "

**verb (VALENCY, TYPE, PREPTYPE, PCTYPE, SSUBCAT, DOSUBCAT, PCSUBCAT, IOSUBCAT):**  
"concrete verb "

---

<sup>10</sup>An approximate translation for these meanings can be found in appendix 2.

The first one accounts for verbs which do not require a prepositional complement, the second one for the ones which do require such a complement. Affixes contain information about the valency of verbs implied: **VALENCY**, about the number of arguments with which the verbs combine (quantitative valency), **TYPE**, about the functions in which appear these constituents and, perhaps, other facultative arguments or functional segments (qualitative valency). In the actual state of the research, these affixes have these values in the metagrammar:

**VALENCY :: zero; one; two; three.**

**TYPE :: p; s; do; sdo; spc; sio; sdoio; sdopc; siopc.**

The values for valency are clear enough: if a verb combines with zero, one, two or three arguments it is assigned the correspondent value. What is more important (it derives from our concept of quantitative valency) is the fact that this value refers to the *undoubtedly* valential constituents (the ones which comply with the left group of characteristics on table I), except for the internal arguments (subject and direct object, which appear, respectively with impersonal and intransitive verbs<sup>11</sup>). The values for type represent the combination of functions implied: p, none, s, subject, do, direct object, sdo, subject and direct object, spc, subject and prepositional complement, sio, subject and indirect object, sdoio, subject, direct object and indirect object, sdopc, subject, direct object and prepositional complement, siopc, subject, indirect object and prepositional complement. This value is assigned taking into account also *doubtful*, *facultative* and *internal* arguments and *central*, but *not valential* constituents (the ones which comply either with the centre group of characteristics in table I or with the right one with a positive value for the feature *central*)<sup>12</sup>. We are aware of the fact that these values do not cover all the syntactic schemes we can find in Spanish. At the actual state of the research, we have decided to begin with the *more frequent* schemes, and, it is again the database which has given us this information:

TABLE II<sup>13</sup>

VALENCY	%	SCHEME	%	EXAMPLES
0	0,1	P	0,1	Llueve
1	15,5	S	93	Ocurren cosas increíbles/ Juan corre
		DO	7	Hay mucha gente/ Hace calor
2	71	SDO	78	Anacleto visitó a su familia
		SPC	16,5	Recurren a la violencia/Van a la escuela/Pórtate bien
		SIO	5,5	A Rosa le gusta el pollo

<sup>11</sup>Examples: *Llueven chuzos de punta* (impersonal verb with internal subject), *Juan corrió una carrera maravillosa* (intransitive verb with internal direct object).

<sup>12</sup>At this moment, we consider the affix structure for values included within VALENCY and TYPE as provisional, not only for the fact that not all possible values are already calculated, but also because we want to express in the metagrammar the relation between them (for example, the fact that if a verb has *valency* two, it cannot have *type* p, s or do).

<sup>13</sup>This table has been taken, with some changes, from García-Miguel (1995: 38).

VALENCY	%	SCHEME	%	EXAMPLES
3	13	SDOIO	59	Le dio un regalo
		SDOPC	36	Puso los libros en el estante /Lo obligó a abandonar
		SIOPC	3	Les preguntó por sus problemas

The affixes *pretype* and *pctype* contain information about the preposition or prepositions a verb requires, and about the type of prepositional complement implied: prepositional object, locative adverbial or modal adverbial.

The rest of the affixes contain information about subcategorization: *ssubcat*, for subject subcategorization, *dosubcat*, for direct object subcategorization, *pcsubcat*, for prepositional complement subcategorization and *iosubcat*, for indirect object subcategorization. All these affixes are nonterminals within the domain of another one, more general: **SUBCAT**. With this affix structure, in verbs (and, of course, predicate functions) and in subject, direct object, prepositional complement and indirect object functions, we can express for each of them their exclusive possibilities of subcategorization, and, at the same time, we do not have to write for each structure which is able to fill in more than one of them (the noun phrase, for example) a different rule for each function which subcategorize such structure as one of its possibilities.

**SUBCAT :: SSUBCAT; DOSUBCAT; PCSUBCAT; IOSUBCAT.**

**SSUBCAT :: structure1; structure2; ...; empty.**

**DOSUBCAT :: structure1; structure2; ...; empty.**

**PCSUBCAT :: structure1; structure2; ...; empty.**

**IOSUBCAT :: structure1; structure2; ...; empty.**

The value *empty*, which does not find any rewriting rule, is assigned to verbs which do not express any of its arguments as the function qualified as *empty*.

Once all this information is coded in lexicon rules, it seems possible to treat it with a very simple and pure syntax. The clause will be rewritten directly as its constituents: the function **PREDICATE**, and as many constituents, to which a function name is given (**SUBJECT, DIRECT OBJECT, PREPOSITIONAL COMPLEMENT, INDIRECT OBJECT, PREDICATIVE COMPLEMENT**), as the verb involved requires. For example, the rewriting rule which describes the structure of a clause whose predicate is a verb of valency *two* and type *sdopc*, with all its constituents realized in the non-marked order would be:

**clause:**

**SUBJECT (SSUBCAT),  
PREDICATE (two, sdopc, PCTYPE, PREPTYPE, empty, SUBCAT,  
DOSUBCAT, PCSUBCAT, IOSUBCAT),  
DIRECT OBJECT (DOSUBCAT, DOTYPE),  
PREPOSITIONAL COMPLEMENT (PCTYPE, PREPTYPE, PCSUBCAT),  
direct object predicate (DOSUBCAT, DOTYPE).**

The problem now is to calculate all the possible permutations of these constituents and write all the resultant rules. Certainly, if we think of the fact that we have, at least, nine combinations of valency and type, the process seems too long. Really, things are not exactly like this, for two reasons:

1. When some function is not realized in a clause with a PREDICATE whose verb has a high valency, its concrete configuration in such case should be already present in the syntax to describe the complete configuration of a clause whose PREDICATE is filled by a verb with a lower valency.
2. Actually, not all permutations come out with a grammatical sequence, and, what is more interesting, with a *frequent* sequence, a common configuration of arguments in the expression of actions, states and processes. It led us to the idea that, at least, in a first approach, we can use the database to restrict the theoretically wide possibilities of linguistic organization of arguments in the clause.

Functions are rewritten as the categories by which they are realized. The affix SUBCAT keeps information about this concrete subcategorization and checks its correctness with the PREDICATE in the rule for the clause, for example a rule for function SUBJECT would be:

**SUBJECT (SSUBCAT):  
noun phrase (SSUBCAT).**

Finally, categories which can appear in these functions are rewritten only once by means of affix SUBCAT (see above in this section).

**noun phrase (SUBCAT):  
alternative 1 (fnan);  
alternative 2 (fnin);  
...  
alternative n (terminal value for SUBCAT).**

#### **4 A practical example with the verb “abandonar”**

The examination of data for the verb *abandonar*, once we have related meanings (see above 2) and verbal schemes and subschemes (see above 2) gave us the following results:

A) With respect to the meanings<sup>14</sup>: from meanings to schemes,

**1 SPD(aU), 48**

**act., S D, 41**  
    **San Dan, 28 (1)**  
    **San Din, 8 (2)**  
    **Sin Dan, 5 (3)**  
**act., S D AD, 1**  
    **San Dan ADin(en), 1 (4)**  
**act., S D SP, 2**  
    **San Dan SPin(a), 1 (5)**  
    **San Din SPin(a), 1 (6)**  
**pas., S, 2**  
    **San, 2 (7)**  
**pas., S A, 1**  
    **San Aan, 1 (8)**  
**pas., S SP, 1**  
    **Sin SPin(a), 1 (9)**

**2, SPD, 1**

**act., S D, 1**  
    **San Din, 1 (10)**

**2b, SPD, 6**

**act., S D, 5**  
    **San Din, 5 (11)**  
**pas., S, 1**  
    **Sin, 1 (12)**

**2c, SPD, 14**

**act., S, 2**  
    **San, 2 (13)**  
**act., SD, 12**  
    **San Din, 12 (14)**

---

<sup>14</sup>Keys and abbreviations for these tables: italic numbers and letters refer to meanings (see 2), arabic numbers represent frequency, numbers in brackets refer to examples on appendix 1, for the rest of capital and small letters see note 8.

**2d, SPD(A), 46**

**act., S D, 40**

**San Din, 38 (15)**

**Sin Din, 2 (16)**

**act., S D AD, 3**

**San Din ADin(en),2 (17)**

**San Din ADin(sobre), 1 (18)**

**pron., S, 2**

**Sin, 2 (19)**

**pas., S, 1**

**Sin, 1 (20)**

**3, SPD, 73**

**act., S D, 73**

**San Din, 62 (21)**

**Sin Dan, 5 (22)**

**Sin Din, 3 (23)**

**San Dan, 3 (24)**

**4, SsePaU, 2**

**pron., S, 1**

**San, 1 (25)**

**pron., S SP, 1**

**San SPcf(a), 1 (26)**

**5, SsePaU, 6**

**pron. S SP, 6**

**San SPin(a), 6 (27)**

**6, SsePaU, 1**

**pron., S SP, 1**

**San SPin(a), 1 (28)**

B) With respect to schemes: from schemes to meanings,

**Act. S, 2**

**2c, 2 (13)**

**Act. S D, 172**

- 1, 41 (1, 2, 3)**
- 2, 1 (10)**
- 2b, 5, (11)**
- 2c, 12, (14)**
- 2d, 40 (15, 16)**
- 3, 73 (21, 22, 23, 24)**

**Act. S D AD, 4**

- 1, 1 (4)**
- 2d, 3 (17, 18)**

**Act. S D SP, 2**

- 1, 2 (5, 6)**

**Pron. S, 3**

- 2d, 2 (19)**
- 4, 1 (25)**

**Pron. S SP, 8**

- 4, 1 (26)**
- 5, 6 (27)**
- 6, 1 (28)**

**Pas. S, 3**

- 1, 1 (7)**
- 2b, 1 (12)**
- 2d, 1 (20)**

**Pas. S A, 2**

- 1, 2 (8)**

**Pas. S SP, 1**

- 1, 1 (9)**

From data, with respect to active voice, we can infer the following general frame which could support actual organization of arguments for the verb *abandonar*. This deduction is based on some observations:

- For all meanings (1, 2, 3), we find the more frequent scheme to be **S D**.
- Two of them (1, 2) can add a prepositional complement (prepositional object in the first case, adverbial in the second one).
- The categories that we find in these functions are noun phrases.
- These noun phrases are animate or inanimate in the functions **S** and **D** (with the only exception of meaning 2, in which direct object is always inanimate), the prepositional complement is always inanimate.

TABLE III

Meanings	Subject	Direct object	Prepositional complement
1	San/Sin	Dan/Din	SPin(a)
2	San/Sin	Din	ADin(en)
3	San/Sin	Dan/Din	
General frame	San/Sin	Dan/Din	SPin(a)/ADin(en)

In the lexicon we should have the following two rules<sup>15</sup>:

**verb (two, sdopc, a, prepositional object, fnan/fnin, fnan/fnin, fnin, empty):**  
**“abandonar”.**

**verb (two, sdopc, de, adverbial, fnan/fnin, fnan/fnin, fnin, empty):**  
**“abandonar”.**

On the other side, with respect to order we have obtained the following data<sup>16</sup>:

Act.

V C, 83  
 S V C, 24  
 V, 24  
 S V, 8  
 V C (AD/SP), 2  
 V S C, 2  
 V S, 2  
 V (AD/SP), 2  
 V C S, 1  
 S V (AD/SP), 1  
 V (AD/SP), 1

<sup>15</sup>They can not be reduced to only one rule, because we have to preserve the strict relation between each preposition and the prepositional complement type it selects.

<sup>16</sup>Relative, interrogative and exclamative clauses have not been included in this study about order, because they imply a fixed order of some elements. Also, as, in the actual state, our interest is mainly concerned with arguments, we have eliminated circumstantials from these schemes. C stands here for direct object, V for verb.

Pron:

**V (AD/SP), 7**  
**S V, 2**  
**V S, 1**

Pas.:

**V A, 2**  
**V S, 1**  
**V (AD/SP), 1**  
**S V, 1**  
**V, 1**

Restricting ourselves to active voice again, the concrete sequences of functions found for the verb *abandonar* can be reduced to the following general frame,

**S V S C S (AD/SP) S,**

in which any constituent, or all of them, can be absent, except, of course for the verb) and only one **S** can occur. It implies the writing of different syntactic rules for each combination **V, S V, V S, S V C, V S C, V C S...** And, even we know that data in this case do not represent all possibilities for Spanish: for example, **C** (or any other constituent different from **S**) is never thematized in first position. Nevertheless, at this state of the research, we have preferred to avoid problems of order related to thematization, because we expect to treat them in the future in a more systematic way than including all possible sequences in which a constituent different from the subject is thematized.

But, even restricting ourselves also in this area, *abandonar* continue to pose some problems. As we already know, any constituent can be absent. Among those, this fact is specially important for the prepositional complement, because, always when we will encounter a verb which subcategorizes different prepositions for different prepositional complements, and if these ones are not realized, we will obtain as much analyses as they can subcategorize prepositions. The solution, at least, for verbs with valency two, which have an optional or doubtful prepositional complement, can be the introduction of a (in the case of *abandonar*) third rule in the lexicon, the basic rule which account for verbs which do not select a prepositional complement (and which describes the basic behaviour of these type of verbs):

**verb (two, sdo, fnan/fnin, fnan/fnin, empty):**  
**“abandonar”.**

In the syntax, we have to separate the rules in which the prepositional complement can appear (they must present it as an *obligatory* member) from the ones in which it does not appears, and, consequently the prepositional selection is not functional any more. It means that, when we will have a clause as the one in example 13, in which only the predicate is realized, the grammar will apply a rule as:

**clause:**

**PREDICATE (two, sdo, FUNCTION, SSUBCAT, DOSUBCAT, IOSUBCAT).**

Instead, *only* when at least the prepositional complement is realized, a rule like the one immediately following (or some of its permutations, with or without subject or direct object) will be applied:

**clause:**

**SUBJECT (SSUBCAT),  
PREDICATE (two, sdopc, PCTYPE, PREPTYPE, empty,  
SSUBCAT, DOSUBCAT, PCSUBCAT, IOSUBCAT),  
DIRECT OBJECT (DOSUBCAT, DOTYPE),  
PREPOSITIONAL COMPLEMENT (PCTYPE, PREPTYPE, PCSUBCAT),  
direct object predicate (DOSUBCAT, DOTYPE).**

We have not still developed a part in the grammar to deal with middle and passive voices. Anyway, as we consider voices as *grammatical means* to change the valency of verbs, we will treat them in a particular module of syntax, which will ultimately call for the same lexical rules as the active voice does. But this is still under development and we cannot, at this stage, enter more deeply in this issue.

## References

Báez, Valerio (1988), *Fundamentos críticos de la gramática de dependencias*, Síntesis, Madrid.

Dekkers, Chris, Cornelis H. A. Koster, Mark-Jan Nederhof, Arend van Zwol, (1992) "The grammar WorkBench: A First Step towards Lingware Engineering", *Proc. of the second Twente Workshop on Language Technology, Memoranda Informatica 92-29*, Twente, pp. 103-115.

Farkas, Ernő, Cornelis H. A. Koster, Peter Köves, Mátyás Naszódi, (1991), "Towards an Affix Grammar for the Hungarian Language", in A. Gyoergy (ed.), *Proceedings CIS'91 conference on intelligent Systems*, John von Neumann Society for Computing Sciences, Budapest, pp. 233-246.

García-Miguel, José M<sup>a</sup>. (1995), *Las relaciones gramaticales entre predicado y participantes*, Universidad de Santiago de Compostela, Santiago de Compostela.

Jiménez Juliá, Tomás (1993) "La caracterización de las unidades sintácticas en español", *Actas del III Congreso de Hispanistas de Asia*, Tokio.

Koster, Cornelis. H. A. (1991) “Affix grammars for Natural Languages” in H. Alblas, B. Melichar (eds.), *Attribute Grammars, applications and systems*, Springer Lecture Notes in Computer Science 545 , Heidelberg, pp. 469-484.

Moreno Cabrera, Juan C. (1987), *Fundamentos de sintaxis general*, Síntesis, Madrid.

Rojo, Guillermo, Tomás Jiménez (1989), *Fundamentos del análisis sintáctico funcional*. *Lalia*, 2, Universidad de Santiago de Compostela, Santiago de Compostela.

Rojo, Guillermo (1983), *Aspectos básicos de sintaxis funcional*, Ágora, Málaga.

\_\_\_\_\_ (1990), “Sobre los complementos adverbiales”, *Jornadas de Filología/Profesor F. Marsá*, Universidad de Barcelona, Barcelona, 153-171.

Santalla del Río, M<sup>a</sup>. Paula (1994), “Spanish noun phrase in AGFL formalism”, published in the AGFL homepage ([www-agfl@cs.kun.nl](http://www-agfl@cs.kun.nl)).

Vázquez Rozas, Victoria (1995), *El complemento indirecto en español*, Universidad de Santiago de Compostela, Santiago de Compostela.

## Appendix 1

1. Por primera vez sentí que me habías abandonado.(SUR 27)
2. [...] uno abandona una cosa [...] (Madrid 132)
3. [...] como si la vida la hubiera abandonado por unos instantes [...] (SUR 61)
4. [...] que abandonó a su hijo de cuatro meses el día 31 de octubre en un descampado de Zaragoza [...] (3VOZ, 18, 6, 1, 3)
5. [...] y abandonándome a la duda de tu verdadera identidad. (SUR, 82)
6. [...] quienes, procedentes de inmuebles contiguos, han abandonado también precipitadamente sus bienes y enseres a la súbita e incontenible marea de aquellos odiosos bichos.(PAISAJES, 163)
7. [...] donde fue abandonada. [...] La mujer agredida [...] (2VOZ, 28, 4, 1, 8)
8. [...] una persona muere [...] Abandonada de su familia [...] (MADRID, 137)
9. [...] dicho departamente decidió impulsar la sección de motoristas, abandonada a su suerte hace años [...] (3VOZ, 35, 5, 1, 4)
10. Parecía un hombre mayor y cansado, sus brazos le caían a lo largo del cuerpo como si los hubiera abandonado. (SUR, 75)
11. Casi en seguida abandoné una idea de vías muertas o socavones abandonados [...] (GLENDA, 50)
12. [...] cuando se descubre que ciertas convenciones son abandonadas por haber dejado de ser valiosas [...] (LING 112)
13. Si abandonamos, el título es de Kankkunen.(2VOZ, 49, 1, 6, 11)
14. La muerte de su padre lo había forzado a abandonar los estudios al término de la escuela secundaria [...] (CRONICA, 15)
15. [...] porque ella, Suzanna Trash, [...] abandonando la forzada postura de karateka que para escucharme había adoptado [...] (LABERINTO, 69)
16. [...] sugiero que la psicolingüística debería abandonar el mentalismo [...] (LING, 97)
17. La ninfa de la cama se levantó y abandonó en el suelo el cesto casi vacío. (JOVENES, 156)
18. Al tiempo, abandona suavemente el cigarrillo sobre la mesa [...] (CAIMÁN,43)
19. Aparte de eso, la agricultura cada día se abandona más. (SEVILLA, 93)
20. En este mismo puerto, además, fueron abandonados 375 kilos de este pescado. (3VOZ, 41, 3, 2, 11)
21. Habíamos abandonado el camino [...] (SUR, 89)
22. [...] aunque no me hayan abandonado hasta el día de hoy los espasmos ni la náusea ni el alarido que siempre se me escapa cuando por la televisión pasan un anuncio de Iberia. (LABERINTO, 28)
23. La nube había abandonado el sol [...] (JÓVENES, 173)
24. El padre los abandonaba, se alejaba de ellos. (JÓVENES, 41)
25. [...] Flora abándonandose con un primer llanto [...] (GLENDA, 99)
26. O tal vez no le inspira confianza alguna abandonarse a serlo con arreglo a la libre decisión del predicado [...] (RATÓN, 45)
27. Pues yo me abandonaba a una quietud absoluta [...] (SUR, 39)
28. La rutinaria admonición de no abandonarse a las instituciones [...] (RATÓN, 103)

## **ARTHUS corpus source texts for the examples**

- Aldecoa, Josefina R., *Por qué éramos jóvenes*, Seix Barral, Barcelona, 1986.
- Buero Vallejo, Antonio, *Caimán*, Espasa-Calpe, Madrid, 1981.
- Bunge, Mario, *Lingüística y filosofía*, Ariel, Barcelona, 1983.
- Cortázar, Julio, *Queremos tanto a Glenda*, Alfaguara, Madrid, 1981 (4<sup>a</sup> edition).
- Esgueva, Manuel, Margarita Cantarero (eds.), *El habla culta de la ciudad de Madrid. Materiales para su estudio*, CSIC (Miguel de Cervantes), Madrid, 1981.
- García Márquez, Gabriel, *Crónica de una muerte anunciada*, Mondadori, Madrid, 1987.
- García Morales, Adelaida, *El Sur (seguido de Bene)*, Anagrama, Barcelona, 1985.
- Goytisolo, Juan, *Paisajes después de la batalla*, Montesinos, Barcelona, 1982.
- La Voz de Galicia*, 22/11/91.
- La Voz de Galicia*, 23/11/91.
- Mendoza, Eduardo, *El laberinto de las aceitunas*, Seix Barral, Barcelona, 1982.
- Pineda, Miguel Ángel de (ed.), *Sociolingüística andaluza 2. Material de encuestas para el estudio del habla urbana culta de Sevilla*, Servicio de Publicaciones, Universidad de Sevilla, 1983.
- Sánchez Ferlosio, Rafael, *La homilía del ratón*, El País, Madrid, 1986.

## Appendix 2

1. [SPD(aU)] *To desert, abandon, leave something or somebody(D), to abandon something or somebody (D) to something (U)* (÷ DALE 1, DRAE 1, DUE 1) [24%].
2. [SPD] *To release something(D), to relax the tension of something (D)* (÷ DUE 2) [1%].
  - 2b. [SPD] *To give up an idea, a project or a plan (D)* (÷DALE 2, DRAE 2, DUE 2) [3%].
  - 2c. [SPD] *To give up something already initiated (D)* (÷ DALE 2, DRAE 2, DUE 2) [7%].
  - 2d. [SPD(A)] *To give up using something(D), to give up something, to leave something aside(D), to give up, to leave something (D) somewhere (A)* (DALE 3, DUE 2) [23%].
  - 2e. [SPD] *To leave the insured person(S) to the attention of the insurer [...] things insured (D) with the scope of being paid* (÷ DALE 5, DRAE6) [0%].
3. [SPD] *To separate from something or somebody(D), to leave a place(D)* (÷ DRAE 3, DUEp 3) [37%].
4. [S seP] *To stop doing something, some effort, to lean back negligently, with laziness* (÷ DALE 4, DRAE 4, DUE 1) [1%].
  - 4b. [S seP] *To decline, to decay, to discourage* (÷ DALE 9, DRAE 10, DUEp 3) [0%].
  - 4c. [S seP] *To stop taking care of one's obligations, one's cleanliness or tidiness* (÷ DALE 7, DRAE 8/9, DUEp 2) [0%].
5. [S seP aU] *To take to something (U), to be ruled by something (U)* (÷DALE 6, DRAE7, DUEp 3) [3%].
6. [S seP aU] *To devote oneself to something or somebody (U), to entrust oneself to something or somebody (U)* (÷ DALE 8, DRAE 5, DUEp 3) [1%].