

PERGA.DOC 900925/901106/910330

## THE LEXICON GRAMMAR OF A LANGUAGE APPLICATION TO FRENCH

Maurice Gross

Université Paris 7  
Laboratoire d'Automatique Documentaire et linguistique  
Institut Blaise Pascal, CNRS

1. Principles
2. Classification of verbs
3. Frozen sentences
4. Support verbs
  - 4.1. Nominalization
  - 4.2 Stand-alone nouns
  - 4.3 Other nouns
  - 4.4 Adverbs
5. Results and conclusions

Lexicon-grammar is a model of syntax limited to the elementary sentences of a natural language. We present its main theoretical features, mostly inferred from large-scale empirical studies aimed at obtaining a large coverage of French. French is the earliest application, but other studies, in varied languages, have also been performed (cf. 5).

### 1. PRINCIPLES

The major principle is that the unit of meaning is not located at the level of the word, but at the level of elementary sentences; we will see how coherent descriptions of natural languages have been built on this basis.

The point of departure was the construction of a transformational grammar of French (M. Gross, 1968, 1975), based on Z.S. Harris' theoretical framework. Adding the requirement of providing a significant coverage of the language led to the development of:

- an extensive list of transformations, and
- a lexicon of verbs (or elementary sentences) to which these transformations applied.

The intended programme consisted in marking each verb (or sentence) with the transformations it underwent. This approach is necessary for obvious reasons, even though there has been little theoretical or practical discussions of the basic phenomena. Consider for example, the traditional notion of transitive verb, or equivalently, the notion of a sentence with a direct object. Attempts to formalize it lead first to the basic structure:

N0 V N1 of which examples are:

- (1) Bob enjoyed all his vacation
- (2) Bob rested all his vacation

The main notations are the following: in a sentence, the noun phrases (NPs) are numbered from left to right starting with the subscript zero: N0 is the subject, N1 is the first complement, N2 is the second complement, etc.

Both sentences (1) and (2) have a direct complement, but only (1) has a direct object. Transformational properties show (define) the difference:

Interrogative pronoun formation:

- (3) --What did Bob enjoy? --All his vacation
- (4) \*--What did Bob rest?--All his vacation

Relative pronoun formation:

- (5) The vacation which Bob enjoyed (was lost time)
- (6) \*The vacation which Bob rested (was lost time)

Passive formation:

- (7) All his vacation was enjoyed by Bob
- (8) \*All his vacation was rested by Bob

Just one of these three properties is enough to discriminate (1) from (2) and hence could provide a basis for a syntactic definition of direct object. But the original intention of the notion of object seems to be also semantic, hence it usually associates meaning and form. Then a question arises: do these three syntactic properties coincide as in the case of to enjoy? Moreover do they coincide with some semantic notion? In fact, if we consider the sentence:

- (9) Bob spent all his vacation sleeping

we observe only two possibilities out of three a priori relevant:

[Relativization]: All the vacation which Bob spent sleeping (was lost time)  
[Passive]: All Bob's vacation was spent sleeping

and one impossibility:

[Wh Question]: \*What did Bob spend sleeping?

Since there is no theoretical reason why the three properties should coincide, they ought to be described independently of each other. Notice that relativization and WH-Question could be correlated; for example, if a common basic pronoun (e.g. something) could be shown to be the source of both pronouns what and which. Such a hypothesis needs to be strengthened, which can only be done through a thorough examination of the verbs. There is no hint whatsoever that Passive forms have a grammatical link to the two other properties.

Consider another example. In our notations, the non-oriented Passive relation can be written:

(10)  $N0 V N1 = N1 \text{ be } V\text{-ed by } N0$

Now, it is largely assumed that the agent complement by  $N0$  can be omitted, that is, the resulting structure  $N1 \text{ be } V\text{-ed}$  always corresponds to an acceptable sentence:

= Bob briefed Max  
= Max was briefed by Bob  
= Max was briefed

However, there are paradigms, such as:

= This assumption underlies Bob's whole discourse  
= Bob's whole discourse is underlied by this assumption  
\*Bob's whole discourse is underlied

where omission of the by-agent is not possible and examples where the same omission has aspectual implications, as can be seen in the pair:

Max is enjoyed      Max is rejoiced

There is no way to detect and describe these patterns of relation, except by looking individually at each verb and marking them.

This situation is fairly general, given a structure  $N0 V W$ , where  $W$  is a sequence of complements: the transformations that apply to this structure are found to be dependent on the choice of the verb  $V$  (hence on the complements  $W$ ).

We can now be more explicit about the form of a lexicon-grammar. First, we describe the syntax of ordinary verbs, which means that we describe verbs in their minimal environment: subjects and essential complements. Separating essential from circumstantial complements is a classical question, for which the traditional answer is largely based on the shape of interrogative pronouns. Considering again examples (1) and (2), we observe the difference:

--\*When did Bob enjoy      All his vacation

-- When did Bob rest            All his vacation

Upon examination of more than 10.000 verbs, delineating a border between essential and circumstantial complements has proved to be a varied enterprise, more dependent on individual verbs than initially thought: tests for characterizing essential complements are highly lexical and tend to apply more to individual verbs or small groups than to broad semantic classes.

The choice of verbs (over 10.000) and of transformational properties (over 500) is an interdependent process performed by a team of linguists who have to carry out acceptability judgments on sentences, which included:

- deciding whether several sentences had the same meaning or transformationally related meanings, (e.g. for example (1), (3) and (7) or for (2) and (9));
- deciding whether two sentences had unrelated meanings (e.g. (1) and (2));
- deciding whether a sentence form was acceptable or not (e.g. for (4), (6), (8)).

The linguistic activity aims:

- (i) at separating the meaning or uses of the morphological verbs,
- (ii) at marking syntactic verbs (i.e. elementary sentences) for the properties they do or do not have.

A crucial requirement in such descriptive work is REPRODUCIBILITY of judgments. Only when a consensus was reached by the team of linguists, was a verb retained for study in the lexicon, and only when the evaluation of a syntactic shape could be reproduced by the team, was this syntactic shape retained as a property (a rule) of the grammar. Judgments are binary: a verb has or does not have a given property. This ideal representation cannot be reached directly. For example, it has been observed in written corpus that intuitions of acceptability underestimate the actual production of forms. The effect has been corrected by a rule of thumb: when acceptability is dubious, the form is considered acceptable.

More precisely, we consider verbs, that is, elementary sentences of the form subject-verb-essential complements, as elementary units for syntactic as well as for semantic composition.

The set of elementary sentences is structured by (Z.S. Harris 1968) transformation rules, construed as equivalence relations between sentences.

## 2. CLASSIFICATION OF VERBS

The principle of all classifications is formal, namely, it is based on reproducible categories such as the traditional parts of speech. Let us first consider a rather simple example: the classification of English frozen adverbial phrases. We are going to use the categories Prep (preposition), Det (determiner), N (noun), Adj (adjective). We distinguish the following types:

- Prep N,
- Prep Det N,
- Prep Det N Prep N, etc.

which lead respectively to the construction of the corresponding tables:

-table PC:

Prep	N
above	all
by	chance
in	effect
for	example
in	fact

- table PDC:

Prep	Det	N
at	this	point
in	this	respect
by	the	way
in	a	way
on	the	whole

- table PCPN:

Prep	Det	N	Prep	Nh	N-h	PDC
at	the	beginning	of	-	+	+
on	-	behalf	of	+	-	-
by	-	comparison	with	+	+	+
-	-	thanks	to	+	+	-
by	-	virtue	of	-	+	-

These last adverbial phrases are semi-frozen: their first part Prep Det N Prep is fixed and listed in the table, whereas the remaining part is a variable noun phrase which is described by means of binary properties; its nouns can be '+' or '-' human (i.e. column Nh), or non human (i.e. column N-h). The second part PrepN can be omitted or not, depending on each expression, this possibility is indicated in the last column named PDC.

Other such tables have to be defined, in order to reach a complete coverage of this family of adverbial complements. For example, the table PAC would contain the structure:

Prep Det Adj N =: by the same token + in the long run + etc.

Let us now describe the classification of verbs, or rather of elementary sentences. The general approach to the construction of classes will be the same as in the preceding example.

All verbs have a subject N0, but their sequence W of essential complements is quite variable. Thus, the examples of verbs discussed above have the following basic declarative structures:

N0 V                                =: Bob rested  
 N0 V N1                            =: Bob enjoyed his vacation  
 N0 V N1 (V1-ing W)<sup>2</sup>        =: Bob spent his vacation sleeping

The notation of the last structure is with a second complement subscripted 2 and made of a verbal structure whose subject is indicated by the superscript 1, that is N1.

The classification is based on several parameters, the main ones being the number of complements and their shape. For French, we had to consider first three basic shapes: direct (i.e. "zero" preposition), indirect with prepositions À and de. It turns out that other prepositions (Prep) introduce few essential complements. We give in table 1 the typology on which the classification of French verbs has been made (J.-P. Boons, A. Guillet, C. LeclÈre 1976a, 1976b; M. Gross 1975; A. Guillet, C. LeclÈre 1990; C. LeclÈre 1990).

Structures	Examples	Numbers of entries
N0 V	Guy fl <sub>ne</sub>	450
N0 V N1	Max boit du vin	3700
N0 V À N1	Luc pense À Jo	350
N0 V de N1	Guy rêve de trains	400
N0 V PrÉp N1	La mer compte pour Max	1000
N0 V N1 N2	On nomme Luc chef	60
N0 V N1 À N2	Guy donne un lit À Max	
900		
N0 V N1 de N2	Max reçoit un lit de Guy	1300
N0 V N1 PrÉp N2	Max pose le lit sur le sol	1700
N0 V À N1 À N2	Le lit sert À Luc À cela	5
N0 V À N1 de N2	Le lit sert À Luc de bureau	10
N0 V À N1 PrÉp N2	Ce texte Équivaut À une insulte pour Jo	30
N0 V PrÉp N1 PrÉp N2	Max s'allie avec Luc contre Guy	200

There are practically no examples with three objects; one candidate is the verb parier (to bet) in the sentence:

Jo a pariÉ cent francs avec Luc que Max partirait  
 (Jo bet one hundred francs with Luc that Max would leave)

## Table 1

A second parameter of the classification is the content of the noun phrases  $N_i$ ; two types have been distinguished:

- sentential, that is accepting one of the forms queS or V-infW (infinitive clause), then, predicative nouns are also possible in general,
- non sentential, that is only nouns.

This classification results in the distribution shown in table 2, where the main structures are displayed with the corresponding number of verbs.

---

The structures are more specific than in table 1. They each define a syntactic class whose description is given in the form of a binary matrix (cf. figure 1).

## Table 2

---

Such a classification has to be completed by the properties that depend on individual verbs. This is done in a uniform way. Each row of the matrix is an entry of the lexicon-grammar, that is, an elementary sentence in the declarative form defining a class (e.g. table 2). Each column is a particular sentence shape, into which the verb may or may not enter. A "+" or a "-" sign is placed accordingly at the intersection of the corresponding row and column. The main properties are:

- distributional properties of Ns, such as human or non human. They are defined by means of interrogative pronouns. Consider the entry N0 influence N1. Both subject and complement are marked +Nhum, +N-hum, as the result of accepting the following sentences:

(What + Who) influences Bob?  
(What + Who) does Max influence?

Columns containing the specifications  $N_i =: Nhum$ ,  $N_i =: N-hum$  correspond to sentence forms such as:

Who V N1?, What do N0 V?

Hence, distributional properties are defined as sentence shapes;

- transformational properties can be seen as sentence shapes associated with the declarative form that defines the class. Hence, the passive transformation is noted as accepted by a verb of the class N0 V N1, when the matrix has a "+" sign in the column that contains the simple form:

N1 be influenced by N0

The general description of verbs can then be visualized as a single matrix of 10,000 rows by 500 columns. There is considerable redundancy in this matrix: for example, when a verb is intransitive, marking "-" the properties of its objects carries no information. Hence, the classes allow a more tractable representation, by means of submatrices of a smaller size: they have an average of 250 rows (verbs) and 25 columns. A sample of a class is given in figure 1.

---

Figure 1

---

### 3. FROZEN SENTENCES

Frozen sentences (we could use alternatively the terminology "compound verbs") include elementary sentences described from a variety of stylistic points of view, but all sharing one feature. They include verb-noun combinations which are not distributional, that is, not semantically predictable:

- idioms: Max took the bull by the horns,
- metaphors: Max burned his bridges,
- clichÉ: The house burst into flames,
- technical: Max integrated the function by parts.

We replace this non-reproducible classification by using distributional criteria. Consider for example the common sentence type:

N0 V N1 Prep N2

In the same way as we subclassified verbs by separating them according to their acceptance of a sentential Ni, we describe the sentences according to the position of the nouns frozen with respect to the verbs; we note Ci the frozen nominal positions. We can then write:

N0 V C1 Prep C2 =: Max took the bull by the horns  
N0 V N1 Prep C2 =: The news took Bob by surprise  
N0 V C1 Prep N2 =: Bob took part in the inquiry  
C0 V C1 Prep C2 =: Money burns a hole in Bob's pocket

The last example should be noted more precisely as follows:

C0 V C1 Prep (N's C)2

since a variable human noun N must be attached to the frozen head noun pocket. Table 3 gives the present state of the description of frozen sentences.

---

In general, the tables named Cx contain verbs different from the "auxiliary" verbs être (to be), avoir (to have), faire (to make). These verbs are described in other tables named respectively Ex (L. Danlos 1979; J. Labelle 1974; A. Meunier 1977, 1981; D. de NÉgromi 1978; R. VivÈs 1983), Fx (J. Giry-Schneider 1978, 1987). Also cf. 4.

### Table 3

---

The description of frozen sentences does not differ from the description of the free sentences discussed in 2. Figure 2 contains an example of a syntactic table which is to be read in the same way as figure 1. Frozen sequences appear explicitly.

---

### Figure 2

---

## 4. SUPPORT VERBS

### 4.1 Nominalizations

We further justify our fundamental principle according to which it is elementary sentences and not words that carry units of meaning. The frozen sentences are clearly a case of meaning not being located in individual words, but in whole sentences. That meaning is not compositional is obviously true for other categories of frozen or compound terms:

- adverbs, such as time and again, by and large,
- compound nouns, from the idiomatic red herring to technical terms such as cathode ray tube or transgenic mouse.

The category of frozen adjectives can also be defined for examples such as:

Bob is well-to-do  
The decision is clear-cut

it does not differ essentially from verbs.

Both of the major categories noun and adverb raise a question about the localization of meaning: these expressions do appear to contain their meaning, hence our basic assumption

would not hold for them. There are however reasons to describe them as elementary sentences.

Consider the relation of nominalization as defined by Z.S. Harris 1964, it is a relation between sentences, not between a sentence and a noun phrase. We have for example:

- (11) This proposal contradicts your ideas  
= (12) This proposal is in contradiction with your ideas

We consider this pair similar to the pair active-passive: the verb has become a noun (in the passive it becomes adjective-like), the preposition with has been introduced (by in the passive form), and an auxiliary verb to be in has been inserted (to be in the passive). Hence Passive and this rule of Nominalization differ formally only by the fact that word order is not affected in this Nominalization. In other nominalizations, the word order can change:

- That Jo left satisfied Bob  
= Bob had the satisfaction that Jo left

Another example is adjective nominalization:

- (13) This office is convenient  
= (14) This office has a certain convenience

Since it is not oriented, the relation (13) = (14) could also be called Adjectivization of a noun, and we can look at Passive forms as adjectivized verbs. We call the auxiliary verbs to be (in, into), to have support verbs. Support verbs may have variants such as in:

- (11) = (15) This proposal (comes in + remains in + stays in + enters into) contradiction  
with you ideas  
= (16) There is a contradiction between this proposal and your ideas  
  
(13) = (17) This office (retains + keeps) a certain convenience  
= (18) This office lost its convenience

where the support verbs to enter, to keep, to lose are not semantically minimal: they carry aspect or other modalities. More complex examples are:

- (19) Her remark puts this proposal in contradiction with your ideas  
(20) Jo's appointment gave some convenience to this office

and to put, to give (together with their causative or agentive subject) are called operator verbs on sentences with support verbs. Since operator verbs apply to elementary sentences, they leave invariant the relations holding in the original sentence: (19) and (20) contain respectively (15) and (16), in some sense. In the same way applying the causative operator N0 put to the sentence:

- (21) The child is at Bob's mercy

where various relations holds between mercy, child and Bob, we obtain:

The accident put the child at Bob's mercy

that contains all the relations of (21). Other operators introduce new relations. For example, the binding operators N0 (have + keep), also applied to (21), lead to:

Bob (has + keeps) the child at his mercy

where the subject of have, keep, must be bound to his:

\*Bob (has + keeps) the child at your mercy

The study of relations between the various parts of elementary sentences has not been much developed, so far in syntax. Lexical decompositions such as those of generative semantics are such attempts. But one is dealing here with a quite general problem that may go beyond the study of operator verbs as exemplified above. The following examples show that the overall syntactic structure does not reflect the basic relations observed between parts of elementary sentences:

Bob attributed the crime to Max  
Bob confessed the crime to Max  
Bob described the crime to Max

Max is the subject of crime with to attribute, Bob is the subject with to confess, Max and Bob have no relation to the crime with to describe. Notice that other lexical items and syntactic structures may be used to express the same relations, such as:

Bob accused Max of the crime  
Bob informed Max of the crime

Support and operator verbs are quite different from ordinary verbs:

- they do not select their subjects and objects in the way ordinary verbs do; thus, it is clear that in (12) and (15), proposal is not selected by to come or to enter, but by the noun contradiction, the same is true for to have, to keep, to lose in (14), (17), (18), where it is convenience that selects the subject office;
- their complement structure is different.

Consider for example the sentences:

(22) Bob opposed John's scheme  
= (23) Bob took opposition to John's scheme  
(24) Bob discussed opposition to John's scheme

(23) is a nominalized form of (12) with support verb to take; (24) contains an ordinary verb, (24) has a very different nature, although its structure is identical to the structure of (23):

- semantically (22) and (23) are synonymous, while in (24) Bob is not the subject of opposition. In fact, although the sequences opposition to John's scheme appear to be identical in (23) and (24), we observe differences with respect to extractions:

It is to John's scheme that Bob took opposition  
\*It is to John's scheme that Bob discussed opposition

Opposition was taken to John's schema  
\*Opposition was discussed to John's schema

#### 4.2 Stand-alone nouns

Support and operator verbs also apply to stand-alone nouns, that is, they are not only auxiliaries in nominalization relations. The following pairs involves a nominalization relation:

Bob (murdered + assassinated) someone  
= Bob committed (a murder + an assassination)

where the support verb is to commit. The following sentence:

(25) Bob committed a crime

has exactly the same properties as the nominalized forms, but crime has no corresponding verb. We call crime a stand-alone noun, with respect to nominalization of verbs. Notice that (25) is to be related to another nominal sentence with support verb to be:

Bob is a criminal

In sentences such as:

Bob's proposal is at variance with John's scheme  
Bob (is + goes) at war against John's scheme

Bob is under pressure  
= Jo's arrival put Bob under pressure

we observe the same properties of the verbs, that is, they are support verbs, and the nouns variance, war and pressure are not (no longer) related to verbs similar in form (to vary, to press, to pressurize). This observation is crucial from a theoretical point of view, since it provides a sentential framework for the description of nouns. We have argued that it is a necessary form of description for various reasons:

- noun complementation which includes shape of prepositions and distribution of their noun arguments is defined at the level of support verbs. When a complex noun phrase is combined with an ordinary verb, as in (24), its properties are derived by systematic rules from the elementary structure with support verb, not from the verbal structure;
- nominalizations, as we extended them, increase the syntactic coverage of equivalence classes of elementary sentences, hence extends the syntactic variety that hides the meaning of sentences. Such equivalence classes provide a new approach to derivational morphology: words will enter a morphological family only if they occur in sentences related by relations of nominalizations and adjectivizations;
- various frozen (idiomatic) forms can only be related by support or operator verbs:

This story has neither head or tail

- = The story is without head or tail
- = Bob cannot make head or tail of this story

- This story broke Bob's heart
- = This story was heart breaking to Bob

It is interesting to observe that ambiguous nouns have different support verbs for each of their meanings: consider the noun order in the following sentences:

- (26) The room is in order
- (27) For Bob to leave at 2 pm is in order
- (28) Bob (is in + took) orders
- (29) The machine (is in running + out of) order
- (30) The books have a certain order on these shelves
- (31) The police kept order
- (32) Bob gave Max a clear order
- (33) Bob placed an order of books with Max

Furthermore, the noun order appears in other forms with other functions and meanings:

- determiner in: (of + on) the order of one thousand,
- adverbial in: till further orders,
- subordinating conjunction in: in order (that he + to) succeed,
- and in many compound terms:

- frozen sentences:
  - Bob called Max to order
  - Bob makes shirts to order
  - Pay to the order of BMC...

- compound nouns: alphabetical order, lexicographic order, order of magnitude, law and order, etc. some of which combine with support verbs.

All the meanings of order are different, but placing order in an elementary sentence with a support verb or in other characteristic contexts eliminates ambiguities.

#### Remark

A seemingly simple rule such as plural formation can be shown to have been overlooked in numerous cases. Consider the plural of order: in sentences (26), (27), (29), (31), order has no plural, in (28) it has obligatory plural, in (30), (32), (33) it may have the plural orders. Such cases of syntactic plurals of nouns are quite numerous, they are described in a natural way at the syntactic level where they co-occur with support verbs.

From an empirical point of view, selecting or discovering the (main) support verb of a given noun has proven to be possible and reproducible for most abstract nouns (i.e. predicative, relational, etc.). The question does not seem to be relevant to concrete nouns, at least in the same sense.

#### 4.3 Other nouns

We have mentioned the problem of the integration of concrete nouns in a lexicon-grammar. The difficulty is in the choice of a support verb. Whereas with abstract nouns, a basic support

verb imposes itself most of the time, with concrete nouns there are too many candidates. Given the noun chair, the question is whether we should introduce it through classificatory sentences such as:

A chair is a seat  
A seat is a piece of furniture  
A chair is a piece of furniture

or whether we should consider a physical description of a chair made in terms of elementary sentences:

A chair has a back  
A chair has four legs, etc.

or a functional description:

A person sits on a chair, etc.

or whether one should enter all this encyclopedic knowledge in the lexicon-grammar at all. These difficulties do not differ essentially from traditional semantic questions about synonymy, hyperonymy, etc., except that they take on a different form when they are stated in terms of elementary sentences rather than in terms of words. For example, the synonymy of words has to be relativized as soon as the words are presented in sentences: we do not discuss semantic relations (e.g. synonymy between two nouns) outside of elementary sentences where they occur, hence we only talk of relations between words embedded into a given set of elementary sentences, hence we compare sentences such as the following, in order to make explicit the situations in which both nouns chair and seat are synonymous:

Take a chair!  
Take a seat!

One sits on a chair  
One sits on a seat  
One sits on a piece of furniture

There are five hundred seats in this concert hall  
There are five hundred chairs in this concert hall

Bob put the chairs around the table  
Bob put the seats around the table

A lexicon of compound nouns has been built (G. Gross 1988, M. Silberztein 1989). This lexicon has been classified according to the internal constitution of the nouns, as shown in table 4.

The intended coverage of this lexicon is of the order of magnitude of 300, 000 entries, which would give it the same level of coverage as the one reached for the electronic lexicon DELA of simple words that has already been built (B. Courtois 1989, 1990). Among compound nouns, there are abstract nouns such as first degree murder, economic future which will have to be described with their support verbs in the lexicon-grammar:

N0 commits a first degree murder  
N0 has a certain economic future

There are also numerous concrete nouns including technical terms: boat basin, electron microscope, resolving power, light-collecting mirror, short range nuclear missile. Compound nouns have been classified according to their internal structure, examples are shown in table 4.

---

Table 4

---

The dictionary DELA of simple words contains 80, 000 word forms as entries. Information about inflections is associated to each of them, which allows for automatic generation of over 700, 000 inflected forms together with the corresponding grammatical properties (B.Courtois1990) and a phonemic representation (E. Laporte 1988). This dictionary is intended to serve as an access key to the lexicon-grammar and to the dictionary of compound nouns: given a word found in a text, a look-up programme extracts the different entries corresponding to this word and the lexical and syntactic properties attached to each of the entries. Then, an analyzer can determine the (single) solution present in the text by exploring the context of the word. This system has various applications to automatic processing of texts (from error-spelling correction to the more ambitious mechanical translation).

The coverage of the DELA dictionary system is comparable to that of current dictionaries, which are incomplete in various respects:

- proper names are missing, and also their derived forms (verbs, nouns, adjectives: to russify, russification, etc.),
- many forms derived from common dictionary entries are absent from all dictionaries, because they are easy to interpret when the canonical entry is known: this is the case for the word unrecommercializability, well formed and easy to interpret by means of the entry to commercialize, the prefixes re-, un- and the suffixes -able, -ability. These gaps are to be filled at the level of the lexicon-grammar, by adding new properties according to derivations such as:

[Passive sentence with modal]:	This product can be commercialized by our firm
[Adjectivization]:	This product is commercializable by our firm
[Prefixation from negation]:	This product is uncommercializable by our firm
[Nominalization]:	This product has a certain commercializability by our firm

where each of these sentence forms constitutes a column of the lexicon-grammar. This is another example (cf. 4.2.) of the syntactic treatment of derivational morphology.

#### 4.4 Adverbs

Z.S. Harris 1986 has proposed the following scheme of derivation for various adverbs, circumstantial complements and most subordinate clauses:

That Bob left occurred yesterday = Bob left yesterday

The sentential subject contains an elementary sentence, that is, with no adverb of time or of place, and the verb is a support verb for adverbs, belonging to a small class (to take place, to happen, etc.). Hence, certain families of adverbs (date, place) are going to be described as elementary sentences of the form:

That S Vsup Adverb

that can be naturally entered into the lexicon-grammar of the language. For other adverbs, the derivation can take a different form, but the principle is the same:

= That Bob entered the smoky room was courageous (of him + on his part)  
= Bob entered the smoky room courageously

where the pronoun or the possessive adjective his must refer to Bob and is the subject of courageous. This constraint does not hold in the following otherwise similar relation:

= That Bob entered the smoky room was unfortunate  
= Unfortunately Bob entered the smoky room

Adverbs or more generally non-object complements have been little studied, and their relations to elementary sentences are far from being understood (for French, cf. M. Gross 1990). As is often the case, frozen forms can be quite revealing about the nature of relations which can be hidden under apparently regular structures. We undertook a description of frozen and compound adverbs, that resulted in the classification given in table 5.

---

Table 5

---

Although this study is primarily about the internal constitution of adverbial phrases, some relations with the sentences where they occur have been studied; for example: coreference relation such as in the examples:

Bob is an active person, in his way  
In his madness, Bob hurt Jo badly

where adverbs have a status different of those in the following pair, where changes of level occur:

Bob told Max in all honesty that he should apologize to Jo  
In all honesty, Max should apologize to Jo

Other detailed studies have been made in the framework of lexicon-grammar: of time adverbials (D. Maurel 1989), of manner adverbials (C. Molinier 1982), of subordinating conjunctions (M. Piot 1978), but many more are needed in order to get a comprehensive picture of the phrases that have been called circumstantial complements, they are much more varied than the classification based on the half dozen WH interrogative pronouns might lead one to believe.

## 5. RESULTS AND CONCLUSION

The lexicon-grammar is stored in a computer data base and maintained on a regular basis: new properties are being added, and new entries are still being introduced, mainly so-called metaphors or figurative meanings of already existing sentences. Given an entry, the "+" signs that appear in the corresponding row (cf. figure 1) provide a syntactic paradigm of the entry. We can ask the following question . How many of these paradigms do we have in the lexicon-grammar? For the 10, 000 verbs, the answer is 8, 000. Hence, there are few verbs which have the same set of constructions. Even more, when we add derivational properties, then no two verbs have the same syntactic paradigm. This observation can be seen as an a posteriori justification of the lexicon-grammar: every verb has to be described on an individual basis.

We have proposed a number of criteria to separate objects from other complements (mainly adverbial). As a consequence, we observed that ordinary verbs had practically never three objects. Verbs with two objects often have a relation between both objects or between the subject and one of the objects (cf. 4.1). With frozen sentences, where often one complement (possibly an adverb) is frozen with the verb, we observe more entries with three complements, but still, they are exceptional, or at any rate much less numerous. For example, in the sentences:

Bob used (the cold)<sup>1</sup> as (a pretext)<sup>2</sup> (for leaving early)<sup>3</sup>  
The cold provides (Bob)<sup>1</sup> with (a pretext + an excuse)<sup>2</sup> (for leaving early)<sup>3</sup>

the three complements are obligatory. We have no explanation for this numerical limitation, nor for another fact we observed, that verbs with two prepositional complements are rare (cf. table 1). We consider such data important, since they have direct structural implications for sentences.

We would like to insist on the fact that in the lexicon, there are many more sentences containing a verbal expressions, that is, which are frozen (i.e. non-compositional), than sentences containing an ordinary verb (i.e. compositional). This should come as a surprise, unless the sentences thought to be compositional turned out to be less so, as further studies tend to show; then truly compositional constructions would be quite rare.

The lexicon-grammar framework together with the requirement of substantial coverage lead to a uniform and consistent description of elementary sentences. Many features have been shown to be lexical, for example the choice of the determiners and modifiers of the subject or object of a verb can be lexical, whether with ordinary or support verbs, compare:

Bob is building a palace      = Bob is building a future palace  
Bob is selling a palace      Bob is selling a future palace

\*Bob has (a + the) courage  
Bob has (a + the) courage we are all missing

Also, the distribution of adverbials appears to obey rules of cooccurrence with sentences. As a consequence, the grammar of complex sentences may be quite different from what has been proposed so far. For example, S.Y. Kuroda 1968 has shown the existence of complex dependencies between determiners and relative clauses, such phenomena could be more general and more complex, that is, more lexical. Thus, lexicon-grammar may have to be extended to non-elementary sentences, as is already the case for frozen complex sentences (M. Gross 1988).

Some of the phenomena mentioned may be language dependent. For example, in English, verb-particle combinations seem to correspond to verb-adverb combinations in French, but too little is yet known about such phenomena to draw any conclusion. One of the main advantages of lexicon-grammar is that it allows comparisons between languages both at the lexical and the syntactic levels. In fact, large segments of lexicon-grammars have already been built, for Italian (A. Elia 1984; A. Elia, E. d'Agostino, M. Martinelli 1981), for Portuguese (Ranchhod 1990; E. Maceido 1984, J. Malaca-Casteleiro, 1981), for Spanish (B. Lamiroy 1983, L. Masso-Pellat 1990, C. Subirats 1987), which already allows for comparisons of Romance languages.

Many classes of constructions have also been described for English (P. Freckleton 1985; P. Machonis 1988; M. Salkoff 1983), for German (F. Caroli 1984; T. Treigg 1977), for Arabic (M. Chad 1988; M. El Hannach 1988), for Korean (Hong Chai-Sîng 1984). All this work confirms the possibility of constructing lexicon-grammars of varied languages in a comparable way.

As mentioned above, the evolution of lexicon-grammar is towards increased structure, because derivational properties are being added. As a consequence, there is a tendency to use roots of words as entries. In fact, this is a tendency to render descriptions more abstract: parts of speech become equivalent and thus lose significance inside a dictionary article. It is one of the features that make feasible comparisons between languages.

## REFERENCES

Boons, Jean-Paul; Alain, Guillet; Christian, Leclère 1976a. La structure des phrases simples en français. I Constructions intransitives, Geneva: Droz, 377p.

Boons, Jean-Paul; Alain, Guillet; Christian Leclère 1976b. La structure des phrases simples en français, II Constructions transitives, Paris: Rapport de recherches du LADL, No6, 85p., tables et index, 58p.

Caroli, Folker 1984. Les verbes transitifs À complément de lieu en allemand, *Linguisticae Investigationes*, Amsterdam-Philadelphia: J. Benjamins B.V. 8.2:225-267.

Chad, Mohammed 1988. SystÈme verbal arabe. RÈgime des constructions transitives, Doctoral Thesis, Université Paris7: LADL.

Courtois, Blandine 1989-1990. Dictionnaire Électronique du LADL pour les mots simples du français. DELAS V06/2, Rapport technique du LADL No20, Université Paris7: LADL, 458p.

Courtois, Blandine 1990. Un système de dictionnaires électroniques pour les mots simples du français, Langue française No87: 11-22, Paris: Larousse.

Courtois, Blandine et Silberztein, Max eds. 1990. Dictionnaires Électroniques du français, Langue française No87, Paris: Larousse.

Danlos, Laurence 1980. ReprÉsentation d'informations linguistiques: les constructions être Prép X, Doctoral Thesis, Université Paris7: LADL.

El Hannach, Mohammed 1988. Syntaxe des verbes psychologiques de l'arabe, Doctoral Thesis, Université Paris-7: LADL.

Elia, Annibale 1984. Le verbe italien, Les complétives dans les phrases à un complément, Bari-Paris: Schena-Nizet, 305p.

Elia, Annibale, Emilio D'Agostino, Maurizio Martinelli 1981. Lessico e strutture sintattiche. Introduzione alla sintassi del verbo italiano, Naples: Liguori.

Freckleton, Peter 1985. Une comparaison des expressions de l'anglais et du français. Doctoral Thesis, Université Paris7: LADL.

Giry-Schneider, Jacqueline 1978. Les nominalisations en français. L'opérateur FAIRE dans le lexique, Geneva: Droz, 353p., tables: 61p.

Giry-Schneider, Jacqueline 1987. Etude de prédicats nominaux en français. Les constructions faire N, Geneva: Droz, 399p.

Gross, Gaston 1988. Degré de figement des noms composés, Langages No 90:57-72, Paris: Larousse.

Gross, Gaston 1989. Les constructions converses du français, Geneva: Droz, 513p.

Gross, Maurice 1968. Grammaire transformationnelle du français. 1-Syntaxe du verbe, Paris: CantilÈne, 188p.

Gross, Maurice 1975. Méthodes en syntaxe, Paris: Hermann, 412p.

Gross, Maurice 1981. Les bases empiriques de la notion de prédicat sémantique, Formes syntaxiques et prédicats sémantiques, A. Guillet et C. Leclère Éd., Langages, No63: 7-52, Paris: Larousse.

Gross, Maurice 1982. Une classification des phrases figées du français, Revue québécoise de linguistique, Vol. 11, No2: 151-185, Montreal: Presses de l'Université du Québec à Montréal.

Gross, Maurice 1988. In Gross, G. and Piot, M. eds. *Syntaxe des connecteurs*, Langue française No77: 151-185, Paris: Larousse.

Gross, Maurice 1990. *Grammaire transformationnelle du français. 3-Syntaxe de l'adverbe*, Paris: ASSTRIL, 670p.

Guillet, Alain; Christian Leclère 1991. *La structure des phrases simples en français. Verbes à complément direct et complément locatif*, Geneva: Droz.

Harris, Zellig S. 1952. *Discourse Analysis*, Language 28: 1-30, Baltimore: The Waverly Press.

Harris, Zellig 1964. *Elementary Transformations*, Philadelphie: University of Pennsylvania, TDAP No54, réimprimé in *Papers in Structural and Transformational Linguistics*, 1970, Reidel: Dordrecht.

Harris, Zellig 1968. *Mathematical Structures of Language*, New York: Wiley, 230p.

Hong, Chaâ-sîng 1984. *La classe des verbes de mouvement en coréen contemporain*, *Lingvisticae Investigationes Supplementa*, Amsterdam-Philadelphia: J.Benjamins B.V., 309p.

Kuroda, Sige-Yuki 1968. *English Relativization and Certain Related Problems*, Language 44:2-1, Baltimore: The Waverly Press.

Labelle, Jacques 1974. *Etude de constructions avec opÉrateur AVOIR (nominalisations et extensions)*, Université Paris VIII: LADL, Doctoral Thesis, 315p.

Lamiroy, Béatrice 1983. *Les verbes de mouvement en français et en espagnol. Etude de syntaxe comparée de leurs infinitives*, *Lingvisticae Investigationes Supplementa*, Amsterdam-Philadelphia: J.Benjamins B.V., 309p

Laporte, Eric 1988. *Méthodes algorithmiques et lexicales de phonétisation du français*, Doctoral Thesis, Université Paris7: LADL, 162p., annexes.

LeclÈre, Christian 1990. *Organisation du lexique-grammaire des verbes français*, Langue française No87, Paris: Larousse.

Maceido-Oliveira, Elisa 1984. *Syntaxe des verbes psychologiques du portugais*, Instituto Nacional de Investigaçào C\_entífica, Centro de L\_inguística da Universidade de Lisboa: Lisbon, 198p.

Machonis, A. Peter 1988. *Support Verbs: An Analysis of be Prep X idioms*, *The SECOL Review*, 12.2:95-125.

Meunier, Annie 1977. *Sur les bases syntaxiques de la morphologie dérivationnelle*, *Lingvisticae Investigationes*, Vol.I, No2: 287-332, Amsterdam-Philadelphia: J. Benjamins B.V.

Meunier, Annie 1981. Nominalisation d'adjectifs par verbes supports, Doctoral Thesis, Université Paris 7: LADL, 282p.

Molinier, Christian 1982. Etude syntaxique et sémantique des adverbes en -ment, Doctoral Thesis, Université de Toulouse-Le Mirail, 416p.

Négroni-Peyre, Dominique de 1978. Nominalisation par être en et réflexivation (admiration, opposition, révolte et rage), *Lingvisticae Investigationes*, Vol. II, No1, Amsterdam-Philadelphia: J.Benjamins B.V., pp.127-164.

Pellat-Masso, Luisa 1990. Une description formelle des expressions figées de l'espagnol, *Mémoires du CERIL* 5:22-290.

Piot, Mireille 1978. Etudes transformationnelles de quelques classes de conjonctions de subordination, Doctoral Thesis, Université Paris7: LADL.

Ranchhod, Elisabete 1983. On the Support Verbs ser and estar in Portuguese, *Lingvisticae Investigationes*, Vol.VII, No 2: 317-353, Amsterdam-Philadelphia: J. Benjamins B.V.

Ranchhod, Elisabete 1990. *Sintaxe dos Predicados Nominais com Estar*, Instituto Nacional de Investigação Científica, Centro de Linguística da Universidade de Lisboa: Lisbon, 477p.

Salkoff, Morris 1983. Bees are swarming in the garden, *Language*, Vol.59, No2: 288-346, Baltimore: The Waverly Press.

Silberztein, Max 1989. Dictionnaires électroniques et analyse lexicale automatique, Doctoral Thesis, Université Paris7, Paris: LADL.

Silberztein, Max 1990. Le dictionnaire électronique des mots composés, *Langue française* No87:71-83, Paris: Larousse.

Subirats-Ruddeberg, Carlos 1987. Sentential Complementation in Spanish, A lexicogrammatical study of three classes of verbs, *Lingvisticae Investigationes Supplementa*, No 14, Amsterdam-Philadelphia: J.Benjamins Pub. Co., 290p.

Treig, Thomas 1977. Complétives en allemand. Classification, Rapport de recherches du LADL, 7:39-203.

Vivès, Robert 1983. Avoir, prendre, perdre: constructions à verbe support et extensions aspectuelles, Doctoral Thesis, Université Paris VIII, LADL, 194p., tables: 40p.