The Argument Structure of Elementary Sentences
Maurice Gross

To cite this version:

HAL Id: halshs-00278304
https://halshs.archives-ouvertes.fr/halshs-00278304
Submitted on 11 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
We present here a global view of the syntactic shapes of the elementary sentences of French. This paper constitutes a synthesis of the numerous observations made on a set of about 12,000 verbs studied in the framework of lexicon-grammar (M. Gross 1975; J.-P. Boons, A. Guillet, C. Leclère 1976; G. Gross 1989; A. Guillet, C. Leclère 1992). The point of view is strictly formal, according to the now classical methodology developed by Z.S. Harris and N. Chomsky. No semantic notion is involved in the descriptive apparatus; in other terms, the metalanguage of the description is built from combinatorial notions applied to sets of words. This metalanguage is entirely derived from the basic concepts that emerged out of Z.S. Harris' work in syntax. Moreover, the approach is systematic, namely for a given phenomenon, one has always attempted to reach a complete coverage of the description in a language.

The empirical basis of syntactic studies is the intuitive recognition that some sequences of words have a distinguished status which has been expressed by the concept of sentence. Thus, the sequence:

This solution pleases a large number of her friends

is perceived as a sentence. It is also the case for the famous examples of Chomsky and Tesnière:

*Colorless ideas dream furiously*

*Le silence vertébral indispose la voile licite*

whose absence of meaning is due to the choice of individual words, a calculated choice that does not allow any consensual interpretation; but both sequences are clearly felt as having a regular syntactic structure. It is this structure which allows us to pronounce them with a smooth intonation and to memorize them easily, whereas the following strings of the same words are pronounceable only as lists of words and are quite difficult to memorize:
On the other hand, sequences of words such as:

- *a large number of her friends*
- *inside the house*
- *as large as a postcard*

are not perceived as sentences. If we can name them as noun phrases or adjectival phrases, it is the result of a thorough grammatical education that led us to analyze them so in the metalanguage of a consciously learned grammar.

The syntactic study of French sentences has a history of several centuries; slowly it has evolved and brought to light a certain number of concepts on which all linguists agree. We recall the main regularities.

To-day, all sentences have the shape:

*Subject-Verb-Complements*

We note this shape:

(1) \[ N_0 \ V \ W \]

where \(W\) is a variable ranging over all complements including an empty one.

Such a formula has various implications we shall now examine.

1. The subject

   The statement:

   A. All sentences have a subject

   is largely verified in French and in English. It is nonetheless worthwhile to remember that a certain number of analyses have to be performed in order to reach the situation A:

   - sentences in the imperative form such as:
     
     - *Put this book on the table*
       
     - *(Pose ce livre sur la table)*
     
     - *Let him put this book on the table*
       
     - *(Qu'il pose ce livre sur la table)*

   are analyzed (M. Gross 1968) by zeroing a sequence such as:

   - *(I request that you = E)* put this book on the table
     
   - *(J'exige que tu = E)* pose(s) ce livre sur la table)
Hence imperative forms do have a basic form with overt subject;
- other zeroing operations of verbs are used to account for the strong intuition of sentence which is triggered by some non verbal sequences, these operations may have a wide range of application, for example the reductions:

- other reductions are restricted, appropriate in Z.S. Harris' terminology:

- sentences or phrases such as:

do not have their subject NO to the left of the verb, but permutation rules relate them to forms that are basic in this respect:

We could point out numerous examples of this type, they are not counter-examples to statement A. But there are also genuine exceptions, frozen sentences such as:

cannot be any longer analyzed by some permutation rule applied for example to:

even in case these regularized forms happen to be their correct etymological source. In the same way, it is difficult to analyze by zeroing the following utterances to which the intuition of full sentence is clearly attached:

True exceptions are not numerous, we have represented for French a few hundred
common ones in the syntactic table EC0 of the lexicon-grammar of French.

2. The complement sequence

The part $N_0 V$ of the structure $N_0 V W$ is then of a great generality. It is not the case for the rest of the structure: $W$, that raises numerous questions stemming from the observation that practically no two verbs of the lexicon (12,000 verbs for French) have the same complements $W$.

In order to clarify the nature of $W$, grammarians traditionally have classified the complements in two main types: object or essential complements that are characteristic of each verb and circumstantial complements that may apply to large sets of verbs and can often be omitted. Both types of complements can take the shape of noun phrases, direct or prepositional, they are noted: $Prep N_i$, where the subscript $i$ indicates their left to right order of occurrence in the sentence, the preposition $Prep$ can be ‘zero’, it is then noted $E$. For French we write:

$$Prep =: E + à + de + dans + sur + pour + etc.$$  

But complements can also be sentential, in which case we write:

$$Prep N_i =: Prep (E + ce) Qu P$$

to outline their content and to indicate that they nonetheless have some of the properties of the ordinary noun phrase. Sentential complements may belong to the type object or they can be circumstantial, in which case they are called subordinate clauses.

This traditional analysis is well motivated but often lacks precision. Among many questions is the fact that one encounters numerous ambiguities that prevent one from distinguishing the various types. For example, circumstantial complements are often subclassified into Time, Place or Manner complements, and these semantic attributes are presented as characteristic of circumstantial complements, but various essential complements and some subjects appear to have these attributes. For example, in the sentences:

La pluie a duré pendant six heures The rain lasted for six hours  
Jo vit en Iran Jo lives in Iran  
Jo se comporte de façon étrange Jo behaves in a strange way

the complements of Time, Place and Manner are essential, whereas in the following sentences they are circumstantial:

Jo a dormi pendant six heures Jo slept for six hours  
Jo a mangé du bon caviar en Iran Jo ate good caviar in Iran  
Jo mange de façon étrange Jo eats in a strange way

In the sentences:

This hotel swarms with Congressmen
This hotel accommodates one thousand people

Bob crossed the lobby

Ten minutes are enough to do it
Bob took ten minutes to do it

the subjects or direct objects are, at least semantically, Place or Time arguments.

There are new methods to cope with such difficulties, we will develop for this purpose Z.S. Harris' theory of support verbs that distinguishes types of verbs that are functionally different (cf. below 4, 5.). Generally speaking, only a thorough description of individual verbs can separate the various types of complements. It is a study of this nature that has been performed on verbs at the Laboratoire d'Automatique Documentaire et Linguistique. A set of 6000 verbs (i.e. 6000 infinitive entries of common dictionaries) was retained and studied. First, semantic distinction led to consider 12000 verbal units instead of 6000. For example the verb voler (one of the 6000 verbs) must be subdivided into two units: voler (to fly) and voler (to steal), which allows a syntactic description of the complement structure:

\[
\begin{align*}
N0 \text{ voler} & =: \text{L'oiseau vole} \quad \text{(The bird is flying)} \\
N0 \text{ voler }N1 \text{ à }N2 & =: \text{Bob a volé un livre à Jo} \quad \text{(Bob stole a book from Jo)}
\end{align*}
\]

in other terms, we have \( W =: E \) for voler-to fly and \( W =: N1 \text{ à }N2 \) for voler-to steal.

The study resulted in a lexicon-grammar of French verbs, namely a set of detailed syntactic tables for the 12000 verbs. Several empirical results derived from this study help us make more precise the variable \( W \).

First the number of essential complements is limited to 2, in other terms one only observes the structures:

\[
\begin{align*}
\bar{W} & =: E \\
W & =: \text{Prep }N1 \\
W & =: \text{Prep }N1 \text{ Prep }N2 \quad \text{(Prep can be }E) \\
\end{align*}
\]

Longer structures:

\( W =: \text{Prep }N1 \text{ Prep }N2 \text{ Prep }N3 \)

are quite rare. The few possible examples are always difficult to analyze, this is the case for the verb parier-to bet in the form:

\[
\begin{align*}
\text{(Bob)0 a parié (dix francs)1 (avec Jo)2 (que Rod viendrait)3} \\
\text{(Bob)0 has bet (ten francs)1 (with Jo)2 (that Rod would come)3}
\end{align*}
\]

where the complements N1 et N3 have some of the features of direct objects, among others passive forms:

\[
\text{Dix francs ont été pariés par Bob que Rod viendrait}
\]
(Ten francs were bet by Bob that Rod would come)

Que Rod viendrait a été parié par Bob avec Jo_
(That Rod would come was bet by Bob with Jo)

No complement sequence of length 4 has been observed so far, the only example we have is the frozen sentence:

(Bob)0 tournera (sept fois)1 (sa langue)2 (dans sa bouche)3 (avant de répondre à Jo)4
(Bob will turn his tongue in his mouth seven times before he answers Jo)

More generally, we have mainly observed sequences of 3 complements when one of them was frozen (cf. the syntactic tables of frozen sentences CPPN, CPPQ, M. Gross 1982).

The possible shapes of $W$ are constrained by the following general observations:

- the preposition 'zero' (i.e. $\text{Prep} =: E$) is the most common and is observed in the two structures:

  $N_O \ V \ N_1$
  $N_O \ V \ N_1 \ \text{Prep} \ N_2$ where $\text{Prep}$ is here different from 'zero'

- structures:

  $N_O \ V \ \text{Prep} \ N_1 \ \text{Prep} \ N_2$ where both $\text{Preps}$ are here different from 'zero'

are rather rare. For example, we found only one example of the structure:

$N_O \ V \ de \ N_1 \ de \ N_2$

=: Bob a hérité (d'une maison)1 (de sa mère)2
(Bob inherited a house from his mother)

and even there, purists recommand to avoid the use of $de$ in the first complement.

The global view we just outlined provides a description of the complexity of each verb, since the number of arguments is a measure of this complexity. However, various linguistic phenomena lead us to correct this view. A first correction will be brought by the study of the content of the arguments of the verbs.

3. The content of arguments

The number and the nature of the arguments depends on each verb. On the whole, the variety of the arguments has turned out to be enormous, but it is possible to create a typology for them, although approximative in some cases. We now present this typology:
(i) Frozen arguments

Some arguments are frozen together with the verb, as in the idiomatic sentences:

(Jo)0 took (the bull)1 by (the horns)2
(Jo)0 a pris (le taureau)1 par (les cornes)2

where two arguments are frozen. The sentence:

Jo a tenu compte de l'intervention de Bob
Jo took into account Bob's intervention

will be analyzed as follows in a first approximation:

N0 V N1 Prep N2 =:
(Jo)0 a tenu (compte)1 de (l'intervention de Bob)2
(Jo)0 took into (account)1 (Bob's intervention)2

Let us now specify the arguments, in order to specify the first complement we write:

N1 =: C1 =: compte, account

The notation C for constant or frozen argument is used in all syntactic positions, that is, C can be subscripted by i ranging from 0 to 4. For the free arguments in positions 0 and 2 we write in the same way:

N0 =: Jo
N2 =: l'intervention de Bob, Bob's intervention

(ii) Free concrete arguments

By concrete nouns, we mean nouns referring to concrete items and which are selected by the verb. For example in the sentence:

N0 V N1 =: Jo mange du pain, Jo eats bread

the verb selects animate or human subjects in N0, and concrete food direct objects in N1.

(iii) Sentential arguments

The preceding sentence form accepts a sentential argument as in:

(Jo)0 a tenu (compte)1 de (ce que Bob interviendrait)2
(Jo)0 took into (account)1 (the fact that Bob intervene)2

we then write symbolically:
N2 := Qu S

S is for sentence, Qu is a subordinating conjunction or a complementizer. More precisely, we have here:

N2 := ce que S + le fait que S
N2 := the fact that S

One question arises immediately: in this classification of argument contents, what is the status of the nouns which are not concrete, that is, where do the nouns appear which we call intuitively abstract? Our example can be used to clarify this point. We have in fact observed:

N2 := l'intervention de Bob + (le fait + ce) que Bob interviendrait
N2 := Bob's intervention + the fact that Bob would intervene

These two specifications of the argument N2 appear to be related, at the same time one is sentential and the other is an abstract noun. The relation is in fact a syntactic one, quite general and which presents various regularities. The noun phrase is derived from the sentence by a nominalization rule involving the notion of support verb (Z.S. Harris 1964, A. Meunier 1977; D. de Négroni 1978; J. Giry-Schneider 1978, 1987; M. Gross 1981; R. Vivès 1983). We can illustrate the relation by means of the following derivation:

Bob intervient      Bob intervened
[Nominalization]  = Bob fait une intervention  = Bob made an intervention
[Relativization]  = l'intervention que fait Bob    = the intervention that Bob made
[Vsup Reduction]  = l'intervention de Bob          = Bob's intervention

This type of transformational rule relates verbs, adjectives and nouns at the level of elementary sentences. For example, we could add to the previous derivation _ lines such as:

[Nominalization]  = Bob est un intervenant       = Bob is an intervener
[Adjectivization] = Bob est interventionniste     = Bob is interventionist

More generally, we have observed that nouns intuitively labelled abstract always enter into elementary sentences with support verbs, independently of possible derivational relations, this the case for the sentences:

*Bob a discours      ?Bob spoke
Bob a fait (une allocation + un discours)  Bob made a speech
*Bob est un discoureur  Bob is a speaker

?Bob talked
Bob gave a talk
?Bob is a talker

*Bob a conférence    = Bob lectured
Bob (fait + donne) une conférence  = Bob gave a lecture
Bob est un conférencier = Bob is a lecturer

Moreover, we can see that some concrete nouns, for example human nouns, also enter into sentences with specific support verbs.

As a consequence of this discussion, we assimilate abstract arguments to sentential ones, but the distinction between abstract and concrete nouns will have to be further refined, since many concrete nouns will have to be treated as abstract ones, at least in certain syntactic positions. Nonetheless one situation should be clear: noun phrases such as Bob's lecture or Bob's intervention which are derived from sentences with support verbs and which can be labelled as abstract are to be considered as sentential. They occur in combination with selectional verbs in any syntactic position where selection of abstract nouns is possible.

To sum up the discussion, we have the following typology of structures and arguments:

\[
N_0 \, V \, (E + \text{Prep} \, N_1 \, (E + \text{Prep} \, N_2 \, (E + \text{Prep} \, N_3)))
\]

\[
N_i =: C + N + \text{Qu} \, P
\]

We will make it more precise, according to observations made in a systematic way for French.

Even at this level of description, several important applications have been realized. One example is the classification of verbs. Just by specifying W we have been able to design a system of about 50 disjoint classes for the 12000 free sentences and of about 30 classes for about 30000 frozen sentences which have been described so far (C. Leclère 1990). Another example is the treatment of families of sentences which intersect the two cases: free and frozen. Consider the following examples:

(1) Jo a (loupé + manqué + raté) le coche
Jo missed the boat

they are clearly frozen: they do not accept any other determiner than the definite article, no plural for coche-boat is allowed, no modifier (adjective, etc.) is accepted by these nouns. On the other hand, the sentences:

(2) Jo a (loupé + manqué + raté) une (occasion + opportunité)
Jo missed (an excellent opportunity + a chance to come back)

are free: the nouns are selected by the verbs, they can be modified in a general way, etc. Clearly (1) and (2) belong to one and the same family of sentences. We mentioned above that we are using the equation:

\[
N_i =: C + N + \text{Qu} \, P
\]

to construct disjoint classes, this same equation shows that in a given syntactic position Ni, one can find phrases that are either frozen, or nominal and free, or
sentential. This is exactly what is happening in (1) and (2). Our requirement that classes be disjoint is a mere convenience which may however introduce some distortions here, since (1) and (2) will be separated in distinct classes, but the reality can be described in a natural way on the basis of our general principles.

We presented a measure of complexity for argument structures, this measure has to be corrected in two ways:

First we have to take into account the nature of the arguments:

- if an argument is frozen it does not count, thus the sentences:
  
  Jo took the bull by the horns
  Jo a pris le taureau par les cornes

  are from a semantic point of view sentences with one argument, as can be seen from the approximate paraphrases:

  Jo acted
  Jo a agi

  Second, we must take into account the nature of the verb: normal or support verb. Nominalization relations such as:

  Bob walked = Bob took a walk

change the number of arguments of sentences without changing their basic meaning. Support verbs are roughly grammatical constants without basic semantic content, they only carry modalities that slightly modify the basic sense of the sentence: aspect, negation, intensity, etc. Hence, counting arguments becomes an ambiguous operation: Do we count the essential noun phrases attached to a given selctional verb or do we count the phrases attached to support verbs? Sentences with support verbs are more explicit with respect to meaning but they are not always available.

4. Modifications of structures by transformations

A transformation such as Passive:

\[ \text{N0 V N1 = N1 be Vpp by N0} \]

leaves invariant both the meaning and the number of arguments. But certain transformations can modify the number of noun phrases attached to a verb, thus introducing a difficulty in the counting process, we just saw that with Nominalizations. Let us discuss this case in more detail. Consider the following nominalization relations:

(1) Bob argues with Jo
(1b) = Bob (is in + enters into) an argument with Jo
(1h) = Bob (has + initiates) an argument with Jo
(1t) = There is an argument between Bob and Jo
(2) Bob reviewed her book  
(2m) = Bob (made + wrote) a review of her book

(3) Bob is nasty  
(3h) = Bob has a certain nastyness  
(3bp) = Bob is of a certain nastyness

(4) The troops attacked the fort  
(4m) = The troops mounted an attack against the fort

As previously observed, introducing a support verb through a nominalization relation can increase the number of arguments. Again the problem arises from the difficulty of equating the number of noun phrases with the number of semantic arguments. But the example (4) raises a new problem. Let us compare (4m) and:

(4d) The troops watched an attack against the fort

Although these two sentences are superficially identical, they differ semantically: (4d) is more complex since it could be expanded into:

(5) = The troops watched an attack of their enemies against the fort

In fact (5) is a complex sentence that must be analyzed as including two elementary sentences: (4) and The troops watched N, where to watch is a selectional verb. But (4m) and (4d) differ in other respects, if we attempt to determine the different noun phrases they include, for example by means of the clefting operation, we observe two complements in (4):

(4m) = It is against the fort that the troops mounted an attack  
      = It is an attack that the troops mounted against the fort

and only one in (4d):

(4d) = *It is an attack that the troops watched against the fort  
       = *It is against the fort that the troops watched an attack  
       = It is an attack against the fort that the troops watched

Thus, this syntactic analysis is in conflict with the semantic interpretation: the elementary sentence has three arguments whereas the complex one has only two. As a matter of fact the situation is even more complicated since (4m) can also be clefted in the following way:

(4m) = It is an attack against the fort that the troops mounted

Hence, the enumeration of the arguments cannot be based on a naive counting of the number of noun phrases. Another situation involving frozen sentences leads to the same conclusion. Let us consider the sentence:

A flash of anger crossed Bob’s eyes
it is built from two noun phrases:

   (A flash of anger)0 crossed (Bob's eyes)1

but the two semantic arguments are not the noun phrases, they are anger and Bob linked by a relation that can be expressed by a support verb:

   Bob has a certain anger

Remark

In all of our examples, the supported noun phrase appeared in a complement position. Although frequent, this situation is not the only possibility and one observes supported noun phrases in the subject position:

   Anger overwhelmed Bob

   = Bob emphasized his results
      = The emphasis is on the results

   Another example of transformation that modifies the number of noun phrases without changing the meaning is the Restructuration operation (M. Gross 1977; A. Guillet, C. Leclère 1981):

   (Dozens of guests)0 are leaving
   = (Guests)0 are leaving (by the dozens)1

   This company is buying up (the stores of our street)1
   = This company is buying up (the stores)1 (in our street)2

5. Adverbs

   The analysis of adverbs proposed by Z.S. Harris 1976 also modifies current views about the notion of argument of a sentence. Consider the following two sentences:

   (1) Jo arrived
   (2) It occurred at noon

   they constitute a discourse, when the pronoun it is interpreted as referring to the first sentence. This leads us to consider that the verb to occur allows sentential subjects. We then write:

   (2a) (That Jo arrived)0 occurred (at noon)2

   Now, starting from the base discourse:

   (1)(2a) =: Jo arrived. That Jo arrived occurred at noon
we pronominalize the subject of to occur in a context where we have a duplication of arguments:

Jo arrived, (which + it) occurred at noon

and we introduce an operation that erases a sequence such as (which + it) occurred. We then obtain:

(3) Jo arrived at noon

Z.S. Harris has motivated this analysis on the following grounds:

- the verb to occur is a support verb, it does not have the usual selectional properties, the supported noun phrases are adverbs, circumstantial complements or subordinated phrases, all functionally equivalent,
- the sentence Jo arrived is elementary, adverbials do not belong to it, they are introduced from other elementary sentences through similar processes.

There are other types of introduction of adverbials, but the basic process is the same, consider for example the sentence:

Jo arrived in a hurry

the adverbial complement in a hurry which is felt as bearing on Jo is introduced by a similar derivation:

(1) Jo arrived
(4) Jo was in a hurry
(1)(4) = Jo arrived, he was in a hurry

and the zeroed sequence will be here he was, also a pronoun bound to an antecedent together with a support verb. The sentence:

(5) Jo arrived in poor shape

will be analyzed in the same way, that is in terms of the two sentences:

Jo arrived
Jo was in poor shape

But consider now the sentence:

(5a) All the people in poor shape arrived late

it has roughly the same semantic content as the preceding sentence, but its analysis will differ in the following way:

(5a) = (6) The people that were in poor shape arrived late

- the sentence People were in poor shape is attached to a noun by a Relativization operation and this rule applies in other syntactic positions, for example in the object
Jo bought a book in poor shape

- the adverb late is introduced by means of the sentence with support verb It occurred late.

We observe that the phrase in poor shape is supported in a common way in the basic form with support verb to be but it may have different functions according to the way it is introduced in more complex sentences: it is an adverbial phrase in (5), it is a noun modifier in (6). We have then three different functions for the same noun phrase:

- a basic function that we introduced: the function supported phrase,
- an adverbial function that can be further refined into subfunctions such as adverbial bearing on a phrase or on a sentence, etc. In all cases, the adverbial phrase has the characteristic property of being movable at any phrase boundary of the sentence structure to which it belongs,
- a modifier function where the supported phrase is an epithet of a noun that cannot be moved out its noun phrase.

Notice that the zeroing operation which applies to the relative clause source of the epithet also reduces a pronoun (that) and the support verb to be.

In conclusion, we think that through a reanalysis of common concepts of traditional grammars, we have considerably gained in precision and in coherence, both at the theoretical level and at the descriptive level. The only new concept introduced is the distinction between selectional verbs and support verbs, but its empirical motivation is beyond discussion.

REFERENCES


Both passive forms can hardly accept the complement sequence of length 3. The arguments are the essential complements and the subject. Provided a finer separation is made between the different meanings of intervention, etc.